Where do you want to go today? -
More observations on daily mobility

Stefan Schönfelder, IVT / ETHZ
Uta Samaga, Technische Universität Dresden

STRC 03 Conference Paper **Updated 24/03/2003**
Session Mobility

Where do you want to go today? – More observations on daily mobility

Stefan Schönfelder   Uta Samaga
IVT / ETHZ   Technische Universität Dresden
Zürich   Dresden

Phone: 01-633 30 92   0049
Fax: 01 633 10 57   0049
e-Mail: schoenfelder@ivt.baug.ethz.ch   uta.samaga@web.de

Abstract

The exploration of the Mobi*drive* six-week travel diary data over the last few years has deepened the understanding of the dynamics, the regularity and the variability in daily travel. The success of the Mobi*drive* project with the implementation of a longitudinal travel survey and the subsequent data analysis, motivated to search for mobility data which covers even longer periods of observation – eventually even capturing the seasonal variations of individual travel demand.

At present, such an innovative long-term travel behaviour data source based on Global Positioning Systems (GPS) information is jointly processed and analysed at the IVT (ETHZ) and ROSO / DMA (EPFL). The Borlänge (Sweden) GPS data covers a monitoring period for privately used cars of up to two years. Travel behaviour and time use research has never had the opportunity to track individuals for longer than some weeks in order to find out about the systematic and the spontaneous part of behaviour.

This paper gives an overview of the current status of the GPS data post-processing work and reports on the suitability of GPS travel data for the analysis of individual activity spaces.

Keywords

Activity spaces – GPS data – travel behaviour – 3rd Swiss Transport Research Conference – STRC 03 – Monte Verità

1. The description and measuring of daily mobility

Transport planning and policy have a notable interest in the mobility patterns of persons and households. Understanding the structures of individual mobility and their connections with the organisation of daily life and the inner-household interaction better allows to design and adjust measures to influence travel behaviour according to the current transport policy priorities. This may include for example the necessary reduction of travel demand in congested areas by land use reorganisation or measures to promote walking or cycling.

For various reasons, the sensitivity of travel behaviour towards supply change (*pull strategies*) as well as demand orientated measures (*push strategies*) has long been tested by collecting and analysing data bases on cross-sectional surveys, only. As a consequence, mobility patterns observed on single days have often been interpreted as optimal decisions of the traveller and as a state of behavioural equilibrium – which is assumed to exist for any point of time and any situation. Travel behaviour research has consistently questioned this working hypothesis (see e.g. Huff and Hanson, 1986; Jones and Clarke, 1988), as daily mobility contains a significant amount of variability and flexibility over time. The stability within travel as well as the deviations from the predominantly routinised structure of daily life can only be revealed by exploring individual panel data, i.e. observations of behaviour of single persons over a prolonged period.

Individual travel data covering periods substantially longer than one week are rare as data collection by means of travel diaries is costly and includes the risk of loosing data precision by fatigue or no-response effects. The 1999-2001 Mobi*drive* project (see Axhausen, Zimmermann, Schönfelder, Rindsfüser and Haupt, 2002) successfully showed that these fears were in the main unjustified. With the implementation of a continuous six-week travel diary as core of the project, a current data set of long-term individual travel behaviour is now available for analysis. The extensive investigation of the data during the last years has led to the development and the adoption of a range of analysis and modelling approaches for daily travel and the stability as well as variability within it (see Zimmermann, Axhausen, Beckmann, Düsterwald, Fraschini, Haupt, König, Kübel, Rindsfüser, Schlich, Schönfelder, Simma and Wehmeier, 2001 for an overview of results). A part of the Mobi*drive* results focusing on locational choice in travel was presented at the 1st STRC (Schönfelder, 2001).

A current analysis stream which is adding to the spatial aspect of the Mobi*drive* data exploration is the visualisation and measuring of human activity spaces (Schönfelder and Axhausen, 2002a; b; Schönfelder and Axhausen, 2003). The identification of revealed individual activity spaces based on a longitudinal observation of trip making and activity performance will increase transport planning's ability to realistically define choice set for destination choice. Furthermore, it allows us to understand the mechanisms of clustering activities around the important pegs of daily life such as home or work.

The Mobi*drive* success in data collection and analysis has increased researchers' interest in accessing long-term travel data bases which eventually cover even longer periods of observation. One technical possibility is the collection of travel behaviour data by in-vehicle or on – person Global Positioning System (GPS) devices. This innovative data collection methodology is promising especially in the field of route choice analysis where exact choice data over prolonged periods is quasi non-existent[1]. The Institute of Transport Planning and Systems jointly with colleagues of the Department of Mathematics at the EPFL is currently exploring the potentials of a vehicle based GPS log data source from Sweden for the purpose of travel behaviour research (Schönfelder, Axhausen, Antille and Bierlaire, 2002). The fully automatically collected data matches the research direction of the Mobi*drive* work mentioned above as it contains movement information for single travellers for up to two years.

This paper gives an overview about both, the recent methodological approaches to capture human activity spaces based on longitudinal travel data and the initial results of the data post-processing steps for the GPS log data. The remainder of the paper is organised as follows: First, a short conceptual introduction into the activity space concept is given in order to outline our underlying research interest. The subsequent chapter reports on the structure of the Borlänge data and the current stage of the necessary data post-processing. Then, a recently developed approach to visualise and to measure activity spaces is applied to the enhanced GPS based travel data. Finally, some notes on the analysis so far and an outlook to the future work conclude the paper.

---

[1] The Danish / European Union research project AKTA / PROGRESS (http://www.progress-project.org) uses a large GPS data set to evaluate road pricing schemes and to detect related changes in route and destination choice.

2. Activity spaces: Background, planning implications and new methodological approaches of measuring

A micro-geographical concept which captures the spatial extent of daily mobility patterns is the *activity space*. The activity space concept – which was developed in parallel with a range of related approaches to describe individual perception, knowledge and actual usage of space in the 1960s and 1970s (see Golledge and Stimson, 1997 for a discussion) – aims to represent the spatial unit which contains the places frequented by an individual over a period of time. Activity spaces are geometric indicators of the *observed or realised* daily travel patterns (see also Axhausen, 2002). This is stressed here as related concepts such as the action space (e.g. Horton and Reynolds, 1971), the awareness space (e.g. Brown and Moore, 1970), the perceptual space (e.g. Dürr, 1979), mental maps (e.g. Lynch, 1984) or space-time prisms (e.g. Lenntorp, 1976) describe the individual *potentials* of travel – based on spatial knowledge, mobility resources, the objective supply of opportunities etc.
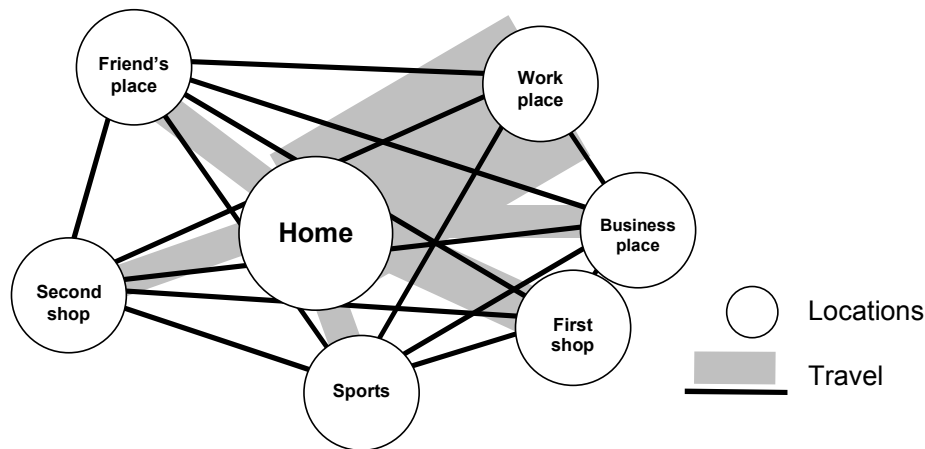
Activity spaces are defined here as a two-dimensional form which is constituted by the spatial distribution of those locations a traveller has personal experience (contact) with. The geometry, size and inherent structure of activity spaces are determined by three determinants (Golledge and Stimson, 1997):

- Home: The position of the traveller's home location, the duration of residence, the supply of activity locations in the vicinity of home and the resulting neighbourhood travel

- Regular activities: Mobility to and from frequently visited activity locations such as work or school

- Travel between and around the pegs: Movements between the centres of daily life travel
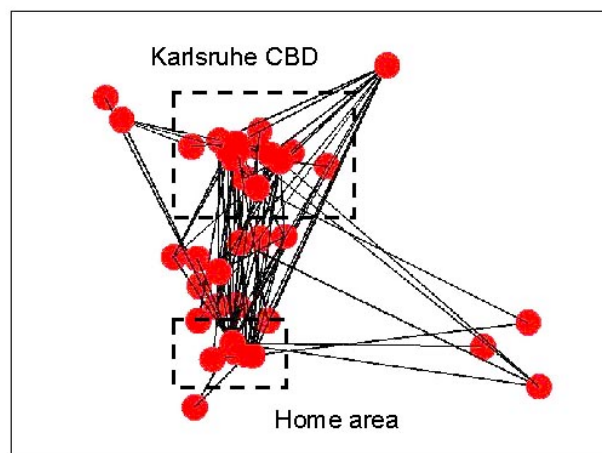
Figure 1 shows a schematic representations of both, the basic concept and a revealed activity space based on the six-week Mobi*drive* travel data.

Figure 1   Simplified activity space representation: Generalisation and actually observed
pattern in Karlsruhe (Mobi*drive*)[*]

a)

b)

a) Adopted from Maier, Paesler, Ruppert and Schaffer (1977) 57
b) Schönfelder and Axhausen, 2003

[*] Dots show observed activity locations of one respondents over six weeks of reporting

In a wider sense, the activity space comprises those locations of which a traveller has personal
experience, as well as those of which the traveller has second hand experiences through fam-
ily, friends, books, films or other media (the knowledge space) (see e.g. Horton and Reynolds,
1971; Dürr 1979 or Goldenberg, Libai and Muller, 2001). In the following, though, activity
space refers only to the first set of locations, those which a traveller has personally visited.

4

For land use as well as transport planning and policy, there exists a range of benefits from a deeper knowledge of individual activity spaces (see e.g. Dangschat Droth, Friedrichs, Heuwinkel and Kiehl, 1980), such as

- better conclusions on the quality and acceptance of the local infrastructure in urban sub-areas, the frequency of usage and the effectiveness of recent infrastructure improvements concerning the change in travellers' behaviour. This may lead to a more demand-tailored planning instead of a focus on pre-defined standards and reference values.

- better assessments of user groups' behaviour and elasticities as well as their affection from general and particular planning measures (e.g. impact opening times, effects of modified infrastructure schemes on time budgets, potentials effects of growing centralisation of facilities)

- improvement of transport modelling techniques, especially by the more realistic design of choice sets for location and route choice.

2.1 Development of approaches based on the Mobi*drive* work

Empirical work on revealed activity spaces, i.e. the physical mapping or enumeration of the places visited by individuals, is rare. Where such work has been done concerning a geometrical representation of personal activity spaces or even the measurement of their sizes, the focus was mostly on travel potentials or opportunities. This was often inspired by the conceptual approaches of space-time geography which puts spatial movement into a context of individual and societal and constraints (Hägerstrand, 1974; Chapin, 1974). Only few studies concentrated on the detailed measurement of individual activity spaces (e.g. Dijst, 1999)[2].

The Mobi*drive* longitudinal data structure which gives us a detailed insight into people's daily travel behaviour has opened up the opportunity for methodological developments and empirical work. As the Mobi*drive* data set contains exact locational data by comprehensive geocoding of most of the reported trips (approximately 40.000 made by approximately 320 respondents), the analysis of the variability in spatial behaviour over time is now possible. The precise locational data was obtained by geocoding the trip destination addresses of all main study trips. The addresses – including home and workplace locations – were transformed into Gauss-Krüger coordinates in a WGS 84 (World Geodetic System) geodetic reference system.

_____

[2] It should be noted that there is a range of studies of spatial behaviour and activity spaces on the aggregate level of sociodemographic groups or zones. Those studies use cross-sectional travel or time-use data.

The methodological development to capture human activity spaces lead to three measures (Figure 2) (see Schönfelder and Axhausen, 2002a; b; Schönfelder and Axhausen, 2003):

- A two-dimensional confidence ellipse (interval) around a suitably chosen centre point.

- The activity space measured by *kernel densities* where again information about the locations visited is used.

- The third approach is based on the idea of *minimum spanning tree (network)*, i.e. the length of the minimum distance routes between the locations visited.
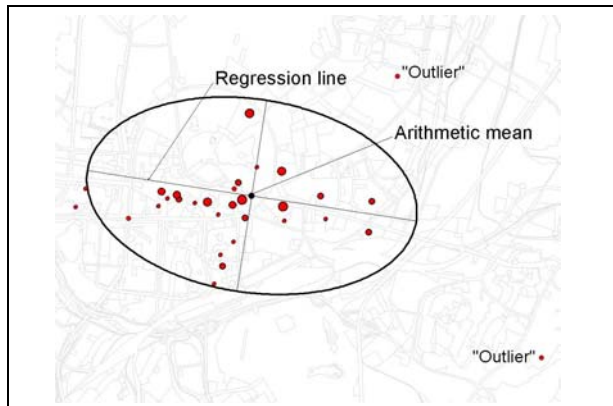
The three approaches to describe the structure and the size of individual activity spaces are models of human behaviour and therefore simplifications of environmental perception and actual decision processes. Nevertheless, they already proved to be powerful for both, visualisation and measuring in initial applications to revealed data. It could be shown that

- the measures are flexible and allow the researchers to chose the input parameters according to the particular analysis interest, such as the interactions between activity location supply and destination choice

- the implementation of the measurement is possible within common GIS software packages.

- the visualisation of examples is straightforward and enables practitioners to gain insight into the travellers' mobility routines.

- the measurement results are consistent with earlier findings on locational choice, such as the positive relationship of the amount of travel and the number of unique locations visited.

Figure 2   Measuring activity spaces: Overview of concepts developed within the Mobi*drive* framework

a) Confidence ellipses

- Basic approach: Probability; smallest possible area in which a defined share of all visited locations is situated

- Measure: Size of area (plus direction of main axis)

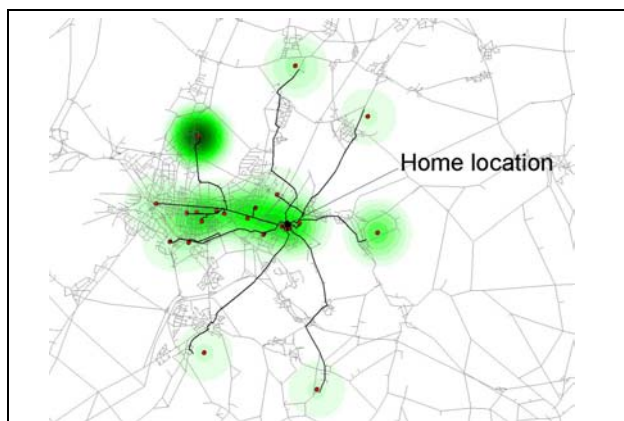- Special feature/quality: Shows dispersion of visited locations

Dots show location and intensity of observed activity locations of one respondent

b) Kernel densities

- Basic approach: Density surface; based on the proximity of activity locations

- Measures: a) Area covered exceeding a certain threshold value, b) "Volume" (sum of all kernel densities calculated)

- Special feature/quality: Represents local clusters / sub-centres within individual activity space

c) Minimum spanning trees (networks)

- Basic approach: Smallest possible geometry based on all observed origin-destination relations

- Measure: a) Length of tree, b) Size of buffered area around the tree indicating potential knowledge spaces

- Special feature/quality: Suitable indicator for the perception of urban space and networks

3. The Borlänge GPS data – an innovative data processing experiment

The visualisation and measuring of revealed activity spaces requires the collection of long-term individual travel data which allows the enumeration of places visited over time. From a methodological point of view, the collection of such data by ordinary travel diary methods remains costly in terms of survey expenses but especially in terms of the burden for the respondents. Travel behaviour research recently raised the question of more suitable survey designs – not the least based on the experiences with the Mobi*drive* data collection and analysis (see Massot, Madre and Armoogum, 2000; Schlich and Schönfelder, 2001).

3.1 GPS data and travel behaviour – a short review

As an additional data collection approach, the use of GPS data for travel behaviour research has been discussed and tested in transport research since mid of the 1990s (see Wolf, Guensler and Bachman, 2001 or Schönfelder *et al.*, 2002 for an overview of studies). The general technical approach used in recent feasibility studies is the combination of mobile GPS data loggers and a Geographical Information System (GIS) (see Draijer, Kalfs and Perdok, 2000 for an example). Principally, vehicles or – in fewer cases – pedestrians or cyclists are equipped with an on-board / on-body data collection system consisting of a GPS receiver, a data storage device with a GIS for mapping all movements and a mobile power supply. For each trip (irrespective of the mode), the respondents switch on the system independently which starts data transmission to the computer (storage) in short intervals of e.g. 10 seconds. After data collection (i.e. the tracking of the travellers), the highly-exact spatial and temporal information is transferred to a conventional PC for processing.

The existing data collection approaches may be categorised as follows (Lee-Gosselin, 2002; Wolf, 2003):

- *GPS based data collection as enhancement of traditional travel diaries*

  In most of the feasibility studies, the portable or in-vehicle GPS / GIS device acts as a supplementary means of collecting exact time, location, route choice data. Hence, the technique substitutes the respective parts of the ordinary travel diary survey and reduces the reporting tasks of the survey respondents. The remaining trip related information, such as trip purpose, number of people travelling together or activity ex-

penses are collected separately either by means of ordinary travel diary forms or electronic data collection devices such as Personal Digital Assistants (PDA).

- *Passive monitoring*

  Within a passive monitoring framework, the travellers are observed automatically without providing any additional information on their trip making (no driver-device interaction). Most of the studies which use passive monitoring are traffic safety driven. The focus of the analysis here is the style of driving respectively the behavioural reaction of the drivers towards external conditions – the rationale for the drive and the activity related to the movement are of a minor interest.

3.2 Borlänge "Rättfart" – the background

Transport psychologists from the universities of Dalarna and Uppsala (Sweden) kindly provided the Institute of Transport Planning and Systems (IVT) with the GPS data set *Rätt Fart* which nicely matches the data requirement for activity space analysis.

The traffic safety project *Rätt Fart* (Right Speed)[3], based in the Middle-Swedish town of Borlänge, is one of the sub-projects of the Swedish National Road Administration initiative approach *Intelligent Speed Adaptation* (ISA). ISA aims to influence car and truck drivers' behaviour by in-vehicle information (see Vägverket, 2000a). The ISA sub-projects are designed to analyse the responses of drivers, ways of integrating interactive technologies into vehicles and the effect of intelligent speed adaptation systems on road safety and the environment.

Rätt Fart in Borlänge had its focus on provision of information for the drivers using GPS devices. The study was conducted from 1999 to 2001 with about 300 private and commercial cars which were equipped with GPS and speed adaptation systems over the period of up to 2 years. Each drive's characteristics such as speed, acceleration, actual time, position etc. was stored internally for analysis in logs every second respectively every tenth second depending to the road link used (see below). The GPS receiver itself did not transmit any signal of its own which prevented external sources to get access to the car's location or other information.

The data logs from the vehicles as well as supplementary data concerning the acceptance of the ISA device are analysed right now in terms of traffic safety by researchers of the two Swedish universities.

_____

[3] See http://www.rattfart.com

The movement file – of which one half with about 100.000 private car trips is ready for analysis – provides accurate trip-specific information such as times, positions, speeds and path choices. The area for detailed monitoring was limited, though, to the town of Borlänge plus some surrounding region – an area with a radius of about 20 km around the town centre of Borlänge.

3.3   Data processing

It is clear that the Borlänge movement data - which is obtained entirely by passive monitoring - lacks a range of essential information ordinary travel surveys usually provide. Apart from the systematic neglect of travelling by modes other than car, there are several data imputation tasks requested in a comprehensive post-processing strategy. This strategy includes the following:

- *Identification of the driver*: The system used in Borlänge did not require the driver to identify himself/herself. With only few exceptions, the private Borlänge vehicles were used by one driver in general, as only households with the same number of drivers and vehicles were recruited. Where uncertainty exists, the detailed records of the driving style (e.g. average speed and acceleration behaviour) may reveal the differences.

- *Identifying unique origins and destinations of travel*: Trip starts and ends are predefined in the Rätt Fart approach by switching on respectively turning off the car (which includes the ISA device). This automatically starts or interrupts the transmission of the drive's characteristics for data storage. However, final positions of drives to the destinations vary significantly due to transmission inaccuracies or simply by using different parking spaces for same destinations over the period of monitoring. Algorithms to categorise unique starting respectively ending points are to be developed.

- *Detecting additional trip ends*: Fortunately, durations during which the vehicle does not move – or in other words, during which zero speed is observed are very rare. Initial investigations of sub-samples showed that both, transmission problems such as "lost" signals and periods of zero-speed concern only about 0.1% of all data logs.

- *Identification of trip purposes:* Together with the identification of the trip-end positions, the assignment of trip purposes for the observed drives is the major post-processing task. The trip purpose is central for travel behaviour analysis as it indicates the reason for travelling and therefor is an important determinant for analyses based on individual utilities (e.g. route choice modelling)

A reasonable data post-processing and the imputation of essential trip attributes such as trip purpose are only possible if additional data sources are integrated in the analysis. The following data sources were acquired:

- *Local / regional road network*

  The available road network for the town of Borlänge comprises all public roads in the monitoring area. The log positions of the GPS trip data are directly connected to and saved as the network link number and sector, but the positions can easily be converted into XY-coordinates. One of the drawbacks of the digital road network is the fact that private roads and minor streets are totally missing. This had also implications for the collection and storage of the trip data as stop positions were not monitored if a link was used which was not included in the data base. In those cases, assumptions about the actual trip ends have to be made based on land-use – trip end matches in a further analysis.

- *Land-use data:*

  In order to conclude the missing purposes from the local land use pattern, point-of-interest data, property and parking space information as well as the official land use plan for the town of Borlänge was acquired.

- *Debriefing data (Sociodemographics of a sub-sample of the test drivers)*

  The GPS data collection period was followed by a debriefing interview (self-administered, paper based, voluntary) which aimed to gain more insight into the sociodemographics of the test drivers and their mobility routines. The requested information included the type of driver (actual test driver or additional driver of the car), sex, age, home address, yearly income, kilometrage of the last year, number of cars in household, main obligatory and leisure activities, main shopping locations, share of non-car travel, access to bicycle, most used modes of getting to particular places, ownership of season ticket and places the test drivers would never go by car.

- *Swedish national travel survey RES*

  The Swedish RES data set provides a rich source of information about travel behaviour of Swedish residents based on a one day recall diary. RES is used as a reference data base to additionally identify missing trip purposed by comparing RES with the structural characteristics of the GPS data.

As the GPS as well as the supporting data is limited in its levels of completeness and exactness, the resulting mobility data set will likely not offer entire comprehensiveness – compared to usual travel diary data. Therefore it needs to define minimum requirements for the quality of the final data set in order to set a framework for the actual analysis steps. The planned structure of the data set and its *quality* of the data is described in following table:

Table 1   Expected structure of the resulting mobility data set

Expected level of availability and precision	Variables
Exact values / high accuracy	Route choice, departure and arrival times, travel times, travel distances, speeds
Acceptable precision	Trip purposes, activity locations (addresses)
Approximations / assumptions	Overall mobility patterns / trip chains including non-car travel
Missing	Activity expenditures, size of company

Analyses implemented so far

This Borlänge data feasibility study is still in an early stage of realisation - nevertheless, several data post-processing steps have been already implemented for a small sub-sample of full-time workers and retirees (N=39) so far:

- *Initial cleaning of raw data*

    The technical shortcomings of the data collection (see above) lead to a large amount of observed trips with unrealistic attributes such as too short trip durations or improbable speeds. The solutions to these problems were already comprehensively discussed elsewhere (Wolf, 2000; Pearson, 2001; Wolf *et al.*, 2001), especially by defining thresholds of (minimum/maximum) durations which initiates elimination rules. As initial thresholds, in this study the following figures were used:

    Delete trips / activities if...

    ... TRIP DURATION (total travel time) smaller than 30 seconds

    ... ACTIVITY DURATION smaller than 120 seconds

    ... observed TRIP DISTANCE is larger than 25 km

    ... calculated average SPEED for the trip is higher than 50 km/h

    Furthermore, the Rättfart data clearly contains journeys outside the Borlänge local monitoring area. Those trips were also eliminated by identifying the last road network link where the vehicle was observed. If this link matches with pre-defined edge links of the GIS road network the trips were erased. This potentially leads to an ambiguous results: On the one hand, it removes stops for the further analysis with no definable / clear destination and activity duration. On the other hand, it "destroys"

complete trip chains which can lead to a misinterpretation of the overall daily activity patterns. This will be shown later and has to be improved in further more sophisticated analysis steps.
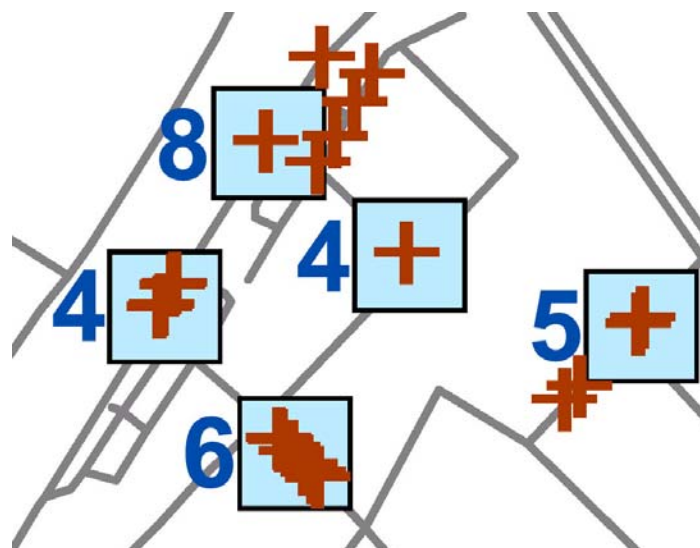
- *Categorisation and filtering*

For most of the travellers, daily mobility differs significantly for weekdays and weekends. This includes for example the fact that potential trip purposes at identical locations may differ according to time or days of week. In order to take these assumptions into account, the stop ends were grouped into the blocks weekdays, Saturdays and Sundays before an assignment of potential trip purposes in the subsequent pre-processing steps.

Filtering of trips which end in known (public) parking lots: As in many cases a parking lot could be directly associated with certain activity purposes such as shopping, a distinction between trips to clearly definable parking facilities and other trips was made.

- *Identification of trip end position:*

A simple way of graphically clustering different stop ends was applied (Figure 3). The distances from each stop position to all others within a radius of 200 m from the respective point were calculated by the GIS. Those stops which have most neighbours (plus the smallest average distance to all other considered points) were classified as cluster centres. All other – less central – stop ends were then assigned to the predefined cluster means. In cases where stop ends were associated with more than one mean point, the nearest cluster centre was chosen as the event destination.

Figure 3   Clustering of observed trip ends (crosses) to unique activity locations (boxes)
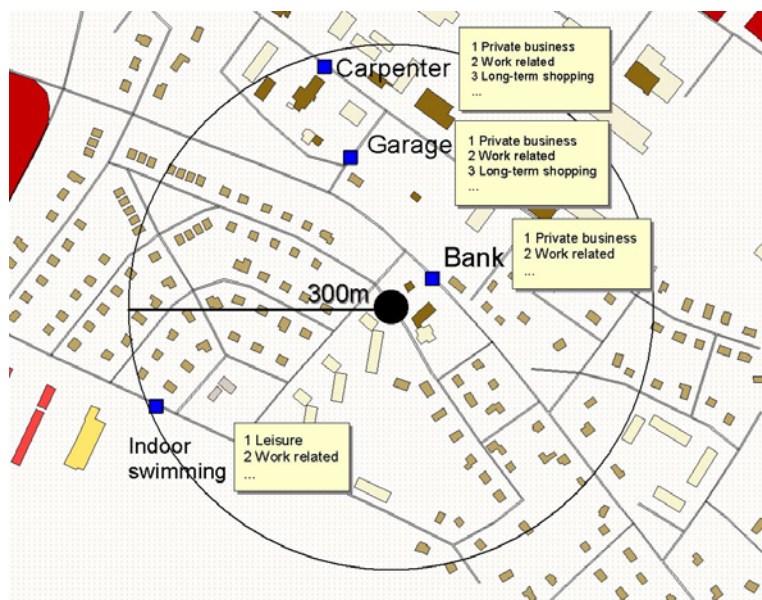
- *Identification of trip purposes:* The assignment of the unknown trip purposes followed a multi-stage approach and focused on a deeper analysis of the cluster centres identified in the preceding step. The approach mainly considered the sociodemographic background information available for the test drivers' and the existing land use data of the town of Borlänge.

  A straightforward identification of the trip purpose was possible when the cluster centre is spatially identical or very close to the driver's household location (<200m). In those cases, the cluster centres as well as the underlying stop ends were assigned with the trip purpose HOME. Although it is possible that other activities are performed in the direct vicinity of home, there is a high probability that the car is driven home if parked nearby. Probable mis-assignments which result from this rigid assumption are accepted.

  As a second step, the cluster centre positions were compared with both, the available point of interest (POI) data such as restaurants, petrol stations etc. and the Borlänge land use pattern (provided in polygons) (see Figure 4). Initially, each POI and land use class was given a certain probability for a potential activity purpose – by a user-defined / consistent tabulation.

Figure 4   Identification of potential trip purposes by land use

For the POI, all cluster centres were buffered by the GIS with a distance of 300 m (distance which is likely to be accepted for a walk from car's parking place to a facility) in order to define a catchment area for the available facilities. If a POI was found within the predefined radius, the related purpose probabilities were assigned to the cluster centre and stored. If even more POI were found, the probabilities were

added up by the distinct purposes with a higher weight for points of interest closer to the cluster centre.

The cluster centre - land use polygon comparison was implemented in a similar way. Again, the cluster centres were buffered (by 200 m radius) to consider the adjoining land use type and the predefined activity purposes in a catchment area. The purpose probabilities (values) of all distinct land uses found were ordered by highest purpose probability and stored.

"Temporal matching": Finally, the structural and temporal characteristics of the stops were compared with the 2000 and 2001 Swedish national travel survey data (Sika Institute, 2001) in order to get information about the trips' rationale. The question behind this approach was: Given a car movement with certain temporal attributes such as a travel time of A and an activity duration of B made by a particular person with the sociodemographic attributes X, Y and Z, which is the most probable trip purpose for this combination (Figure 5)? To find reasonable activity purposes for the Borlänge GPS data, a multi-dimensional table was created out of the Swedish travel data as the comparative data base. The tabulation included the variables sex and car availability of the traveller, his/her occupation status, day of week, trip starting time and the activity duration and yielded for each of the table cells the mode for the activity purpose. This value was assigned to each of the stops as an additional potential purpose.

Figure 5   Exemplary identification of trip purposes by exploring the structural characteristics of the reference data

Sex	Occupation status	Car availability	Weekday	Activity start time	Activity duration [min]		Most probable activity purpose
M	Self-employed	Always	Monday	20.00	60	→	Home
M	Self-employed	Always	Monday	20.00	90	→	Leisure
M	Self-employed	Always	Monday	20.00	150	→	Leisure
M	Self-employed	Always	Monday	20.00	180	→	Leisure
M	Self-employed	Always	Monday	20.00	210	→	Leisure
M	Self-employed	Always	Monday	20.00	240	→	Leisure
M	Self-employed	Always	Monday	20.00	270	→	Leisure
M	Self-employed	Always	Monday	20.00	300	→	Leisure
M	Self-employed	Always	Monday	20.00	330	→	Home
M	Self-employed	Always	Monday	20.00	360	→	Home
M	Self-employed	Always	Monday	20.00	420	→	Leisure
M	Self-employed	Always	Monday	21.00	30	→	Private business
M	Self-employed	Always	Monday	21.00	60	→	Leisure
M	Self-employed	Always	Monday	21.00	90	→	Daily shopping
M	Self-employed	Always	Monday	21.00	150	→	Leisure
M	Self-employed	Always	Monday	21.00	180	→	Leisure
M	Self-employed	Always	Monday	21.00	210	→	Work related

Final assignment of most probable activity purpose: The preceding analysis and comparison steps yielded a range of purpose probability assignments which had to be consolidated in one final activity rationale for each cluster centre. This was done in six hierarchical steps where the group of clusters to be assigned were reduced after each stage by the ones already categorised. The single steps were:

- Clusters which could be associated with the travellers home address were automatically categorised as HOME.

- For fulltime workers, the purpose WORK was assigned if a) cluster centre was the second most frequented of all, b) the comparison with the national travel data was positive for the purpose 'work' and c) the activity took place on a weekday (Monday to Friday)

- Clusters which yielded the same purpose probability by both, land use / POI and temporal characteristics comparison received the agreed purpose.

- If there existed a difference between the POI / land use based purpose assignment and the temporal matching, the POI / land use categorisation was preferred. It was deviated from this rule if the national travel survey tabulation yielded the activity purpose PICK UP / DROP OFF which is principally independent of any land use at the point of activity. In this case, the temporal matching was given priority.

- If there was no clear POI / land use assignment possible, the ultimate purpose assignment followed the categorisation based on the national travel survey grouping.

The steps were entirely implemented in the ARCINFO GIS environment using the ARC Macro language (AML) (see Samaga, 2003). The whole procedure was designed to allow easy modification of the imputation framework by adding more sophisticated identification steps such as for example *discriminant analysis* for the temporal matching or a Bayesian updating of the probabilities.

The resulting sample data base

The small test sample which was extracted for the initial imputation of the missing trip information consists of 28 fulltime workers and 11 retirees (Table 2). As sociodemographic information such as exact home and work address, temporal routines etc. was not available for all Rätt Fart survey participants, those test drivers were chosen who provided sufficient information for the imputation work.

Table 2   Test sample characteristics

Characteristics	Minimum	Maximum
First / last date of monitoring	9 October 2000	14 August 2001
Number of monitored days (including immobile days)	24	231
Mobile days	19	208
Observed trips	66	1185

The imputed GPS trip data does not entirely correspond with available cross-sectional travel data which is not illogical if considering the unique longitudinal structure of the Rätt Fart data, the limitations of the small test sample and the ad-hoc extraction of the reference data base. Whereas the number of car trips is consistent with the information provided by the Swedish national travel survey (RES), the Borlänge test drivers made considerably shorter drives – concerning the trip distance as well as the trip duration (Table 3). At this stage of the analysis, it is difficult to assess if the GPS data collection method yields systematically different results compared to ordinary paper based travel diaries or if the recruited test drivers in fact show a dissimilar travel behaviour including a different structure of daily car usage. Further detailed investigations will give more insights.

Table 3   Post-processed Borlänge GPS data base compared with 2000/2001 RES data (Swedish national travel survey) [*]: Selected mobility charcateristics

Variable	Retirees				Fulltime workers			
	RES		Borlänge GPS		RES		Borlänge GPS	
Mean number of daily trips (Std.)	3.8	(3.8)	4.3	(2.8)	4.8	(5.7)	3.9	(2.4)
Mean daily trip distance [km] (Std.)	24.6	(33.7)	16.3	(12.5)	32.0	(45.3)	13.2	(11.3)
Mean daily trip duration [min] (Std.)	44.1	(55.2)	31.8	(22.4)	58.7	(82.1)	24.4	(18.2)

[*] Local car trips made by respective groups; Sample sizes: RES weighted by sex and age, N(Fulltime workers) = 1516, N(Retirees) = 440; Borlänge: N(Fulltime workers) = 28, N(Retirees) = 11

The shares of the single purposes (10 categories) principally show the same pattern as the (weighted) Swedish reference data. The main problems are the differences in the shares of the PRIVATE BUSINESS, WORK RELATED BUSINESS and DAILY SHOPPING purpose. This may be due to many things. For example, the areal land use information and the available point of interest repertoire which are important bases for the trip purpose assignment procedure allow an ambiguous and partly fuzzy categorisation of the detected trip ends. Buildings with a certain land use categorisation may be visited by the test drivers because of several reasons – for example, a bank can act as work place, a place to do private business or a location where family members may be picked up or dropped off. Even if integrating the temporal characteristics of the respective trip (e.g. start time or activity duration), the trip purpose assignment may be misleading in some cases. Furthermore, the land use data base as well as the points of interests available appears to be insufficient at this stage and have to be completed in the next imputation steps. Finally, it should be noted that the test sample is extremely small and there contain considerable biases.

Table 4   Initial Borlänge GPS trip purpose assignment compared with 2000/2001 RES data
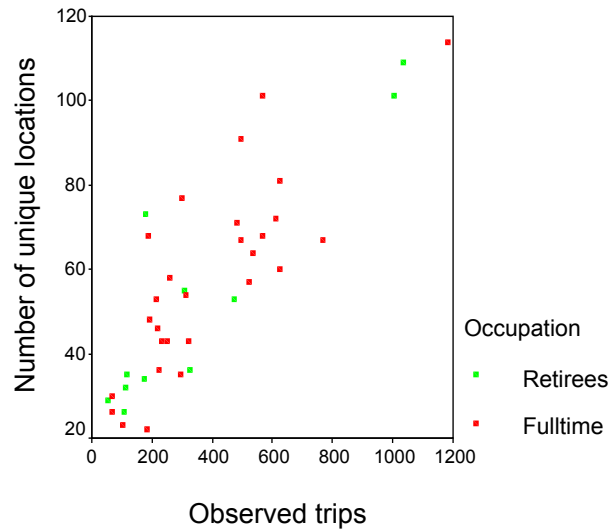(Swedish national travel survey data): Trip purpose shares [%]

Trip purpose	Retirees		Fulltime workers	
	RES	Borlänge GPS	RES	Borlänge GPS
Pick up / Drop off	6.8	7.0	8.9	8.6
Private business	4.6	10.0	3.7	8.2
Work related	0.1	9.8	8.3	5.7
School	-	0.1	0.0	0.1
Work	0.3	-	16.0	10.2
Daily shopping	12.4	4.4	6.1	1.5
Long-term shopping	8.6	7.3	5.7	7.9
Leisure	20.7	23.6	10.6	20.4
Other	5.2	-	4.1	-
Home	41.3	37.8	36.6	37.4

[*] Local car trips made by respective groups; all days; RES weighted by sex and age,
N(retirees) = 1516, N(fulltime workers) = 440; Borlänge: N(Fulltime workers) = 28,
N(Retirees) = 11

Another consistency check of the trip purpose assignment can be done by having a closer look
at the daily level of activity performance. Figure 6 shows the aggregated activity pattern of a
selected Borlänge test driver over the days of the monitoring period. Each dark box represents
the performance of the given activity category or immobility on the respective monitoring day
(ordered by the day of the observation period). Here, it does not matter if the activity was per-
formed once or several times per day. It can be seen that the weekday blocks with the activity
WORK were clearly identified for this test driver which here matches with his/her description
of daily life and the common working days assumptions. This visual analysis reveals also the
weaknesses of the initial imputation process and the quality of the GPS data itself: As there
are a considerable amount of mobile days for which there was no way back home identified –
indicating incomplete daily trip chains- there seems to be a mistake in the initial data cleaning
(see above), the trip purpose assignment or the consistency of the data collection and storage
itself. For some of the test drivers, about 25% of the monitored days do not have a complete

trip chain with a final trip to home. Further efforts have to be done to eliminate this short-coming.

Figure 6   Activity pattern of one selected Borlänge test driver over the monitoring period (Full-time worker)

Nevertheless, the imputed Borlänge trip data confirms earlier findings of the Mobi*drive* analysis work on long-term spatial mobility. This is true for both, the number of unique locations the test drivers go to and the amount of variety seeking in locational choice. The share of trips to unique locations was previously unknowable, as cross-sectional surveys cannot provide a credible estimate of this parameter. The available long-term travel data now permit an impression of this aspect of spatial choice behaviour (Figure 7). If the number of unique locations grows consistently with the number of trips, then variety seeking, for its own sake, becomes a credible explanation of these choices.
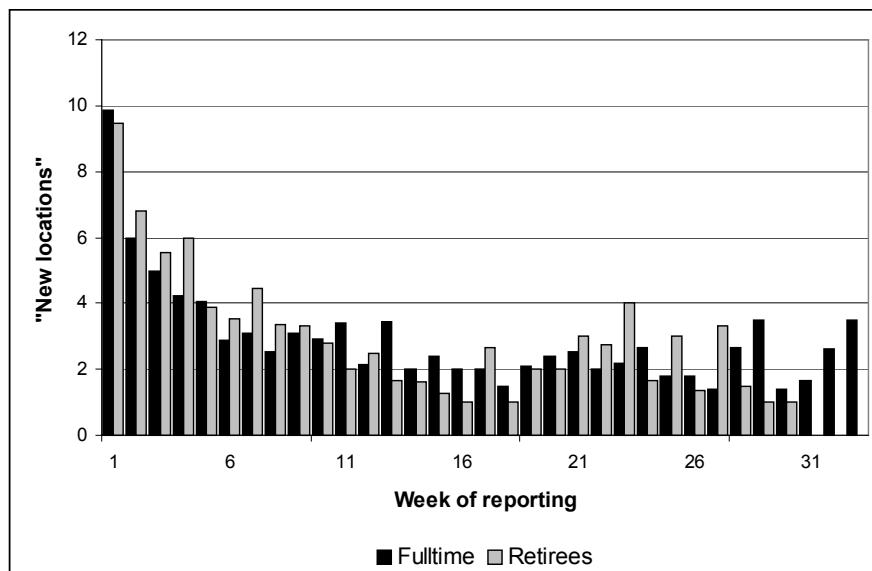
Figure 7   Relationship between amount of travel and unique locations visited

The Borlänge data indicates that there seems to be an almost unlimited number of places peo-
ple know, because even after half a year of reporting there are still places the test drivers "dis-
cover" as new (i.e. never visited before) destinations (Figure 8).

Figure 8   Variety seeking in locational choice – Amount of locations which have been never
visited before over the monitoring period

Summing up these findings briefly, two aspects should be stressed:

1) The post-processing of the passively monitored trip data which aims to provide essential information for travel behaviour analysis has so far provided mainly consistent but also few ambiguous results. There is need for a refinement of the methodology and especially for an enlargement of the reference data base for identifying the trip purposes (e.g. further point-of-interest data).

2) Although the data base is restricted to local car travel only and there is space for methodological improvements, the resulting data base already shows potentials for intense spatial and temporal investigation of daily life travel. As Mobi*drive*, the data set reveals the variability and stability in trip making and activity performance over time.

4. Applications of the activity space concepts to the post-processed Borlänge data

Turning finally to our actual research interest, i.e. the description of realised activity spaces, a recently developed measure is applied here to the imputed GPS travel data. As an example, the kernel density approach is chosen which considers activity space as an area of activity / trip intensity (density).

The intensity area is estimated using a *kernel density* approach. The basic process behind the estimation of kernel densities is a transformation of a point pattern (such as the set of activity locations visited) into a continuous representation of density in a wider area. Generally spoken, the estimation is an interpolation or smoothing technique which generalises events or points to the area they are found in. The interpolation then leads to a calculation of a value for any point, cell or sub-region of the entire area which characterises the density.
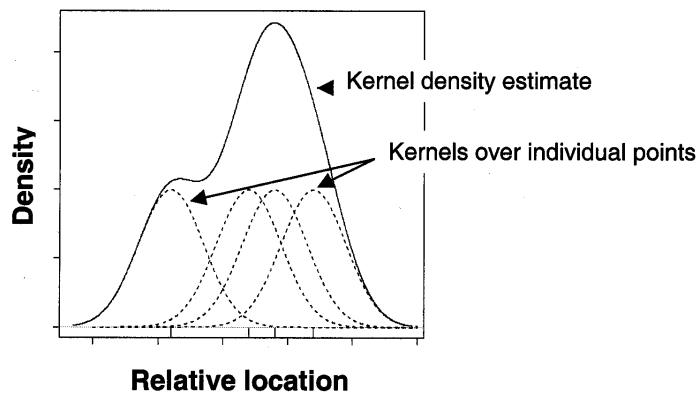
Kernel densities have been already applied successfully to large cross-sectional data sets (Kwan, 2000; Buliung, 2001). Modern GIS applications include tools to calculate such density measures effectively[4]. The density values are assigned to the cells according to the kernel densities estimated for the underlying point pattern.

For the actual density estimation, a variety of approaches exist (for overviews see Silverman, 1986 or Fotheringham, Brunsdon and Charlton, 2000). Probably the most common approach is the *fixed kernel method* (also applied here). A symmetrical – variably distributed – *kernel function* is placed over each data point (Figure 9). For all locations in the entire area – not only for the data points – the overlapping values are summed which yields the density or intensity estimate. This automatically leads to a smoothing of the surface where the level of smoothness depends on the *bandwidth* of the kernel function which is analogous to the width of ordinary histogram boxes. The bandwidths may be varied according to the necessary degree of smoothness – with greater smoothing at bigger bandwidths or values of the smoothing parameter. The GIS finally may represent the resulting estimates for all grid cells as a continuous surface.

---

[4] ArcInfo was also used in this analysis.

Figure 9   Kernel estimate showing individual kernels

Source: Adopted from Silverman (1986) 45

Considering a grid structure in which single points are substituted by grid cells, the base kernel density is given by:

$$\lambda(s) = \sum_{d_i < \tau} K\left(\frac{d_i}{\tau}\right)$$

with

λ density estimate at grid point s

τ bandwidth or smoothing parameter

K kernel function (to be further specified)

d_i distance between grid point s and the observation of the ith event

The kernel function K itself may have different forms such as normal, triangular or quartic. The results do not differ significantly as long as the distribution is symmetrical. In the following, a quartic kernel function (see Mitchell, 1999 for details) – is used which leads to the following kernel density

$$\lambda(s) = \sum_{d_i < \tau} K \left\{ [\frac{3}{\tau^2 \pi}][1 - \frac{d_i^2}{\tau^2}]^2 \right\}$$

A particularity of the quartic function – e.g. compared to a normal distribution – is that outside the specified bandwidth τ, the function is per definition set to zero – with implications for the behavioural model. This means that activity locations outside a specified radius do not contribute to the density estimation of the particular point (cell) in space. In other words, a quartic distribution of the kernel function adds weight to locations closer to the centre of the bandwidth than those further apart (see NedLevine and Associates, 1999 for characteristics of the different kernel forms).

Figure 10 shows a visualisation of kernel densities for one Rätt Fart test driver. The observation period covers 4 months form April to August 2001. It shows nicely the variation in size and structure of the locational choice patterns over the week with a large activity space during the weekdays, similar densities for weekdays and Saturdays in the city centre area and around the home location as well as the reduced local travel intensity on Sundays.

Figure 10   Visualisation example: Activity space compared by days of the week[*]

[*] Retiree, male, 71years; observation period: app. 4 months; grid cell size: 500*500 meters, activity locations weighted by frequency of visit, bandwidth: 1000 m

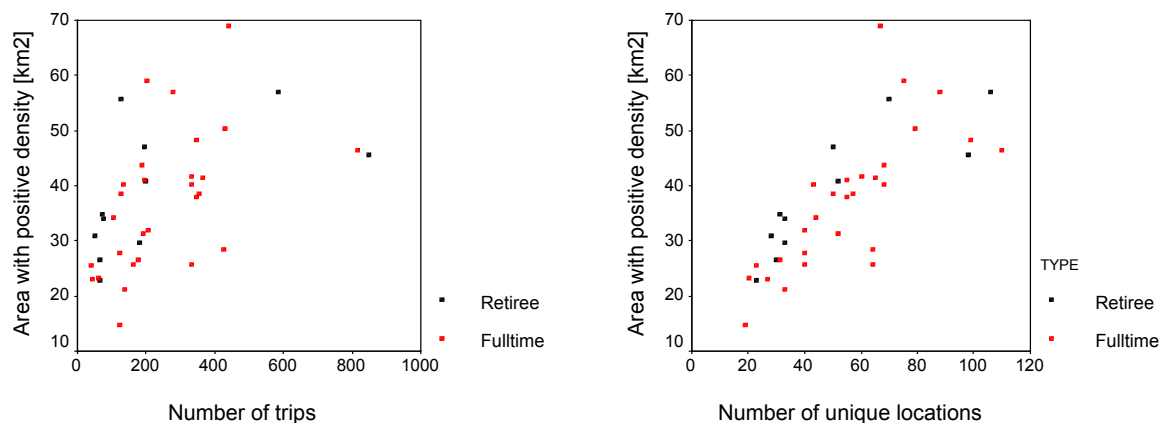Turning to the measuring method, the size of the activity space can be measured as:

- the number of cells for which the density exceeds a certain threshold (i.e. > 0)

- the intensity of space usage given by the sum of the densities for all grid cells.

In the Mobi*drive* analysis, the latter measure turned out to be problematic if – as shown in the figure above - the activity locations are "weighted" by the frequency of visit. In this case, the results highly correlate with the trip frequencies, i.e. the overall amount of mobility on the

particular days. One could ask if an indicator for the size of individual activity spaces is useful which tells us that the structures of spatial mobility are tied to the pure amount of travel? At the same time, though, the measure and the outcome of the investigation strongly confirms our expectations. It indicates that the usage as well as the up-to-date knowledge or urban space is a function of the amount of contact a traveller has.

Figure 11 shows the calculation results based on the first measure, i.e. aggregated area of all cells for which the density exceeds a certain threshold 0. The results are related to the amount of trips and the number of unique locations over the GPS monitoring period. For the calculation of the kernel densities a GIS grid cell size of 500*500 meters was chosen with a kernel bandwidth (radius) of 1000 meters. The results again indicate that there is a strong relationship between the amount of local travel and the size of the activity space. Due to the small sample size there is no consistent tendency to be found for the differences between the two sociodemographic groups. The issue of activity space size influencing factors, such as sociodemographics or location of households with the urban area will be highlighted in future investigations.

Figure 11   Area with positive kernel density (locations not weighted by number of trips to places visited) [km^2]

5. Concluding remarks

The results of the initial data processing steps have shown that the usage of the Borlänge GPS data set for travel behaviour analysis is promising but connected with a range of inherent data difficulties. The challenges for the future methodological refinement within the post-processing procedure are:

- a more sophisticated methodology to beforehand detect and improve the structural inaccuracy of the available GPS data, i.e. the elimination of falsely reported stops and the imputation of missing trips (e.g. to home)

- the development of more sophisticated approaches to assign trip purposes to the stop ends of the GPS monitored trips; this includes the completion of the reference data base (land use, point-of-interest data etc.), the incorporation of individual trip making and activity performance regularities as a further step of the multi-stage approach and

Nevertheless, it seems worthwhile to consider the GPS data base as input data in models of spatial behaviour. The research work adds substantially to the expertise on the implementation of passive monitoring for travel behaviour surveys. From a methodological point of view, the usage of the longitudinal data base for mobility analysis is unique and will enrich travel behaviour research's tool box on data processing and analysis.

Furthermore, the Borlänge data set will provide us with unique findings on the variability of revealed human activity spaces. Even considering the limitations of the data set with car travel only, the knowledge about spatial navigation and usage will be significantly improved. This is also true for the issue of regularity in locational choice which has not yet been discussed deeply in travel behaviour research. Apart from the geometrical representation of human activity spaces and the measuring of their sizes, the following analysis will concentrate more deeply on the inherent structures such as the activity and travel times within subregions of the activity space, the clustering of activities around the pegs of daily mobility and the interactions of activity densities with infrastructure supply.

6. References

Axhausen K.W., A. Zimmermann, S. Schönfelder, G. Rindsfüser and T. Haupt (2002), Observing the rhythms of daily life: A six-week travel diary, *Transportation,* **29** (2) 95-124.

Axhausen, K.W. (2002) A dynamic understanding of travel demand: A sketch, *Arbeitsbericht Verkehrs- und Raumplanung*, **119**, IVT, ETH, Zürich.

Brown, L.A. and E.G. Moore (1070) The intra-urban migration process: a perspective, *Geografiska Annaler*, **52**(B) 1-13.

Buliung, R.N., Spatiotemporal Patterns of Employment and Non-Work Activities in Portland, Oregon, ESRI International User Conference, San Diego, California, July, 2001.

Chapin, F.S. (1974) *Human Activity Patterns in the City*, Wiley & Sons, New York.

Dangschat, J., W. Droth, J. Friedrichs. D. Heuwinkel and K. Kiehl (1980) Aktionsräumliche Forschung – Ergebnisse und Planungsrelevanz, *ARL-Arbeitsmaterial*, **45**, 3-14, Akademie für Raumforschung und Landesplanung, Hannover.

Dijst, M. (1999) Two-earner families and their action spaces: A case study of two dutch communities, *GeoJournal*, **4** (3) 195-206.

Draijer, G., N. Kalfs and J. Perdok (2000) Global Positioning System as data collection method for travel research, *Transportation Research Record*, **1719**, 147-153.

Dürr, H. (1979) Planungsbezogene Aktionsraumforschung – Theoretische Aspekte und eine Pilotstudie, *Beiträge der Akademie für Raumforschung und Landesplanung*, **34**, Hannover.

Fotheringham, A.S., C. Brunsdon und M. Charlton (2000) *Quantitative Geography*, Sage, London.

Goldenberg, J., B. Libai and E. Muller (2001) Talk of the network: A complex systems look at the underlying process of word-of-mouth, *Marketing Letters*, **12**(3) 209-221.

Golledge, R.G. and R.J. Stimson (1997) *Spatial Behavior*, The Guilford Press, New York/London.

Hägerstrand, T. (1970) What about people in regional science?, *Papers of the Regional Science Association*, **24** (1) 7-21.

Horton, F. und D.R. Reynolds (1971) Effects of Urban Spatial Structure on Individual Behaviour, *Economic Geography*, **47**, 36-48.

Huff, J.O. und S. Hanson (1986) Repetition and variability in urban travel, *Geographical Analysis*, **18** (2) 97-114.

Jones P.M. and M.I. Clarke (1988) The significance and measurement of variability in travel behavior, *Transportation*, **15** (1) 65-87.

Kwan, M.-P. (2000) Interactive geovizualization of activity-travel patterns using three-dimensional geographical information systems: A methodological exploration with a large data set, *Transportation Research C*, **8**(2), 185-203.

Lee-Gosselin, M. (2002) Some reflections on GPS-supported travel survey methods in an increasingly ICT-rich environment, Presentation at Workshop on ICT, Innovation and the Transport System, Arlington, January 2002.

Lenntorp, B. (1976) Paths in space-time environment: a time geographic study of possibilities of individuals, *Lund Studies in Geography, Ser. B. Human Geography*, **44**, Department of Geography, The Royal University, Lund.

Lynch D. (1960) *The Image of the City*, MIT-Press, Cambridge.

Maier, J., R. Paesler, K. Ruppert and F. Schaffer (1977) *Sozialgeographie*, Westermann, Braunschweig.

Massot, M.-H., J-L. Madre and J. Armoogum (2000) Monthly frequency versus previous day description of trips, What information is needed on urban mobility? Paper prepared for the 9th International Association of Travel Behaviour Research Conference, Gold Coast/Australia, July 2000.

Mitchell, A. (1999) *The ESRI Guide to GIS Analysis, Volume 1: Geographic Patterns & Relationships*, ESRI Press, Redlands.

Ned Levine and Associates (1999) CrimeStat (Version 1.0), A spatial statistics program for the analysis of crime incidents locations, The National Institute of Justice, Washington.

Pearson, D. (2001) Global Positioning System (GPS) and travel surveys: Results from the 1997 Austin Household Survey, Paper presented at the Eighth Conference on the Application of Transportation Planning Methods, Corpus Christi, Texas, April 2001.

Samaga, U. (2003) Entwicklung von GIS-Funktionen zur Analyse und Visualisierung GPS-basierter Mobilitätsdaten. Diplomarbeit an der Technische Universität Dresden, Technische Universität, Dresden.

Schlich, R. and S. Schönfelder (2001) Zwei Sichten der Aktivitätenteilnahme: Tagebücher und Aktivitätenhäufigkeitsbögen, *Stadt Region Land*, **71**, 147-164.

Schönfelder, S. (2001) Some notes on space, location and travel behaviour, Paper presented 1. Swiss Transport Research Conference, Ascona, March 2001.

Schönfelder, S. and K.W. Axhausen (2002a) Measuring the size and structure of human activity spaces: the longitudinal perspective, *Arbeitsbericht Verkehr- und Raumplanung*, **135**, IVT, ETH, Zürich.

Schönfelder S. and K.W. Axhausen (2002b) Activity spaces: Measures of social exclusion? *Arbeitsbericht Verkehrs- und Raumplanung*, **140**, Institut für Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.

Schönfelder, S. and K.W. Axhausen (2003) On the variability of human activity spaces, *Arbeitsbericht Verkehrs- und Raumplanung*, **149**, Institut für Verkehrsplanung und Transportsysteme (IVT), ETH Zürich, Zürich.

Schönfelder, S., K.W. Axhausen, N. Antille and M. Bierlaire (2002) Exploring the potentials of automatically collected GPS data for travel behaviour analysis – A Swedish data source, in J. Möltgen, and A. Wytzisk (Eds.) GI-Technologien für Verkehr und Logistik, *IfGIprints*, **13**, Institut für Geoinformatik, Universität Münster, Münster, 155-179.

Sika Institut (2001) RES 2000, Den nationella resundersökningen, Statistika Centralbyrån, Stockholm.

Silverman, B. W. (1986) *Density Estimation for Statistics and Data Analysis*, Chapman and Hall, London.

Vägverket (2000a) ISA Intelligent Speed Adaptation, Vägverket, unpublished, Vägverket, Borlänge.

Wolf, J. (2000) Using GPS data loggers to replace travel diaries in the collection of travel data, Dissertation, Georgia Institute of Technology, School of Civil and Environmental Engineering, Atlanta, Georgia.

Wolf, J. (2003) Tracing people and cars with GPS and diaries: Current experiences and tools, Presentation at the IVT-Seminar, ETH, Zürich, Februrary 2003.

Wolf, J., R. Guensler and W. Bachman (2001) Elimination of the travel diary: An experiment to derive trip purpose from GPS travel data, Presentation at TRB 80th Annual Meeting of Transportation Research Board, Washington.

Zimmermann, A., K.W. Axhausen, K.J. Beckmann, M. Düsterwald, E. Fraschini, T. Haupt, A. König, A. Kübel, G. Rindsfüser, R. Schlich, S. Schönfelder, A. Simma and T. Wehmeier (2001) Mobi*drive*: Dynamik und Routinen im Verkehrsverhalten: Pilotstudie Rhythmik, Report to Bundesministerium für Forschung und Technologie, PTV AG, Institut für Verkehrsplanung, Transporttechnik, Strassen- und Eisenbahnbau (ETH Zürich) and Institut für Stadtbauwesen (RWTH Aachen), Karlsruhe, Zürich and Aachen.