ELSEVIER
Available online at www.sciencedirect.com
SciVerse ScienceDirect
Procedia - Social and Behavioral Sciences 39 (2012) 450 – 462
Procedia
Social and Behavioral Sciences

The Seventh International Conference on City Logistics

GPS data analysis for understanding urban goods movement

Pascal Pluvinet[a], Jesus Gonzalez-Feliu[a*], Christian Ambrosini[a]

[a]*Laboratoire d'Economie des Transports, 14 avenue Berthelot 69007 Lyon, France*

Abstract

This paper aims to investigate the contribution of GPS survey techniques for urban goods movement characterization and diagnosis, more precisely the implementation and application issues related to the introduction of real-time data transmission procedures and phone tools with integrated GPS devices. We propose a GPS-based data collection method for urban freight route characterization using a Smartphone application. After testing and calibrating the data processing tool, we analyze the main results on a baseline of about 900 rounds, with the R software. This analysis allows us to define the characteristics of the overall routes as well as the environmental impacts linked with the categories of roads: urban highways, main roads and residential streets. Moreover, the study shows that the environmental behavior of the driver is connected with the main activity of the carriers. The complementarity between GPS and traditional urban freight surveys is finally discussed.

Keywords: Urban goods movement; survey techniques; deliveries; GPS

1. Introduction

Urban freight is one of the most challenging aspects of urban planning. In the last decades, several data collection methods have been developed, and many models derived from these methods have been proposed in literature [2][14]. Classical surveys are the most popular methods, as the data collected by this means are a good information support and can have many uses, as for example: modeling, quality and performance evaluation, descriptive statistics, decision support, transport operator feedback and marketing issues [1][3]. However, traditional surveys have their limits, and ensuring the quality of data

* Corresponding author: Tel.: +33-0-47272 6444; fax: +33-0-472726448.
E-mail address: jesus.gonzales-feliu@let.ish-lyon.cnrs.fr

1877-0428 © 2012 Published by Elsevier Ltd. Selection and/or peer-review under responsibility of 7th International Conference on City Logistics
Open access under CC BY-NC-ND license. doi:10.1016/j.sbspro.2012.03.121

implies an increasing cost both in time and money. Another type of method is related to a GPS-based data collection, which is an appropriate way to make automatic data collection for vehicle routes and stops, but does not give a global view of the freight flows and operations [1][2]. Although very popular in person transport, this second category of methods has not been yet much developed for urban goods and only preliminary studies have been recently carried out [7][12].

Concerning GPS data recording, several works deal with this technique in person transport. We have to note that the goals of GPS data collection are not always similar, and three main uses can be identified. The first one allows providing information on probing vehicles to test a technology, or a methodology in order to track the vehicles and collect all wished information. These works usually deal with few vehicles and a small quantity of data. The second one concerns data collection for model calibration and testing. In this case, very popular in city logistics, the number of vehicles can be limited and the data obtained is easily identifiable and analyzable. The third one is less used and related to data collection surveys for UGM characterization. This last category needs an adaptation to the measured object (not necessarily the route but sometimes directly related to pickup/delivery operations) and has to be connected with some individuals (mainly the companies or the nature of freight). Moreover, the quantity of data is important (route surveys can contain thousands of routes).

Traditional surveys are also implemented for route characterization [3]. They involve thousands of routes but the non-response rates and the quantity of incomplete questionnaires can be important (about 15% of questionnaires filled-up correctly [10])

Therefore, the aim of this paper is to study the potential of GPS-based data collection in urban goods transport and discuss the ways of substituting and/or complementing traditional surveys. Firstly, an overview of main survey methods for urban goods understanding is presented. Secondly, a first analysis about benefits and drawbacks of GPS surveys is carried out. Therefore, a methodology of GPS data collection for urban goods in medium urban areas is proposed. After analyzing the surveyed data, we propose a set of indicators, and make a comparison between routes surveyed with and without GPS that completes the preliminary findings. In conclusion, the possible cases for substitution or complementarity of GPS and classical surveys are addressed.

2. Overview and motivation

Many policy makers are reliant on urban goods movement surveys to make a diagnosis and determine policy approaches for urban freight transport [1]. This provides a little insight into relevant factors like goods and service flows, the specific purposes of commercial vehicle trips, the involved establishment and the supply chain decisions related to these trips, among others.

Traditionally policy makers consider traffic counts as the only surveying method for urban freight policy-making support. In the last years, many urban freight transport studies attempted to go beyond vehicle traffic counts [1], and several types of surveys have been proposed. Urban Goods Movement (UGM) data can be collected using different techniques that can be summarized into the following list of data collection techniques: - Establishment survey; - Commodity flow survey; - Freight operator survey; - Driver survey; - Roadside interview survey; - Vehicle observation survey; - Parking survey; - Vehicle trip diaries; - GPS survey; - Suppliers survey; - Service provider survey.

In addition, vehicle traffic counts are commonly used in conjunction with the above techniques to give complementary information. Among them, two main types of surveys can be used for route reconstruction. Driver surveys are used to gather data about the driver's overall trip pattern, as well as information about pick-up and delivery operations where the survey takes place. They are usually conducted face to face at the establishments that receive pickups/deliveries, with driver interviewed after carrying out work before they drive away. They allow collecting information like: time spent, loading and parking locations or methods for moving the goods from vehicle, among others. Vehicle trip diaries are used to collect detailed information about the activities of a single vehicle, usually over a single day or a few days. They provide data about the exact locations served, route details, arrival and departure times, time spent for pickup and delivery operations or type of goods, among others.

In this section, we first try to complete the works of Allen and Browne [1] in order to see how GPS-based surveys and GIS representation tools may be generalized and combined with classical surveys or even substituting them. To do this, we take into account several criteria such as the total survey costs, the data quality (routes, delivery points …) or more accurate measures concerning precise paths, instantaneous speeds and accelerations. Secondly, we go on with identifying the GPS possibilities in order to estimate more accurately a set of data about routes. Moreover, we will also analyze the limits of GPS uses. It is important to note that collected data in city logistics planning can feed behavioral models to help public and private decision-makers [2].

GPS data collection can also be an alternative to these surveys. Although many researches have been made in the field of personal trip GPS data collection, the application of these techniques to UGM has been less extended [1]. As seen above, these techniques have been used in many cases on small data sets for model calibration, but few UGM surveys have taken into account GPS data collection methods with the aim of doing a generalization. We briefly present these works below.

2.1. A brief state of the art

In Canada, McCabe *et al.* [12] present the main results of a pilot Commercial Vehicle Survey for the Region of Peel, located just west side of the City of Toronto. This survey consisted in collecting commodity, modal choice, and commercial vehicle movement data from a sample of approximately 600 shippers and a sample of their drivers. The work presents two complementary techniques used: a mail-out/ mail-back survey and a combined mail-out/ mail-back survey with a GPS-supplement. The main objectives of the survey was to provide specific data about shipments, trips and rounds for Peel region, as well as to describe the behavioral and economic processes related to the movement of commercial vehicles. Moreover, it aimed at acting as a pilot for a large-scale shipper-based survey of commercial vehicle movements.

The survey was conducted from October 2006 to May 2007 and 597 completed shipper surveys were obtained (response rate of 25.3 %), providing detailed information on trip generation rates, firm characteristics and inbound and outbound shipments. 86 driver surveys were returned and a sub-sample of 42 GPS surveys were conducted by installing GPS units for seven days in the vehicles of the participating trucks, in order to track the routes taken, identify stops and collect engine information (as idle time, engine load, fuel consumption and incidence of rapid decelerations). The route tracker was placed on the dashboard of the vehicle and recorded time-stamped locations, engine data for the vehicle every few seconds, and stored those data points every 500 m and 5 min, or when the vehicle stopped for at least 5 min.

The two techniques where compared and the main results of this work show that driver surveys are difficult to carry out because of the number of incomplete questionnaires returned by the drivers. To counterbalance this problem, GPS-supplements help to complete them, but several privacy questions may appear if they are generalized (not all drivers accept to be tracked by a GPS device). Moreover, this method needs the double effort of providing both a questionnaire and a GPS data collection device, increasing the costs of the survey. In conclusion, the authors make list of some specific difficulties (technical, logistical, administrative complexity …) when a GPS component is associated with a survey. In the end, the lessons of the experience are going to feed future studies on GPS survey methodology, but the current method is not yet extended to national survey campaigns.

In Australia, Greaves and Figliozzi [7] give account of a GPS survey of commercial vehicles, undertaken in mid-2006 as a pilot study in the Greater Melbourne region (a spatial area of 7,700 km² and 3.6 M inhabitants), in order to support a major update of freight data and modeling capabilities in the metropolitan region. In this survey, the data collection is passive because once the GPS device is installed in the vehicle; no human intervention is required afterwards, minimizing truck driver's interventions. Four kinds of advances in the field of urban freight data collection were expected: (*a*) to describe implementation issues with the data collection, (*b*) to detail the algorithms used to process the raw GPS data into meaningful travel and trip information, (*c*) to present a discussion about pilot survey data tour results, and (*d*) to discuss potential uses and limitations of GPS technology in urban freight modeling and planning.

It has been very important to assure the companies that confidentiality and competitiveness would not be compromised all along the survey. For the pilot study, one week's worth of GPS data were collected for 30 trucks (i.e. 210 truck-days). A GEOSTATS® in-vehicle GPS device was installed, drawing power from the cigarette lighter and capable of storing second-by-second information for several weeks. The antenna was placed on the dashboard or on the roof of the vehicle to maximize the likelihood of maintaining a good line of sight to the satellites. Then data were downloaded using software relating to the GPS device. At last, raw GPS data were processed into discrete trips to feed significant analyses. At this stage, several difficulties occurred, especially given that GPS data are provided as a continuous stream, there is the issue of how to identify trips ends correctly. Another problem is the loss of data that occurs at the beginning of a trip because of the time required to lock onto sufficient satellites (four or more) to infer the correct position, or the degradation of GPS data during a trip. This occurs when there is not a clear line to pick up the satellite information. In spite of such problems, various iterative adjustments led to fair solutions. The final output of the processing was a trip-level summary file for further analysis, a second-by-second file of all the trips, and a daily map showing the trip information.

The authors point out that the pilot survey was not based on a random sample, so it is not an unbiased representation of truck travel and trip characteristics in Melbourne. However, a summary of the data suggests the average tour in the pilot survey is approximately 184 km long and lasts approximately 8.5 h. Of this time, approximately two thirds of the time is spent on the road at an average speed of 33 km/h, with the remaining third stopped (loading, unloading, refueling, etc.). These averages hide a great variability. Another result of interest shows the impact of the time of the day on travel time. The highest average and median speed are observed in the evening and night hours (between 18:00 and 5:30). The lowest average and median speeds are observed during the morning and evening peaks (7:30 to 9:30 and 15:30 to 18:00). The highest proportion of trip starts occurs in the morning peak. Almost 80 % of the trips take place during 8 h, between 7:30 and 15:30.

In the end, the authors present some concluding remarks about the potential applications of passive GPS data in urban freight surveys. Thus, GPS data collection can be successfully used to build precise origin-destination matrices broken down by time of day. Combining GIS sociodemographic data and GIS truck trip data, can be invaluable for trip generation models (constructing trip length distributions). Furthermore, GPS freight routes can be effectively combined with weight data collection in order to track payloads. Another potentially useful application of the GPS-based data collection is to develop speed-time profiles as inputs to fuel consumption and emissions models. A significant limitation of non-intrusive GPS collected data is the lack of behavioral data. As with passenger data collection, GPS data collection methods can be a great complement to commercial vehicle and business surveys. However, the authors conclude it is unlikely that GPS-based data collection will completely replace traditional data collection methods, among others because of sampling issues that require traditional methods for validation purposes.

In South Africa, Joubert and Axhausen [9] and Joubert [8] present a novel methodology using GPS-based data collection, in order to infer commercial vehicle activities. A private service provider to the transport industry offered vehicle tracking and fleet management services. Another company made a data set available, acquired through its own system, containing the detailed GPS log for commercial vehicles for six months (January-June 2008). A total of 31,053 vehicle files were identified, representing approximately 1.5 % of the national heavy and light delivery vehicle population. It is important to note that the vehicles represented in the data set are possessed by customers subscribing to vehicle tracking and fleet management services of the company, so that a bias selection is possible.

Within the context of our paper, one of the noteworthy contributions of this South-African analysis is methodological. The methodology makes a distinction between vehicles performing the majority of their activities within a studied area (within-traffic) and those performing activities across larger geographic areas (through-traffic). The authors propose a method to extract commercial vehicle activities and activity chains from raw GPS data, based on actual observed vehicle movement. So knowing the characteristics of commercial activity chains allows then to generate and simulate a synthesized population of commercial vehicles in the future. Combining commercial and person movement in a single model provides improved decision support for policy and infrastructure evaluation.

The data set file used was consisting of : - vehicle identifiers, not traceable to the original customer for confidential reasons; - seconds-based Unix timestamps; - longitude and latitude values; - vehicles status identifier used by the company for fleet management purposes; - speed values.

Ignition-related triggers reported in the GPS data were used to identify activity start (i.e. vehicle off) and activity end (i.e. vehicle start) times. In order to eliminate any false starts or false stops, a threshold for minimum activity duration was determined. A thorough analysis led to consider that 5 h were a fair threshold to distinguish major activities (indicating start and end-points of the activity chains) and minor activities (those making up the links of the activity chains, as periods to collect or deliver goods).

Finally, the authors remind the readers of some facts: business location density plays an important role in freight affecting congestion and the related environmental and fuel consumption effects. So knowing where activities take place (where business and freight stakeholders are located) and what the inter-activity distances are, may help in future to evaluate and compare the activities against underlying land use data. In turn, land use analysis may provide guidelines in predicting future freight activities based on strategic land use master plans prepared by local authorities. Moreover, the activity locations, along with the activity chain characteristics allow the generation of a representative synthesized population of freight agents. Using a large-scale agent-based transport simulator enables to incorporate freight movement in dynamic traffic simulations, in order to provide better decision support on transport infrastructure. However, the paper does not make mention of some drawbacks in conducting urban freight surveys, based only on GPS data collection.

Although there are few other GPS-based freight surveys in practice [1][11], they have been proposed for preliminary studies and have not been generalized. Moreover, these works focus on data collection, but the data processing and analysis issues are less studied. These surveys suggest some avenues of research, in order to facilitate and complete data collection in the field of urban freight movement. Generally, most of the works on UGM GPS data collection are not directly related to other survey methods. Indeed, only McCabe *et al.* [12] propose this type of approach as a supplement to classical in-depth surveys, and Allen and Browne [1] formulate not much detailed comparisons based on current works, which are essentially focused on GPS data collection feasibility. Therefore, we propose below a GPS-based approach that is focused on both data collection and processing. Moreover, this method will be applied to real data collection and compared to classical driver and trip board surveys.

3. GPS data collection method

In this section, we propose a GPS survey methodology to be tested on real cases, i.e. several carriers from the cities of Bilbao (Spain) and Lyon (France). The chosen technology is a mobile phone with an integrated GPS that has been gifted to lorry drivers in exchange of their collaboration for the tests.

3.1. Goals of the data collection method

This data collection system has been developed and tested in the context of the European project FREILOT [6], which main goal is the deployment of four technological solutions to improve urban freight distribution after testing their efficiency in different pilot studies. The main goal of the data collection method is to estimate the fuel consumption and the environmental impacts of infrastructural solutions on trucks that do not have specific on-board devices. Moreover, the effects of speed and acceleration have to be estimated; therefore, it is important to collect real time data through very small time slots. For these reasons, it is important to define a standard data collection protocol and use appropriate tools in order to collect the needed data in the same way for each individual.

3.2. Implementation issues

There are many GPS data loggers. Most of them must be connected to a computer to retrieve data. For example, the connection by Bluetooth systems has been chosen in other surveys because of its easy use [12]. These loggers, often used in probing vehicles, are less relevant for surveys with plenty of vehicles. In fact, human resources are needed to collect and transfer the data after each journey. This last decade, a new technology has been developed in the mobile telephony sector: the Smartphone. This technology permits to get GPS positions, while being connected to the Internet. In addition, some companies are now specialized in making Smartphone applications. For all these reasons, we have adopted this new technology for GPS data collection (Gertek, a company specialized in electronic services for parking and freight transportation is a partner of the FREILOT project, having already a GPS position recording application. This application has been cheaply adapted to the project's needs). We have uploaded a specific application in the Smartphone and have distributed the latter to transport operators' partners. The application works in several steps. Before the journey, the driver switches on the Smartphone and enters his ID, the vehicle ID number and the company name (already entered in the database). The existing data log file is automatically sent to an FTP server after logging out because this procedure has been considered more efficient and less constraining for the drivers. During the journey, the phone collects the GPS positions every two seconds. Therefore, each file corresponds to one journey. The nomenclature of the file's name contains the date, the company name and the city.

In order to implement the survey, two main issues have to be taken into account. The first concerns the cost, mainly related to the survey's deployment. Today, it costs around 100 € to get a Smartphone connected to the network. The plan subscription is around 20 € per month, depending on the operator, and the quantity of data. The main cost is undoubtedly the creation of the appropriate Smartphone's software. In our case the software was already uploaded thanks to our partner who had a commercial solution easily adaptable to our survey. For our survey, about 50 Smartphone have been distributed and will be used for a year. Considering both data analysis and other indirect costs, we can state that this kind of survey entails a unitary cost of about 400 € per truck tracked, throughout a one-year period. The second issue concerns privacy. According to the law, no detailed records of GPS data must be stored permanently. Moreover, the drivers can feel they are spied, so this method may affect their practices. To deal with these two questions, we have implemented a procedure that combines the data so that each route would be defined without a connection to the exact truck and driver. At any moment, the available information recorded is anonymous and refers only to the company and the type of truck (according to the standard private car unit conversion table, which four types of vehicles are taken into account: private cars and tourism vehicles, light commercial vehicles, simple trucks and semi-articulated and articulated trucks). It is not possible to find information about the driver in the data log files once adjusted. Moreover, the drivers are carrying Smartphone and not classical data collection devices. This approach is more informal and familiar, so the data collection process, after a small period of adaptation, does not seem to be biased a lot. In some cases, the Smartphone has been promised to the driver once the survey has been completed. This gift makes the driver more involved and improves the collaboration between all the actors.

3.3. The data set and the processing

The data set consists of text files which correspond to a specific route (to each file, a GPS ID, a vehicle ID, a driver ID and a route ID have been associated). In each file, the following data are collected every two seconds: GPS Time, Latitude and GPS Longitude, travelled distance and GPS speed. Before providing results, a data processing algorithm using the R language has been implemented (http://www.r-project.org/). The modeling data analysis will be carried out with this software.

- Firstly, distances and speeds will be recalculated to check the accuracy of GPS efficiency.
- Secondly, we can identify some errors into files:
  - Repetition of a same point: When the GPS system loses connection with the satellites, the same position is repeated several times. This error can be tracked easily because the calculated distance between two points is equal to zero. When the system reconnects with the satellites, the distance is usually unrealistic. We can correct it by interpolation of GPS positions.
  - Speed or acceleration problems: Sometimes the speed or the acceleration is unrealistic. We identify these problems through a logical field without correcting them at the time. These cases are very rare (less than 0.01 %) but they can alter the results if they are not taken into account.
- Thirdly, we identify the delivery stops. The criterion to consider the delivery stops are a speed less than 3 km/h and duration greater than 150 seconds. It permits to exclude stops caused by traffic lights.
- The fourth step is to affect a type of road to each GPS point. For this purpose, we used the OpenMapStreet data and aggregate streets in three groups: motorway, main road and residential street. The affectation is made with a GIS software named PostGIS, which looks for each recorded point the corresponding street.

The last step is to estimate the instantaneous fuel consumption and pollutants emissions (CO2, CO, NOx and hydrocarbons, noted HC) using the application proposed by the Comprehensive Modal Emission Model (CMEM). The University of California Riverside for microscopic emissions modeling [4] has initially developed this tool. This model predicts second-by-second tailpipe emissions and fuel consumption based on different operations from in-use vehicle fleet. This model (one of its most important features) uses a physical, power-demand approach based on a parameterized analytical representation of fuel consumption and emissions. More precisely, the entire fuel consumption and emissions process is broken down into components that correspond to physical phenomena associated with vehicle operation and emissions production. Each component is modeled as an analytical representation consisting of various parameters. They characterize the process. These parameters vary according to the vehicle type, the engine, the emission technology, and the level of wear. For the FREILOT project evaluation, it is important to take into account speeds, accelerations and vehicle parameters such as mass. The initially developed model was taking into account only light vehicles. It has then been extended to commercial vehicles, like vans and trucks [5].

The proposed survey is expected to collect a big quantity of data. The amount of data collected, then analyzed, can explain the lack of GPS survey. Indeed, the computers are just efficient enough to analyze this data without memory problems. This information has been treated with specific software tools, which need a broad knowledge in programming. The software R (64 bits version) seems well adapted for such a type of analysis because it is not limited by RAM memory, while the language is complete and flexible and the runs are quick.

4. Main results

In this section, we present the main results of the data analysis. The GPS devices have been deployed on different vehicles belonging to three types of carriers: hostelry logistics, grocery distribution and express couriers. The data collection period covers 7 months, from July to December 2010. Using the proposed data reading procedure, we can identify 911 routes (corresponding to 6,607,605 points). This survey involves 14 companies, 65 vehicles, and 140 drivers.

The results suggested below are based on descriptive statistics. Moreover, we propose different types of aggregation, by focusing respectively on the composition and characteristics of the routes, the driver's behavior per company, the categories of road and the time of the day.

First, we can describe the routes as shown in Table 1. We report respectively the mean and the median values for the overall set of routes, then the first and the third quartiles.

Table 1. Main characteristics of the tours

	Mean	Median	First quartile	Third quartile
Number of points	7,253	6,291	4,182	9,189
Number of delivery stops	21	15	8	25
Duration of a delivery stop (min)	17	14	12	19
Length of a route (km)	57	54	32	76
Duration of a route (hour)	4.27	3.70	2.45	5.38
CO_2 (g/km)	1,040	1,020	715	1,243

We observe an average route length of 57 km, and an average duration of about 4 h 16' (corresponding to the half of a working period). The number of delivery stops is about 21 and the average duration of these stops is 17 minutes. If we analyze the median, we observe that these values are a little lower, but remain close to the average values. If we compare to the results of traditional surveys [2][3][13][14], we observe that we are in the same order of magnitude. However, the comparison can be difficult because the companies of GPS survey are not well sampled and representative of all the types of establishments.

Concerning CO_2 emissions the average and median values are 30% higher than those estimated in France with traditional surveys (Routhier *et al.*, 2009). If we look at the CO_2 harmonic mean emissions for each company, we observe several differences between them (Fig. 1).

If we aggregate the companies by type (Table 2), we observe that grocery distribution, made by large trucks (more than 12 t) is the most consuming company type per kilometer. This is due to the type of vehicle used. However, the environmental impacts are reported here to each travelled kilometer. To estimate the real impact of the company, in terms of CO_2 emission, the unitary emissions have to be weighted to take into account the average length of each company's routes.

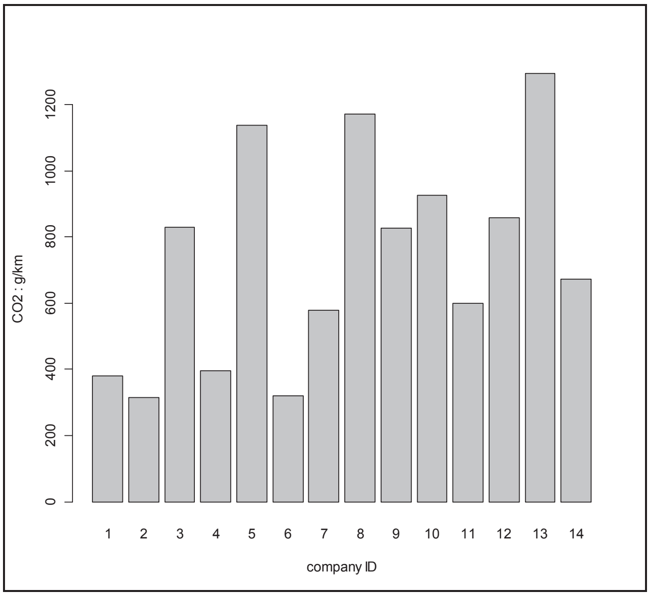

Fig. 1. Harmonic mean of CO2 emissions per km and per company

Table 2. Mean unitary CO_2 emissions and standard deviation for each category of transport company

	Mean CO_2 emissions (g/km)	Standard deviation
Grocery	1,063	177
Inns	684	294
LTL Transport	488	102

By focusing on the type of road, and excluding the time for delivery stops, we can observe different patterns. Most of the transportation time occurs on residential roads (58% of the total time, excluding the pickup/delivery stops). In these streets we observe a reduction of the harmonic mean speed and a big increase of CO_2 emissions, this second fact is linked to the speed and acceleration behavior of drivers.

Table 3. Statistics per type of road (delivery stops are not considered for speed estimation)

	Part of the duration	Part of the distance	Harmonic Mean of speed	CO_2 emissions per truck
Residential	58 %	29 %	7 km/h	1,047 g/km
Main road	22 %	25 %	16 km/h	813 g/km
Motorway	20 %	46 %	33 km/h	761 g/km

Moreover, we can identify the streets which are the most used by the trucks. Main roads and motorways are the roads which structure the itinerary of each trip. Nevertheless, the drivers are most of time travelling on residential streets. This fact shows clearly the importance in using vehicles appropriate to the urban space of the cities. Similarly, we can represent on a map the CO_2 emissions on each type of street (Fig. 2). The traffic in residential streets is very polluting because of numerous accelerations due mainly to traffic lights and crosswalks.

Fig. 2. CO_2 emissions per street

Finally, combining time and geo-location, this system allows a better understanding of truck flows all day long. Below Fig. 3 shows the starting points of the journeys. We observe that most of rounds occur between 5 a.m. and 7 a.m. Moreover, in terms of tour before noon, the activity represents about ¾ of the total tours.

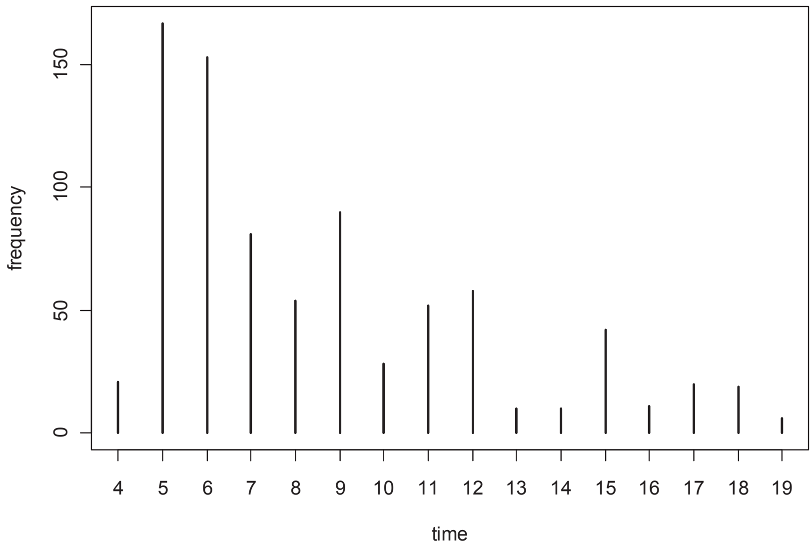

Fig. 3. Distribution of tour start points per hour

5. Lessons and generalization

After having presented the most significant results of our GPS data collection and analysis procedure, we now can discuss the possibilities of application of this type of approach, with respect to traditional surveys. The main aspects that we want to focus on are the following:

- *Quantity and quality of collected data*. GPS data are relatively easy to collect on a large scales, especially when using new technologies (like the Smartphone applications) which facilitate automatic data transmission. However, the accuracy of these data is not always complete for the goals of the survey. For instance, if we want to estimate speeds, accelerations, paths, fuel consumption and environmental impacts, the results presented above show that GPS data allow estimating them correctly. On the contrary, at present time, satellite technology is not efficient enough to give a good approximation concerning data positioning. Therefore, the technology needs still improvements in order to estimate the precise stop location (irregular or regular on-road parking) or to distinguish a pickup or delivery stop from a signal stop or other unexpected events related to traffic congestion.
- *Freight-related information*, like loading rates, type of goods or other practice issues, cannot be obtained by a GPS-based technique. Moreover, transport plans are not enough standardized and detailed in order to extract this information (during the tests, data about loads and pickup/delivery locations were asked to carriers, but each company often provided different file formats; so that some of them could not be automated). It could be interesting to improve the Smartphone software by interfacing it with a form, enabling the driver to enter information about company delivered, weight and volume of goods, parking conditions among others. However, the most practical way to make the driver enter this information is to limit the options and to provide, when possible, a pre-defined menu to make the driver fill in the information efficiently.
- *Data sampling*. Whatever survey method used (GPS, driver, trip diaries, etc.), it is important to ensure a good sampling.
- *Costs*. GPS surveys are not extremely expensive in the light of the collected information, especially related to the non-response rates. For example, in French surveys, a mail-out/ mail-back approach was used and the non-response rate was between 15 % and 40 %, according to the urban area surveyed [13]. With a GPS-based survey, all information is collected and the number of removed rounds (for data accuracy reasons) is less than 10 %. However, the cost of GPS-survey data collection is proportional to the number of data loggers installed (about 300 €/device, in our case) but after calibrating and testing the data processing applications, the data processing and analysis can be carried out by only one person in a reduced time period.
- *Implementation limits*. Two main limits occur: The first one is the acceptability of the carriers, mainly for confidentiality reasons (for example, the fear that confidential information about their trips can be used by a competitor) and the second one concerns the drivers' privacy, who may think the data logger is used by the transport carriers to track them.

From these issues, we can conclude on the following table (Table 4). In the table, we state some avenues of research, in order to facilitate and complete data collection in the field of urban freight movement. Indeed, each method presents several limits. We think important to combine them in order to meet the survey targets: GPS-based techniques, more precise data can be obtained about a number of parameters, as routes, speeds, energy consumption, … whereas current mail-out/mail-back and face-to-face surveys give invaluable information about the behavior of the stakeholders and of the logistics schemes.

Table 4. Comparison of different survey techniques for route characterization

	Driver survey	Vehicle diary	Establishment survey	Delivery space observation	GPS survey
Length of route	+	++			+++
Duration of route	+	++			+++
Number of stops	++	+++			++
Location of stops	++	+++			++
Duration of stops	+	++			+
Type of operation	++	+	+	+++	
Nature of the freight	+	+	+++	+	
Type of package	++	+	+++	+++	
Type of carrying tools	+++	+	+++	+++	
Usage of truck equipment	+++	+++	++	+++	
Nuisance to the environment		+			+++
Parking infractions	+	+		+++	
Stopping maneuvers	+			++	
Data collection costs	+++	+	+++	+++	++
Data processing costs	++	+++	++	++	+
Non-response rates	++	+++	++	++	+
Privacy issues	++	++	++	-	+++

6. Conclusion

This paper presents a GPS-based data collection technique for urban goods surveys, more precisely for route characterization. The technique has been implemented and tested in the city of Bilbao (Spain). The main results show that several data like route length, number of stops, fuel consumption and CO_2 emissions can be correctly estimated from GPS data. Several aggregation approaches can also be undertaken (for example, by company, type of road or time period). The suggested method uses a Smartphone application that offers a larger flexibility than techniques used currently. Indeed, this application allows online data transmission and makes its use by the driver easier. With respect to traditional survey techniques, GPS data gives a better estimation of the parameters mentioned above. Nevertheless, current mail-out/mail-back and face-to-face surveys remain essential to in-depth studies concerning a good knowledge of the behavior of the stakeholders and of the logistic chains. Indeed, it seems to be an appropriate way to understand well goods vehicle flows. Finally, the combination of several survey techniques (mail-out/mail-back, face-to-face and GPS-based) is essential to find relations between urban logistics and land use, in order to improve efficiently local authorities' decisions.

Acknowledgements

This research has been carried out related to the FREILOT project (Urban Freight Energy Efficiency Pilot), partially financed by the European Commission, in the context of its ICT Policy Support Program (ICT PSP). The authors want also to acknowledge Fernando Zubillaga, from MLC ITS Euskadi, for his help deploying the GPS devices and gathering all the information related to transport companies, as well as Irantzu Bores, from GERTEK, for the technical support and the adaptation of the Smartphone application for GPS data collection.

References

[1] Allen J, Browne M. *Review of survey techniques used in urban freight studies*. Research Report. University of Westminster, UK; 2008.

[2] Ambrosini C, Patier D, Routhier JL. Urban freight establishment and tour-based surveys for policy-oriented modeling. *Procedia Social and Behavioral Sciences* 2010; 2(3): 6013-6026.

[3] Ambrosini C, Routhier JL. Objectives, methods and results of surveys carried out in the field of urban freight transport: An international comparison. *Transport Reviews* 2004; 24(1), 57–77.

[4] Barth M, An F, Norbeck J, Ross M. Modal emissions modeling: A physical approach. *Transportation Research Record* 1996; 1520: 81-88

[5] Barth M, Scora G, Younglove T. Modal emissions model for heavy-duty diesel vehicles. *Transportation Research Record* 2004; 1880: 10-20.

[6] Blanco R, Garcia E, Novo I *et al. FREILOT – Evaluation methodology, Report D 4.1*. FREILOT project, Brussels; 2010.

[7] Greaves SP, Figliozzi, MA. Collecting commercial vehicle tour data with passive global positioning system technology. Issues and potential applications. *Transportation Research Record* 2008; 2049, 158–166.

[8] Joubert JW. A provincial comparison of commercial vehicle movement. *Proceedings of the 29th Southern African Transport Conference*, 16-19 August 2010, Pretoria, South Africa, 9p.

[9] Joubert JW, Axhausen KW. Inferring commercial vehicle activities in Gauteng, South Africa. *Journal of Transport Geography* 2009, in press, doi: 10.1016/j.jtrangeo.2009.11.005, 10p.

[10] LET. *Transports des marchandises en ville. Enquête quantitative réalisée à Bordeaux*. Final Report LET-DRAST-ADEME-CUB; 1997.

[11] LET. *Transports des marchandises en ville. Résultats de l'enquête pilote réalisée à Lyon*. Final Report LET-PREDIT-ADEME; 2010.

[12] McCabe S., Roorda M, Kwan K. Comparing GPS and non-GPS survey methods for collecting urban goods and service movements. In: *Proceedings of the 8th International Conference on Survey Methods in Transport: Harmonisation and Data Comparability*, Annecy, France, 25-31 May 2008.

[13] Patier D, Routhier JL. Livraisons de marchandises en ville. *Transport, Environnement, Circulation* 1997; 145, 8-14.

[14] Patier D, Routhier JL. Une méthode d'enquête du transport de marchandises en ville pour un diagnostic en politiques urbaines. *Les Cahiers Scientifiques du Transport* 2009; 55: 11-38.

[15] Routhier JL, Traisnel JP, Gonzalez-Feliu J, Henriot F, Raux C. *ETHEL II: Energie, Transport, Habitat, Environnement, Localisations*. Paris: GICC-ADEME; 2009.

[16] Taniguchi E, Thompson RG, Yamada T. Data collection for modeling, evaluating and benchmarking city logistics schemes. In: Taniguchi E, Thompson RG, editors. *Recent advances in city logistics*, Amsterdam: Elsevier; 2006, p. 1-14.