

COMPARISON OF TRAVEL DIARIES GENERATED FROM SMARTPHONE DATA AND DEDICATED GPS DEVICES

10TH INTERNATIONAL CONFERENCE ON SURVEY METHODS IN TRANSPORT:
LEURA, AUSTRALIA, NOVEMBER 16-21, 2014

Lara Montini, IVT, ETH Zürich, CH-8093 Zürich

Sebastian Prost, AIT, Business Park Marximum, A-1110 Wien

Johann Schrammel, AIT, Business Park Marximum, A-1110 Wien

Nadine Rieser-Schüssler, IVT, ETH Zürich, CH-8093 Zürich

Kay W. Axhausen, IVT, ETH Zürich, CH-8093 Zürich

ABSTRACT

The goal of this paper, is to provide further insight into the usability of smartphones and dedicated GPS devices for collecting travel survey data at this point in time. For analysis, GPS and accelerometer time series of 31 PEACOX study participants are available, who were tracked simultaneously with smartphones and dedicated devices for 8 weeks. Meaningful diaries can be extracted from both data sources. If high resolution data is needed however, results suggest that dedicated GPS devices are still relevant, as they do not have battery issues thus more data is recorded and furthermore data quality is more stable.

1 INTRODUCTION

In transportation research, GPS tracks are used amongst other data sources to construct travel diaries. This data is mostly collected using dedicated GPS devices that respondents have to carry it with them as a special commitment during the tracking period. A promising new source for GPS data are smartphones (see e.g. Gould (2013)) as they became both equipped with GPS and

other sensor functionality in the last years and, as opposed to dedicated devices they are carried by potential respondents in any case. The advantage is this double use as a consequence they are less likely to be left at home in comparison to dedicated GPS devices. However, double use also implies an important challenge especially with regard to battery life. Another difficulty for survey use of smartphones is that usually different operating systems, brands and types with antennas of differing quality have to be covered.

The goal of this paper is to provide further insight into the usability of smartphones and dedicated GPS devices for collecting travel survey data at this point in time. Data quality and usage of the two device types are compared. Further, travel diaries generated from the two data sources are compared. For analysis, GPS and accelerometer time series of 31 study participants are available, who were tracked simultaneously with smartphones and dedicated devices for 8 weeks. The data was collected as part of the PEACOX project (www.project-peacox.eu), in which a personalised journey planner application for smartphones that encourages ecological behavior is developed. In the app, position data is collected in the background in order to generate travel diaries, these are then used to personalize route suggestions. In this paper data is analysed that was collected in the the second field trial of the app, that took place in Vienna and Dublin from the beginning of August until beginning of October 2014.

The paper is structured as follows. First, smartphone applications used in the context of the PEACOX project, that is the journey planning app as well as a prompted recall app are presented. In the subsequent section the field trial is described. Section 4 outlines the differences in the travel diary construction for the different device types. In the following results are reported, including quantitative analysis as well as the subjective perceptions of users. To conclude an interpretation of results and an outlook on continuing work is given.

2 STUDY CONTEXT: THE PEACOX PROJECT AND APPLICATIONS

The main focus of PEACOX is the potential influence of the journey planning app including its persuasive components on users' travel behaviour and attitudes towards mobility. The data analysed in this paper is a side product of the PEACOX journey planning application. This application is a prototype and tested in field trials, which enabled us to enrich the data with questionnaires, a prompted recall tool and by giving participants dedicated GPS loggers (MobiTest GSL).

Following the journey planning as well as the prompted recall applications are introduced.

2.1 Journey planning app

The journey planning app allows the user to perform a multi-modal search for a route, which is tailored to the user's individual preferences and behavior patterns. In general, it works like a common journey planner, that is an origin and a destination is specified and then possible routes are suggested. When routes are requested in PEACOX the available alternatives are enriched with emission information (Alam and McNabola, 2012). The enriched results are then ranked and personalised by the recommender engine (Bothos et al., 2012). Recommendations are partially based on the trip history that is detected from recorded GPS and accelerometer data. Selected eco-friendly route options are promoted by adding a persuasive message (Figure 1(a)). When clicking on a route all details are displayed, that is walking-, driving- and waiting times, public transport line and schedule. After that the routes can also be viewed on the map and as results are multi-modal different transport modes are distinguished by colours (Figure 1(b)).

PEACOX is a persuasive advisor, potential persuasion strategies are discussed in Prost et al. (2013b), the following are tested in the second prototype. First, users are kept interested in the topic by challenges where they can compete against other users (Figure 1(c)). An emission ranking where CO₂ emissions of all users are shown aims at promoting eco-friendlier modes. As it is hard for a car commuter to compete with someone working in walking distances, the user's improvement compared to his own previous behaviour is measured as well and represented by a growing or shrinking tree (Figure 1(d)).

The journey planner is implemented as a smartphone application for the Android platform version 4.0. and higher. Concerning data storage, all data is transmitted via mobile internet every few minutes to a central database. GPS and accelerometer data pose the biggest problems because of the huge numbers, especially accelerometer data had to be backed up and deleted every day so that only two days of data are saved in the database at a time, otherwise querying the database (PostgreSQL) became slow. During the field trial 1.5 to 4 GB of accelerometer data were saved per day. More implementation details are given in Artukovic et al. (2013).

2.2 Prompted recall app

The main function of this app is to allow participants to review the collected trip history, and to provide manual corrections. One goal of the PEACOX study is to observe behavioural changes during the 8 week field trial due to the journey planning app. The travel diaries are one data source to analyse such changes, it is therefore important that in particular the chosen transport modes are known. Having corrected travel diaries is also extremely valuable as quality of the

automatically produced diaries can be tested and further improved.

To obtain corrections from survey participants a prompted recall tool was developed. This tool supports participants in recalling transport modes and activity types, by presenting GPS tracks based on the automatically generated diary. Most commonly prompted recall surveys are web-based as shown in Auld et al. (2009) and as e.g. implemented by Montini et al. (2013) for a GPS-based survey in Zurich.

For the second PEACOX field trial, a clearly laid out and simple to use user interface was developed. As PEACOX is a smartphone application, all participants have one and are also experienced in using it. Therefore, the prompted recall tool was developed for smartphones as well.

Screenshots of the 'trip diary' app, as it was known to the users, are shown in Figure 2. First, the users have to select the day they want to correct. This is done in the menu (Figure 2(a)). All days with data that have not been corrected are listed there. To finish corrections the checkbox 'I have reviewed this day' has to be checked on the main screen (Figure 2(b)).

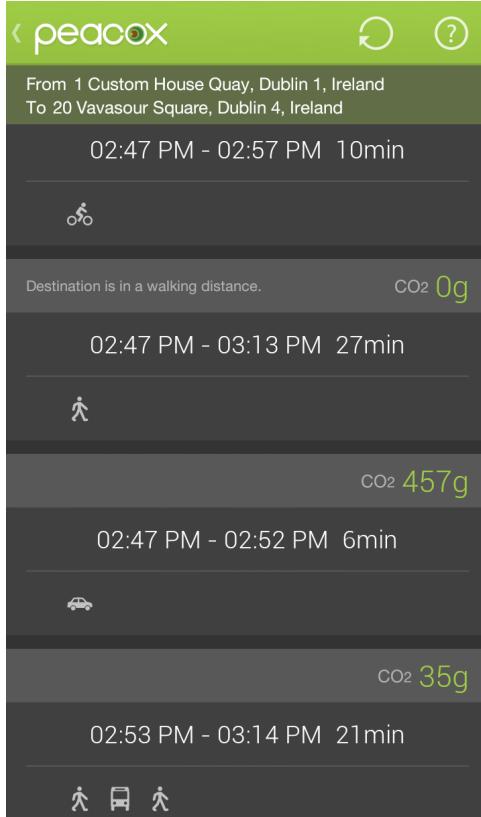
The main screen of the prompted recall tool shows the map as well as the automatically generated travel diary as a list (Figure 2(a)), which also includes editing possibilities for transport mode (Figure 2(c)) and activity type. The different transport modes are represented by different colours and all activity types have distinct icons. This supports processing the information displayed on the map, it particularly helps linking the map with the diary at the bottom. On the right of the top bar, the number of activities is indicated, and one can quickly flip through them by clicking on the arrows. The map is automatically zoomed to the selected activity. If no stages are detected a checkbox 'I stayed home all day' is shown instead of the diary list. Unfortunately, it was realised only after the field trial, that in the list of days that are not corrected yet, only days were included where some data was available. Consequently, most days without data were not confirmed by users.

Users can leave a comment for every day they correct (Figure 2(d)). This is important because users can write down if something is unclear, or if trips or activities are missing.

The app shows very private data and is therefore login protected, the login is the same as for the PEACOX journey planner.

Figure 1: Screenshots of the journey planning app

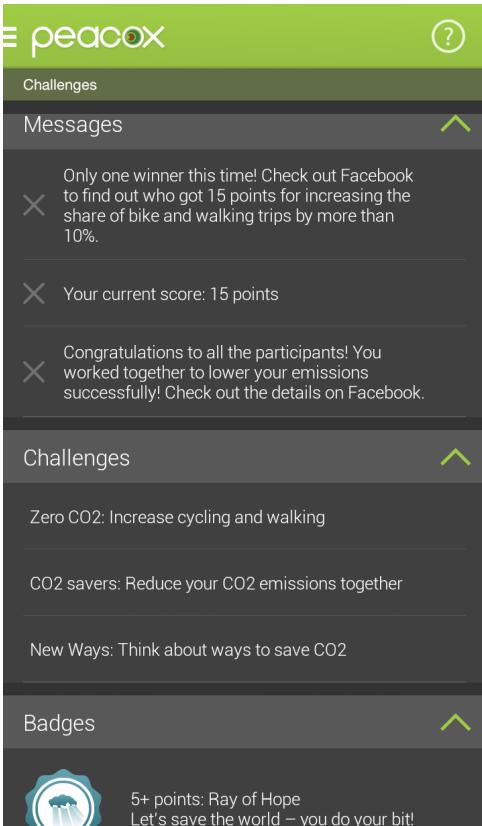
(a) Route: suggestions with CO2 information



(b) Route: visual representation



(c) Persuasion: challenges and badges



(d) Persuasion: growing tree

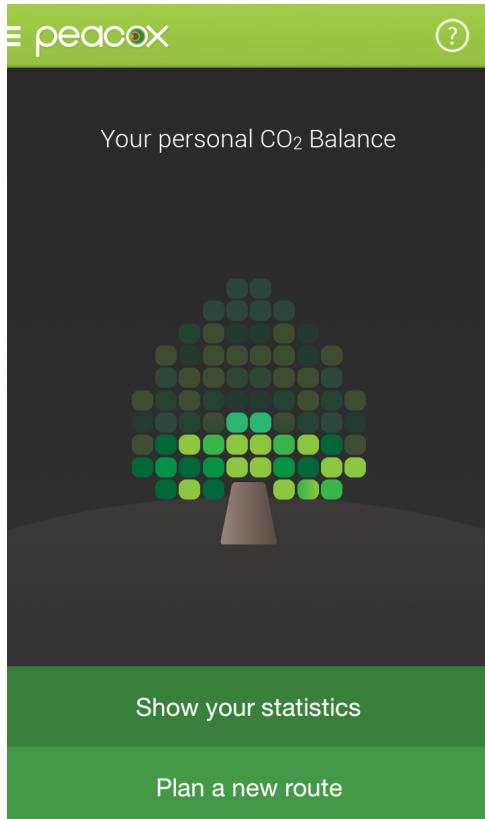
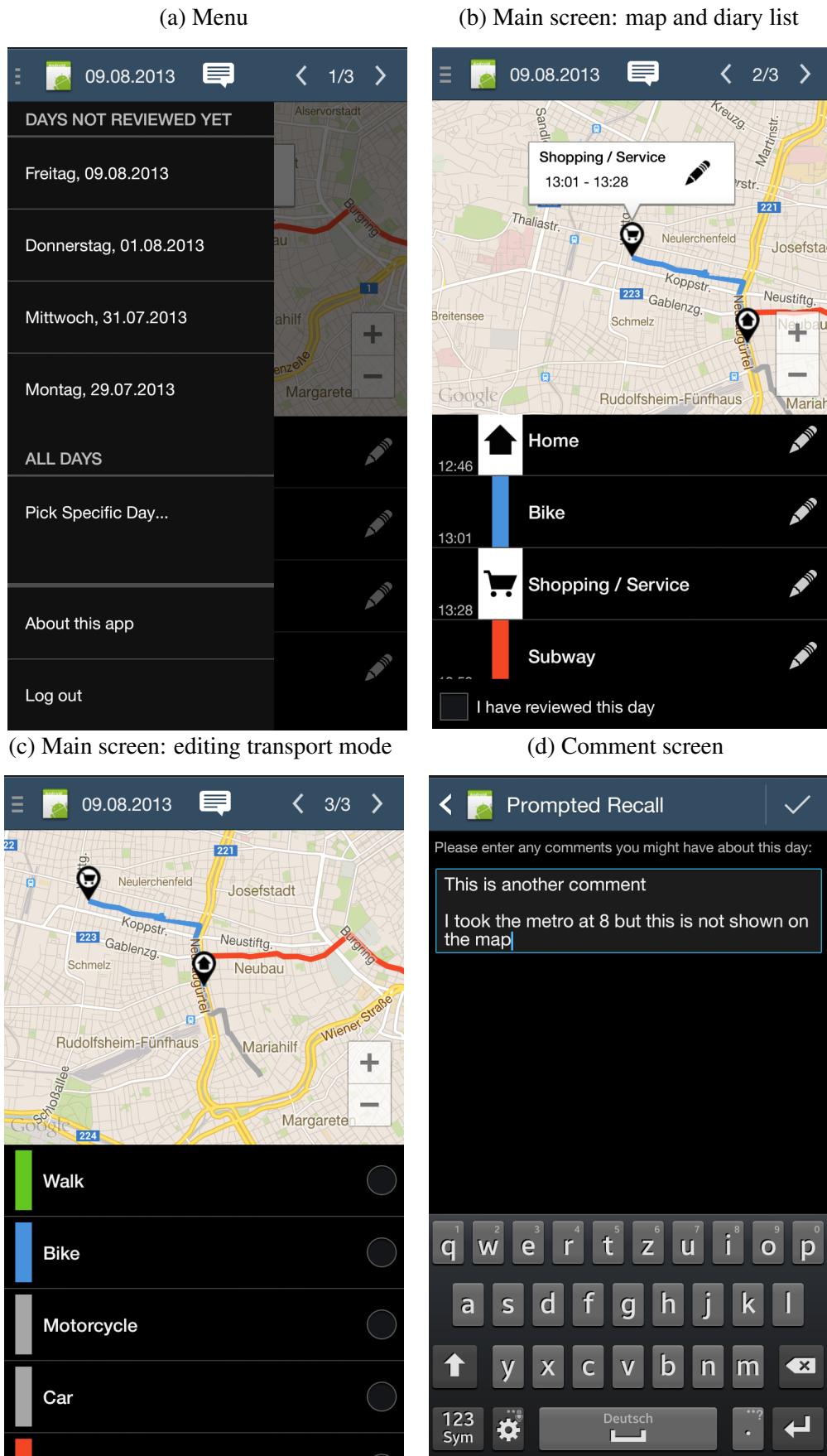


Figure 2: Screenshots of the prompted recall app



3 FIELD TRIAL

In 2013 the first version of the trip planning application was tested in a first field trial in Vienna Prost et al. (2013a). The trial allowed to identify serious issues with the GPS and accelerometer logging on the Smartphones, which were removed for the second trials.

The second field trial was undertaken in 2014 from August 11th to October 4th in Vienna, Austria as well as in Dublin, Ireland where the field trial started one week later.

Altogether 37 test users (20 in Vienna, 17 in Dublin) participated in the field trial and tested the application on their own smartphones for eight consecutive weeks. The application accessed the built-in sensors of the smartphone and logged their GPS as well as accelerometer data. Additionally, participants were equipped with a dedicated high-precision GPS positioning and logging device. As part of the trial users were also asked to manually monitor their logged data (based on the smartphone GPS) using the prompted recall diary, and to provide corrections to validate the automatically generated travel diaries.

Participants

Participants were recruited from a database of people interested in taking part in usability and user experience studies and by open calls for participation promoted in university lectures and by use of university mailing lists. Prospective participants had to fill in a screening questionnaire, and only were recruited in case they fulfilled the following predefined criteria: age 18 or older, living and working/studying in the test area (Vienna respectively Dublin metropolitan area), users of an Android smart phone (running Android OS 4.0 or newer) for at least three months, must have an associated data plan with a minimum of 500 MB per month, and during the eight weeks of trial plan to be absent (e.g. holiday outside of the study regions) for no more than one week.

Overall, the recruitment aimed at including a balanced representation of relevant mobility types (car users, cyclists, pedestrians, users of public transport) as well as demographic characteristics such as sex and education. This recruiting strategy resulted in the following sample:

Age Average age is 33, whereas the oldest participant is 69 and the youngest 19.

Sex 14 female, 23 male

Occupation 16 participants are employed, 12 are students, 4 are unemployed or retired, 3 are self-employed and 2 are on parental leave

Main transportation means 6 users reported to mainly use car or motorbike, 6 use bicycles,

11 public transport, 5 walking and 9 are missing

Usage of journey planning app 8 participants never used a journey planning app prior to the study, 29 did

Procedure

After agreeing to take part in the trial, participants were invited to a introductory workshop which focused on instructing the users on the trial procedure, explaining the functionality and handling of the devices and apps and how participants were expected to use them. Participants were instructed to carry the devices around at all times, to enable GPS-positioning and logging on their smartphones and to regularly charge the devices.

During the field trial, after about three and six weeks of usage qualitative in-depth interviews concerning the usage of the app, the experiences made and their influence on transport mode decisions were conducted with most users unfortunately, some were not reachable.

At the end of the trial participants were invited to focus groups that concentrated on collecting and reflecting the experiences made by the users during the trials.

Besides these face-to-face interactions with the participants also online questionnaires were sent three times during the trial: at the beginning, in the middle and at the end. The questionnaires focused on demographic data, mobility behavior and attitudes towards different transportation means and environmental issues. From the second onwards also questions regarding the usage and experience of the apps were included.

Data collection

As described above two different approaches to collect data were used. The dedicated GPS was a MobiTest GSL device (MGE DATA (2012)), GPS data was collected with 1 Hz and accelerometer data with a frequency of 10 Hz. Data was stored locally on the device, and after the end of the trials when the participants handed back the devices the data was downloaded and made accessible for analysis. Regarding the smartphone data, as participants used their own devices, the sample consists of a variety of models with a dominance of Samsung devices as shown in Table 1.

Position data was collected in the background by the PEACOX app. GPS data was collected with a frequency of 1 Hz and uploaded to the server every minute. Accelerometer data was specified to use the standard frequency of the sensor which is usually set to 5 Hz, data was uploaded every 70 seconds. Dedicated programming of the app made sure that the logging

Table 1: Smartphone types used in Field Trial

Smartphone type	Nr devices
Samsung Galaxy S3	7
Samsung Galaxy S2	6
Motorola Moto G	3
Samsung Galaxy Nexus 2	2
Samsung Galaxy Nexus 4	2
Samsung Galaxy S3 mini	2
Sony Xperia Z1	2
Samsung Galaxy Note 2	1
Samsung Galaxy S4	1
Samsung Galaxy S4 mini	1
Alcatel One Touch 4030x	1
Huawei Ascend Y330	1
LG Nexus 5	1
LG P760 Optimus L9	1
UTime U100	1
Vodafone 875 Smart mini	1
Not reported	4

process was not stopped by the Android Task management, and that all available location information sources (GPS and network) were used for acquiring position information.

4 TRAVEL DIARY GENERATION

To process GPS and accelerometer data the software package POSDAP (2012) is used. The three most relevant steps when creating travel diaries are:

1. *Cleaning of raw data* GPS points are filtered when too few satellites are in view or accuracy measures are bad.
2. *Identification of activities and trips* is mainly based on point clouds, signal gaps and changes in the accelerometer signal if mode is changed to or from walk.
3. *Identification of transport mode and activity type* this is either done using a fuzzy rule or a random forest classifier.

The configuration of the routines was calibrated on data collected with the same dedicated GPS loggers that are also used in this survey (MobiTest GSL). The training data was collected in and around Zurich for up to one week by 150 different people (Montini et al., 2013).

Following the differences in processing are described for the three travel diary types: (1) uncorrected diaries from smartphone data, (2) corrected diaries from smartphone data and (3) uncorrected diaries from dedicated device data.

Uncorrected diaries from smartphones

The diaries evaluated in this paper are the ones actually presented to the participants. They were created every night during the field trial.

A random forest classifier for activity type identification is learned new every day incorporating three data sources: (1) the data set collected in Zurich (around 7000 observations), (2) data collected in the first field trial I (425 observations) and, (3) all data that is collected and corrected during the second field trial. Using the freshly corrected data is the reason why the activity type classifier has to be updated on a daily basis. As it was shown in Montini et al. (2014), the distance to home and work locations are important but these locations are not known for PEACOX users, therefore both locations are learned as fast as possible: if corrected data is available, the location most often annotated as home and as work are saved for that person. Otherwise, if GPS data was collected but no corrections are available, a classifier that does not use distance to home and work is used to classify all activities, the locations predicted to be home and work are then used to extract an approximation of these two locations. Using these approximations, distance to home and work can be calculated and classification is run again using a classifier that takes advantage of distance to home and work.

After two thirds of the field trial configuration of the processing routines was changed, because many stages were detected within point clouds. Hence, the detection of point clusters was relaxed (radius for clouds increased from 10 to 35 meters) and the duration criteria were increased (minimum stage duration 3 minutes instead of 1 minute). Unfortunately, trip purpose detection stopped working due to an error when loading the freshly corrected data into a new classifier, this error went unnoticed for 3 weeks (day 22 to 41 of the field trial) . The trip detection was rerun for all these days, but dutiful participants who corrected the diaries reliably were burdened unfortunately even more.

At first, mode detection was implemented as fuzzy rule system. For the last third of the field trial period it was replaced by a random forest classifier. This classifier apart from considering the commonly used speed and accelerometer variables also included knowledge of the self-reported mode shares.

Corrected diaries from smartphones

The corrected diaries are heavily based on the uncorrected ones, as users were not allowed to change start and end times or adding new activities. But users could add the flags 'no activity' and 'no trip', therefore stages are merged if 'no activity' was in between and accordingly activities are merged as well. Further, the transport mode and trip purposes corrections by participants are considered. For this paper, no further corrections were made by the researchers.

Uncorrected diaries from dedicated devices

The data collected by dedicated devices is processed all in one run, after the field trial. The processing used the configuration of the first few weeks of the field trial. For trip purpose and mode detection random forest classifiers were learned on the training data from Zurich.

5 RESULTS

In this section results on the usage of smartphones to collect travel survey data are reported. First, an analysis of the differences in quality, both with regard to basic data collection and the imputation of trip diaries is shown. In the second part the users' perception, experience and usage patterns are reported. It is important to remember that all comments made by the users relate to data collected by smartphone, as no data from the dedicated devices was accessible to the users during the trials or used for the calculation of the trip diaries.

5.1 What the data shows

First, it is analysed if user actually corrected their trip diaries and how much they had to correct. Following, data quality and travel diaries generated from smartphone data and from dedicated devices are analysed.

User corrections

In total 10322 stages were detected. Users made corrections in 41 % of all cases. Unfortunately, there is evidence that even more corrections are needed. First, only 51 % of the days were confirmed by users. Second, 6.5 % of all detected stages are marked with mode 'unknown'. And third, 24 % of all stages were corrected to be no trips at all, and of those 68 % are stages detected as bike. Even when removing all 'no trip'-stages a bike share of 27.5 % remains which seems too high, even though the study encourages green behaviour. The remaining reported mode shares are 41.4 % walking, 12.8 % car, 5.8 % bus or tram, 0.9 % rail and 2.8 % metro.

Trip segmentation was reconfigured during the trial (day 39 after start in Vienna) and also the mode identification was changed. The changes had a positive effect, after the changes slightly less stages were classified as 'no trip' (22 % compared to 25 % before). Overall, the share of correctly identified modes increased from a very low 52 % to 75 %.

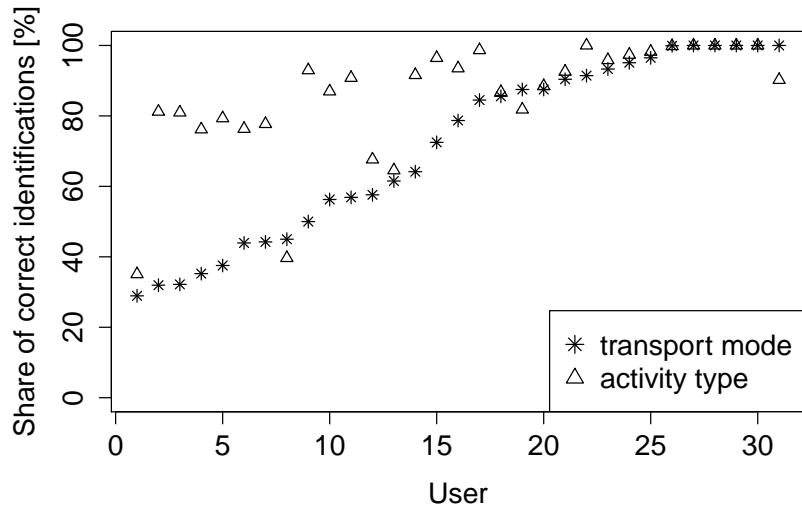
Activity type detection was not perceived very well by users when asked about it, this was probably highly influenced by the three-week interruption where the classifier did not run, as number-wise it performed better than mode detection. Less than 20 % of all activities are corrected by users. Performance stayed stable when comparing week 2 and 3 with week 6 and 7 of the field trial. Of all corrections 4 % are declared as 'no activity'. After removing those and merging activities with 'no trip' in between the following shares are reported: 27.9 % being home, 27.2 % leisure activities, 15.0 % mode transfers, 12.1 % work or education, 10.3 %

shopping, 1.1 % business activities, 0.9 % picking up someone and 5.5 % unspecified stops. The travel diaries extracted from the dedicated devices have 5 % more activities. The main difference is in detected activity types as more than 3-times more mode transfer points and less than half as many home, work and leisure activities are detected.

In Figure 3 the detection accuracies are shown per user, ordered by the share of correctly detected transport modes. Six user have 100 % accuracy which indicates, that no corrections were made. Again it is clearly visible that activity type detection performed better than mode detection which was either more influenced by the quality of the segmentation, or participants tend to correct modes but not activities.

Figure 4 shows that 30 % of corrections were done within one day, and the majority within one week. But several entries were corrected more than 3 weeks after collecting the data.

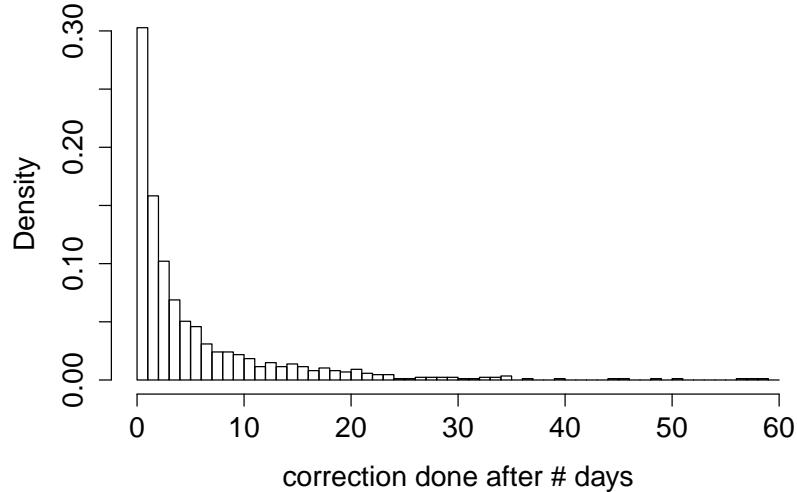
Figure 3: Detection accuracy per user (based on corrections of smartphone-based diary)



Comparison of GPS data quality and device usage

In order to compare usability of dedicated devices and smartphones for mobility studies the detected movement duration is chosen over the number of trips. The problem of number of trips is that it relies on both the trip segmentation and the activity type detection, as stages are merged into trips if inbetween a mode transfer point is detected. Summing up the duration, on the other hand, should result in similar total durations even if activities are not classified correctly and also if numbers of stages differ.

Figure 4: Number of days passing before corrections are done



Following two assumptions about differences between smartphones and dedicated devices are tested:

1. Is the data quality of dedicated devices better and more stable than that of smartphones?
2. Are more days covered by smartphones as they are less likely to be forgotten at home?

To get a proxy for data quality the sampling frequency is used. Both device types are specified to use a sampling frequency of one GPS point per second. The sampling frequency was computed for all detected stages (movement segments). In Figure 5 the average sampling is shown per user ordered by smartphone sampling rate. It can be seen that for the smartphones the sampling is generally lower and in between phones there are greater differences in sampling rate than in between devices. This confirms our assumption and is therefore not surprising, but the different frequencies have to be considered when configuring GPS track segmentation routines. In this Figure also the types of smartphones used most in the study are highlighted, it is shown that sampling frequencies also differ within smartphone types, but it is unclear if similar to the dedicated devices or to a greater extent.

Figure 6 shows the movement detected by smartphones, the corrected movement ('no trip', 'no activity' flags) as well as duration detected using data of the dedicated devices for 4 different users. Further, the color indicates whether the user corrected the diary or not and also if s/he flagged stayed home, it can be seen that this is not the case due to the error in the trip diary that

these days were not listed under 'days to be reviewed'.

In Figure 6(a) the positive effect of the before-mentioned configuration change is visible as movement detected for smartphone drops clearly after day 39, and is then similar to the dedicated device. But it can also be seen, that for some users the original configuration was good (Figure 6(c)) and for other users the change in configuration had no clear effect (Figure 6(b)). Figure 6(d) shows a user who collected a reasonable amount of data. One and half weeks are well corrected, including comments in the trip diary but, after no data was collected in the third week, the trip diary app seems to be forgotten. For all users it can be seen that more or less no corrections were made in the last week.

Figure 7 shows for each user on how many days movement was detected for both devices, for one device or for no device. First, on many days no movement was detected at all (yellow). Approximately half of the users from Vienna did not move for a maximum of 10 days which is above expectations for an 8-week-field-trial, but not too much. In any case, there is definitely movement, that was not captured by any of the devices. Interestingly, more movement is captured with devices, which contradicts the assumption that smartphones cover more days. The feedback of users, given in detail in the next subsection, suggests that smartphones are not forgotten at home, but the app is turned off to save battery. Restarting the app when moving again is easily forgettable.

It is also obvious, that the quantity of collected data differs for the two cities. Potential reasons for this are that, first, some of the Vienna users already participated in the first field trial. Second, the main survey team was located in Vienna which maybe induced more commitment in users living there. And third users from Dublin were younger.

Figure 5: Average GPS point frequency after cleaning for detected stages (after cleaning GPS data)

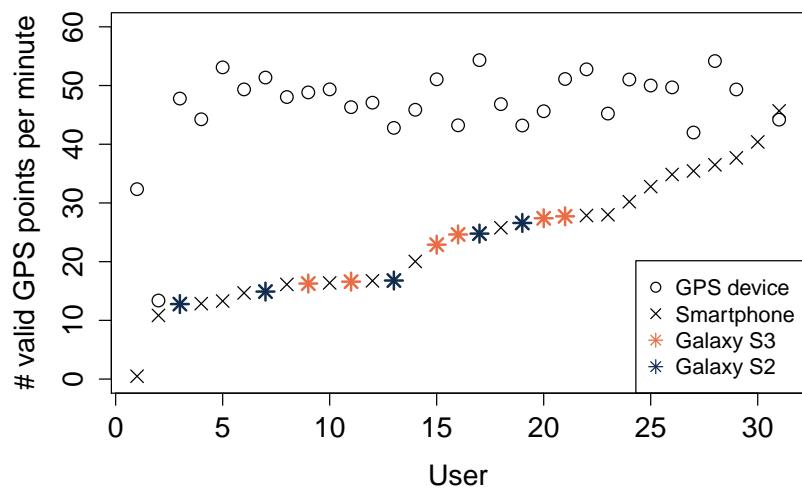


Figure 6: Detected movement for each day for some sample users

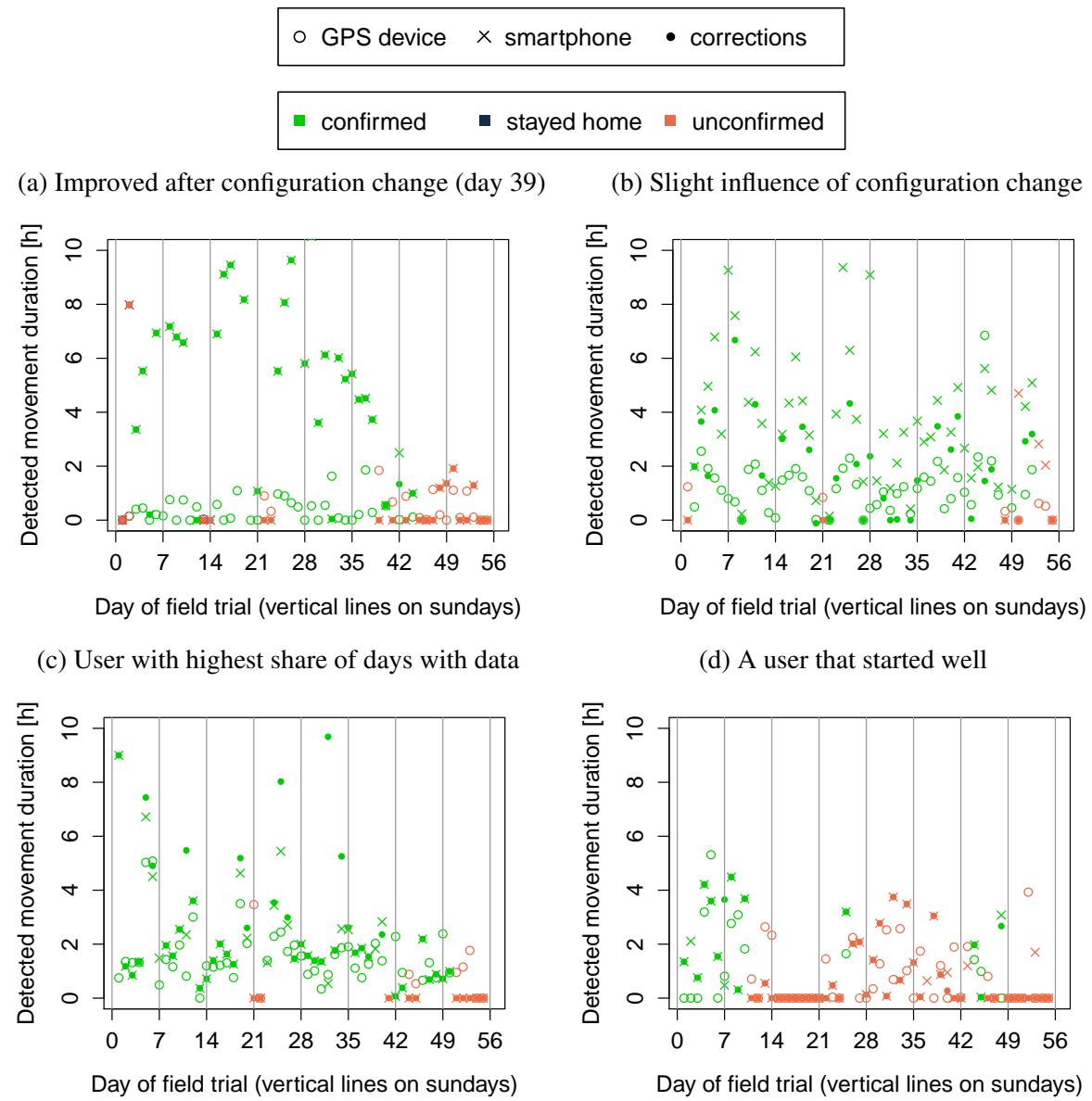
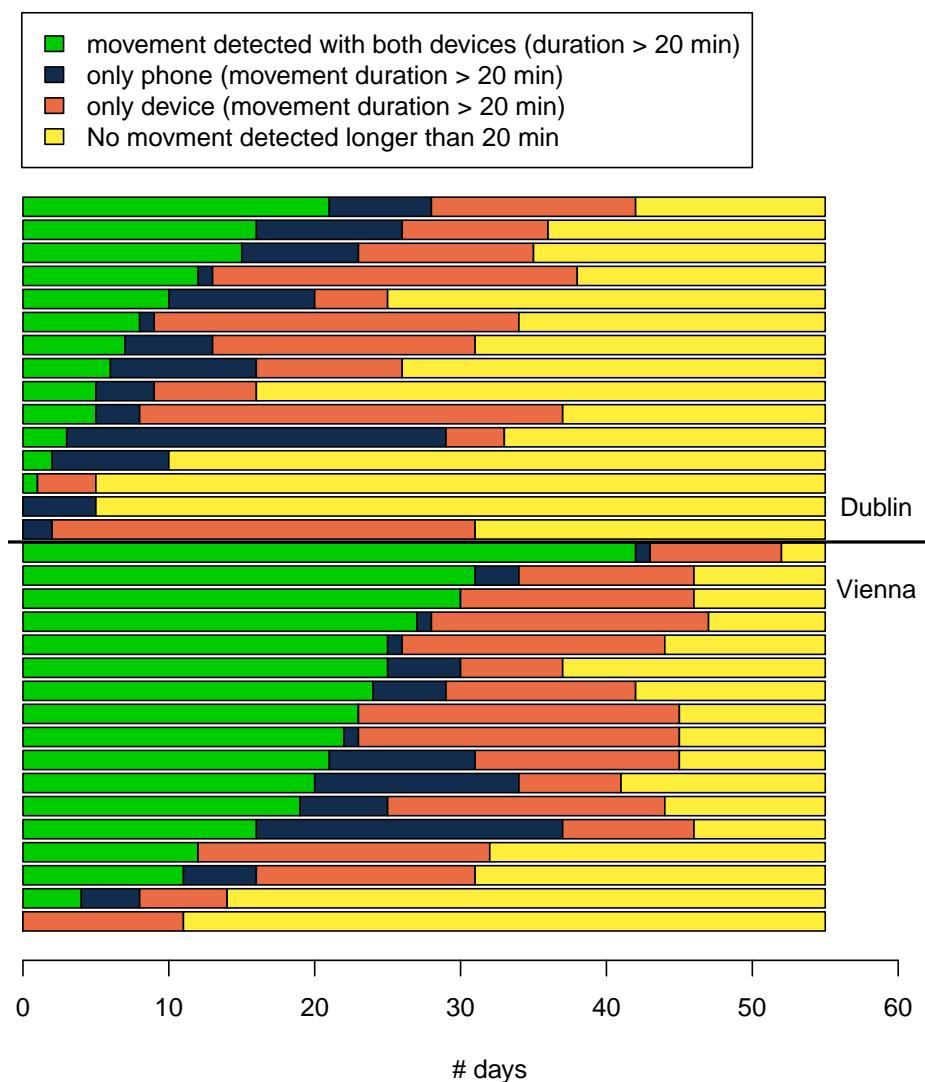


Figure 7: Daily detected duration per person.



5.2 What the users say

Quality of GPS logging

A few users reported that GPS points are missing, which sometimes resulted in strange renderings of routes as straight lines across the city. Several users reported that full days of data are missing. This could be related to the fact that GPS coverage was bad or that data did not get uploaded to the PEACOX server due to connection problems. Also it could be that data was not logged, as a few users reported that logging was not always working even though they opened the app.

Usage

During the introductory workshops it was recommended to have a look at the trip diary app every day as memory is still fresh and it is therefore easier to assess if recorded trips are correct. Despite this request, only a few users stated explicitly during the interview that they use, or at least try to use the prompted recall app every day. Some use it once a week. Others did not use it at all within the first two weeks of the trial. That corrections are not made immediately was also clearly shown above in Figure 4. Some users reported that they could not log in or that no data was showing up.

Quality of trip segment detection

In the previous section it could be seen in the data that there were problematic stage detections within point clouds. This was of course noticed by users, and they complained that the segmentation created a lot of short segments, sometimes just 20 meters long, which is cumbersome to fix manually. This effect was generally reported in situations where users stated they were not actually moving at all. At least one user reported a lot of quick interchanges between different modes within a few minutes, including public transport. Plausibility check are needed to avoid such unrealistic results.

Quality of trip mode and purpose detection

The users' assessment of the quality of the trip mode and purpose detection varied. A few users reported that they had no or almost no wrong trips, some reported a share of around 50 % to 75 % correct modes and activity types. Others again stated that it is very inaccurate and they have to fix a lot. The users estimation correspond well with corrections they made as shown in Figure 3.

Users reported that car trips are detected well, whereas cycling was often detected instead of

walking, public transport, or other. One person reported that neither metro, train nor car was recognised.

Usability and User Experience

Overall users were pleased with the handling of the trip diary app. They described it as easy to use and user friendly. At least one user found the app also interesting for private use to check on the routes one did during a day. For another user on the other hand it was not clear that the diary is prepopulated with trips. A few users reported minor bugs and usability issues, some of them could be solved during the field trials.

Problems with the data quality of course also affected the user experience. Users were informed at the start of the study that detection accuracy rates of 60-80 % can be expected. However, at least one user expressed disappointment with the app, s/he would have expected more accurate results.

In some instances the predefined activity categories were not clear to users. There were a few grey areas, e.g. is going for lunch shopping, leisure, or other? Additional explanations or more activity types to chose from might be beneficial.

Battery drain issues

A common issue with constant GPS logging is a considerable drain on the mobile phone's battery. In order to reduce impact on battery life, a scheduling mechanism was implemented that stopped any logging activity between 22:00 at night and 06:00 the next morning. However, several users reported that this did not work and that they had to turn the logging off manually at night.

In the introductory workshop, participants were advised to keep GPS antenna, the Google location services, WiFi and the PEACOX sensor logging on whenever possible. They should only turn off the background logging service, if the battery was low. Most participants stated that they followed these instructions. Users were already concerned about battery in the introductory workshops, the consensus was that it is acceptable to turn off logging when not moving for some time, e.g., at home or in the office as long as logging is turned on again when leaving the building. Several users implemented this strategy. Still, some users had to turn the logging off a few times (generally 1-3 times), in particular when the phone was used more intensively for other purposes or they knew that they will need the phone later and no power outlet is available.

Also at the start of the study, participants were given the advice to always carry a charger with

them and recharge their phone whenever possible, e.g. in the office (many users did that) or in the car (users reported that this was not always sufficient, depending on the length of the trip). One user even regularly asked customers permission to charge the phone there. Two users used a second battery or a mobile charging device.

Generally battery drain was described as quick and a problem especially when at a place with no charging options, such as on the go or outdoors. Some users reported that they 'could watch the percentages dropping', another user described an up to '2 % drop per minute' when actively using the phone. A few users had to charge their phone several times a day, or even almost constantly plugged in. But still, there were a few users that they did not have to recharge their phone during the day, only once, or only if they use it a lot for other purposes.

One user dropped out of the study because of the battery problems, as the phone was needed for job reasons and s/he was on the move a lot without opportunity to charge the phone.

6 CONCLUSION AND OUTLOOK

It was shown that, as expected, sampling frequencies of smartphones are lower and more diverse than if the same dedicated GPS device is used. In general the potential quality of data collected with smartphones is sufficient e.g. to detect routes, but if different smartphones are used the calibration of the detection routines is a major challenge. Especially, if it is important to collect information about short routes as well, which is often one of the justifications to use GPS in travel surveys. In that case from a researchers perspective it is better to detect more and let users delete the wrongly detected trips and activities, which is much easier than adding trips. Following that reasoning it might even be an option to detect 'no trip' and 'no activity'. This was actually done in this field trial, the activity type classifier was learned on the actual field trial data which includes these options.

Further, data suggests that for most users more data is collected with dedicated devices, as user take them with them slightly more reliably than smartphone apps are turned on. This is due to the heavy use of battery of smartphone logging, which also renders such applications impractical beyond a dedicated study setup. For the sake of the study (and the financial compensation of 150 Euro) users were willing to accept annoyances like carrying a charger with them and the occasional flat battery. However, for large-scale, long-term data collection it is very unlikely users are willing to compromise. Therefore, research to optimise energy consumption of smartphone movement data collection is necessary. Because of these issues - at the moment - if high resolution data is needed dedicated devices are still relevant.

It was also shown, that when letting users correct diaries, uncertainties remain and one is never sure if all is really corrected correctly. On the one hand, many entries are not confirmed at all, on the other hand trip entries are confirmed as being corrected, where obviously too much movement was detected. Also for mode and activity type detection several users made no changes at all, which is very suspicious. Therefore, it is often required that data is processed again or cleaned manually after the end of the survey. The problem-indicators mentioned above can be used as starting points where to clean. In this study particularly, several configuration will be tested for the different smartphone traces, to remove the many outliers that were often detected as cycling. This is necessary before using the data for behavioural analysis but is also educational in itself in order to understand differences between data sets.

Despite the problems with the quantity of corrections, implementation of this task as smartphone application instead of a (paper) diary was successful. Users found the application easy to use. However, depending on the quality of the data, this can introduce additional workload to fix many false detections of trips or activities. While users expressed interest in reviewing their trips to learn about their own mobility habits, they have high expectations in terms of quality of detection. In order not to discourage user activity in correcting detections, it needs to be carefully considered how much workload is put onto users. In any case, an easy to use user interface is essential.

The difference between cities in terms of number of days with movement data, showed that even with passive collection methods of high resolution data a lot of personal contact and effort by survey organisers is needed to ensure high quality and completeness of the generated travel diaries.

The data of the field trial, will be used to analyse changes in behaviour. Further, mode and route choice models could be estimated, as PEACOX is a journey planning app and all suggested as well as the actually selected routes are logged and even tracked.

7 ACKNOWLEDGEMENTS

This research was funded by the European Union as part of the project "PEACOX - persuasive advisor for CO2-reducing cross-modal trip planning" within the Seventh Framework Programme (FP7).

REFERENCES

Alam, S. and A. McNabola (2012) Peacox – door to door emissions models, *Deliverable, D3.1*, PEACOX project, Trinity College Dublin, Dublin.

Artukovic, M., B. Abunaim, L. Montini, E. Bothos, T. Tvrzsky and A. McNabola (2013) Peacox – server and client implementation, *Deliverable, D6.2*, PEACOX project, Fluidtime Data Services GmbH, Vienna.

Auld, J., C. Williams, A. K. Mohammadian and P. Nelson (2009) An automated GPS-based prompted recall survey with learning algorithms, *Transportation Letters*, **1** (1) 59–79.

Bothos, E., D. Apostolou and G. Mentzas (2012) Recommending eco-friendly route plans, paper presented at the *Workshop on Recommendation Technologies for Lifestyle Change*, Dublin, September 2012.

Gould, J. (2013) Cell Phone Enabled Travel Surveys: The Medium Moves the Message, in J. Zmud, M. E. H. Lee-Gosselin, M. A. Munizaga and J. A. Carrasco (eds.) *Transport Survey Methods: Best Practice for Decision Making*, 51–70, Emerald, Bingley.

MGE DATA (2012) MobiTest GSL, webpage, May 2012, <http://www.mgedata.com/de/hw-und-sw-produkte/custom-produkte/mobitest>.

Montini, L., N. Rieser-Schüssler and K. W. Axhausen (2013) Field Report: One-Week GPS-based Travel Survey in the Greater Zurich Area, paper presented at the *13th Swiss Transport Research Conference*, Ascona, April 2013.

Montini, L., N. Rieser-Schüssler, A. Horni and K. W. Axhausen (2014) Trip purpose identification from GPS tracks, *Transportation Research Record*, **2405**, 16–23.

POSDAP (2012) Position Data Processing, webpage, <http://sourceforge.net/projects/posdap/>.

Prost, S., K. Röderer, J. Schrammel, E. Bothos, L. Montini and S. Alam (2013a) Peacox – field trials I report, *Deliverable, D7.4*, PEACOX project, Center for Usability Research and Engineering, Vienna.

Prost, S., J. Schrammel, K. Röderer and M. Tscheligi (2013b) Contextualise! personalise! persuade!: a mobile hci framework for behaviour change support systems, paper presented at the *MobileHCI '13*, 510–515, Munich, August 2013.