Designing Travel Behaviour Change interventions: A spatiotemporal perspective

by
Justin Tycho van Dijk

Dissertation presented for the Degree of Doctor of Philosophy in Transport Economics in the Faculty of Economics and Management Sciences, at Stellenbosch University

The financial assistance of the National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the author and are not necessarily to be attributed to the NRF.

Supervisor: Professor Stephan Krygsman
Co-Supervisor: Professor Hannelie Nel

December, 2017

Declaration

By submitting this thesis electronically, I declare that the entirety of the work contained therein is my own, original work, that I am the sole author thereof (save to the extent explicitly otherwise stated), that reproduction and publication thereof by Stellenbosch University will not infringe any third party rights, and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Signature: Justin Tycho van Dijk
Date: December, 2017

Declaration by the candidate

With regard to Chapter 4, the nature and scope of my contribution were as follows:

Nature of contribution	Extent of contribution (%)
Background research, conceptualised idea for research, research, data analysis, lead author writing up of article.	75%

The following co-authors have contributed to Chapter 4:

Name	E-mail-address	Nature of contribution	Extent of contribution (%)
Tom de Jong	t.dejong@uu.nl	Data preparation, conceptualised idea for research, data analysis, external advisor.	25%

Declaration by co-author

The undersigned hereby confirm that:

1. the declaration above accurately reflects the nature and extent of the contributions of the candidate and the co-authors to Chapter 4.
2. no other authors contributed to Chapter 4 besides those specified above, and
3. potential conflicts of interest have been revealed to all interested parties and that the necessary arrangements have been made to use the material in Chapter 4 of this dissertation.

Signature	Institutional affiliation	Date
Declaration with signature in possession of candidate and supervisor.	Utrecht University, the Netherlands	July 23, 2017

Abstract

Against the background of unprecedented growth in private vehicle ownership and the entrenchment of the private car in everyday life, the past decades have seen a growing and ongoing academic and policy debate on how to encourage individuals to change to more sustainable ways of travelling; for instance, with voluntary travel behaviour change (VTBC) interventions. VTBC interventions aim to alter travel behaviour by providing information. In recent years, a large body of research has focused on the evaluation of the effectiveness of these programmes. However, no consensus has been reached on the question of whether a broad implementation of VTBC programmes is effective in stimulating people to use more sustainable ways of travelling. This dissertation argues that location-aware technologies, particularly GPS-enabled smartphones, could potentially augment the research on VTBC interventions. Smartphones can not only source data (such as place and time of travel or activity) but can also provide individuals with real-time information, feedback, and suggestions for alternative behaviour or travel options.. However, between sourcing the data and relaying feedback to individual commuters, significant research is required on how to obtain, clean, and interpret the data, as well as on how to account for individual spatiotemporal accessibility.

GPS data need to be collected and analysed systematically; especially in the context of evaluating the effectiveness of VTBC interventions in which effect sizes are known to be small and inconspicuous. As such, the translation of raw GPS trajectories into activity episodes and the best estimation of a travelled route are pivotal. Methods of activity recognition were explored with advanced machine learning algorithms, and two approaches for identifying travelled routes were proposed. Furthermore, it was demonstrated how spatiotemporal measurements could aid the design of VTBC interventions. Attention was drawn to the time-geographical concepts of activity spaces and potential path areas. Based on the examination of GPS tracks with different two-dimensional operationalisations of activity spaces, it was found that the density of opportunities within an activity space is related to the size of the activity space: larger activity spaces have lower densities of opportunities than smaller activity spaces. This may suggest that individuals who have a low opportunity density are less likely to respond to external stimuli and/or awareness programmes than individuals who have a high opportunity density. In turn, potential path areas were used to establish to what extent individuals have different spatiotemporal opportunities that will enable behavioural change in travel and activity.

The findings indicate that location-aware technologies hold great potential to supplement transport geographical-research. Moreover, the results show that the incorporation of spatiotemporal measurements is crucial to consider for the design, implementation, and evaluation of VTBC interventions. The added value of seemingly new technologies, such as GPS, is that they can be easily integrated into a larger spatiotemporal framework of analysis. However, one has to be careful not to consider GPS as a panacea, because GPS data and technology also have some drawbacks. Careful consideration should go into application development, sample selection, site selection, and data imputation.

Opsomming

Teen die agtergrond van 'n ongekende groei in privaatvoertuigeienaarskap en die volledige integrering van die private motor in die alledaagse lewe, het die afgelope dekades 'n groeiende en volgehoue akademiese- en beleidsdebat gesien oor hoe om individue aan te moedig om oor te skakel na meer volhoubare maniere van reis, met onder meer intervensies om vrywillige reisgedragverandering (VRGV) te stimuleer. VRGV-intervensies poog om reisgedrag te verander deur inligting aan motorgebruikers te verskaf. 'n Groot deel van die huidige navorsing fokus op die evaluering van die effektiwiteit van hierdie programme. Daar is egter steeds geen konsensus oor of 'n breë implementering van VRGV-programme effektief is om mense te stimuleer om meer volhoubare maniere van reis te gebruik nie. Hierdie proefskrif beweer dat plekbewuste tegnologieë, veral Globale Posisionering Stelsel (GPS)-geaktiveerde slimfone, moontlik die navorsing oor VRGV -intervensies kan ondersteun. Slimfone kan nie net data versamel (soos plek en tyd van reis of aktiwiteite) nie, maar kan ook individue intyds van inligting voorsien en terugvoering en voorstelle gee vir alternatiewe gedrag en vervoeropsies. Tussen die verkryging van die data en die terugvoering van inligting aan pendelaars, word heelwat navorsing benodig oor hoe om die data te verkry, skoon te maak en te interpreteer, asook hoe om rekening te hou met individuele ruimtelike-temporale toeganklikheid.

GPS-data moet versamel en sistematies ontleed word, veral in die konteks van die evaluering van die effektiwiteit van VRGV intervensies waarin effekgroottes dikwels klein en onopsigtelik is. As sodanig is die verwerking van rou GPS-trajekte in aktiwiteite-episodes en die beste beraming van roetekeuse noodsaaklik. Metodes van aktiwiteitsherkenning is ondersoek met gevorderde masjienleeralgoritmes en twee benaderings vir die identifisering van gekose roete is voorgestel. Verder is getoon hoe ruimtelike-temporale metings die ontwerp van VRGV-intervensies kan help. Aandag is gevestig op die tyd-geografiese konsepte van aktiwiteitsruimtes en potensiële roete areas. Op grond van die ondersoek van GPS-data met verskillende twee-dimensionele operasionele aktiwiteitsareas, is bevind dat die digtheid van geleenthede binne 'n aktiwiteitsruimte verband hou met die grootte van die aktiwiteitsruimte: groter aktiwiteitsruimtes het laer digtheid van geleenthede as kleiner aktiwiteitsruimtes. Dit kan impliseer dat individue wat 'n lae geleentheidsdigtheid ervaar in hulle aktiwiteitsruimte, minder geneig sal wees om op eksterne stimuli en/of bewusmakingsprogramme te reageer as individue wat 'n hoë geleentheidsdigtheid het. Op hul beurt is potensiële roete areas gebruik om vas te stel tot watter mate individue oor verskillende ruimtelike-temporale geleenthede beskik wat gedragsverander (vervoer en aktiwiteite) moontlik maak.

Die bevindings dui daarop dat plekbewuste tegnologie groot potensiaal bied om vervoer- en aktiwiteitsgedragnavorsings aan te vul en verder uit te bou. Verder toon die resultate dat die inkorporering van ruimtelike en temporale metings van kritieke belang is vir die ontwerp, implementering en evaluering van VRGV-intervensies. Die bykomende waarde van nuwe tegnologieë, soos GPS, is dat hulle maklik in 'n groter ruimtelike-temporale raamwerk van analise geïntegreer kan word. Navorsers moet egter versigtig wees om nie GPS as 'n wondermiddel te beskou nie want GPS-data en tegnologie het ook nadele. Omsigtigheid behoort aan die dag gelê

te word ten opsigte van toepassingsontwikkeling, steekproef seleksie, terreinkeuse en data-afleidings.

Acknowledgments

The past three and half years have been turbulent. At times, I did not see how I would ever be able to finish all the work. Too much to do, too little time. Fortunately, I have had the privilege to spend this time in this beautiful country, and I will be forever grateful for the opportunity I have had. However, undertaking a research project in a country so different than my own has not always been easy, and at times it has been very emotional. As such, this dissertation would not have been possible without the help and support of the people around me.

First of all, I would like to thank my promotor, Dr Stephan Krygsman. Your input, many comments, and support were invaluable during my studies. Also, Dr Tom de Jong deserves to be acknowledged. Tom was not only my supervisor during my Research Master, and the one who got me in contact with Stellenbosch University in the first place but also became a friend over the years. Many thanks for your support, suggestions, contributions, and guidance throughout the years. The many lunches and dinners, often accompanied by a good glass of wine, whenever you were in Stellenbosch, make up many good memories. Let us continue this in the future. Also, I would like to extend my gratitude to the Graduate School of Economic and Management Sciences as well as the National Research Foundation of South Africa for their financial support, without which all of this would not have been possible.

Second of all, I want to thank all of my colleagues in the Graduate School, many of whom became my friends, and the many other people who made my time here special. Thank you Hlokoma and Mike for the talks, the many, many, beers, and all the great times we spent together. It has not always been easy, but your friendship has made this journey a lot more comfortable. Also, thank you Spencer, Love, Trevan, Jaco, Marlies, Bart, and Lungi, for your friendship and the many good times. In this list, also my non-local friends should not be forgotten. Wilbert, even though you are facing the burden of your PhD, your long-distance consultation has been invaluable. I am grateful to know that distance has never impacted on our friendship. Also, thanks to Dajo, Jeroen and Wouter, who not only had to listen to me complain during our many Skype conversations but who also made every visit to the Netherlands (or Germany for that matter) unforgettable.

Third of all, I want to thank my family. Thank you, mama, papa, Sean, and Lenore, for believing in me. I know we have been missing each other a lot, but thank you for your unwavering support of my dream to pursue my PhD studies in South Africa. Enkosi. Baie dankie. Dankjewel. Lastly, I owe a very special thanks to my girlfriend, Dumisile, who had to share me with my project and my ambitions. Thank you for your love. Thank you for your support. Thank you for your understanding. Ndiyakuthanda kakulu.

Table of contents

List of figures

List of tables

Part I
Introduction

People are travelling further and for longer, car ownership has increased while car occupancy has dropped, and the perceived indispensable nature of cars combined with habitual car use means that cost is not a strong deterrent of car use, it is crucial therefore to identify alternative ways to encourage changes in travel behaviour.

Howarth and Polyviou (2012: 766)

Chapter 1. The sustainability of mobility

Against the background of unprecedented growth in private vehicle ownership and the entrenchment of the private car in everyday life, the past decades have seen a growing and ongoing academic and policy debate on how to encourage individuals to change to more sustainable ways of travelling. More recently, researchers have started to build on so-called location-aware technologies (LAT), exploring innovative methods to more accurately capture, visualise, and analyse individual spatiotemporal travel patterns – information that can be used to formulate strategies to accommodate the increasing demand for transport vis-à-vis growing environmental and societal concerns. This dissertation aims to expand on this research by using GPS-enabled smartphones for the collection of individual travel data in South Africa. The first section of this chapter provides the background to the research problem, while section two details the theoretical framework adopted throughout the rest of the study. The main objective and research questions are outlined in section three. The overall outline of the dissertation is provided in section four.

1.1 Background and research gaps

Mobility is a central aspect of everyday life. In its simplest form, human mobility refers to the movement of individuals from location A to location B. This can be a relocation from one city to another city, as well as a trip from home to work (Cresswell, 2006).[1] Transport systems provide the physical nodes and linkages that facilitate this mobility. However, transport systems and road networks in many cities around the world are under pressure as a result of unparalleled growth in private vehicle ownership and increasingly complex and fragmented travel patterns (Dimitriou, 2011; Howarth & Polyviou, 2012; Gwilliam, 2013; Salonen, Broberg, Kyttä & Toivonen, 2014; Florida, 2017). Particularly in urban areas, this is problematic because it leads to problems such as congestion, accidents, road decay, and reduced accessibility. In Sub-Saharan Africa, these problems are expected to get even worse, because the demand for transport in urban regions is projected to grow exponentially in the coming years as a result of rapid urbanisation. At the same time, significant investments in public transport services are absent, and the lion's share of transport funding currently caters for private transport (Candiracci, Schlosser & Allen, 2010; Pojani & Stead, 2015).[2]

In South Africa's urbanised areas, the pressure on the transport system is exacerbated by, amongst other factors, low-density housing, the concentration of employment in city centres, growth in disposable income, the absence of an adequate public transport system, and the remnants of apartheid planning policies (Lucas, 2011). The combination of these factors manifests itself in severe daily congestion and traffic gridlock (Gwilliam, 2003). In South Africa's

---

[1] Some authors have argued that this basic understanding of mobility should rather be called movement because it only focuses on displacement, whereas it ignores power structures, embodied experiences, and meaning (cf. Cresswell, 2006; Urry, 2007; Bissell, 2010; Middleton, 2010). However, in the remainder of this dissertation, mobility and movement will be used interchangeably to denote movement of individuals.

[2] There are some notable exceptions, such as the bus rapid transit systems introduced in Cape Town (MyCiTi), South Africa, and in Dar es Salaam (DART), Tanzania.

most congested city, Cape Town, for instance, commuters on average incur an additional 70 percent of travel time during the morning peak hour; as such, Cape Town is ranked at number 48 in terms of worldwide traffic congestion levels (TomTom International, 2017).

Not only does the growth in private vehicle ownership put a strain on the transport system, but it also has severe environmental effects at both a local and a global level. While at local level air pollutants of fossil-fuel-based vehicles come with health risks, at global level private vehicles contribute to climate change through greenhouse gas emissions (Gwilliam, 2013; IPCC, 2014). In fact, in 2010, the transport sector as a whole accounted for 14 percent of the worldwide anthropogenic greenhouse gas emissions (IPCC, 2014). In the decade from 2000 to 2010, the transport sector also accounted for 11 percent of the increase in these emissions. Because of these negative externalities, it is safe to state that humankind's current mobility is not environmentally or economically sustainable. As such, governments and researchers throughout the world have started to recognise the need to curtail demand for private road transport (Taylor & Ampt, 2003), and agree that "technological breakthroughs [alone] are not going to provide the silver bullet for the mitigation of climate change and energy security threats caused by the transport sector" (Stradling & Anable, 2008: 195).

The realisation that increasing road infrastructure and improvements in car technology are not sufficient to address the transport problems around the world has led to the idea that transport planning should shift from supply-side to demand-side passenger transport planning (Behrens & Del Mistro, 2010). In a number of cities, there are currently even plans to completely ban car usage (World Economic Forum, 2017); however, in most cases, travel demand management (TDM) strategies have been put forward as a tool to achieve a change in people's travel behaviour (see Kitamura, Fujii & Pas, 1997; Gärling & Schuitema, 2007; Cairns, Sloman, Newson, Anable, Kirkbridge & Goodwin, 2008; Stradling & Anable, 2008; Schönfelder & Axhausen, 2010). TDM is an umbrella term for all interventions that aim to modify travel behaviour in favour of more socially, environmentally, and economically sustainable alternatives (Taylor, 2007). Examples of TDM interventions include operational interventions (e.g. dynamic traffic information systems), physical interventions (e.g. dedicated bus lanes), financial interventions (e.g. road pricing), organisational interventions (e.g. flexible work-hours), and informational interventions (e.g. individual travel plans).

Within the domain of TDM, a useful distinction can be made between 'hard strategies' and 'soft strategies'. Hard strategies, on the one hand, focus on restraining and managing through regulations and economic disincentives, often according to the user-pays principle. However, "the perceived indispensable nature of cars combined with habitual car use means that cost is not a strong deterrent of car use" (Howarth & Polyviou, 2012: 766). Also, more often than not economic disincentives suffer from a lack of public acceptability (Gärling, Eek, Loukopoulos, Fujii, Johansson-Stenman, Kitamura, Pendyala & Vilhelmson, 2002; Eriksson, Garvill & Nordlund, 2006). Soft strategies, on the other hand, aim to realise a behavioural change using information provision, for instance by providing households with personalised information concerning the sustainability of their travel patterns as well as suggestions for possible future travel alternatives (Eriksson *et al.*, 2006). Taylor and Ampt (2003: 165) state that: "These new approaches have attempted to work with individuals in a community rather than impose measures on that community, tapping into the desire of a significant minority of people to make and be seen to make their own contributions to improving the environment and reducing resource consumption".

The strategies to manage travel demand on the soft side of the continuum are typically referred to as voluntary travel behaviour change (VTBC) programmes. These are interventions "where the objective of the program[me] is to allow people to choose to change travel behaviour rather than to expect or force reactions in response to external stimuli or pressures" (Taylor & Ampt, 2003: 171). Several studies have focused on evaluating the effectiveness of these programmes (cf. Cairns *et al.*, 2008; Bonsall, 2009; Brög, Erl, Ker, Ryle & Wall, 2009; Chatterjee, 2009; Stopher, Clifford, Swann & Zhang, 2009; Zhang, Stopher & Halling, 2013; Meloni & Sanjust, 2014; Meloni, Sanjust Di Teulada & Spissu, 2016), however, the results have been mixed. Whereas Seethaler and Rose (2009), for instance, found that the large-scale TravelSmart® project in Melbourne, Australia, did not lead to a significant reduction in average vehicle kilometres travelled, Zhang *et al.* (2013) found that participants were willing to reduce private car use in the TravelSmart® project in Western Adelaide, Australia. Success was also achieved with a 'mobility management' intervention in the city of Bihari, Japan, in which monthly newsletters and questionnaires led to an increase in the use of an experimental bus service (Taniguchi & Fujii, 2007). Although the increasing pressure of private vehicle ownership on the transport system and the effects of road transport on climate change are more challenging than ever, the current scientific debate on VTBC interventions contains several areas that require further research.

Firstly, most VTBC intervention evaluations rely heavily on self-reported data and as such are contingent on the respondent's ability to accurately remember his or her daily movements and activity participation, as is the case in, for instance, retrospective travel surveys and activity-travel diaries. The fact that the evaluation of a travel programmes may be socially charged may affect the accuracy of these data even further (Stopher & Greaves, 2006; Bonsall, 2009; Stopher *et al.*, 2009; Bamberg, Fujii, Friman & Gärling, 2011; Meloni & Sanjust, 2014; Salonen *et al.*, 2014). Furthermore, the possibility of a VTBC intervention has not yet been investigated in the context of a developing country; even though in South Africa, for instance, road passenger transport is highly inefficient and unsustainable (Walters, 2008, 2013). This raises the question of whether VTBC interventions could be effective in stimulating more sustainable ways of transport in a developing country with limited opportunities for public transport, sprawling urban forms, and inefficient private transport modes.

Secondly, technological developments in the field of location-aware technologies (LAT), Global Positioning Systems (GPS) in particular, have greatly enhanced opportunities to collect accurate data on human spatiotemporal behaviour (Shen & Stopher, 2014). These data need to be collected and analysed systematically to be intelligible for transport researchers and policy makers. Yet, little is known on how to effectively deploy LAT in stimulating a behavioural change where a high level of accuracy is a necessary, but not the only condition. This requires the development of methods and tools for data collection that can capture the spatiotemporal attributes of activity and trip behaviour (Meloni & Sanjust, 2014). Moreover, the challenges inherent to mobile data collection techniques include not only harnessing the tools to obtain geo-referenced data, but also the development of new skills sets for cleaning, analysing, and interpreting these data (Van Dijk & Krygsman, 2016).

Thirdly, quantifiable spatial measurements that draw attention to individual accessibility and spatiotemporal access to transport and activity opportunities have not been employed in the design of VTBC strategies. However, as argued by Meloni and Sanjust (2014), travel patterns first have to be analysed and understood before a behavioural change intervention can be tailored to fit the individual context. The research presented in this dissertation addresses these gaps by exploiting the possibilities of innovative data collection techniques, and provides new evidence for understanding and analysing travel behaviour to aid in the design of a VTBC intervention from within a transport-geographical framework.

1.2 Theoretical framework

There is a consensus amongst many scholars that people predominantly travel with a specific purpose, rather than for the intrinsic value of travel.[3] Travel is thus a means to participate in different activities that take place in geographically dispersed locations. As Gärling *et al.* (2002: 59) summarise it: "It has become increasingly evident that travel results from choices people make that are both interdependent and dependent on desires and obligations to participate in activities." This has led to the understanding that the daily life of an individual is embedded in a set of complex temporal and spatial arrangements that determine their engagement in activities; something which was acknowledged for the first time by Swedish geographer Törsten Hägerstrand (1970) in his seminal exploration of time geography.

Time geography describes the life of an individual as a continuous path through time and space (Hägerstrand, 1970). This path is constituted of movements through space and activities localised in space; both of which have a temporal component. Although an individual's movements stem from their desire to engage in an activity, their movements are subject to at least three different constraints: capability constraints, coupling constraints, and authority constraints. The first constraint refers to physiological and cognitive issues; for instance, humans need sleep, shelter, and food. The second constraint is faced when planning social interactions: different individual space-time paths have to be matched or "coupled". The third constraint stems from norms, rules, and laws that act upon space and time. For instance, you do not necessarily have access to a location because of the separation of public and private spaces. Furthermore, time regimes, like the trading hours of a supermarket, also exert their influence on one's mobility (Hägerstrand, 1970; Thrift, 1977; Schwanen & Kwan, 2008). Together, as Schwanen and Kwan (2008: 1363) argue "these constraints can mitigate but also reinforce one another's impacts on activity participation and travel behaviour".

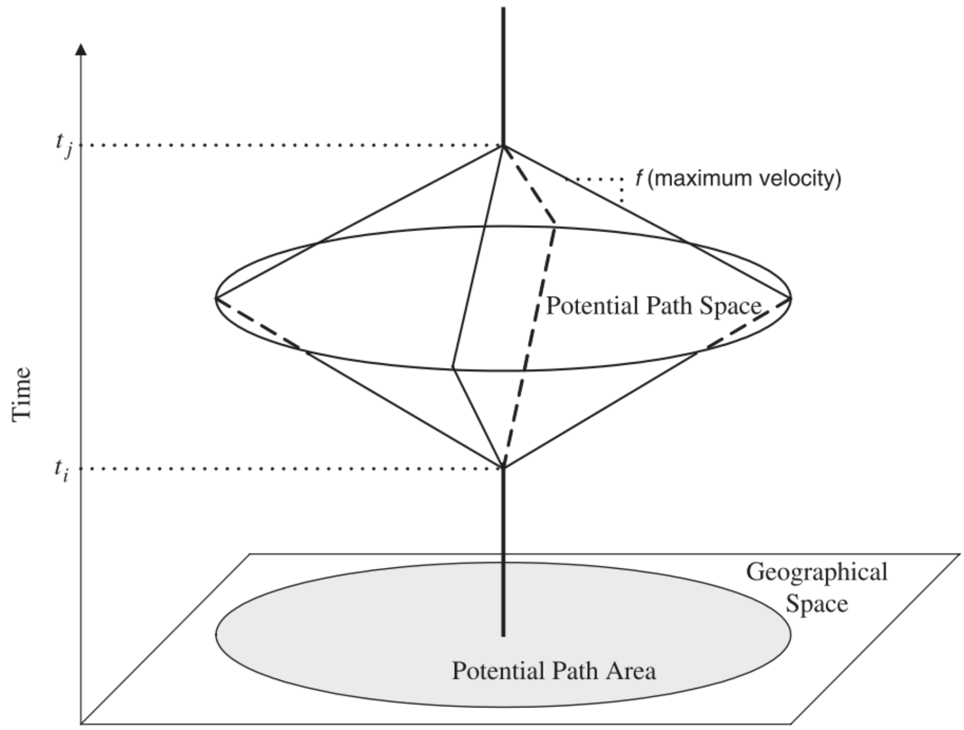
When it comes to designing policies to reduce car use, it is crucial to consider the bandwidth of alternative opportunities an individual has; a lack of feasible alternatives for activity destinations would inhibit a possible change in activity pattern and participation (Dijst, De Jong & Van Eck, 2002). The quantification of this individual accessibility to opportunities from a time-geographical perspective is accomplished with space-time accessibility (STA) measures. STA measures are based on one of the key concepts of time geography: the space-time prism. The space-time prism is the graphical representation of all possible paths an individual can take

[3] There are several critiques on this conceptualisation of mobility (see, for example, Urry, 2004, 2007; Cresswell, 2006).

within a given time budget between two primary activity locations, such as home and work; the (average) speed of the mode of transport being used; and the network distance between these activities (Neutens, Schwanen & Witlox, 2011).

As shown in Figure 1.1, the two-dimensional projection of the three-dimensional space-time prism, known as the potential path area (PPA), represents the area that an individual can potentially visit given the time available in between two primary activities (Dijst *et al.*, 2002; Schwanen & De Jong, 2008). The extent of the PPA, therefore, is equal to an individual's spatiotemporal accessibility to opportunities. (Note that once one moves towards the outer edge of the PPA, the time one can spend at that activity location diminishes until it eventually reduces to zero.) While a car typically offers a flexible, fast, and individualised way to maximise PPA, and, thus the number of activity locations within reach, it is hardly ever the most sustainable choice (Krygsman, 2004; Salonen *et al.*, 2014). It follows that this framework strongly suggests that the consideration of individual accessibility be a necessary condition of any policy aimed at stimulating a behavioural change towards more sustainable modes of travel.

Figure 1.1 | The space-time prism and the potential path area (PPA)



Source: Miller (2005: 20)

1.3 Research aim and research questions

With the ever-increasing demand for transport, the growth in the number of on-the-road private vehicles, and the reduction in the number of passengers per automobile, the need to effectively manage demand for private transport is more important than ever before. This is reflected in a wide range of studies on the effectiveness of VTBC interventions (see Gwilliam, 2003; Taylor & Ampt, 2003; Cairns *et al.*, 2008; Haq, Whitelegg, Cinderby & Owen, 2008; Bonsall, 2009; Brög *et al.*, 2009; Zhang *et al.*, 2013; Meloni *et al.*, 2016). However, no consensus has been reached on the question of whether a broad implementation of VTBC programmes is effective in stimulating people to use more sustainable ways of travelling. This dissertation argues that location-aware technologies, particularly GPS-enabled smartphones, could potentially augment the research on VTBC interventions (cf. Brög *et al.*, 2009; Stopher *et al.*, 2009; Meloni *et al.*, 2016; Sanjust di Teulada, Meloni & Spissu, 2017). Smartphones can not only source data, they can ultimately also provide individuals with real-time information, feedback, and suggestions. However, between sourcing the data and relaying feedback to individual commuters, significant research is required on how to obtain, clean, and interpret the data, as well as on how to account for individual spatiotemporal accessibility. The aim of this research is therefore

> To evaluate the potential of location-aware technologies and the analysis of spatiotemporal behaviour for the design of a voluntary travel behaviour change (VTBC) intervention in South Africa.

This aim will be addressed by answering the following four research questions:

> RQ1: What are the current insights and research gaps on the effectiveness of VTBC interventions?

Several studies have evaluated the effectiveness of VTBC interventions, yet to date, the outcomes remain ambiguous. Overall, there seems to be some evidence to suggest that soft transport measures have the potential to reduce private vehicle usage and stimulate public transport ridership. However, in recent years, concerns have been raised in relation to the methodology of some of these evaluation studies (cf. Graham-Rowe, Skippon, Gardner & Abraham, 2011; Melia, 2015). With this first question, we aim to give an overview of the current state-of-the-research, and highlight possible gaps.

> RQ2: To what extent are current GPS data processing techniques suitable for analysing the effectiveness of a VTBC intervention?

Travel data collection methods have a long history in the field of transport planning. Since the first face-to-face interviews in the 1950s in the United States, the field has undergone significant changes with the introduction of telephone-assisted interviews in the 1970s and the implementation of computer-assisted interviews in the 1980s (Shen & Stopher, 2014; Xiao, Juan & Zhang, 2015). Arguably one of the most defining changes arrived in the late 1990s, when Global Positioning System (GPS) technology was introduced as a method for augmenting or (partly) substituting traditional travel surveys methods (see Wolf, 2000; Bohte & Maat, 2009; Chen & Kwan, 2012; Nitsche, Widhalm, Breuss & Maurer, 2012; Nitsche, Widhalm, Breuss, Brändle & Maurer, 2014; Feng & Timmermans, 2013; Shen & Stopher, 2013, 2014; Shoval, Kwan, Reinau & Harder, 2014). While GPS technology may have opened new avenues for collecting accurate data on human spatiotemporal behaviour (Shen & Stopher, 2014), GPS data need to be collected and analysed in a systematic way to be useful for transport researchers – especially in evaluating the effect of a VTBC intervention, in which effect sizes are known to be small

(Richardson, Seethaler & Harbutt, 2004; Stopher & Greaves, 2007; Stopher *et al.*, 2009). Important is the translation of raw GPS trajectories into activities, and the best estimation of a route travelled by a user. This is not a trivial task, as GPS faces several limitations regarding its accuracy, especially indoors and in dense urban areas. Accordingly, over the past years, a wide variety of methods have been employed to analyse GPS data to, for instance, infer activity locations and trip purposes with discrete hidden Markov models (e.g. Nitsche *et al.*, 2014), kernel densities (e.g. Thierry, Chaix & Kestens, 2013), and supervised machine learning techniques, such as Bayesian belief networks, random forest classifiers, and decision trees (e.g. Feng & Timmermans, 2015). With this second research question, we specifically address the issue of how to recognise activity points and reconstruct actual routes travelled from raw GPS trajectories with the evaluation of a VTBC intervention in mind.

> RQ3: To what extent are GPS-enabled smartphones suitable for acquiring high-resolution space-time data in South Africa?

Since the seminal work on the topic by Jean Wolf (2000), researchers have started to use GPS data in transport and mobility research. However, few implementations of mobile technologies for activity-travel data collection have been documented in Sub-Saharan Africa (notable exemptions are Krygsman & Nel, 2009; Siedner, Lankowski, Tsai, Muzoora, Martin, Hunt, Haberer & Bangsberg, 2013; Venter & Joubert, 2013, 2014). In addition, despite calls for experimenting with the acquisition of high-resolution space-time data by means of GPS devices and employing GPS-enabled smartphones for data collection (for instance Bonsall, 2009; Chatterjee, 2009; Stopher *et al.*, 2009), to date few studies have used these techniques to assess the effectiveness of a VTBC intervention. To answer this research question, we present the results of an experimental study on the reliability and feasibility of passively collecting high-resolution spatiotemporal data on activity and travel behaviour using GPS-enabled smartphones with a purposely designed tracking application.

> RQ4: To what extent can space-time accessibility measures be integrated into modelling the feasibility of a VTBC intervention?

Measures of travel tend to reduce activity-travel behaviour to static representations, such as the number of trips, vehicle-kilometres travelled, and person-kilometres travelled. While these are useful and widely employed measures, truly understanding individual responses to transport policies requires an understanding of individual spatial behaviour. Particularly, an individual's spatiotemporal access to opportunities and travel alternatives play a central role in their response to transport policies (Kingham, Dickinson & Copsey, 2001). Yet, the consideration of these opportunities seems to have been largely neglected in the design of travel behaviour change interventions. Whereas the lack of consideration of spatiotemporal opportunities may be partly attributed to the limitations of traditional data collection tools such as paper-based surveys, technological developments in the domain of location-based services offer new opportunities to assess individual accessibility. With this fourth research question, we elaborate on this limited knowledge and try to address how new technologies and spatiotemporal measurements could aid the design of a VTBC intervention.

1.4 Dissertation outline

To meet with the stated aim of this dissertation and answer the research questions stated in the previous section, this study adopts an article based approach and presents the findings of one theoretical and five empirical articles. Each article is presented as a chapter, and the empirical chapters are either published, or accepted for publication, in a peer-reviewed journal (Chapters 4 and 6), published in peer-reviewed conference proceedings (Chapters 5), presented at an international conference (Chapter 7), or prepared for submission to a peer-reviewed journal (Chapter 3). These chapters were included as is with only slight changes to the introduction and conclusions where necessary. The article-based nature of the chapters results in some repetition across the different chapters.

Table 1.1 summarises the outline of this dissertation, including how the chapters relate to the research questions. Because the focus is on evaluating the potential of location-aware technologies and the analysis of spatiotemporal behaviour for the design of VTBC interventions, the body of this dissertation is divided into three main parts. Part II gives an overall overview of the literature concerning VTBC interventions. Part III focuses on methodological issues concerning GPS data imputation vis-à-vis VTBC interventions. Part IV moves to the analysis of GPS data collected at Stellenbosch University, and explores avenues on how the analysis of spatiotemporal accessibility could be used in the design of travel behaviour change interventions.

A literature search formed the basis for addressing Research Question 1, which is the focus of Chapter 2. As several systematic reviews and meta-analyses have been published in recent years, the review is not a full systematic review. Research Question 2 was answered based on artificial GPS data and a GPS dataset collected by Utrecht University, the Netherlands, that was made available for analysis. Chapter 3 explores some of the limitations and challenges of identifying activity locations from raw GPS data and uses advanced machine learning algorithms applied to artificial GPS data to identify activity points. Chapter 4 proposes two map-matching algorithms for reconstructing actual routes travelled.

Research Questions 3 and 4 are answered based on GPS data that were gathered by means of a purposely designed smartphone application, named Tracklog on the Google Play Store (for Android devices) and Tracklogging on the Apple App Store (for iOS devices). In Chapter 5, a small pilot study is used to examine the reliability and feasibility of passively collecting high-resolution spatiotemporal data on activity and travel behaviour using GPS-enabled smartphones. Chapters 6 and 7 use data collected during a research project on the travel patterns of staff and students on both the main campus and the Tygerberg satellite campus of Stellenbosch University, South Africa. In these chapters, Research Question 4 is answered using a combination of activity space analysis and potential path area analysis, supplemented with open-source GIS information. The answers to the various research questions are combined in Chapter 8, which discusses the findings, provides policy recommendations, draws attention to some limitations of this study, and suggests future directions for research.

Table 1.1 | Dissertation outline

Part	Chapter	RQ and topic	Data sources	Analyses
Part II Literature	2	Research Question 1, voluntary travel behaviour change interventions and their effectiveness	Google Scholar, ScienceDirect, Scopus	Literature review, identifying methodological shortcomings in existing VTBC literature
Part III Methodological contributions	3	Research Question 2, GPS data imputation, imputing activity points from raw GPS trajectories	Artificial GPS data with varying noise levels	Machine learning algorithms to categorise GPS points into *move* and *stay* points
	4	Research Question 2, GPS data imputation, reconstructing routes travelled from raw GPS trajectories	GPS data set made available by Utrecht University	Development of two GIS-based map-matching algorithms to reconstruct routes travelled
Part IV Empirical results	5	Research Question 3, GPS data collection using GPS-enabled smartphones	Pilot project with GPS-enabled smartphones (Stellenbosch University)	Visualisation of collected GPS tracks
	6	Research Question 4, relationship between activity spaces and willingness to consider alternative modes of transport	Mobility study with a household travel survey and GPS-enabled smartphones (Stellenbosch University), open source GIS data	Spatial analyses using individual activity spaces (with different conceptualisations) in a GIS-environment
	7	Research Question 4, relationship between spatiotemporal accessibility to opportunities and willingness to consider alternative modes of transport	Mobility study with a household travel survey and GPS-enabled smartphones (Stellenbosch University), open source GIS data	Accessibility analyses using potential path areas
Part V Conclusion	8	Overall aim, overall findings, limitations, and recommendations	Research findings presented throughout the dissertation	Synthesise research findings

References

Bamberg, S., Fujii, S., Friman, M. & Gärling, T. 2011. Behaviour theory and soft transport policy measures. *Transport Policy*. 18(1):228–235.

Behrens, R. & Del Mistro, R. 2010. Shocking habits: Methodological issues in analyzing changing personal travel behavior over time. *International Journal of Sustainable Transportation*. 4(5):253–271.

Bissell, D. 2010. Passenger mobilities: Affective atmospheres and the sociality of public transport. *Environment and Planning D: Society and Space*. 28(2):270–289.

Bohte, W. & Maat, K. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*. 17(3):285–297.

Bonsall, P. 2009. Do we know whether personal travel planning really works? *Transport Policy*. 16(6):306–314.

Brög, W., Erl, E., Ker, I., Ryle, J. & Wall, R. 2009. Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*. 16(6):281–292.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbridge, A. & Goodwin, P. 2008. Smarter choices: Assessing the potential to achieve traffic reductions using 'soft measures'. *Transport Reviews*. 28(5):593–618.

Candiracci, S., Schlosser, C. & Allen, H. 2010. *A new perspective: Sustainable mobility in African cities*. Nairobi, Kenya: United Nations Human Settlements Programme (UN-HABITAT).

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*. 16(6):293–305.

Chen, X. & Kwan, M.-P. 2012. Choice set formation with multiple flexible activities under space–time constraints. *International Journal of Geographical Information Science*. 26(5):941–961.

Cresswell, T. 2006. *On the move: Mobility in the modern Western world*. 1st ed. New York: Routledge.

Van Dijk, J. & Krygsman, S. 2016. Mobile technologies for data collection in sub-Saharan Africa: An outlook to the future of mobility research. In *#CelebrateGeography: Proceedings of the Centenary Conference of the Society of South African Geographers*. Stellenbosch: SunMedia. 180–189.

Dijst, M., De Jong, T. & Van Eck, J.R. 2002. Opportunities for transport mode change: An exploration of a disaggregated approach. *Environment and Planning B: Planning and Design*. 29(3):413–430.

Dimitriou, H.T. 2011. Transport and city development: Understanding the fundamentals. In H.T. Dimitriou & R. Gakenheimer (eds.). *Urban transport in the developing world: A handbook of policy and practice*. Cheltenham: Edwar Elgar Publishing. 8–39.

Eriksson, L., Garvill, J. & Nordlund, A.M. 2006. Acceptability of travel demand management measures: The importance of problem awareness, personal norm, freedom, and fairness. *Journal of Environmental Psychology*. 26(1):15–26.

Feng, T. & Timmermans, H.J.P. 2013. Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies*. 37:118–130.

Feng, T. & Timmermans, H.J.P. 2015. Detecting activity type from GPS traces using spatial and temporal information. *European Journal of Transport and Infrastructure Research*. 15(4):662–674.

Florida, R. 2017. *What drove the driving downturn?* [Online], Available: https://www.citylab.com/transportation/2017/04/what-drove-the-driving-downturn/518601/?utm_source=SFFB [2017, April 07].

Gärling, T. & Schuitema, G. 2007. Travel demand management targeting reduced private car use: Effectiveness, public acceptability and political feasibility. *Journal of Social Issues*. 63(1):139–153.

Gärling, T., Eek, D., Loukopoulos, P., Fujii, S., Johansson-Stenman, O., Kitamura, R., Pendyala, R. & Vilhelmson, B. 2002. A conceptual analysis of the impact of travel demand management on private car use. *Transport Policy*. 9(1):59–70.

Graham-Rowe, E., Skippon, S., Gardner, B. & Abraham, C. 2011. Can we reduce car use and, if so, how? A review of available evidence. *Transportation Research Part A: Policy and Practice*. 45(5):401–418.

Gwilliam, K. 2003. Urban transport in developing countries. *Transport Reviews*. 23(2):197–216.

Gwilliam, K. 2013. Cities on the move - Ten years after. *Research in Transportation Economics*. 40(1):3–18.

Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional Science Association*. 24(1):6–21.

Haq, G., Whitelegg, J., Cinderby, S. & Owen, A. 2008. The use of personalised social marketing to foster voluntary behavioural change for sustainable travel and lifestyles. *Local Environment*. 13(7):549–569.

Howarth, C.C. & Polyviou, P. 2012. Sustainable travel behaviour and the widespread impacts on the local economy. *Local Economy*. 27(7):764–781.

IPCC. 2014. *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Core Writing Team, R.K. Pachauri, & L.A. Meyers (eds.). Geneva, Switzerland: IPCC.

Kingham, S., Dickinson, J. & Copsey, S. 2001. Travelling to work: Will people move out of their cars. *Transport Policy*. 8(2):151–160.

Kitamura, R., Fujii, S. & Pas, E.I. 1997. Time-use data, analysis and modeling: Toward the next generation of transportation planning methodologies. *Transport Policy*. 4(4):225–235.

Krygsman, S.C. 2004. *Activity and travel choice(s) in multimodal public transport systems.* Published doctoral dissertation. Utrecht: Utrecht University.

Krygsman, S.C. & Nel, J.H. 2009. The use of global positioning devices in travel surveys - A developing country application. In *Proceedings of the 28th Southern African Transport Conference (SATC 2009)*. Pretoria: Southern African Transport Conference. 108–118.

Lucas, K. 2011. Making the connections between transport disadvantage and the social exclusion of low income populations in the Tshwane Region of South Africa. *Journal of Transport Geography*. 19(6):1320–1334.

Melia, S. 2015. Do randomised control trials offer a solution to 'low quality' transport research? In *Proceedings of the 47th Annual UTSG Conference*. London: Universities Transport Studies Group. [Online], Available: http://www.utsg.net/web/index.php?page=annual-conference [2016, April 03].

Meloni, I. & Sanjust, B. 2014. Using a GPS active logger to implement travel behavior change programs. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 325–340.

Meloni, I., Sanjust Di Teulada, B. & Spissu, E. 2016. Lessons learned from a personalized travel planning (PTP) research program to reduce car dependence. *Transportation*. 44(4):1–18.

Middleton, J. 2010. Sense and the city: Exploring the embodied geographies of urban walking. *Social & Cultural Geography*. 11(6):575–596.

Miller, H.J. 2005. A measurement theory for time geography. *Geographical Analysis*. 37(1):17–45.

Neutens, T., Schwanen, T. & Witlox, F. 2011. The prism of everyday life: Towards a new research agenda for time geography. *Transport Reviews*. 31(1):25–47.

Nitsche, P., Widhalm, P., Breuss, S. & Maurer, P. 2012. A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys. In Vol. 48. *Procedia - Social and Behavioral Sciences*. Oxford: Elsevier. 1033–1046.

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N. & Maurer, P. 2014. Supporting large-scale travel surveys with smartphones - A practical approach. *Transportation Research Part C: Emerging Technologies*. 43:212–221.

Pojani, D. & Stead, D. 2015. Sustainable urban transport in the developing world: Beyond megacities. *Sustainability*. 7(6):7784–7805.

Richardson, A.J., Seethaler, R.K. & Harbutt, P.L. 2004. Design issues for before and after surveys of travel behaviour change. *Transport Engineering in Australia*. 9(2):103–118.

Salonen, M., Broberg, A., Kyttä, M. & Toivonen, T. 2014. Do suburban residents prefer the fastest or low-carbon travel modes? Combining public participation GIS and multimodal travel time analysis for daily mobility research. *Applied Geography*. 53:438–448.

Sanjust di Teulada, B., Meloni, I. & Spissu, E. 2017. The influence of activity-travel patterns on the success of VTBC. *International Journal of Urban Sciences*. In press.

Schönfelder, S. & Axhausen, K.W. 2010. *Urban rhythms and travel behaviour: Spatial and temporal phenomena of daily travel*. 1st ed. Farnham, United Kingdom: Ashgate Publishing, Ltd.

Schwanen, T. & De Jong, T. 2008. Exploring the juggling of responsibilities with space-time accessibility analysis. *Urban Geography*. 29(6):556–580.

Schwanen, T. & Kwan, M.-P. 2008. The internet, mobile phone and space-time constraints. *Geoforum*. 39(3):1362–1377.

Seethaler, R. & Rose, G. 2009. Using odometer readings to assess VKT changes associated with a voluntary travel behaviour change program. *Transport Policy*. 16(6):325–334.

Shen, L. & Stopher, P.R. 2013. A process for trip purpose imputation from Global Positioning System data.

Transportation Research Part C: Emerging Technologies. 36:261–267.
Shen, L. & Stopher, P.R. 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*. 34(3):316–334.
Shoval, N., Kwan, M.-P., Reinau, K.H. & Harder, H. 2014. The shoemaker's son always goes barefoot: Implementations of GPS and other tracking technologies for geographic research. *Geoforum*. 51:1–5.
Siedner, M.J., Lankowski, A., Tsai, A.C., Muzoora, C., Martin, J.N., Hunt, P.W., Haberer, J.E. & Bangsberg, D.R. 2013. GPS-measured distance to clinic, but not self-reported transportation factors, are associated with missed HIV clinic visits in rural Uganda. *AIDS*. 27(9):1503–1508.
Stopher, P.R. & Greaves, S.P. 2006. Guidelines for samplers: Measuring a change in behaviour from before and after surveys. *Transportation*. 34(1):1–16.
Stopher, P.R. & Greaves, S.P. 2007. Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice*. 41(5):367–381.
Stopher, P.R., Clifford, E., Swann, N. & Zhang, Y. 2009. Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy*. 16(6):315–324.
Stradling, S. & Anable, J. 2008. Individual transport patterns. In 1st ed. R. Knowles, J. Shaw, & I. Dochterty (eds.). *Transport Geographies: Mobilities, flows and spaces*. Oxford: Blackwell Publishing Ltd. 179–195.
Taniguchi, A. & Fujii, S. 2007. Promoting public transport using marketing techniques in mobility management and verifying their quantitative effects. *Transportation*. 34(1):37–49.
Taylor, M. 2007. Voluntary travel behavior change programs in Australia: The carrot rather than the stick in travel demand management. *International Journal of Sustainable Transportation*. 1(3):173–192.
Taylor, M. & Ampt, E. 2003. Travelling smarter down under: Policies for voluntary travel behaviour change in Australia. *Transport Policy*. 10(3):165–177.
Thierry, B., Chaix, B. & Kestens, Y. 2013. Detecting activity locations from raw GPS data: A novel kernel-based algorithm. *International Journal of Health Geographics*. 12(1):14.
Thrift, N. 1977. *An introduction to Tme-Geography*. Vol. 13. *Concepts and techniques in modern geography*. Norwich: Geo Abstracts Ltd, University of East Anglia.
TomTom International. 2017. *TomTom traffic index measuring congestion worldwide*. [Online], Available: https://www.tomtom.com/en_za/trafficindex/list [2017, March 28].
Urry, J. 2004. The "system" of automobility. *Theory, Culture & Society*. 21(4–5):25–39.
Urry, J. 2007. *Mobilities*. 1st ed. Malden, MA: Polity Press.
Venter, C. & Joubert, J. 2013. Use of multisource Global Positioning System data to characterize multiday driving patterns and fuel usage in a large urban region. *Transportation Research Record: Journal of the Transportation Research Board*. 2338:1–10.
Venter, C. & Joubert, J. 2014. Tax or Toll? GPS-based assessment of equity impacts of large-scale electronic freeway tolling in Gauteng, South Africa. *Transportation Research Record: Journal of the Transportation Research Board*. 2450:62–70.
Walters, J. 2008. Overview of public transport policy developments in South Africa. *Research in Transportation Economics*. 22(1):98–108.
Walters, J. 2013. Overview of public transport policy developments in South Africa. *Research in Transportation Economics*. 39(1):34–45.
Wolf, J. 2000. *Using GPS data loggers to replace travel diaries in the collection of travel data*. Published doctoral dissertation. Atlanta, Georgia: Georgia Institute of Technology.
World Economic Forum. 2017. *From Oslo to Paris, these major cities have plans to go car-free*. [Online], Available: https://www.weforum.org/agenda/2017/02/these-major-cities-are-starting-to-go-car-free [2017, March 28].
Xiao, G., Juan, Z. & Zhang, C. 2015. Travel mode detection based on GPS track data and Bayesian networks. *Computers, Environment and Urban Systems*. 54:14–22.
Zhang, Y., Stopher, P. & Halling, B. 2013. Evaluation of south-Australia's TravelSmart project: Changes in community's attitudes to travel. *Transport Policy*. 26:15–22.

Part II
Literature

Travel behaviour schemes have demonstrated significant potential in delivering sustainability at the local level, yet these schemes do not fully recognise that travel decisions are unique and are made at the individual level embedded in the specific context within which the traveller is located.

Howarth and Polyviou (2012: 778)

Chapter 2. Assessing voluntary travel behaviour change interventions: A literature review

Abstract

This chapter discusses the existing knowledge on voluntary travel behaviour change (VTBC) interventions. Notwithstanding the seemingly positive results of VTBC interventions in reducing car use, three issues seem to surface. The first issue is that in the majority of VTBC studies weak research designs have been used; as such it is difficult to draw definite conclusions about their effectiveness. The second issue is related to the precision of the measurement instrument. In many cases, the evaluation of a VTBC policy has relied heavily on traditional activity-travel diaries. Yet these instruments may not be capable of picking up subtle changes in travel behaviour. The third issue is that in both the design and the assessment of travel demand management (TDM) policies, travel behaviour is often described using static measures. Quantifiable spatial measurements have not been employed. Based on this review, two complementary lines of research are suggested: (1) field experiments from within the framework of a randomised controlled trial, and (2) field experiments in which dynamic user feedback is provided by GPS-enabled smartphones. The latter suggestion also urgently calls for integrated methods that can deal with activity reconstruction and route reconstruction of raw GPS measurements, and methods that consider the temporal and physical characteristics of the (built) environment.

Keywords

Transport demand management; Voluntary travel behaviour change; GPS technology

2.1 Introduction

Transport systems and road networks in many urban regions are under pressure as the result of a rapid increase in private vehicle ownership and increasingly complex travel patterns (Axhausen, Zimmermann, Schönfelder, Rindsfüser & Haupt, 2002; Mokhtarian, Salomon & Handy, 2006; Järv, Ahas & Witlox, 2014). The negative externalities of these trends include daily congestion and traffic gridlock, but also environmental degradation. In an effort to limit car use, policy makers and researchers have shifted their attention from supply-side measures to innovative demand-side measures (see, for example, Kitamura, Fujii & Pas, 1997; Gärling & Schuitema, 2007; Stradling & Anable, 2008). In the last decade in particularly, scholars have started to explore the opportunities of less coercive approaches in the context of transport travel management (TDM) approaches, typically known as voluntary travel behaviour change (VTBC) interventions, i.e. interventions "where the objective of the program[me] is to allow people to choose to change travel behaviour rather than to expect or force reactions in response to external stimuli or pressures" (Taylor and Ampt, 2003: 171). It is this group of soft measures that this review focuses on.

Over the years, a number of researchers have sought to determine the effectiveness of VTBC programmes (cf. Cairns, Sloman, Newson, Anable, Kirkbridge & Goodwin, 2008; Bonsall, 2009; Brög, Erl, Ker, Ryle & Wall, 2009; Chatterjee, 2009; Stopher, Clifford, Swann & Zhang, 2009; Zhang, Stopher & Halling, 2013; Sanjust di Teulada, Meloni & Spissu, 2017). The outcomes have been mixed. Seethaler and Rose (2009), for example, reported on a large-scale project carried out in Melbourne, Australia, in which the anticipated reduction in average vehicle kilometres travelled (VKT) was not realised. Zhang *et al.* (2013), on the other hand, communicated positive results regarding the willingness of the participants to reduce private car use in the TravelSmart® project in Western Adelaide, Australia. Furthermore, in a review of the results of several evaluation studies, Bonsall (2009: 307) claims that several studies reported on a reduction in vehicle kilometres travelled (VKT), a decrease in the number of trips made by car, and an increase in the number of trips made by public transport or non-motorised forms of transport.

Notwithstanding the positive results, these outcomes should be interpreted carefully; in recent years, concerns have been raised in relation to the methodology of some of these evaluation studies. Bonsall (2009) mentions, amongst other issues, inadequate sample sizes, unrepresentative samples, a focus on short-term effects only, and a lack of appropriate control groups (see also Stopher & Greaves, 2006; Bamberg, Fujii, Friman & Gärling, 2011). Also, it has been argued that habitual behaviour is hard to influence (Gärling & Axhausen, 2003) and that individuals will never change their behaviour on a voluntary basis by informing them about their responsibility. In fact, according to Hardin (1968: 1247), "responsibility is a verbal counterfeit for a substantial *quid pro quo*. It is an attempt to get something for nothing." This claim is substantiated by Gärling and Schuitema (2007) who question the idea of VTBC interventions being effective at all without the support of more coercive measures.

Given the negative externalities associated with private car use against the backdrop of rapid motorisation and unprecedented urbanisation in the developing world, understanding VTBC interventions as a means of managing transport demand is essential. Section 2 of this chapter provides a brief overview of TDM and contextualises VTBC interventions. Section 3 takes a closer look at travel behaviour and the notion of travel behaviour change itself. Section 4 explores the state of the current research and tries to highlight the gaps in the current body of knowledge. Section 5 concludes with possible implications for further research when it comes to designing, implementing, and assessing VTBC interventions.

2.2 Voluntary travel behaviour change interventions in context

Most scientists nowadays agree on the anthropogenic nature of climate change. The transport sector contributes to climate change through the emission of greenhouse gases, particularly through the combustion of fossil fuels. In fact, in 2010 the transport accounted for 14 percent of worldwide greenhouse gas emissions (IPCC, 2014). Furthermore, private car use is the "second biggest contributor to greenhouse gas emissions in the transport sector" (Chapman, 2007: 357). Whilst fuel consumption is quickly depleting our natural resources, car usage causes social and economic problems such as congestion, noise pollution, and accidents (Kingham, Dickinson & Copsey, 2001; Loukopoulos, Jakobsson, Gärling, Schneider & Fujii, 2004; Eriksson, Garvill & Nordlund, 2006; Howarth & Polyviou, 2012; Van Wee, 2014). Besides the impact on the environment, transport-induced pollution directly affects public health by its relation to diseases such as asthma, bronchitis, and, through lack of exercise, obesity (Banister, 2008; Howarth & Polyviou, 2012; Sietchiping, Permezel & Ngomsi, 2012; Van Wee, 2014; Van Wee & Ettema, 2016).

The negative externalities of private transportation are aggravated by a drastic growth in private vehicle ownership and a decrease in vehicle occupancy, particularly in urban areas (Saleh

& Farrell, 2007; Howarth & Polyviou, 2012; Gwilliam, 2013; Järv *et al.*, 2014). Not only in the developed world, but also in the developing world the growth of private car ownership goes hand in hand with a decrease in vehicle occupancy rates. In Cape Town, South Africa, for instance, both the absolute and the relative number of registered vehicles have increased in the past decade. In addition, it was found that amongst the employees of several large firms in the Central Business District, "single occupancy vehicles (SOVs) accounted for 79% of employee private car use, (which in turn accounted for 59% of the modal split)" (Behrens, Adjei, Covary, Jobanputra, Wasswa & Zuidgeest, 2015: 413). Furthermore, while developed countries are still responsible for 70 percent of the worldwide transport related greenhouse emissions, the relative contribution of upcoming economies such as China, India, and South Africa is expected to grow significantly in the next decades as a consequence of these developments (Hickman & Banister, 2010).

In the developing world, the need for adequate transport demand management, combined with integrated land and transport planning, is even more pronounced. In Sub-Saharan Africa, for instance, as one of the fastest urbanising areas in the world, the demand for transport will increase drastically in the coming decades (Candiracci, Schlosser & Allen, 2010). Often against the backdrop of an already inefficient transport system, the transport crisis not only threatens public health and causes environmental degradation; it also jeopardises urban mobility and thus the primary product of a transport system: accessibility (Candiracci *et al.*, 2010). Accessibility is crucial for both economic and social development because it provides people with access to other people, goods, jobs, and urban amenities.

It is due to the current environmental crisis, as well as the transport crisis in many areas around the world, that TDM became a component of transport planning. In South Africa, for example, the National Land Transport Act (Act 5 of 2009) requires municipalities to formulate and implement TDM techniques in their transport planning. It was recognised that, whereas technological advancements may help reduce some externalities of private vehicle use, some of the issues (such as congestion) can only be resolved if private vehicle use is reduced (Wall, Brög, Erl, Ryle & Barta, 2011). Originating in the United States in the 1970s, the concept of TDM was introduced to exploit the current capacity of transport infrastructure by improved management. This was a reaction to a declining funding base, as well as a response to environmental concerns. In addition, the oil crisis around the time, in combination with a high automotive dependency, spurred policy makers to re-evaluate transport planning (Meyer, 1999). As Meyer (1999: 576) states: "In its broadest sense, transportation-demand management is any action or set of actions aimed at influencing people's travel behaviour in such a way that alternative mobility options are presented and/or congestion is reduced".

Although TDM in general aims to change people's or households' travel behaviour, it is also an umbrella term for a wide range of strategies. Within the domain of TDM, these strategies can be categorised in many ways. One distinction is between hard (or structural) strategies and soft (or psychological) strategies. An example of the first category could be a change to the physical environment, or the usage of economic disincentives such as taxation. An example of the second category could be providing travellers with information about the negative externalities and latent costs of their travel behaviour. Another distinction is between push and pull measures. Here, push measures refer to policies to steer people away from the automobile, for instance by taxes, as opposed to pull measures, which refer to policies to make alternative travel modes more attractive (Eriksson *et al.*, 2006). See Table 2.1 for examples.

Table 2.1 | Examples of TDM measures ranked from push to pull

	Travel demand management instruments
Push	Taxation of cars and fuels
	Closure of city centres for car traffic
	Road pricing
	Parking control
	Avoiding major new road infrastructure
	Teleworking
	Land use planning encouraging shorter travel distances
	Traffic management reallocating space between modes and vehicles
	Park and ride schemes
	Improved public transport
	Improved infrastructure for walking and biking
	Public information campaign about the negative effects of driving
Pull	Social modelling where prominent public figures use alternative travel modes

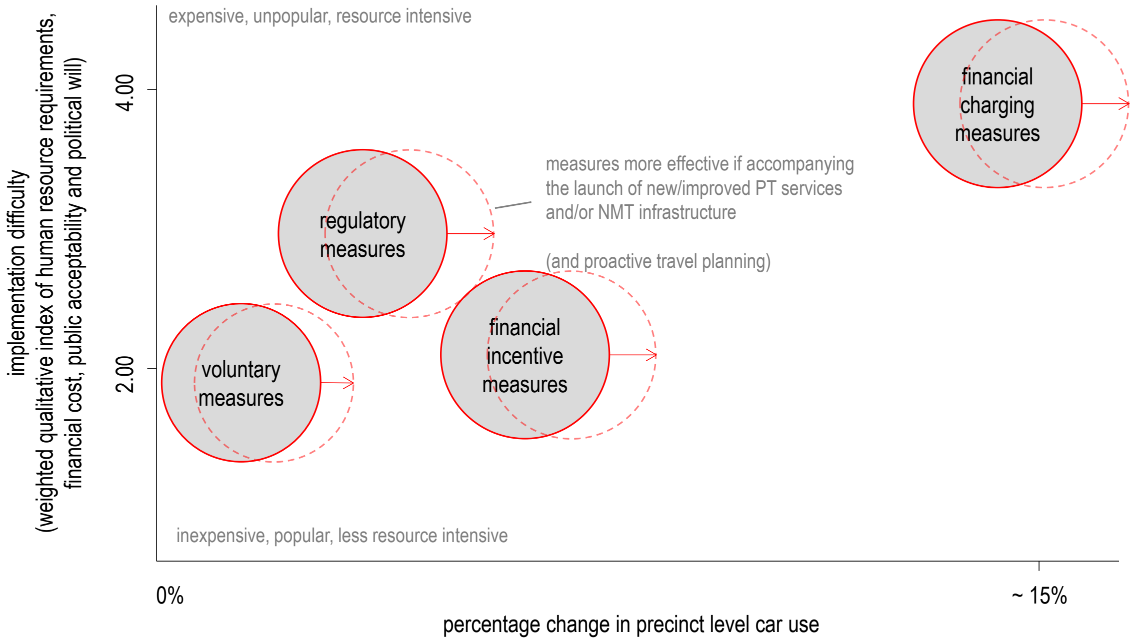
Source: Gärling *et al.* (2002: 60)

The great diversity in terms of measures and policies within the domain of TDM also means that they differ in terms of cost of implementation, effectiveness, efficiency, and public acceptability (Gärling *et al.*, 2002; Eriksson *et al.*, 2006; Wall *et al.*, 2011). Especially public acceptability, or political feasibility, plays an important role in whether or not a measure will actually contribute to a change in travel behaviour. For instance, there seems to be evidence that push measures should count on lower public acceptability than pull measures. In the context of road pricing, Eriksson *et al*. (2006) argue that an increase in the perceived level of infringement on car users' freedom to travel is directly related to a decrease in acceptability. It is on the other side of the continuum that one may find the less coercive measures, referred to, amongst other terms, as soft transport policy measures or VTBC interventions (Bamberg *et al.*, 2011). Cairns *et al.* (2008: 594) add to this that soft refers both to the response of the traveller as well as the nature of the intervention itself. This means that soft transport measures focus on changing travellers' current behaviour using psychological and economic incentives rather than changing the built environment itself through infrastructural investments (Taylor, 2007; Bamberg *et al.*, 2011; Ho, Mulley, Tsai, Ison & Wiblin, 2017).

VTBC interventions come in many forms. Some established examples of VTBC interventions include travel plans that encourage commuting employees not to use their private car; school travel plans that encourage parents not to bring their children to school by private car; ride-sharing schemes; and travel awareness campaigns (Bamberg & Möser, 2011). Well-known examples of schemes that have been promoted and implemented are the TravelSmart® programme in Australia, the Smarter Choices programme in the United Kingdom, and several travel feedback programmes in Japan. These schemes typically include a number of travel demand management instruments in order to stimulate both individuals and households to reconsider their travel behaviour (Chatterjee & Bonsall, 2009; Zhang, Stopher & Halling, 2011).

Not only do VTBC interventions have a higher acceptability amongst the general public because of their voluntary nature; they are also said to have short-term benefits such as an 'immediate' reduction in greenhouse house gas emissions. Moreover, as shown in Figure 2.1, their cost-benefit ratio when compared to large infrastructural developments is potentially attractive. The figure also suggests that voluntary measures are not only less expensive, but are also more acceptable by the public as compared to, for example, financial charging measures such as a congestion charge. In addition, as VTBC interventions aim to stimulate more sustainable modes of travel, they may also come with significant health benefits by their stimulation of more active forms of travel (Graham-Rowe, Skippon, Gardner & Abraham, 2011; Zhang *et al.*, 2011; Sietchiping *et al.*, 2012). Sustainable mobility thus not only holds benefits for the environment, but also for individual physical well-being (Banister, 2008). As such, present-day definitions of TDM include not only the management of travel demand, but more specifically refer to those interventions that aim to promote more socially, economically, and environmentally desirable forms of transport (Taylor, 2007).

Figure 2.1 | Implementation difficulty versus effectiveness of TDM measures



Source: Behrens *et al.* (2015: 415)

2.3 Understanding travel behaviour change

As the goal of soft transport policy measures is generally to alter current behaviour rather than to alter the built environment (Bamberg *et al.*, 2011), an understanding of (spatial) behaviour is necessary for designing and implementing policies that can effectively reduce congestion and improve transport service delivery (Abdulazim, Abdelgawad, Nurul Habib & Abdulhai, 2013). Most understandings of travel behaviour consider the notion that mobility is a crucial aspect of everyday life because it facilitates activity participation. This implies that travel is a derived demand and that people do not travel for its intrinsic value (De Dios Ortúzar & Willumsen, 2011).; rather, the need to travel is rooted in the desire or necessity to participate in activities at geographically dispersed locations (see for a critique on this line of thought Cresswell, 2006; Urry, 2007; Duyvendak, 2011).

Because travel is a derived demand, it is not surprising that in transport studies attention has shifted from trip-based approaches to activity-based approaches in attempts to understand human travel behaviour. The departure point of these activity-based approaches is the idea that time-use decisions are fundamental to human decision making, and that time cannot merely be reduced to the cost of activity participation – something which trip-based approaches have failed to acknowledge (Ettema & Timmermans, 1997; Bhat & Koppelman, 1999; Arentze & Timmermans, 2000; Schönfelder & Axhausen, 2010). These activity-based approaches, on the other hand, recognise that the daily life of an individual is embedded in a complex spatiotemporal context. This means that activity-travel patterns are a result of an individual's decision on how to allocate their time (Bhat & Koppelman, 1999: 120).

Törsten Hägerstrand, in his seminal work on time geography, also places time-use decisions at the core of human activity-travel behaviour. Time geography recounts the life of an individual as a continuous path through time and space (Hägerstrand, 1970). This path is constituted by actual movements through space, and activities localised in space; both of which have a temporal component. However, humans are confined in their movements in three ways, as identified by Hägerstrand (1970): capability constraints, coupling constraints, and authority constraints. Capability constraints refer to physiological and cognitive issues such as the need for sleep, shelter, and food. Coupling constraints relate to multi-actor activities for which multiple individual space-time paths have to be merged; examples are work, meeting up with a friend, and attending a meeting. Authority constraints pertain to rules and norms that act upon space and time; for instance, areas with restricted access, or the trading hours of a supermarket. These constraints are often encountered in conjunction, either reducing or reinforcing each other's impact on one's spatial behaviour (Schwanen & Kwan, 2008).

Acknowledging the spatiotemporal context of activity and travel behaviour, however, is only one side of the story. Individual preferences also play a role. The private car not only fulfils an instrumental role in facilitating travel between activity locations, but also frequently holds a symbolic and affective value (Gatersleben, 2014). Whereas the spatial layout of a region may predominantly dictate the mode choice in areas with few transport alternatives, psychological motivations play a crucial role in most travel decisions. Howarth and Polyviou (2012: 769) argue that:

> A fundamental human need for mobility ultimately drives the demand for travel; but it also has an intrinsic positive utility and is valued both for its necessity and recreational qualities, not just for being a means to reach a destination. Consequently, changing travel behaviour is not just a question of encouraging modal shifts; this requires a deep understanding of psychological processes that lead to travel behaviours in the first place.

Together, this means that travel behaviour is derived from the spatial structure of the area in which this behaviour takes place; the socio-economic position of the person; their lifestyle; and their personal preferences (Van Acker, Van Wee & Witlox, 2010; Van Acker, Goodwin & Witlox, 2016; Chen, Ma, Susilo, Liu & Wang, 2016). To effectively change travel behaviour, all four factors need to be taken into consideration.

An influential theory in understanding travel behaviour change, amongst others (see for instance Schwanen, Banister & Anable, 2012), is the Theory of Planned Behaviour (Gatersleben, 2014). In short, this theory suggests that a certain behaviour comes into existence as a result of the expected consequences of the behaviour (behavioural beliefs), the level of acceptability of the behaviour by others (normative beliefs), and the ease of executing the behaviour (control beliefs). Based on these three perceptions, an actor decides whether a behaviour is either favourable or unfavourable, which in turn leads to a behavioural intention. Behavioural change is, therefore, understood as a reaction to a change in these beliefs which, in turn, leads to a behavioural intention. After the behavioural intention has been formed, an actor is expected to carry out this behaviour when the opportunity arises to do so (Bamberg, Azjen & Schmidt, 2003; Fujii, 2009). As such, many programmes aimed at establishing a behavioural change use a combination of feedback and feedforward information. In the context of individual transport planning, feedback often provides participants with information on their current behaviour, for instance on its negative consequences, whilst feedforward suggests new travel information, for example, public transport timetables (Fujii, Bamberg, Friman & Gärling, 2009). This means that these programmes aim to influence an actor's beliefs in favour of the preferred behaviour, which could result in a behavioural response (Howarth & Polyviou, 2012).

Travel behaviour is the outcome of a complex decision-making process, but this does not mean that an individual balances the (perceived) costs and benefits of a certain transport mode before every trip. This is because behavioural intentions, once formed, tend to be stable over time. In other words, there is a role for habitual behaviour in the domain of mode choice, route choice, and departure time (Gärling & Axhausen, 2003). This implies that if one wants to influence travel behaviour, old habits need to be broken (for instance, by stimulating people to reconsider their choices), and new habits need to be established (Thøgersen, 2014). Bamberg *et al.* (2003) add to this that future travel behaviour normally does not change after the initial decision-making processes, provided that the circumstances remain relatively stable.

The amalgamation of these different factors, ranging from an individual's spatiotemporal context to the formation of travel habits, demands a better understanding of the process of change itself, as change is more a process than it is an event (Prochaska, 1991; Prochaska, Velicer, Rossi, Goldstein, Marcus, Rakowski, Fiore, Harlow, Redding, Rosenbloom & Rossi, 1994; Gatersleben & Appleton, 2007). As Prochaska (1991: 805) stated: "People do not change chronic behaviours quickly. (…) [or] discretely." Instead, Prochaska argues that there are different stages within a behavioural change, ranging from having no intention to change (pre-contemplation) and the actual change (action) to sustaining the change (maintenance), as shown in Table 2.2. Although this model has been predominantly applied in psychology with regard to problematic behaviour like smoking, Gatersleben and Appleton (2007) successfully applied it to examine individual attitudes towards cycling in order to find out how people can be persuaded to use bicycles more. They found that individuals in different stages of Prochaska's transactional model of behavioural change require different strategies. Moreover, Bamberg *et al.* (2011) used elements of this model to design a theoretical base for soft transport measures.

Table 2.2 | Stages of Prochaska *et al.*'s transactional model of behavioural change

	Characteristics	Change Strategy
Pre-contemplation	Unaware of problems, no intention to change	Increase general problem awareness
Contemplation	Aware of problems, thinking about change	Motivate, encourage specific action
Prepared for action	Intention to change in next 6 months	Assist in developing specific plans
Action	Action being taken	Feedback, social support, reinforcement
Maintenance	Has maintained action for 6 months or more	Reminders, feedback, social support

Source: Gatersleben and Appleton (2007: 304)

What the transactional model of behavioural change makes clear is that behavioural change is directly linked to an individual's decision-making process. Nonetheless, the relationship between individual actions and structural outcomes should not be overlooked. The moment an individual changes to a more sustainable behaviour as a consequence of, for instance, a VTBC policy, it is important that the effect of this change be 'locked in' in the system (Behrens & Del Mistro, 2010). Locking in the effects of a successful TDM implementation goes beyond breaking an individual's old habits and establishing new habits, as we are now talking about a system-wide change. For instance, if within a transport system someone switches from his/her private vehicle to public transport, yet at the same time someone else switches from public transport to a private vehicle, there is no effective system-wide change (Saleh & Farrell, 2007). This situation can arise, for instance, when many people would change to public transport and because of that other users would start using the road space that freed up. Therefore, for a TDM policy to succeed, the number of persons switching to public transport should be maximised and the number of persons switching to private vehicle usage should be minimised or contained. This is referred to as 'asymmetric churn'. Aggregate or system-wide change is the result of asymmetry in 'churning' individual decisions (Behrens & Del Mistro, 2010: 256). A system that appears stable could, therefore, be the result of individuals making reciprocal changes in their travel behaviour.

2.4 Voluntary travel behaviour change interventions: What do(n't) we know?

Soft transport policy measures aim to achieve an 'asymmetric churn' by motivating car users to make a voluntary switch to a more sustainable mode of travel (Richter, Friman & Gärling, 2011). Although there is no consistent definition of soft transport policy measures, a number of interventions have in the past been classified as soft (Cairns, Sloman, Newson, Anable, Kirkbride & Goodwin, 2004; Möser & Bamberg, 2008). These include: "(1) workplace travel plans, (2) school travel plans, (3), personalised travel planning, (4) travel awareness campaigns, (5), public transport information marketing, (6) car clubs, (7) car sharing schemes, (8) teleworking, (9) teleconferencing, and (10) home shopping" (Möser & Bamberg, 2008: 12). According to a quantitative meta-analysis of 141 evaluation studies conducted by Möser and Bamberg (2008), the first five types of policies have been most frequently implemented. Despite differences in the quality of the evaluation studies, the authors found an overall effect size of 15 percent. This indicates that soft transport policy measures can be effective in reducing single-occupant private vehicle usage (Möser & Bamberg, 2008). Yet there are several issues related to the quality of these evaluations (cf. Stopher & Greaves, 2006; Möser & Bamberg, 2008; Brög *et al.*, 2009; Fujii *et al.*, 2009; Richter *et al.*, 2011). These issues can be classified into three overarching themes: (1) the study design, (2) the data collection instrument, and (3) the measurement units. The next section describes the state of the research on VTBC and subsequently at the issues regarding the quality of these evaluations.

2.4.1 The effectiveness of voluntary travel behaviour change interventions

Although soft transport policies cover a wide variety of measures, the common denominator is the aim of stimulating households and individuals to voluntarily switch to more sustainable modes of travel. Because some form of information provision is present in all cases, these types of measures are known as travel feedback programmes (for instance in Japan), personal travel planning (PTP) programmes, and VTBC programmes (e.g. in the United Kingdom and Australia) (Richter *et al.*, 2011). These terms are therefore used interchangeably. The techniques differ with respect to whether they motivate travel behaviour change, whether they provide customised information, whether they request setting goals of changing travel behaviour, and whether they request plans to be made for how to change travel behaviour (Fujii *et al.*, 2009: 44).

Taylor and Ampt (2003) reviewed two interventions that have been implemented in Australia: Travel Blending® in Adelaide and IndiMark™ (Individualised Marketing) in Perth and Brisbane. As Australia has one of the world's highest greenhouse gas emissions per capita (Taylor, 2007) and faces major congestion problems, the primary aim of these programmes was to reduce private vehicle use in urban areas. After two fairly successful pilots with 50 households in Sydney and 100 households in Adelaide (Rose & Ampt, 2001), a much larger trial of Travel Blending® targeting 900 households was implemented in Adelaide. Travel Blending® started with the distribution of travel diaries, after which participants were provided with personalised information on their travel behaviour. Using a second round of travel diaries, it could subsequently be assessed to what extent the participants had changed their behaviour. In the Adelaide trial, the participants managed on average to reduce their vehicle kilometres travelled (VKT) by 21 percent. When this was measured again after six months, it was found that the reduction had been sustained, and that in some cases the VKT had been reduced even further.

The IndiMark™ intervention, as implemented in Perth and Brisbane, was tailored to individuals with the strongest propensity to change. The aim of the IndiMark™ intervention was to stimulate the use of more sustainable modes of transport such as walking, cycling, and public transport, whilst decreasing the number of car trips and VKT. In Perth, approximately 400 households were contacted of which 150 households expressed interest in reducing their car use. In Brisbane, a sample of 1,080 households was drawn that was subsequently divided into a control group and a test group. Of the 455 households that were requested to participate in the test group, 196 households were targeted with the intervention. (118 households withdrew their participation, and 98 households that already frequently used environmentally friendly modes were filtered out by the researchers). The participants that expressed interest to participate were consulted through home visits and telephone calls. Eventually, an evaluation study was conducted to establish to what extent the participants had adjusted their travel behaviour. In Perth, the trial led to a reduction of 10 percent in 'car as driver' trips, and a 14 percent reduction in VKT. In Brisbane, a reduction of 10 percent in private car trips was recorded, alongside an increase in both public transport usage and cycling trips (Taylor & Ampt, 2003). In a follow-up study twelve months after the IndiMark™ implementation in Perth, it became clear that the effects on car use and VKT had been sustained and had become even more pronounced (Taylor, 2007). Interestingly, however, in a panel evaluation study on an IndiMark™ intervention in a sub-area of Melbourne containing roughly 800 households, Seethaler and Rose (2009) came to the conclusion that a year after its implementation there was no significant difference in VKT and that the variability in VKT should instead be attributed to socio-demographic factors.

Fujii and Taniguchi (2006) reviewed the effectiveness of ten travel feedback programmes In Japan. Some of these were aimed at the workplace, for instance, in Toyonaka and Kanazawa, whilst others focused on the school environment, for instance, in Sapporo and Izumi. Although different techniques were used and the programmes had different objectives, the common feature of these interventions was that they all included a form of personalised information. Overall it was found that the interventions led to a reduction in greenhouse gas emissions and an increase in public transport usage. In Toyonaka, for example, a 10 percent reduction in car use was achieved among 100 workers of one company. In Kanazawa, on the other hand, a total of 100 employees working in 10 different companies were targeted, and an increase in bus use of 30 percent and an increase in bicycle use of 10 percent was realised. In Sapporo, the interventions led on average to a reduction of 15 percent in CO^2 emissions among 150 students and their families. Amongst 200 students and their families in Izumi, a similar reduction in CO^2 emissions was achieved. In Obihiro, where circa 15,000 households were targeted, the outcomes of a mobility management project were also quite successful with a 100 percent increase in bus ridership. Whilst in the intervention group a noticeable increase in bus ridership was found, this effect was not present in the control group. In addition, it was found that a spin-off effect was present: bus users would recommend the bus to others; suggesting a potential role for mouth-to-mouth advertising (Taniguchi & Fujii, 2007).

In a large-scale review carried out by the University of Leeds on the effectiveness of VTBC programmes around the world, the results were inconclusive. On the one hand, it was found, based on examples from Australia, the United Kingdom and Chile, that behavioural effects after an intervention were not sustained. On the other hand, some implementations of IndiMark™ in German cities resulted in a decrease in VKT with a concurrent increase in public transport use (Taylor, 2007). Similar success stories surfaced in the 'Smarter Choices' study, which reviewed the results of 24 project evaluations in the United Kingdom. The authors suggest that 'Smarter Choice' measures have the potential to lead to a significant decrease in the reduction of traffic levels, with an even stronger effect on urban peak hour traffic (Anable, Kirkbride, Sloman, Newson, Cairns & Goodwin, 2004; Cairns *et al.*, 2004, 2008). Chatterjee (2009) also reported positive results, with an average reduction in car use of 11 percent when reviewing the effectiveness of personal travel plans in eight areas in the United Kingdom. In Valencia, Spain, smaller reductions in private vehicle usage (5.3 percent) were reported in a two-wave panel study; although as a result of panel attrition the number of respondents who participated in both surveys (118) was quite low (García-Garcés, Ruiz & Habib, 2016).

Notwithstanding these successes, Howarth and Polyviou (2012: 771) suggest that:

> The provision of information alone is not sufficient to ensure behaviour change is maintained as decisions are also influenced by personalities, attitudes, norms, beliefs as well as by the degree of adequate support at the systems level to ensure travel choices are actually able to materialise.

A number of studies that relied on more than just information provision contribute to the idea that information provision alone may not be as effective as is often assumed. In the York Intelligent travel survey, for instance, incentives played a role in stimulating the use of the bus over a period of six months, and an increase in public transport ridership was achieved. Yet, the role of the incentives, in this case, was crucial as the behavioural change was not maintained six months after that when the incentives had stopped. It has to be noted that from the initial 242 individuals in the intervention group, 167 individuals responded to the questionnaire after six months and only 81 individuals answered to the final questionnaire after twelve months (Haq, Whitelegg, Cinderby & Owen, 2008). Others have suggested that social norms also play an important role (Kim, Fujii & Lee, 2013; Zhang, Schmöcker, Fujii & Yang, 2015; Hiselius & Rosqvist, 2016; Skarin, Olsson, Roos & Friman, 2017).

Overall, there seems to be a fair amount of evidence in support of the idea that soft transport measures have the potential to reduce private vehicle usage and stimulate public transport ridership. However, there are several limitations to indicate that this is too optimistic. For instance, it has been suggested that if non-coercive measures are not implemented side by side with coercive measures, behavioural change is unlikely (Gärling & Schuitema, 2007). Some scholars have argued that soft and hard transport measures should therefore rather complement each other (cf. Gärling & Schuitema, 2007; Behrens *et al.*, 2015). Furthermore, the success of an intervention may not solely depend on a willingness to change. In fact, when Kingham *et al.* evaluated employee perceptions (*n = 962*) of modal choice of the work commute in two large companies in the United Kingdom in 2001, it turned out that many employees were willing to make a change. However, the respondents also indicated that one of the most important barriers to change was the negative perception of the available alternatives (Kingham *et al.*, 2001: 159). This shows that changing travel behaviour is a complicated exercise. On a disaggregate level, individuals have to be able to accommodate a change in behaviour given their spatiotemporal context. Yet, the personal context of individual traveller is not always recognised (Howarth & Polyviou, 2012). Furthermore, the notion of the 'asymmetric churn' plays a role. A behavioural change by one individual or a group of individuals does not guarantee a system-wide change. As Bonsall (2009: 312) summarised it: "despite a decade of experience with PTP [Personal Travel Planning] we still cannot be sure whether it works."

2.4.2 Methodological uncertainties: Research design and validity

Assessing the effectiveness of a VTBC intervention poses three challenges. Firstly, it requires both ex-ante and ex-post measurements. Secondly, it requires data on the number of trips and average daily kilometres travelled. Thirdly, large sample sizes are required to accurately determine the often small changes in travel behaviour (Richardson, Seethaler & Harbutt, 2004; Stopher & Greaves, 2007; Stopher *et al.*, 2009). Given these challenges, the literature suggests two methodologies: a repeated cross-sectional study, or a longitudinal panel study. Although both methodologies have their advantages and disadvantages, the consensus is that a panel study is the preferred option because smaller sample sizes can be used as a result of a smaller variance between the samples. Using panel data thus makes it possible to curb the sampling error whilst still reducing the sample size (Richardson *et al.*, 2004; Stopher & Greaves, 2006; Brög *et al.*, 2009). This is crucial, because, as argued by Stopher and Greaves (2006: 3): "There will be some level of change that might be detected in two successive samples that is not measuring change over time, but is simply reflective of the inherent variability." This implies that, given a certain level of precision (for instance, an error level of +/- 1 percent with a confidence of 95 percent), much larger sample sizes are necessary in the case of a cross-sectional design than in a panel design. Moreover, conclusions drawn from a properly designed longitudinal panel study have a stronger statistical base (Richter *et al.*, 2011).

Not only do studies with a repeated cross-sectional design require larger samples, they also disregard temporal changes and are susceptible to regression to the mean. The regression-to-the-mean fallacy entails that extreme measurements in the pre-intervention tend to be closer to the mean when they are measured again in the post-intervention. Because of this, the effects of an intervention may be overestimated (Torgerson & Torgerson, 2008). As Behrens and Del Mistro (2010: 259) formulate it in the context of travel behaviour: "The disadvantage of this method [repeated cross-sectional design] is an inability to identify or explain intra-personal change, and the large (and expensive) samples necessary to account for variability between the separate sequential samples and to enable robust inferences". In fact, in a study by Seethaler and Rose (2009) in which no sustained effect of a TravelSmart® intervention was found, it is suggested that the initial positive outcome may be explained by a regression to the mean. A way to avoid these problems is by implementing a different study design, such as a randomised controlled trial (RCT), in which participants are randomly allocated to either control or intervention groups (Torgerson & Torgerson, 2008).

Where cross-sectional designs can only partly account for confounding factors by controlling for different environmental and socio-demographical variables in statistical analysis, in an RCT these issues are not of concern, as latent and unmeasurable variables are also randomised. As Torgerson and Torgerson (2008: 28) state: "Randomisation ensures that, on average, the two or more groups that are formed are similar in all variables that will affect the outcome." This implies that any difference between control and intervention group can be attributed to the intervention that has been administered, yielding a very high internal validity. Naturally, in order to ensure a high external validity as well, a random sample of sufficient size and from an appropriate population has to be drawn that sufficiently accounts for self-selection biases (Torgerson & Torgerson, 2008; Bonsall, 2009; Richter *et al.*, 2011).

When considering the research designs of previous evaluation studies on VTBC interventions, Möser and Bamberg (2008) found that in all of the 141 studies they analysed, a study design was implemented in which the car use of a single group was measured before and after the intervention. This poses questions regarding inferred causality. Furthermore, the external validity of these studies could also be questioned, as inferences were drawn from biased or non-representative samples (Bamberg *et al.*, 2011). Because of this, Möser and Bamberg (2008) conclude that there is at the moment no solid empirical evidence that a broad implementation of VTBC programmes will actually reduce car use. According to Bamberg *et al.* (2011), the outcomes of the meta-analysis conducted by Fujii *et al.* (2009) were more robust because a control or comparison group was incorporated in the studies evaluated there. However, the number of evaluated studies was small, and most of the studies did not have a large representative sample.

In a more recent study on the available evidence on the effectiveness of personalised travel feedback programmes, Graham-Rowe *et al.* (2011) evaluated 77 studies and grouped them according to their methodological quality. Of the 77 studies, only twelve had a high methodological quality: six studies using an experimental design such as a randomised controlled trial; four studies using a quasi-experimental design, in which control groups were used, but not randomised; and two studies using a cohort-analytic method in which control groups were used, but in which the intervention was not managed by the researchers. For 48 out of the 77 studies, the methodology was considered to be weak, leading to the conclusion that these results should not be used to make inferences about the effectiveness of the interventions. When looking only at the evidence from those studies classified as having a high methodological quality, the results indicate that car use can be reduced, but that the effects are modest. In addition, it is suggested that interventions are most effective when the participants already have a strong positive attitude towards changing behaviour, or when they recently moved or got a new job and have yet to establish new travel patterns. Nonetheless, there is also criticism against randomised controlled trials, such as that it can actually create additional biases (see Melia, 2015).

2.4.3 Methodological uncertainties: Precision of measurement instruments

Even though travel behaviour can be considered as habitual, people's daily travel patterns display a great deal of intrapersonal variability (Hanson & Huff, 1988; Stopher, Clifford & Montes, 2008; Stopher, Kockelman, Greaves & Clifford, 2008). For a precise evaluation of the effectiveness of a VTBC intervention, larger samples consisting of multi-day travel data are required that account for spatial and temporal differences, socio-demographic diversity, variety in travel modes, and differentiation in trip purposes (Richardson, 2003; Richardson *et al.*, 2004; Taylor, 2007; Bonsall, 2009; Nitsche, Widhalm, Breuss, Brändle & Maurer, 2014). Obtaining disaggregate multi-day travel data through traditional methods such as paper-based survey instruments is a complex endeavour as a result of their high respondent burden, the underreporting of trips, and costly administrative processes (Behrens, 2004; Stopher *et al.*, 2009; Jariyasunant, Carrel, Ekambaram, Gaker, Kote, Sengupta & Walker, 2011; Nitsche *et al.*, 2014; Prelipcean, 2016). More cost-effective, computer-assisted methods do not solve all limitations associated with traditional methods either. That is because these methods are also contingent on the respondent's ability to accurately remember his/her movements and activities. Moreover, because evaluation programmes often measure vehicle kilometres travelled (VKT) and in-vehicle time, travel diaries alone may not be suitable for gathering these types of data.

The use of diaries, both trip-based and activity-based (for the differences between these approaches see Stopher, 1992), is well-established in the transportation field. Since the seminal work of Jean Wolf (2000), however, researchers have started to experiment with acquiring high-resolution space-time data for research purposes by means of GPS devices (cf. Bohte & Maat, 2009; Chen & Kwan, 2012; Nitsche, Widhalm, Breuss & Maurer, 2012; Feng & Timmermans, 2013, 2016; Shen & Stopher, 2013; Shoval, Kwan, Reinau & Harder, 2014). GPS offers several potential advantages for researchers to collect precise data on trip origin, trip destination, trip duration, trip timing, and route choice. In addition, GPS devices can be used to corroborate the quality of other data sources, such as activity and trip diaries (Wolf, 2000).

According to Kelly *et al.* (2013: 455), "studies using self-reported journey duration over multiple journeys and days are likely to be substantially over-estimating the travel time or exposure to active and sedentary travel". Although the use of GPS in the context of acquiring data on individual travel behaviour still has to deal with a number of drawbacks, the techniques show much potential as (partial) substitutes for conventional paper diaries (Du & Aultman-Hall, 2007; Bohte & Maat, 2009; Schuessler & Axhausen, 2009; Nitsche *et al.*, 2014). Because of the possibilities to accurately collect spatiotemporal data, a number of authors have recommended the usage of GPS devices, both in the context of assessing VTBC interventions and in the context of travel data collection in general (Stopher, Clifford, *et al.*, 2008; Bohte & Maat, 2009; Bonsall, 2009; Chatterjee & Bonsall, 2009; Krygsman & Nel, 2009; Behrens & Del Mistro, 2010; Richter *et al.*, 2011; Shen & Stopher, 2014; Chai, Chen, Liu, Tana & Ma, 2014; Feng & Timmermans, 2014; Meloni & Sanjust, 2014; Shafique & Hato, 2015; Meloni, Sanjust Di Teulada & Spissu, 2016). Moreover, GPS is nowadays available on most smartphones. Accordingly, Nitsche *et al.* (2012) argue that there is much potential to automatically reconstruct trips from data collected by means of a smartphone application.

Other than their ubiquity, smartphones have at least four advantages to augment current travel data collection methods. Firstly, the fact that an application can be programmed makes data collection very flexible. Also, researchers can have real-time access to the data and can therefore monitor the data collection. Secondly, developing an application is very cost-efficient. With a properly designed data collection exercise using smartphones, face-to-face interaction between the research team and their participants can be minimised. Thirdly, smartphones are a very unobtrusive way to collect data from respondents (Zhou & Golledge, 2007; Bohte & Maat, 2009; Chen, Gong, Lawson & Bialostozky, 2010). Many people carry their smartphones with them continuously. The respondent burden of the data collection can therefore be kept to a minimum (Raento, Oulasvirta & Eagle, 2009), especially when the GPS application runs only in the background. Finally, smartphone penetration is increasing all around the world (Herrera, Work, Herring, Ban, Jacobson & Bayen, 2010).

Although travel data collection with smartphones is still in its infancy, in recent years a number of scholars have experimented with developing applications for exactly that purpose. Some researchers have experimented with a smartphone application to automatically reconstruct its user's trips (cf. Li, Dai, Sahu & Naphade, 2011; Nitsche *et al.*, 2012, 2014;

Abdulazim *et al.*, 2013; Ferrer & Ruiz, 2014). Others have integrated a prompted recall survey in which the user was requested to update an automatically generated activity diary (for example Cottrill, Pereira, Zhao, Dias, Lim, Ben-Akiva & Zegras, 2013; Jianchuan, Zhicai, Guangnian & Xuemei, 2014). The focus of most of these experiments was to automatically reconstruct an activity diary, for instance by neural networks, machine learning, and other data classification techniques, although smartphone applications have also been implemented in estimating the velocity of traffic flows (Herrera *et al.*, 2010). Where the majority of the experiments have used small sample sizes so far, large-scale experiments have been done in Singapore and the Netherlands, among other countries (Cottrill *et al.*, 2013; Geurs, Thomas, Bijlsma & Douhou, 2015; Zhao, Ghorpade, Pereira, Zegras & Ben-Akiva, 2015).

Specific to measuring the effectiveness of a VTBC programme, the usage of smartphones has the potential to not only be involved in the data collection itself, but also in administering the intervention. Because smartphone owners are always connected to the rest of the world, researchers have the opportunity to directly communicate with their respondents in ways not previously possible. Based on the two-way stream of information facilitated by smartphones, it not only becomes possible to analyse an individual's spatial behaviour, but it also becomes possible to give users feedback on their travel behaviour and the negative externalities this behaviour causes. For example, "[c]arbon dioxide emissions associated with transport can be considered a latent attribute of any trip, as it very unlikely that it will be considered by an individual in the normal course of events" (Brazil & Caulfield, 2013: 100). However, if individuals are aware of these latent costs, they may actually try to modify their behaviour towards more sustainable modes or activity schedules (Froehlich, Dillahunt, Klasnja, Mankoff, Consolvo, Harrison & Landay, 2009; Raento *et al.*, 2009; Jariyasunant *et al.*, 2011). In this way, the research instrument and research intervention could blend together into one (Froehlich *et al.*, 2009).

The possibility of continuous measurement and provision of immediate feedback makes smartphones an attractive data collection instrument. This is exactly what happened with a small-scale pilot on phones running the UbiGreen mobile application (Froehlich *et al.*, 2009). Using the sensing capabilities of the mobile phone as well as self-reported data on their travel behaviour, users were provided icon feedback on their environmental impact in terms of greenhouse gas emissions. A sequence of icons represented how 'green' their transport mode was, as well as what other benefits the transport mode had, for instance, that it saved money compared to driving. Similarly, researchers from the University of California, Berkeley recruited 28 participants that were requested to install a smartphone application on their devices capable of giving personalised feedback and statistics on the participant's travel behaviour. Originating from the idea of the 'Quantified Self' movement, i.e. the idea that individuals want to quantify a variety of their behavioural characteristics such as calorie intake and sleeping patterns, the idea of the 'Quantified Traveler' was coined. The aim was to provide participants with a direct return on their efforts in travel data collection. On a feedback website, participants could see not only the trips they made, but also other trip characteristics such as the trip duration, trip cost, CO_2 emissions, and the number of calories burnt. It was found, through a before-and-after attitudinal survey, that the feedback had increased the participants' awareness of the impact of their transport footprint (Jariyasunant *et al.*, 2011).

Another application that fits into the self-tracking movement is the CO2GO application. Although only tested in an experimental setting and mostly focused on trip diary reconstruction and transport mode recognition, CO2GO can provide relatively accurate CO_2 emissions estimates by using a combination of different smartphone sensors (Manzoni, Maniloff, Kloeckl & Ratti, 2011). The European SUNSET (Sustainable Social Network Services for Transport) project takes self-tracking even further and "motivate[s] users on a personal level to change their mobility behaviour" with their tripzoom application (Bie, Bijlsma, Broll, Cao, Hjalmarsson, Hodgson, Holleis, van Houten, Jacobs, Koolwaaij, Kusumastuti & Luther, 2012: 125). In the same way as the 'Quantified Traveler', the tripzoom application is capable of providing feedback on personal mobility (SUNSET Consortium, 2016). However, tripzoom also incorporates incentives and access to social networks to stimulate cycling behaviour. Based on individual interest, these incentives can be money, time, or information. Examples include earning credit by avoiding rush hour, or by setting goals and stimulating the user to adhere to these goals. In addition, tripzoom incorporates an element of 'gamification', which can serve as a supporting framework (Haq *et al.*, 2008; Kormos, Gifford & Brown, 2015). Users can share their trip statistics with other users, as well as with friends, enabling comparisons (Bie *et al.*, 2012; SUNSET Consortium, 2014).

Aside from the personalised feedback and unique social networking element, tripzoom's technology has also been implemented in experimental settings in Enschede (the Netherlands), Leeds (the United Kingdom), and Gothenburg (Sweden) (Groenewolt, Meeuwissen & Bijlsma, 2014; Hjalmarsson, Karlsson & Van Amelsfort, 2014; SUNSET Consortium, 2014; Thomopoulos, Grant-Muller & Hodgson, 2014). The results indicate that most of the participants did not significantly change their travel behaviour, although small changes in morning peak hour avoidance and increased public transport ridership were reported. In addition, some evidence suggests that participants took bicycles more often. The technology is currently being used in a number of projects in the Netherlands aiming at stimulating more sustainable travel behaviour, particularly in projects aimed at stimulating cycling behaviour (SUNSET Consortium, 2016).

2.4.4 Methodological uncertainties: Measurement issues

Important variables in assessing the effectiveness of VTBC interventions include the exact distances and routes travelled on an individual level. In fact, because the effects of a VTBC intervention can be expected to be small compared to more coercive measures, it is of crucial importance to measure travel behaviour as accurately as possible. This means that an evaluation should be designed in such a way that only the measured variation in travel behaviour that is a result of the intervention should be attributed to the intervention. Something that has been largely overlooked in studies so far is the measurement metric itself, although there seems to be a consensus that transport behaviour should be considered from within the context of the individual activity behaviour and activity schedule from which it is derived (Kitamura, 1988; Meloni & Sanjust, 2014).

Steered by physiological, coupling, and authority constraints, as per Hägerstrand's time-geographical framework, time-use decisions are the main drivers behind individual spatial decision making and behaviour (Bhat & Koppelman, 1999). Because of this, it is important that data collection instruments be able to capture the sequence as well as the attributes of activities and trips (Meloni & Sanjust, 2014). The same applies to types of trips and whether the trip is bounded or non-bounded, which in turn depends on the type of activity. Whereas bounded trips are fixed in both time and space (such as trips to work), non-bounded trips are more flexible in terms of time, space, or both (such as leisure activities). Other trips may be positioned in the middle of this continuum, in the sense that the time of the activity is rigid, but the location is variable (Vilhelmson, 1999; Naess, 2006). Despite the wide recognition of the importance of these temporal and sequential characteristics of activity and travel behaviour, both on the implementation side and on the evaluation side of VTBC programmes, the measures deployed seem to be trying to capture something extremely dynamic and fluid with static measures. Howarth and Polyviou (2012: 778) accurately summarise this as follows:

> Travel behaviour schemes have demonstrated significant potential in delivering sustainability at the local level, yet these schemes do not fully recognise that travel decisions are unique and are made at the individual level embedded in the specific context within which the traveller is located.

Although metrics such as number of trips and VKT are widely accepted and useful measures in describing travel behaviour, in the context of measuring and stimulating behavioural change they may not be able to tell the entire story, as some people may express a willingness to change but are not actually able to (Kingham *et al.*, 2001). This is because "people's daily activity-travel is constrained by the spatiotemporal availability of alternatives for activity destinations" (Ren, Tong & Kwan, 2014: 330). In other words, one's accessibility should be considered. The analysis of individual accessibility from a time-geographical perspective has led to the development of so-called space-time accessibility (STA) measures (cf. Neutens, Delafontaine, Schwanen & van de Weghe, 2012). STA measures are based on one of the key concepts of time geography: the space-time prism. This prism is the graphical representation of all possible paths an individual can take within a given time budget between two primary activities, the (average) speed of the mode of transport being used, and the network distance between these activities (Neutens, Schwanen & Witlox, 2011; Long & Nelson, 2013).

Whereas the space-time prism is a three-dimensional object, as shown in Figure 1.1, it can be projected onto a two-dimensional surface, commonly referred to as the potential path area (PPA). The PPA represents the area which an individual can potentially visit given the time available in between two primary activities within his/her actual action space (Dijst, De Jong & Van Eck, 2002; Schwanen & De Jong, 2008). The extent of the PPA is therefore equal to an individual's accessibility to opportunities. Even though the PPA is still a static way of representing movement, it includes the temporal element of travel, and is therefore an important tool to quantitatively analyse movement data. In fact, the PPA is frequently employed as a spatial range method in the quantitative assessment of human movement (Long & Nelson, 2013). Another way of describing movement data in the same class of spatial range methods, although more temporally naïve, is by means of spatial polygons, such as the minimum bounding area of a set of points. A third option would be to consider an 'activity space' or 'action space' (Järv *et al.*, 2014). In essence, the identification of some form of a spatial polygon comes down to a PPA. However, the temporal or sequential component of a movement is not explicitly considered.

Whereas the PPA seems to be more suitable for designing or analysing VBTC interventions, simpler methods should not be disregarded. The notion of the activity space, for instance, can still provide useful information, as "the ellipse represents the conceptual (perhaps minimal) area over which we know the individual is willing or able to engage in activities" (Newsome, Walcott & Smith, 1998: 362). As such, it could be a useful and less computationally intensive way of identifying intervening opportunities within an individual's spatiotemporal context. Even the ratio of the minor versus the major axis of the ellipse could already provide some insights (Dijst, 1995; Newsome *et al.*, 1998). Whereas long and flat ('stretched') ellipses probably have less intervening opportunities, the number of intervening opportunities in round ellipses is likely to be very different. If this were to be combined with an identification of the actual routes travelled, which would not only provide an accurate estimate of for instance the vehicle kilometres travelled, this could suggest intervening opportunities in the vicinity of preferred travel routes.

2.5 Directions for future research

It is due to the current environmental crisis, as well as the transport problems in many areas around the world, that travel demand management has become part of transport planning. As transport travel management became more popular, and slowly shifted its focus from only reducing congestion to moving people to more sustainable modes of transport, VTBC programmes or soft transport measures, came into existence. VTBC programmes are the programmes in which private car users are not forced by external policies or disincentives, but are encouraged to make their own decision to change their travel behaviour, for instance, through personalised feedback and personalised information campaigns. Even though soft transport measures have been widely implemented, there is still a lot of uncertainty surrounding the effectiveness of these programmes. This chapter has set out to explore the current body of knowledge on VTBC interventions.

Throughout the last decade, VTBC programmes have been implemented in several countries, particularly Australia, Japan, and the United Kingdom. Many of these programmes have yielded positive results, with reductions in private vehicle usage of around 10 percent. Notwithstanding these results, three issues have surfaced. The first is that the cross-sectional research designs that have been most frequently employed in the evaluation studies of VTBC programmes may not be suitable to properly identify the effects of such an intervention, because of the presence of large sampling errors and the lack of appropriate control groups. The second issue is the precision of the measurement instruments. In many cases, evaluations of VTBC programmes have relied heavily on traditional activity-travel diaries or on prompted-recall methods collected during short periods of time. As the expected changes in VTBC interventions are often small, these instruments may have been too blunt to pick up these subtle changes. The third issue is that the evaluation studies have focused mostly on the number of car trips and the vehicle kilometres travelled; as such they do not properly account for the context in which travel behaviour occurs.

Based on the identified issues, two lines of research are needed to address the question of whether VTBC interventions work and, if so, in which contexts. The first line of research should focus on field experiments that include a randomisation of participants to experimental and control groups. To account for the variability of travel behaviour, these should preferably be longitudinal studies or, at the very least, panel studies that cover multiple days of travel. Another option here would be to target the intervention at populations or locations that have ample opportunities for participants to change their behaviour so that circumstances and causality can be tested. The second line of research that is proposed should focus on field experiments that take full advantage of more precise measurement instruments such as GPS-enabled smartphones. Moreover, it should be explored whether there are more complete measures to analyse travel behaviour, for instance, by turning to the large literature on space-time accessibility measures or 'spatial range methods' that have been developed to been developed to analyse movement data quantitatively. This is not to suggest that the traditional measures like the number of trips and VKT should be abandoned; they should rather be supplemented.

Although these two research lines may seem to push into two different directions, they strengthen each other, especially when considering recent advancements in the field of location-aware technologies. With a steady rise in smartphone penetration all around the world, the sensing capabilities of smartphones can be further exploited as an instrument to give participants immediate, personalised feedback on their travel behaviour. The possibilities to incorporate highly personalised feedback into personalised incentives, as well as to include elements of social recognition, can aid longitudinal data collection. Especially in the realm of mobile technologies such as GPS, suggestions in the existing literature call for integrated methods that can deal with activity reconstruction and route reconstruction of raw GPS measurements that take both the temporal and physical characteristics of the (built) environment into account.

References

Abdulazim, T., Abdelgawad, H., Nurul Habib, K.M. & Abdulhai, B. 2013. Using smartphones and sensor technologies to automate the collection of travel data. *Transportation Research Record: Journal of the Transportation Research Board*. 2383:44–52.

Van Acker, V., Van Wee, B. & Witlox, F. 2010. When transport geography meets social psychology: Toward a conceptual model of travel behaviour. *Transport Reviews*. 30(2):219–240.

Van Acker, V., Goodwin, P. & Witlox, F. 2016. Key research themes on travel behavior, lifestyle, and sustainable urban mobility. *International Journal of Sustainable Transportation*. 10(1):25–32.

Anable, J., Kirkbride, A., Sloman, L., Newson, C., Cairns, S. & Goodwin, P. 2004. *Smarter choices: Changing the way we travel. Case study reports*. London: Department for Transport.

Arentze, T. & Timmermans, H.J.P. 2000. *ALBATROSS: A learning-based transportation oriented simulation system*. Eindhoven: European Institute of Retailing and Services Studies.

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

Bamberg, S. & Möser, G. 2011. Please Mr. Brög, give us your data! Reply to the comment of Wall, Brög, Erl, Ryle, & Barta on our paper "The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence". *Journal of Environmental Psychology*. 31(3):270–271.

Bamberg, S., Azjen, I. & Schmidt, P. 2003. Choice of travel mode in the Theory of Planned Behavior: The roles of past behavior, habit, and reasoned action. *Basic and Applied Social Psychology*. 25(3):175–187.

Bamberg, S., Fujii, S., Friman, M. & Gärling, T. 2011. Behaviour theory and soft transport policy measures. *Transport Policy*. 18(1):228–235.

Banister, D. 2008. The sustainable mobility paradigm. *Transport Policy*. 15(2):73–80.

Behrens, R. 2004. Understanding travel needs of the poor: Towards improved travel analysis practices in South Africa. *Transport Reviews*. 24(3):317–336.

Behrens, R. & Del Mistro, R. 2010. Shocking habits: Methodological issues in analyzing changing personal travel behavior over time. *International Journal of Sustainable Transportation*. 4(5):253–271.

Behrens, R., Adjei, E., Covary, N., Jobanputra, R., Wasswa, B. & Zuidgeest, M. 2015. A travel behaviour change framework for the City of Cape Town. In *Proceedings of the 34th Southern African Transport Conference (SATC 2015)*. Pretoria: Southern African Transport Conference. 412–430.

Bhat, C.R. & Koppelman, F.S. 1999. A retrospective and prospective survey of time-use research. *Transportation*. 26(2):119–139.

Bie, J., Bijlsma, M., Broll, G., Cao, H., Hjalmarsson, A., Hodgson, F., Holleis, P., van Houten, Y., et al. 2012. Move Better with tripzoom. *International Journal On Advances in Life Sciences*. 4(3):125–135.

Bohte, W. & Maat, K. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*. 17(3):285–297.

Bonsall, P. 2009. Do we know whether personal travel planning really works? *Transport Policy*. 16(6):306–314.

Brazil, W. & Caulfield, B. 2013. Does green make a difference: The potential role of smartphone technology in transport behaviour. *Transportation Research Part C: Emerging Technologies*. 37:93–101.

Brög, W., Erl, E., Ker, I., Ryle, J. & Wall, R. 2009. Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*. 16(6):281–292.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbride, A. & Goodwin, P. 2004. *Smarter Choices - Changing the Way We Travel*. London: Department for Transport.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbridge, A. & Goodwin, P. 2008. Smarter choices: Assessing the potential to achieve traffic reductions using 'soft measures'. *Transport Reviews*. 28(5):593–618.

Candiracci, S., Schlosser, C. & Allen, H. 2010. *A new perspective: Sustainable mobility in African cities*. Nairobi, Kenya: United Nations Human Settlements Programme (UN-HABITAT).

Chai, Y., Chen, Z., Liu, Y., Tana & Ma, X. 2014. Space-time behavior survey for smart travel planning in Beijing, China. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 79–90.

Chapman, L. 2007. Transport and climate change: A review. *Journal of Transport Geography*. 15(5):354–367.

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*. 16(6):293–305.

Chatterjee, K. & Bonsall, P. 2009. Special issue on evaluation of programmes promoting voluntary change in travel behavior. *Transport Policy*. 16(6):279–280.

Chen, X. & Kwan, M.-P. 2012. Choice set formation with multiple flexible activities under space–time constraints. *International Journal of Geographical Information Science*. 26(5):941–961.

Chen, C., Gong, H., Lawson, C. & Bialostozky, E. 2010. Evaluating the feasibility of a passive travel survey collection in a complex urban environment: Lessons learned from the New York City case study. *Transportation Research Part A: Policy and Practice*. 44(10):830–840.

Chen, C., Ma, J., Susilo, Y., Liu, Y. & Wang, M. 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies*. 68:285–299.

Cottrill, C.D., Pereira, F.C., Zhao, F., Dias, I., Lim, H.B., Ben-Akiva, M. & Zegras, P. 2013. Future Mobility Survey - Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board*. 2354:59–67.

Cresswell, T. 2006. *On the move: Mobility in the modern Western world*. 1st ed. New York: Routledge.

Dijst, M. 1995. *Het elliptisch leven: Actieruimte als integrale maat voor bereik en mobiliteit - modelontwikkeling met als voorbeeld tweeverdieners met kinderen in Houten en Utrecht*. Published doctoral dissertation. Utrecht: Koninklijk Nederlands Aardrijkskundig Genootschap.

Dijst, M., De Jong, T. & Van Eck, J.R. 2002. Opportunities for transport mode change: An exploration of a disaggregated approach. *Environment and Planning B: Planning and Design*. 29(3):413–430.

De Dios Ortúzar, J. & Willumsen, L.G. 2011. *Modelling transport*. 4th ed. J. De Dios Ortúzar & L.G. Willumsen (eds.). Chichester, West Sussex, United Kingdom: Wiley.

Du, J. & Aultman-Hall, L. 2007. Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues. *Transportation Research Part A: Policy and Practice*. 41(3):220–232.

Duyvendak, J.W. 2011. *The politics of home: Belonging and nostalgia in Europe and the United States*. 1st ed. London: Palgrave Macmillan.

Eriksson, L., Garvill, J. & Nordlund, A.M. 2006. Acceptability of travel demand management measures: The importance of problem awareness, personal norm, freedom, and fairness. *Journal of Environmental Psychology*. 26(1):15–26.

Ettema, D. & Timmermans, H.J.P. Eds. 1997. *Activity-based approaches to travel analysis*. Oxford: Pergamon.

Feng, T. & Timmermans, H.J.P. 2013. Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies*. 37:118–130.

Feng, T. & Timmermans, H.J.P. 2014. Multi-week travel surveys using GPS devices: Experiences in the Netherlands. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile Technologies for Activity-Travel Data Collection and Analysis*. Hershey, Pennsylvania: IGI Global. 104–118.

Feng, T. & Timmermans, H.J.P. 2016. Comparison of advanced imputation algorithms for detection of transportation mode and activity episode using GPS data. *Transportation Planning and Technology*. 39(2):180–194.

Ferrer, S. & Ruiz, T. 2014. Using smartphones to capture personal travel behavior. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 171–186.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B. & Landay, J.A. 2009. UbiGreen: Investigating a mobile tool for tracking and supporting green transportatoin habits. In *Proceedings of the 27th international conference on human factors in computing systems - CHI 09*. New York: ACM Press. 1043–1052.

Fujii, S. 2009. Retrospectives and perspectives on travel behavioral modification research: A report of the "Behavior Modification" workshop. In R. Kitamura, T. Yoshii, & T. Yamamoto (eds.). *The expanding sphere of travel behaviour research: Selected papers from the 11th International Conference on Travel Behaviour Research*. Bingley, UK: Emerald Group Publishing Limited. 439–446.

Fujii, S. & Taniguchi, A. 2006. Determinants of the effectiveness of travel feedback programs - A review of communicative mobility management measures for changing travel behaviour in Japan. *Transport Policy*. 13(5):339–348.

Fujii, S., Bamberg, S., Friman, M. & Gärling, T. 2009. Are effects of travel feedback programs correctly assessed? *Transportmetrica*. 5(1):43–57.

García-Garcés, P., Ruiz, T. & Habib, K.M.N. 2016. Effect of travel behaviour change programmes on time allocated to driving. *Transportmetrica A: Transport Science*. 12(1):1–19.

Gärling, T. & Axhausen, K.W. 2003. Introduction : Habitual travel choice. *Transportation*. 30:1–11.

Gärling, T. & Schuitema, G. 2007. Travel demand management targeting reduced private car use: Effectiveness, public acceptability and political feasibility. *Journal of Social Issues*. 63(1):139–153.

Gärling, T., Eek, D., Loukopoulos, P., Fujii, S., Johansson-Stenman, O., Kitamura, R., Pendyala, R. & Vilhelmson, B. 2002. A conceptual analysis of the impact of travel demand management on private car use. *Transport Policy*. 9(1):59–70.

Gatersleben, B. 2014. Psychological motives for car use. In T. Gärling, D. Ettema, & M. Friman (eds.). *Handbook of sustainable travel*. Dordrecht: Springer Netherlands. 85–94.

Gatersleben, B. & Appleton, K.M. 2007. Contemplating cycling to work: Attitudes and perceptions in different stages of change. *Transportation Research Part A: Policy and Practice*. 41(4):302–312.

Geurs, K.T., Thomas, T., Bijlsma, M. & Douhou, S. 2015. Automatic trip and mode detection with move smarter: First results from the Dutch mobile mobility panel. *Transportation Research Procedia*. 11:247–262.

Graham-Rowe, E., Skippon, S., Gardner, B. & Abraham, C. 2011. Can we reduce car use and, if so, how? A review of available evidence. *Transportation Research Part A: Policy and Practice*. 45(5):401–418.

Groenewolt, B., Meeuwissen, M. & Bijlsma, M. 2014. *SUNSET Sustainable Social Network Services for Transport Deliverable D7.2 "Living Lab Report Enschede"*. Enschede: SUNSET Consortium.

Gwilliam, K. 2013. Cities on the move - Ten years after. *Research in Transportation Economics*. 40(1):3–18.

Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional Science Association*. 24(1):6–21.

Hanson, S. & Huff, O.J. 1988. Systematic variability in repetitious travel. *Transportation*. 15(1–2):111–135.

Haq, G., Whitelegg, J., Cinderby, S. & Owen, A. 2008. The use of personalised social marketing to foster voluntary behavioural change for sustainable travel and lifestyles. *Local Environment*. 13(7):549–569.

Hardin, G. 1968. The tragedy of the commons. *Science*. 162(3859):1243–1248.

Herrera, J.C., Work, D.B., Herring, R., Ban, X., Jacobson, Q. & Bayen, A.M. 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies*. 18(4):568–583.

Hickman, R. & Banister, D. 2010. Low-carbon transport in a developed megalopolis: The case of London. In W. Rothengatter, Y. Hayashi, & S. Wolfgang (eds.). *Transport moving to climate intelligence: New chances for controlling climate impacts of transport after the economic crisis*. New York: Springer. 41–52.

Hiselius, L.W. & Rosqvist, L.S. 2016. Mobility management campaigns as part of the transition towards changing social norms on sustainable travel behavior. *Journal of Cleaner Production*. 123:34–41.

Hjalmarsson, A., Karlsson, M. & Van Amelsfort, D. 2014. *SUNSET Sustainable Social Network Services for Transport Deliverable D7.4 "Living Lab Report Gothenburg"*. Gothenbug: SUNSET Consortium.

Ho, C., Mulley, C., Tsai, C.-H., Ison, S. & Wiblin, S. 2017. Area-wide travel plans - Targeting strategies for greater participation in green travel initiatives: A case study of Rouse Hill Town Centre, NSW Australia. *Transportation*. 44(2):325–352.

Howarth, C.C. & Polyviou, P. 2012. Sustainable travel behaviour and the widespread impacts on the local economy. *Local Economy*. 27(7):764–781.

IPCC. 2014. *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Core Writing Team, R.K. Pachauri, & L.A. Meyers (eds.). Geneva, Switzerland: IPCC.

Jariyasunant, J., Carrel, A., Ekambaram, V., Gaker, D.J., Kote, T., Sengupta, R. & Walker, J.L. 2011. *The Quantified Traveler: Using personal travel data to promote sustainable transport behavior*. (UCTC-FR-2011-10). Berkeley: University of California Transportation Center. [Online], Available: http://www.uctc.net/research/papers/UCTC-FR-2011-10.pdf [2015, November 12].

Järv, O., Ahas, R. & Witlox, F. 2014. Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*. 38:122–135.

Jianchuan, X., Zhicai, J., Guangnian, X. & Xuemei, F. 2014. Smartphone-based travel survey: A pilot study in China. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 209–223.

Kelly, P., Krenn, P., Titze, S., Stopher, P. & Foster, C. 2013. Quantifying the difference between self-reported and Global Positioning Systems-measured journey durations: A systematic review. *Transport Reviews*. 33(4):443–459.

Kim, J., Fujii, S. & Lee, B. 2013. Strategies to promote sustainable mobility management incorporating heterogeneity. *International Journal of Sustainable Transportation*. 7(2):107–124.

Kingham, S., Dickinson, J. & Copsey, S. 2001. Travelling to work: Will people move out of their cars. *Transport Policy*. 8(2):151–160.

Kitamura, R. 1988. An evaluation of activity-based travel analysis. *Transportation*. 15(1–2):9–34.

Kitamura, R., Fujii, S. & Pas, E.I. 1997. Time-use data, analysis and modeling: Toward the next generation of transportation planning methodologies. *Transport Policy*. 4(4):225–235.

Kormos, C., Gifford, R. & Brown, E. 2015. The influence of descriptive social norm information on sustainable transportation behavior: A field experiment. *Environment and Behavior*. 47(5):479–501.

Krygsman, S.C. & Nel, J.H. 2009. The use of global positioning devices in travel surveys - A developing country application. In *Proceedings of the 28th Southern African Transport Conference (SATC 2009)*. Pretoria: Southern African Transport Conference. 108–118.

Li, M., Dai, J., Sahu, S. & Naphade, M. 2011. Trip analyzer through smartphone apps. In *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.* Chicago: ACM. 537–540.

Long, J.A. & Nelson, T.A. 2013. A review of quantitative methods for movement data. *International Journal of Geographical Information Science*. 27(2):292–318.

Loukopoulos, P., Jakobsson, C., Gärling, T., Schneider, C.M. & Fujii, S. 2004. Car-user responses to travel demand management measures: Goal setting and choice of adaptation alternatives. *Transportation Research Part D: Transport and Environment*. 9(4):263–280.

Manzoni, V., Maniloff, D., Kloeckl, K. & Ratti, C. 2011. *Transportation mode identification and real-time CO2 emission estimation using smartphones. How CO2GO works*. Cambridge, Massachusetts, USA.

Melia, S. 2015. Do randomised control trials offer a solution to 'low quality' transport research? In

Proceedings of the 47th Annual UTSG Conference. London: Universities Transport Studies Group. [Online], Available: http://www.utsg.net/web/index.php?page=annual-conference [2016, April 03].

Meloni, I. & Sanjust, B. 2014. Using a GPS active logger to implement travel behavior change programs. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 325–340.

Meloni, I., Sanjust Di Teulada, B. & Spissu, E. 2016. Lessons learned from a personalized travel planning (PTP) research program to reduce car dependence. *Transportation*. 44(4):1–18.

Meyer, M.D. 1999. Demand management as an element of transportation policy: Using carrots and sticks to influence travel behavior. *Transportation Research Part A: Policy and Practice*. 33(7–8):575–599.

Mokhtarian, P.L., Salomon, I. & Handy, S.L. 2006. The impacts of ICT on leisure activities and travel: A conceptual exploration. *Transportation*. 33(3):263–289.

Möser, G. & Bamberg, S. 2008. The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence. *Journal of Environmental Psychology*. 28(1):10–26.

Naess, P. 2006. Accessibility, activity participation and location of activities: Exploring the links between residential location and travel behaviour. *Urban Studies*. 43(3):627–652.

Neutens, T., Schwanen, T. & Witlox, F. 2011. The prism of everyday life: Towards a new research agenda for time geography. *Transport Reviews*. 31(1):25–47.

Neutens, T., Delafontaine, M., Schwanen, T. & van de Weghe, N. 2012. The relationship between opening hours and accessibility of public service delivery. *Journal of Transport Geography*. 25:128–140.

Newsome, T.H., Walcott, W.A. & Smith, P.D. 1998. Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and space dimensions. *Transportation*. 25(4):357–377.

Nitsche, P., Widhalm, P., Breuss, S. & Maurer, P. 2012. A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys. In Vol. 48. *Procedia - Social and Behavioral Sciences*. Oxford: Elsevier. 1033–1046.

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N. & Maurer, P. 2014. Supporting large-scale travel surveys with smartphones - A practical approach. *Transportation Research Part C: Emerging Technologies*. 43:212–221.

Prelipcean, A.C. 2016. *Capturing travel entities to facilitate travel behaviour analysis: A case study on generating travel diaries from trajectories fused with accelerometer readings*. Published licentiate thesis. Stockholm, Sweden: Royal Institute of Technology (KTH).

Prochaska, J. 1991. Assessing how people change. *Cancer*. 67:805–807.

Prochaska, J.O., Velicer, W.F., Rossi, J.S., Goldstein, M.G., Marcus, B.H., Rakowski, W., Fiore, C., Harlow, L.L., et al. 1994. Stages of change and decisional balance for 12 problem behaviors. *Health Psychology*. 13(1):39–46.

Raento, M., Oulasvirta, A. & Eagle, N. 2009. Smartphones: An emerging tool for social scientists. *Sociological Methods & Research*. 37(3):426–454.

Ren, F., Tong, D. & Kwan, M.-P. 2014. Space–time measures of demand for service: Bridging location modelling and accessibility studies through a time-geographic framework. *Geografiska Annaler: Series B, Human Geography*. 96(4):329–344.

Richardson, A.J. 2003. Temporal variability of car use as an input to design of before and after surveys. *Transportation Research Record: Journal of the Transportation Research Board*. 1855:112–120.

Richardson, A.J., Seethaler, R.K. & Harbutt, P.L. 2004. Design issues for before and after surveys of travel behaviour change. *Transport Engineering in Australia*. 9(2):103–118.

Richter, J., Friman, M. & Gärling, T. 2011. Soft transport policy measures: Gaps in knowledge. *International Journal of Sustainable Transportation*. 5(4):199–215.

Rose, G. & Ampt, E. 2001. Travel blending: An Australian travel awareness initiative. *Transportation Research Part D: Transport and Environment*. 6(2):95–110.

Saleh, W. & Farrell, S. 2007. Investigation and analysis of evidence of asymmetric churn in travel demand models. *Transportation Research Part A: Policy and Practice*. 41(7):691–702.

Sanjust di Teulada, B., Meloni, I. & Spissu, E. 2017. The influence of activity-travel patterns on the success of VTBC. *International Journal of Urban Sciences*. In press.

Schönfelder, S. & Axhausen, K.W. 2010. *Urban rhythms and travel behaviour: Spatial and temporal phenomena of daily travel*. 1st ed. Farnham, United Kingdom: Ashgate Publishing, Ltd.

Schuessler, N. & Axhausen, K.W. 2009. Processing raw data from Global Positioning Systems without additional information. *Transportation Research Record: Journal of the Transportation Research Board*. 2105:28–36.

Schwanen, T. & De Jong, T. 2008. Exploring the juggling of responsibilities with space-time accessibility analysis. *Urban Geography*. 29(6):556–580.

Schwanen, T. & Kwan, M.-P. 2008. The internet, mobile phone and space-time constraints. *Geoforum*. 39(3):1362–1377.

Schwanen, T., Banister, D. & Anable, J. 2012. Rethinking habits and their role in behaviour change: The case of low-carbon mobility. *Journal of Transport Geography*. 24:522–532.

Seethaler, R. & Rose, G. 2009. Using odometer readings to assess VKT changes associated with a voluntary travel behaviour change program. *Transport Policy*. 16(6):325–334.

Shafique, M.A. & Hato, E. 2015. A review of travel data collection methods. In *Proceedings of the International Conference on Civil Engineering and Applied Mechanics (ICCEAM 2015)*. Paris: World Academy of Science, Engineering and Technology. 1906–1909.

Shen, L. & Stopher, P.R. 2013. A process for trip purpose imputation from Global Positioning System data. *Transportation Research Part C: Emerging Technologies*. 36:261–267.

Shen, L. & Stopher, P.R. 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*. 34(3):316–334.

Shoval, N., Kwan, M.-P., Reinau, K.H. & Harder, H. 2014. The shoemaker's son always goes barefoot: Implementations of GPS and other tracking technologies for geographic research. *Geoforum*. 51:1–5.

Sietchiping, R., Permezel, M.J. & Ngomsi, C. 2012. Transport and mobility in sub-Saharan African cities: An overview of practices, lessons and options for improvements. *Cities*. 29(3):183–189.

Skarin, F., Olsson, L.E., Roos, I. & Friman, M. 2017. The household as an instrumental and affective trigger in intervention programs for travel behavior change. *Travel Behaviour and Society*. 6:83–89.

Stopher, P.R. 1992. Use of an activity-based diary to collect household travel data. *Transportation*. 19(2):159–176.

Stopher, P.R. & Greaves, S.P. 2006. Guidelines for samplers: Measuring a change in behaviour from before and after surveys. *Transportation*. 34(1):1–16.

Stopher, P.R. & Greaves, S.P. 2007. Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice*. 41(5):367–381.

Stopher, P.R., Clifford, E. & Montes, M. 2008. Variability of travel over multiple days: Analysis of three panel waves. *Transportation Research Record: Journal of the Transportation Research Board*. 2054:56–63.

Stopher, P.R., Kockelman, K., Greaves, S.P. & Clifford, E. 2008. Reducing burden and sample sizes in multiday household travel surveys. *Transportation Research Record: Journal of the Transportation Research Board*. 2064:12–18.

Stopher, P.R., Clifford, E., Swann, N. & Zhang, Y. 2009. Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy*. 16(6):315–324.

Stradling, S. & Anable, J. 2008. Individual transport patterns. In 1st ed. R. Knowles, J. Shaw, & I. Dochterty (eds.). *Transport Geographies: Mobilities, flows and spaces*. Oxford: Blackwell Publishing Ltd. 179–195.

SUNSET Consortium. 2014. *Sustainable social network services for transport*. Enschede: SUNSET Consortium.

SUNSET Consortium. 2016. *Tripzoom*. [Online], Available: http://en.tripzoom.nl/ [2016, April 08].

Taniguchi, A. & Fujii, S. 2007. Promoting public transport using marketing techniques in mobility management and verifying their quantitative effects. *Transportation*. 34(1):37–49.

Taylor, M. 2007. Voluntary travel behavior change programs in Australia: The carrot rather than the stick in travel demand management. *International Journal of Sustainable Transportation*. 1(3):173–192.

Taylor, M. & Ampt, E. 2003. Travelling smarter down under: Policies for voluntary travel behaviour change in Australia. *Transport Policy*. 10(3):165–177.

Thøgersen, J. 2014. Social marketing in travel demand management. In T. Gärling, D. Ettema, & M. Friman (eds.). *Handbook of sustainable travel*. Dordrecht: Springer Netherlands. 113–129.

Thomopoulos, N., Grant-Muller, S. & Hodgson, F. 2014. *SUNSET Sustainable Social Network Services for Transport Deliverable D7.4 "Living Lab Report Leeds"*. Leeds: SUNSET Consortium.

Torgerson, D.J. & Torgerson, C.J. 2008. *Designing randomised trials in health, education and the social sciences*. New York: Palgrave Macmillan.

Urry, J. 2007. *Mobilities*. 1st ed. Malden, MA: Polity Press.

Vilhelmson, B. 1999. Daily mobility and the use of time for different activities. The case of Sweden. *GeoJournal*. 48(3):177–185.

Wall, R., Brög, W., Erl, E., Ryle, J. & Barta, F. 2011. A response to: Möser, G., & Bamberg, S. (2008). The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence. Journal of Environmental Psychology, 28, 10-26. *Journal of Environmental Psychology*. 31(3):266–269.

Van Wee, B. 2014. The unsustainability of car use. In T. Gärling, D. Ettema, & M. Friman (eds.). *Handbook of sustainable travel*. Dordrecht: Springer Netherlands. 69–83.

Van Wee, B. & Ettema, D. 2016. Travel behaviour and health: A conceptual model and research agenda. *Journal of Transport & Health*. 3(3):240–248.

Wolf, J. 2000. *Using GPS data loggers to replace travel diaries in the collection of travel data*. Published doctoral dissertation. Atlanta, Georgia: Georgia Institute of Technology.

Zhang, D., Schmöcker, J.-D., Fujii, S. & Yang, X. 2015. Social norms and public transport usage: Empirical study from Shanghai. *Transportation*. 43:869–888.

Zhang, Y., Stopher, P. & Halling, B. 2011. A study of the perceived effectiveness of TravelSmart tools for travel behaviour change. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice*. 20(2):23–29.

Zhang, Y., Stopher, P. & Halling, B. 2013. Evaluation of south-Australia's TravelSmart project: Changes in community's attitudes to travel. *Transport Policy*. 26:15–22.

Zhao, F., Ghorpade, A., Pereira, F.C., Zegras, C. & Ben-Akiva, M. 2015. Stop detection in smartphone-based travel surveys. *Transportation Research Procedia*. 11:218–226.

Zhou, J. (J) & Golledge, R. 2007. Real-time tracking of activity scheduling/schedule execution within a unified data collection framework. *Transportation Research Part A: Policy and Practice*. 41(5):444–463.

Part III
Methodological contributions

The use of GPS technologies to collect activity-travel data and reduce respondent burden ultimately depends on the accuracy of imputation algorithms.

Feng and Timmermans (2016: 192)

Chapter 3. Identifying activity-travel points from GPS-data with multiple moving windows

Van Dijk, J.T., 2017, Identifying activity-travel points from GPS-data with multiple moving windows. *Submitted for publication. Under review.*

Abstract

Where GPS technology can precisely register the spatiotemporal elements of activity-travel behaviour, travel characteristics need to be imputed from the data. As such, throughout the last decade, a plethora of methods has been developed for identifying trips, activities, and travel modes from raw GPS trajectories. However, rule-based methods that use dwell time for the classification of activity and trip episodes may disregard short activities. This chapter therefore describes a set of machine learning algorithms to identify activity and travel episodes on a point-by-point basis. To account for the underlying spatial structure of the data, attribute information is supplemented by moving multiple spatial windows over preceding and succeeding points to derive local densities. Because ground truth is essential to evaluate the accuracy of the algorithms, a set of 200 artificial GPS activity-travel sequences with varying noise levels was generated to be used for testing the algorithms. The results indicate that especially the random forest algorithm can lead to high accuracies. Moreover, local densities are important variables for the accuracy of the classifiers.

Keywords

Activity-travel detection; GPS; local densities; machine learning

3.1 Introduction

Because of the potential to obtain high-resolution data, several researchers have called for GPS data collection methods to be integrated into research on the effectiveness of voluntary travel behaviour change (VTBC) interventions (cf. Chatterjee, 2009; Chatterjee & Bonsall, 2009; Richter, Friman & Gärling, 2011; Chai, Chen, Liu, Tana & Ma, 2014; Meloni & Sanjust, 2014). While GPS technology can precisely register the spatiotemporal elements of activity-travel behaviour, travel characteristics – such as activity locations, number of trips, travel mode, vehicle kilometres travelled, and person kilometres travelled – need to be imputed from the GPS data. Accordingly, throughout the last decade, a plethora of methods has been developed to extract information from raw GPS trajectories. These methods range from deterministic (rule-based) methods to advanced machine learning algorithms (Shen & Stopher, 2014).

Because VTBC interventions are known to be less effective than more coercive transport demand management measures (Richardson, Seethaler & Harbutt, 2004; Stopher & Greaves, 2007; Stopher, Clifford, Swann & Zhang, 2009), accurate data on activity-travel behaviour are essential to evaluate whether an intervention has been successful. As such, when using GPS data, travel characteristics need to be imputed with a high level of accuracy. However, rule-based methods that use dwell time for classification activity and trip episodes may disregard short activities. This is problematic because trip points are often identified as the inverse of the GPS data points that not have been identified as activity points (or vice versa) (Zhou, Yu & Sullivan, 2016). As an alternative, it has been suggested that machine learning algorithms are potentially more effective in handling GPS data imputation (Feng & Timmermans, 2016).

Despite the importance of imputation algorithms for activity and trip classification, little is known about their relative performance. One explanation for this is that “spatial settings have a direct bearing on the discriminatory power of GPS information” (Feng & Timmermans, 2016: 180), which makes comparisons of different algorithms across spatial settings difficult. In addition, GPS data imputation methods require ground truth validation to assess their performance (Bohte & Maat, 2009). Another issue with GPS data imputation methods is that they often use attribute values to classify each data point individually. However, GPS points are part of a space-time path that is strictly ordered by time (Miller, 2005). This ordering imposes a structure on the data, such that point *n* is spatially related to its preceding and its succeeding points. The usage of attribute values (e.g. speed, acceleration, position accuracy) ignores this underlying spatial structure.

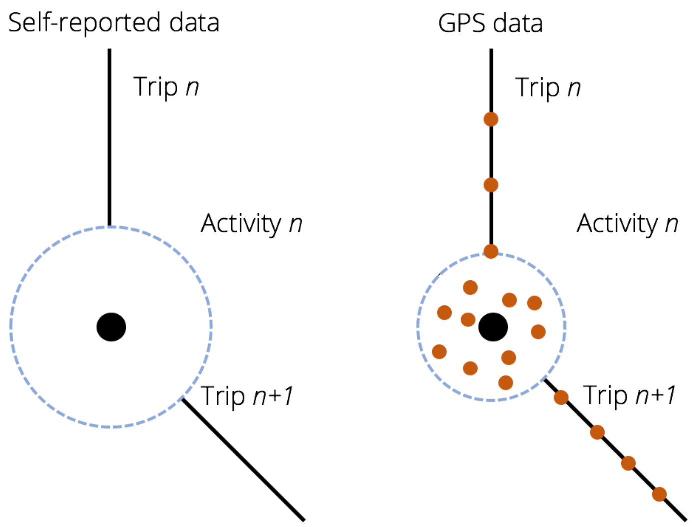
This chapter describes and systematically compares the relative performance of four machine learning algorithms to classify GPS points into activity points and travel points. (Note that activity type recognition and travel mode imputation are out of the scope of this chapter.) Three rule-based methods are used as a baseline comparison. Instead of collecting GPS data and requesting respondents to annotate their data to obtain ground truth, we use artificially created GPS tracks with varying noise levels to assess the relative performance of the methods. To account for the underlying spatial structure of the data, attribute information is supplemented with information derived from moving multiple spatial windows over preceding and succeeding points.

3.2 Identifying activity-travel points from GPS data

Travel data collection methods can roughly be classified into two, not mutually exclusive, methods. The first method uses self-reported data, such as data collected through telephone-assisted interviews, computer-assisted-interviews, and pen-and-paper interviews. The second method relies on passively collected data, such as data collected through call-detail records (CDR) and GPS data (Jianchuan, Zhicai, Guangnian & Xuemei, 2014). Over the past two decades, there has been a dramatic increase in the use of GPS for collecting activity-travel data. GPS surveys have been undertaken in, for instance, the Netherlands, Austria, Australia, Israel, Switzerland, South Africa, the United States, and the United Kingdom (cf. Bohte & Maat, 2009; Krygsman & Nel, 2009; Schuessler & Axhausen, 2009; Nitsche, Widhalm, Breuss & Maurer, 2012; Feng & Timmermans, 2013, 2016; Shen & Stopher, 2013; Shoval, Kwan, Reinau & Harder, 2014). The main reason for the incorporation of GPS devices in travel surveys is their ability to precisely register location and time. Whereas self-reported surveys suffer from underreported trips, omitted activities, overestimations of trip lengths, rounding of temporal information, and many other issues, GPS offers a superior alternative for gathering data on the spatiotemporal dimensions of travel behaviour (Stopher & Greaves, 2007; Feng & Timmermans, 2016).

Although GPS may have a lot of potential in addressing the inaccuracies found in self-reported data, GPS itself also comes with different inaccuracies. Figure 3.1 illustrates these differences. On the left-hand side of the figure, a hypothetical self-reported trip is visualised. The blue-dotted circle indicates a spatiotemporal confidence interval; we can assume some uncertainty regarding the accuracy of the precise activity location, as well as the exact end time of trip *n* and start time of trip *n + 1*. On the right-hand side of the figure, a hypothetical GPS trajectory is visualised. Again, the blue-dotted circle represents the spatiotemporal confidence interval; although the confidence interval is potentially smaller, there is still uncertainty concerning the exact demarcation of the activity location and activity duration. This suggests that the 'trick' to any GPS imputation method is to reduce the size of this confidence interval, both in space and in time without identifying false positives (for example, waiting for a traffic light, being stuck in traffic, etc.) and at the same time without disregarding short stops. Ideally, every point in the data set should be evaluated against a set of criteria and subsequently categorised into a *stay* or a *move*.

Figure 3.1 | Spatiotemporal confidence interval self-reported data versus GPS data



The segmentation of GPS data into activity and travel episodes is often the first step in a more elaborate process of identifying activity types and transport modes (Feng & Timmermans, 2014), although some scholars have tried to identify activity types and transport modes simultaneously (for instance Feng & Timmermans, 2013, 2016). The methods that have been proposed in the literature vary from deterministic (rule-based) algorithms to more advanced learning algorithms. Yet, trip and activity identification is most frequently done with deterministic rule-based algorithms (Shen & Stopher, 2014). The rules often use the fact that activities episodes, as opposed to travel episodes, are characterised by very low speeds and clusters of GPS points. Wolf (2000), for example, uses a method that registers an activity if GPS speeds are low (< 5 km/h) for at least 120 seconds. Although the 120 second time threshold is often used, it is prone to missing short activities, and lacks an empirical base (Shen & Stopher, 2014). Auld, Williams, Mohammadian and Nelson (2009), on the other hand, use both a time threshold and a distance threshold to identify clusters of points that could indicate an activity. A similar approach is taken by Schuessler and Axhausen (2009), who try to identify activity locations from raw GPS trajectories in a two-step approach. In the first step, an activity is flagged if the speed is below a threshold value (< 3.6 km/h) for at least 120 seconds. In the second step, a local point density is established. Others combine a larger number of rules to remove noise and identify trips simultaneously (Bohte & Maat, 2009).

Rather than using deterministic rules, a second category of TI uses clustering algorithms to find clusters of points that may be indicative of an activity. Examples are *k*-means clustering (Ashbrook & Starner, 2003), grid-based density clustering (Sterkenburg, Pierik & De Vries, 2012), and kernel-based methods (Schoier & Borruso, 2011; Thierry, Chaix & Kestens, 2013; Wan & Lin, 2013). The disadvantage of these methods is that they require a high measurement frequency to be effective. Also, short activities are likely to be missed. The kernel-based algorithm of Thierry *et al.* (2013), for example, is tested with a minimum activity duration of 5 minutes.

A third category of TI moves to more flexible and advanced statistical methods (Feng & Timmermans, 2016). Examples of these include the fuzzy classification method proposed by Wan and Lin (2016) and the use of relational Markov networks by Laio, Fox and Kautz (2005). Nowadays, more advanced methods, particularly methods belonging to the domain of machine learning, are gaining popularity. The main reason for its popularity is that "instead of making strict assumptions about the data, machine learning models learn to represent complex relationships in a data-driven manner" (Hagenauer & Helbich, 2017: 273).

3.3 Time-based local densities with multiple moving windows

Classification algorithms typically aim to minimise in-group variation and maximise between-group variation to discriminate between episodes (i.e. moves and stays) using attribute values such as speed, acceleration, the number of satellites insight, the distance between consecutive points, and the distance to the nearest road centre line. However, GPS points form part of a space-time path that is strictly ordered by time (Miller, 2005). This ordering imposes a structure on the data, such that point *n* is spatially related to its preceding and its succeeding points. This suggests that the presence of nearby points in Euclidian space may be indicative of an activity, while the absence of nearby points may be indicative of travel. Because classification algorithms typically use attribute values to classify each input data point individually, some structure should be imposed onto the input data to preserve the relationship between the individual data point, the preceding data point(s), and the succeeding data point(s).

One way of preserving a data structure within an individual data point is by deploying a moving window. In a study on travel mode recognition with smartphone sensor data, Shafique and Hato (2016) use a moving average to smoothen the acceleration of a data point by averaging the acceleration of each point over a number of its preceding and succeeding data points. Feng and Timmermans (2016) also use the moving window concept on a number of attribute variables for their imputation algorithms. Schuessler and Axhausen (2009), on the other hand, deploy a spatial moving window that counts the proportion of the 30 preceding and 30 succeeding points that fall within a 15-metre buffer to acquire a local point density. Against this background, we propose to combine these two moving window concepts by moving a window over the attribute values, and by moving multiple spatial windows to identify local point densities.

Figure 3.2 | Identifying local densities with a moving spatial window (buffer)

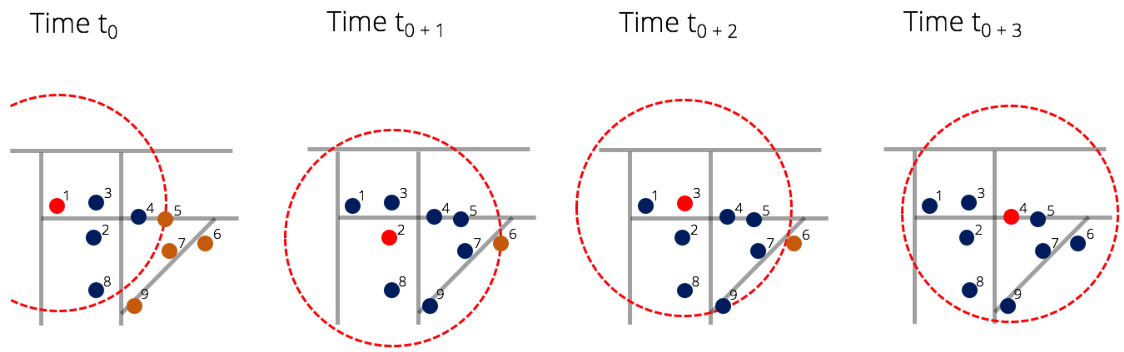


Figure 3.3 | Identifying local densities with a moving spatiotemporal window (local point density)

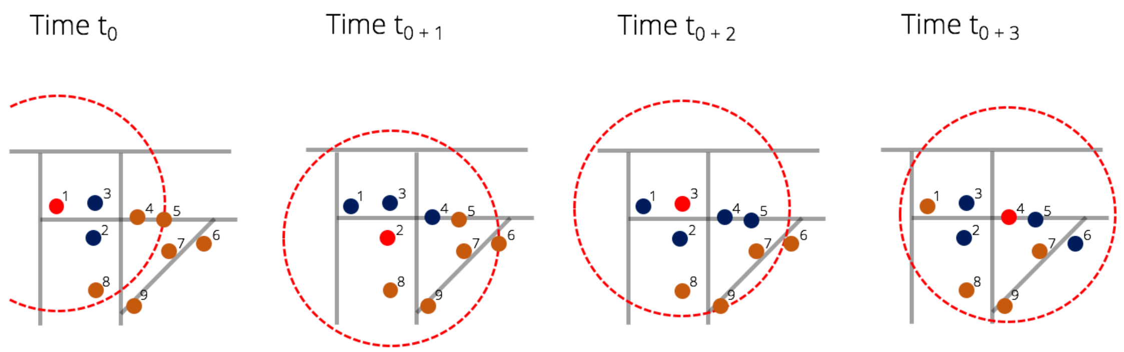


Figure 3.2 illustrates how the identification of local densities with a moving spatial window works. The points in the figure represent hypothetical GPS points; point x_1 being the first point and x_9 the last point in the sequence. The spatial window is represented by the red dotted circle. At time t_0, the window starts at point x_1 (red) and registers four points (blue) (x_2, x_3, x_4, x_8) within the spatial window. In the subsequent times (time t_{0+1} until time t_{0+3}), the spatial window (buffer) moves with the XY location of the point under consideration as its centroid. Figure 3.3 extends the spatial window with a temporal component. Instead of counting all points that fall within the buffer at the position of the window, it only counts the two preceding and two succeeding points within the buffer to establish a local point density. In other words, it only counts the number of points that fall within the time window, even if the points fall within the spatial search of the buffer. This is important, because individuals may visit the same activity location twice. If one were to deploy only a spatial moving window, it could lead to incorrect results.

3.4 Simulating GPS activity-travel tracks

A major issue with GPS data imputation is the necessity of a ground truth to test whether an imputation algorithm correctly categorises GPS points into activity (stay) points and trips (move) points (Bohte & Maat, 2009). Feng and Timmermans (2016), for instance, addressed this issue by corroborating trip details with face-to-face interviews with their participants (*n* = 8). As they stated in an earlier work: "Since the accuracy and quality of imputed activity-travel data cannot be ensured without the validation of respondents, a prompted recall instrument is necessary for any GPS-based travel survey. (…) It should be realised, however, that prompted recall data are not error free" (Feng & Timmermans, 2014: 116). However, this is not only very time-consuming, but also constrains the number of data points to be used for analysis. Moreover, to deploy advanced machine learning algorithms, a large number of labelled data is required for training purposes (Prelipcean, Gidófalvi & Susilo, 2017). As such, Thierry, Chaix and Kestens (2013) propose a different approach: instead of collecting actual GPS data, artificial GPS tracks were created. Using their method and parameters as a basis, 100 artificial GPS tracks with a 30-second measurement frequency and 100 artificial GPS tracks with a 60-second measurement frequency were created. (See Appendix 3.A for the steps involved.)

With a measurement frequency of 30 seconds, the procedure resulted in roughly 235,000 labelled data points for the 100 simulated activity-travel sequences. With a measurement frequency of 60 seconds, the procedure yielded almost 110,000 data points. Furthermore, to simulate data sets with high levels of noise, three extra data sets for both measurement frequencies were generated by sampling the data points of each GPS track at 75 percent, 50 percent, and 25 percent. The whole process resulted in eight data sets with GPS points. After generating the tracks, the local point densities for every point were calculated with a buffer of 50 metres, 100 metres, 150, and 250 metres. In addition, the moving window concept was applied to calculate the time-based local densities with a temporal constraint of ten minutes in both directions. In a final step, the points in each data set were randomly assigned to a training data set and a test data set with a 70/30 split. The training data set is used to train the classifiers to recognise patterns. In turn, the test data set is used to assess how well the algorithms perform on data points that were not part of the learning process, i.e. how well the algorithms predict new examples.

3.5 Activity-travel recognition with learning algorithms

To impute the activity and travel trajectories of the artificial activity-travel GPS tracks on a point-by-point basis, four supervised machine learning algorithms were selected: naive Bayes (NB), boosted C5.0, support vector machine (SVM), and random forest (RF). Naive Bayes is a well-known and relatively simple generative machine learning algorithm. Using Bayes' theorem, NB tries to estimate the probability of a data point belonging to a certain class. The model assumes that the probability of the output class Y is equal to value y_k, given the probability of X (Feng & Timmermans, 2016: 184–185):

(1) $$p(Y = y_k \mid X)$$

NB further assumes that all attributes used for the classifier are conditionally independent of one another. For independent variables X_1 through X_n, the classifier is trained with the following probability model:

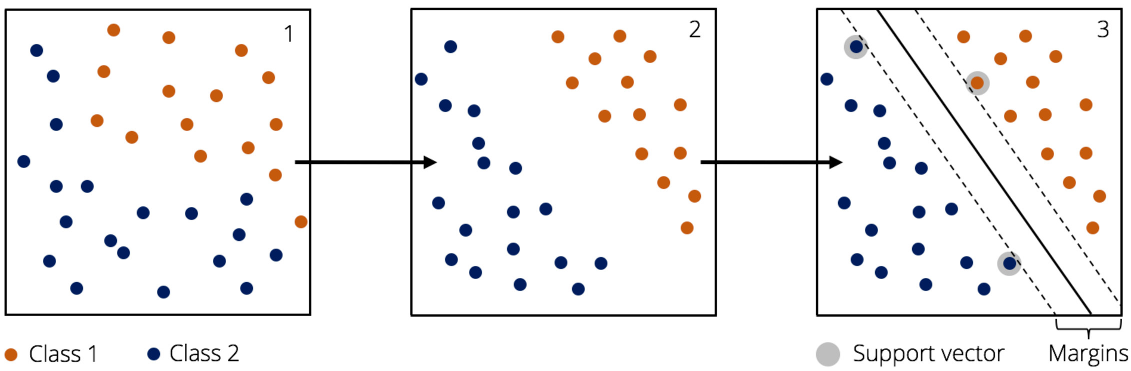
(2) $$p(X_1, \dots, X_n \mid Y) = p(X_1 \mid Y) p(X_2 \mid Y) \dots p(X_n \mid Y)$$

Using Bayes' theorem, the probability of Y taking on value y_k given X for the data points in the test data set then becomes:

$$p(Y \mid X_1, \dots, X_n) = \frac{p(Y = y_k)p(X \mid Y = y_k)}{p(X_1, \dots, X_n)} \tag{3}$$

In contrast to NB, support vector machine (SVM) is a discriminative classifier that uses a linear model for a (non-linear) binary classification problem. An SVM transforms the input data using a kernel function to make them linearly separable. The algorithm then finds the best hyperplane to separate the data by searching for the representation of the largest separation between the two classes, i.e. the margin. The margin is thus defined by the points that lie closest to it, i.e. the support vectors (Witten & Frank, 2005). The maximum-margin hyperplane is the line that lies halfway in between the margins. Figure 3.4 illustrates this process. For the training data, with the data points in this case being labelled either as a stay or as a move, the SVM model "is a representation of the examples [i.e. the training data] as points in space, mapped so that the examples of the separate categories are divided by a clear gap that is as wide as possible" (Feng & Timmermans, 2016: 187). For the data points in the test data set, the SVM projects the new points into the new space, and predicts to which category they belong using the maximum-margin hyperplane as decision boundary.

Figure 3.4 | Transformation of data and creation of maximum-margin hyperplane



Just like the SVM, the C5.0 algorithm also tries to draw a decision boundary. However, it does this by growing a decision tree. The nodes of the tree represent the variables and values on which the algorithm splits the data. To decide on which attribute to split the data, entropy is used to calculate information gain. If, for instance, one of the attributes perfectly splits the data into the available classes (in this case move or stay), the entropy is zero. When, for instance, a split of the attribute would result in two groups of equal sizes, the entropy is one. Given a set of training data, $T = (t_1, t_2, \dots, t_n)$, with for each training example a set of attributes, $(x_{1,i}, x_{2,i}, \dots, x_{n,i})$, with x_j as the attributes and t_i as the class of the example, entropy is derived as follows (Feng & Timmermans, 2016: 187–188):

$$entropy(j \mid \bar{t}) = \frac{|t_j|}{|\bar{t}|} \log \frac{|t_j|}{|\bar{t}|} \tag{4}$$

$$(5) \qquad entropy(\bar{t}) = -\sum_{j=1}^{n} \frac{|t_j|}{|\bar{t}|} \log \frac{|t_j|}{|\bar{t}|}$$

In turn, the information gain is based on the decrease in entropy after splitting on attribute x_j. So, the entropy before the split on attribute x_j is compared to the entropy after the split on attribute x_j.

$$(6) \qquad Gain(\bar{t}, j) = entropy(\bar{t} - entropy(j|\bar{t}))$$

The algorithm calculates the information gain for all attributes, and splits on the attribute with the highest information gain (i.e. the largest decrease in entropy). This process continues recursively until no further splits are necessary. Finally, the leaves of the tree represent the class labels. We grew a full tree and used five boosting iterations to create an ensemble of trees. In every boosting iteration, the resulting decision tree is tested against the training data. After this, the algorithm tries to improve the misclassified class labels from the previous iteration by growing a new tree. This process continues until the maximum number of boosting iterations has been reached. The final model uses a majority vote amongst the trees to classify the test data (Feng & Timmermans, 2016; Hagenauer & Helbich, 2017).

The last algorithm, random forest (RF), is also a tree-based ensemble method. However, instead of growing one tree with several trees to boost the initial classifier, RF grows an ensemble of trees (i.e. a forest) simultaneously, using bootstrap sampling. The result is that every tree is grown on a different subset of the training data. The randomness aspect is introduced at the nodes of the trees, where the features that are used to calculate the next split in the tree are randomly selected from the feature space (i.e. the input variables). The result is a forest with trees that have been grown using varying features on different subsets of the training data. For prediction, every new data point is fed into every tree of the forest. A majority vote amongst the trees decides on the final classification (Breiman, 2001; Hagenauer & Helbich, 2017). We grew a forest with 500 trees that used three random variables at each split.

To systematically compare the relative performance of the four selected machine learning algorithms, three deterministic rule-based methods were also selected to classify the GPS tracks into activity points and travel points. The first method (RULE1) was based on the study by Bothe and Maat (2009), who use a distance and time threshold to separate trips from activities. In their method, they flag an activity if someone remained in the same location for at least 180 seconds. The second method (RULE2) was based on a study by Schuessler and Axhausen (2009), who use a two-step approach. In the first step, a point is classified as an activity if the speed is lower than 3.6 km/h for at least 120 seconds. In the second step, point densities are considered. The point density is calculated by counting the number out of the 30 preceding and succeeding GPS points that fall within a pre-defined buffer. If at least 25 percent of these points fall within this buffer, and this lasts for at least 300 seconds, a point is also classified as an activity. Because of the high levels of noise introduced in the GPS tracks, we used a radius of 400 metres to assess if points are in the same location (RULE1) or fall within the buffer (RULE2). This value corresponds with twice the maximum standard deviation of the introduced noise levels for activity points. The third method (RULE3), is a variation on the second method. Instead of using the speed criteria, the first step now classifies a point as an activity if the distance to the nearest road segment is larger than 10 metres for at least 120 seconds.

3.6 Comparison of activity-travel recognition algorithms

All machine learning algorithms were executed in R (R Core Team, 2016) and the following R packages were used to run the classification algorithms: *e1071* (Meyer, Dimitriadou, Hornik, Weingessel & Leisch, 2017), *C50* (Kuhn, Weston, Coulter & Culp, 2015), and *randomForest* (Liaw & Wiener, 2002). Besides the moving windows, a number of other variables were used to train and test the algorithms. Table 3.1 gives an overview of all variables used.

Table 3.1 | Attribute variables for activity-travel imputation

Variable	Description
SPEED	speed
AVGSPEED	average speed within time window
MAXSPEED	maximum speed within time window
ACC	acceleration
AVGACC	average acceleration within time window
MAXACC	maximum acceleration within time window
DIRCHANGE	direction change as compared to preceding point
NEARDIST	distance of point to the nearest road centre line
BUF50	number of points within a 50-metre buffer
BUF100	number of points within a 100-metre buffer
BUF150	number of points within a 150-metre buffer
BUF250	number of points within a 250-metre buffer
POINTS50	number of points within a 50-metre buffer within time window
POINTS100	number of points within a 100-metre buffer within time window
POINTS150	number of points within a 150-metre buffer within time window
POINTS250	number of points within a 250-metre buffer within time window

All machine learning algorithms were first trained on the training data set, and subsequently tested on the test data set. For the evaluation of the relative performance of the different algorithms, two measures were selected to assess the level of agreement between the predicted values of the data points in the test set and the actual values in the test set: overall accuracy and the kappa value. The overall accuracy is the overall percentage of data points that the algorithms correctly classified as being either a stay or a move. The kappa value gives a more conservative estimate of the accuracy. With p_o as the observed agreement and p_e as the expected agreement based on prior class probabilities, kappa is defined as:

(7) $$kappa\ k = \frac{p_o - p_e}{1 - p_e}$$

A kappa value of 1 indicates perfect agreement between the predicted class labels and the actual class labels, whereas a kappa value of 0 indicates complete disagreement.

3.6.1 Overall accuracy algorithms

Table 3.2 presents the overall accuracy and the kappa values for the test data sets for all algorithms, including the rule-based methods, on the data sets that were generated with both the 30-second and 60-second measurement frequencies. The first column shows the overall accuracy of the model for the 30-second measurement frequency data. For all machine learning classifiers, the accuracy of the full model was above 95 percent. However, the rule-based algorithms did not perform that well, with only RULE3 giving a score of 76 percent. The second column gives the kappa value. The results show that among the machine learning classifiers, NB had the lowest kappa value, with 91.1 percent. The boosted C5.0, SVM, and RF yielded similar kappa values (98 percent). The kappa statistic for the rule-based algorithms, on the other hand, was significantly weaker. RULE1 in particular gave very poor results, and RULE3 classified just over half of the data points correctly.

Table 3.2 | Accuracy measures machine learning models for identifying activity-travel points

	30 seconds		60 seconds	
	Accuracy	Kappa	Accuracy	Kappa
NB	0.956	0.911	0.912	0.825
SVM	0.994	0.988	0.989	0.977
C5.0	0.993	0.986	0.988	0.976
RF	0.994	0.987	0.990	0.979
RULE1	0.547	0.064	0.589	0.146
RULE2	0.685	0.402	0.643	0.326
RULE3	0.765	0.546	0.800	0.610

The third and fourth column of Table 3.2 report on the accuracy and the kappa of the different models, but this time the tracks with a measurement frequency of 60 seconds are used. Lowering the measurement frequency seems to have affected the NB classifier, whose kappa statistic dropped from 91.1 percent to 82.5 percent. The other machine learning classifiers did not seem to be affected as much by the lower measurement frequency, and they only showed a minor reduction in accuracy and kappa values. A possible explanation for this is that the NB classifier is a relative simple classifier as compared to the others. Whereas NB classifies the data once, C5.0 and RF, for example, create multiple instances of its classifier, by growing multiple trees, for the final classification. The rule-based methods were performing slightly better. Both RULE1 and RULE2 had higher kappa values for the data with a 60-second measurement frequency than for the data with a 30-second measurement frequency. For both measurement frequencies, the rules that incorporated a moving window (RULE2 and RULE3) had better results.

The differences in overall accuracy and the kappa values, between the rule-based methods and the machine learning algorithms, are not entirely surprising. First, the rule-based methods only use a number of the variables to classify the individual points, whereas the machine-learning algorithms take advantage of the entire set of variables. For instance, a change of direction (DIRCHANGE), could be indicative of a stay point because it is unlikely that when someone is moving, s/he would suddenly move in the opposite direction. However, if someone is stationary, sudden direction changes can occur as a result of drift or signal reflection. Second, rule-based methods are more rigid (Hagenauer & Helbich, 2017). Machine-learning algorithms try to understand complex relationships in a data-driven manner, and, as such, they are capable of designing more complicated rules using a large set of input features.

To more realistically represent the noise levels and measurement gaps that exist in real GPS data, all the models were run on the sampled data sets as well. After this, the attribute variables were, where necessary, re-calculated and the data were again separated into a training data set and a test data set with a 70/30 split before applying the different methods. The kappa values of these analyses are shown in Table 3.3. For both measurement frequencies, there seemed to be a negative relationship between the data quality and the kappa values for the NB classifier, as it only classified 83.7 percent of the points correctly on a 25 percent random sample of the 60-second measurement frequency tracks. For the other classifiers, however, the results again do not deviate that much. RF, for example, still gave a kappa value of 96.3 for a 25 percent random sample of the 60-second measurement frequency tracks.

Table 3.3 | Kappa values machine learning models for identifying activity-travel points

	30 seconds			60 seconds		
	75%	50%	25%	75%	50%	25%
NB	0.936	0.940	0.900	0.836	0.839	0.837
SVM	0.984	0.979	0.965	0.971	0.962	0.952
C5.0	0.981	0.974	0.962	0.966	0.958	0.949
RF	0.983	0.979	0.968	0.971	0.965	0.963
RULE1	0.095	0.121	0.198	0.169	0.209	0.361
RULE2	0.411	0.441	0.530	0.357	0.400	0.521
RULE3	0.581	0.623	0.691	0.627	0.616	0.645

Unlike with most of the machine learning classifiers, the measurement frequency did seem to affect the rule-based methods. Interestingly, for all rule-based methods, the kappa statistic increased when the track quality decreased. However, this was most likely caused by the fact that with reduced data quality, the clusters of data points during an activity become smaller. It seems that because of the high noise levels that were introduced for the activity locations, the rules tended to overclassify activity points as a move. This is illustrated in Table 3.4 and Table 3.5, in which the confusion matrices for all rule-based algorithms, for all data sets, are presented. A confusion matrix shows the correctly classified instances versus the incorrectly classified instances. The predicted values are shown over the rows and the actual values are shown over the columns. As such, the light-grey shaded areas represent the number of cases in which the rule correctly classified the point as being either a stay or a move. So, for example, RULE1 classified for the full data set with a 30-second measurement frequency, 11,977 move points correctly as a move and 26,388 stay points correctly as a stay (Table 3.4). However, it also classified 19,667 points as a stay, whilst the points were in fact a move. Similarly, it classified 12,166 points as a move, whereas these points should have been classified as a stay – leading to a sensitivity of 37.9 percent and a specificity of 68.4 percent.[1]

Table 3.4 | Confusion matrices rule-based methods 30-second measurement frequencies

	100%	Move	Stay	**75%**	Move	Stay
RULE1	Move	11,977	12,166	Move	9,043	8,320
	Stay	19,667	26,388	Stay	14,842	20,731
RULE2	Move	31,621	22,080	Move	23,768	16,254
	Stay	23	16,474	Stay	117	12,797
RULE3	Move	31,538	16,405	Move	23,558	11,098
	Stay	106	22,149	Stay	327	17,953

	50%	Move	Stay	**25%**	Move	Stay
RULE1	Move	5,751	4,811	Move	3,030	1,849
	Stay	10,041	14,686	Stay	4,916	7,853
RULE2	Move	15,601	10,154	Move	7,779	4,124
	Stay	181	9,343	Stay	167	5,578
RULE3	Move	15,264	6,166	Move	7,442	2,246
	Stay	518	13,331	Stay	504	7,456

[1] See Appendix 3.B for the confusion matrices of the machine learning algorithms. See Appendix 3.C for the sensitivity, specificity, the positive predictive value, and the negative predictive value for both the rule-based methods and the machine learning algorithms.

Table 3.5 | Confusion matrices rule-based methods 60-second measurement frequencies

	100%	Move	Stay	**75%**	Move	Stay
RULE1	Move	5,888	4,789	Move	4,365	3,217
	Stay	8,558	13,270	Stay	6,578	10,474
RULE2	Move	14,006	11,180	Move	10,441	7,846
	Stay	440	6,879	Stay	502	5,845
RULE3	Move	13,809	5,852	Move	10,140	3,856
	Stay	637	12,207	Stay	803	9,835

	50%	Move	Stay	**25%**	Move	Stay
RULE1	Move	2,893	1,838	Move	1,956	808
	Stay	4,270	7,226	Stay	1,735	3,736
RULE2	Move	6,767	4,713	Move	3,461	1,801
	Stay	396	4,351	Stay	230	2,743
RULE3	Move	6,420	2,405	Move	3,218	993
	Stay	743	6,659	Stay	473	3,551

3.6.2 Variable importance machine learning algorithms

Because we are also interested in the effectiveness of incorporating multiple moving windows into the machine learning algorithms, the importance of these variables for the accuracy of the models needed to be established. However, as Hagenauer and Helbich (2017: 276), argue: "[assessing variable importance] remains a complex task due to interactions and correlations among the variables. Seemingly irrelevant variables may become important only in the context of others, while redundancies between variables may lead to an overestimation of importance". One way of assessing the importance of each variable is by randomly permutating the data in the test data set for each variable and comparing the outcomes to the overall accuracy of the model (Breiman, 2001; Hagenauer & Helbich, 2017). For each model, we firstly estimated the overall kappa values of the algorithm. The model accuracy was subsequently compared to a new instance of the predictor on the test data, where the values of one of the variables were randomly reshuffled. This was repeated for all variables.

Figure 3.5 shows the kappa values of the four machine learning algorithms with the permutated variables on the Y-axis for the tracks with the 30-second measurement frequency. In case of NB, the figure reveals that both the spatial window (with different sizes) and the spatiotemporal window (with different sizes) are important to the accuracy of the model. Permutating the data in BUF250, for instance, led to a decrease in kappa from the initial 91.1 percent to 62.0 percent (for the full data set). This suggests that the spatial windows are by far the most important variables to classify the points. In fact, this observation holds not only for the full data set, but also for the various sampled data sets. In addition, the importance of BUF250, when compared to the other buffer sizes, suggests that the size of the buffer matters because it captures more information. Besides the spatial windows, the distance to the nearest road centre line also plays an important role in the accuracy of the classifier. These results indicate that, as expected, the probabilities of a point being part of a stay increase with a larger number of points in its vicinity. Similarly, the importance of the distance to the nearest road centre line can be explained by the fact that most moves, especially when travelling by bike or by private car, are restricted by the network.

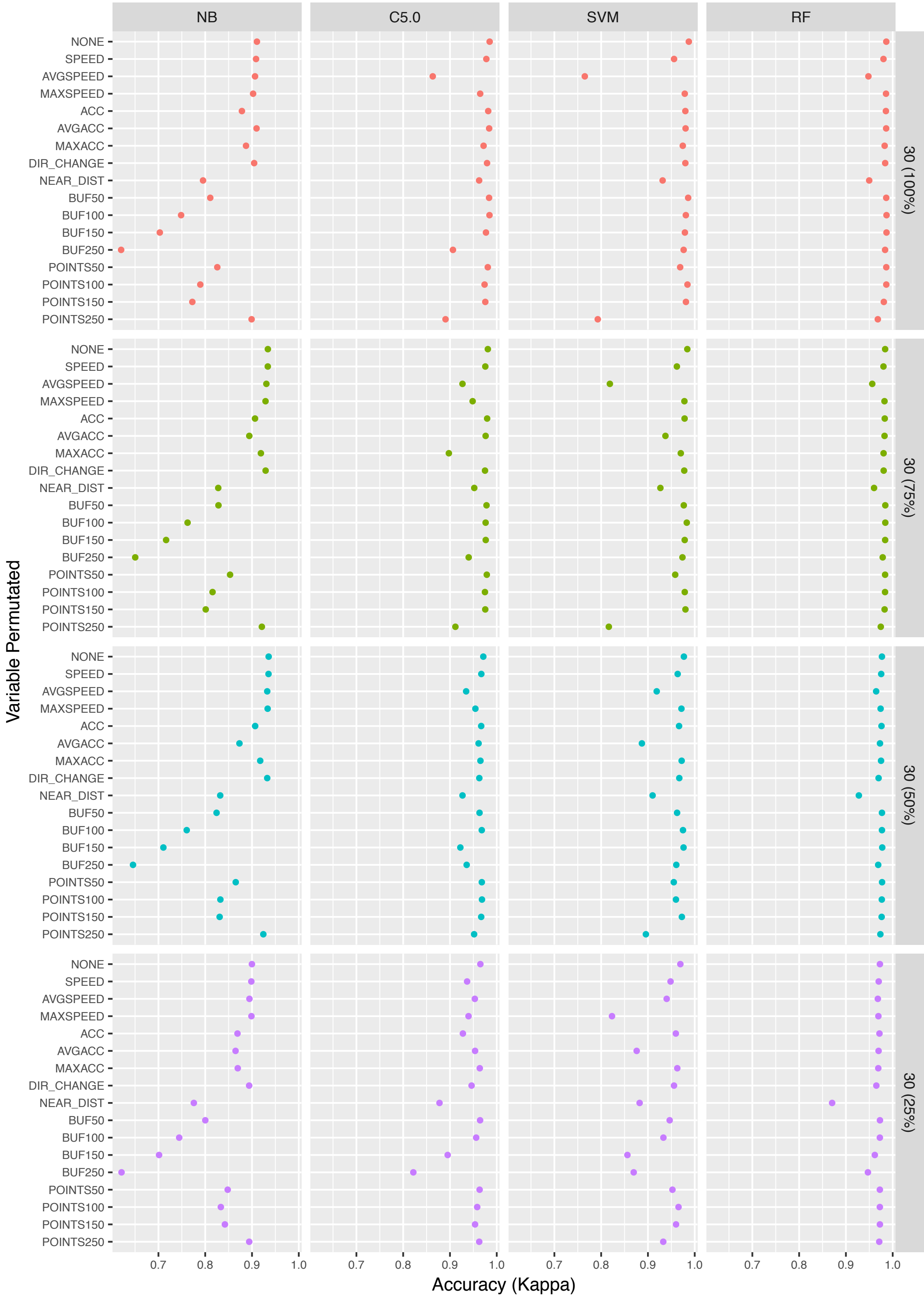
The variable importance of C5.0 looks quite different than the variable importance of NB, as only BUF250 and POINTS250 led to a large reduction in accuracy. Although the spatial window and the spatiotemporal window still lead to a reduction in the accuracy of the classifier, their impact on the entire classification is less pronounced than in case of NB. This may be explained by the process underlying the C5.0 algorithm. Because the model resulting from the first iteration of the algorithm is first tested against the training data and subsequent iteration try to improve the misclassified class labels, it is likely that C5.0 is less dependent on a single variable. In addition, the average speed came up as important classification variable. The importance of average speed can be explained by the fact that the artificial GPS activity-travel sequences were generated with a variety of travel modes (e.g. walking, cycling, and driving a car at various speeds) and, thus, it is likely to have a high entropy-based information gain.

The SVM classifier shows a different pattern than NB and C5.0, with two variables that seem crucial: the average speed and the 250-metre buffer within a time window. Interestingly, the importance of these two variables reduced significantly with increasingly noisy data. Instead, the maximum speed, as well as the 250-metre buffer and 150-metre spatial buffer, gained importance. The underlying algorithm of SVM could explain some of these differences, because the reduction of the data quality may lead the algorithm to more efficiently separate the data and maximise the maximum-hyperplane. In case of RF, on the other hand, the most important variable to surface was the distance to the road network. Yet, it does not affect the overall accuracy by much. The RF classifier, thus, appears to be the most robust. This may not be surprising because RF grows its decision trees at every split on three randomly picked variables, and is therefore less likely to depend on one particular variable.

Figure 3.6 shows the kappa values for the four machine learning algorithms with the permutated variables of the 60-second interval track data. Although the patterns show a similar trend as that of the 30-second interval data, there are some differences between Figure 3.5 and Figure 3.6 for some of the classifiers. For NB, for instance, the most important variables were largely the same. As such, although the overall accuracy is decreasing, the probabilities of the individual features do not seem to be affected. For the boosted C5.0, on the other hand, the distance to the road network now surfaced as important. Moreover, for the full data set, the permutation of BUF250 changed the kappa value from 97.6 percent to 83.3 percent. SVM also shows an interesting trend, in which the importance of BUF250 and POINTS250 seems to switch with increasing noise in the data. Whereas POINTS250 turned out to be the most important variable on the full data set, BUF250 took over this role in the 25 percent sample data set. Lastly, the distance to the road network again played an important role in constructing the RF classifier. In addition, when compared to the 30-second measurement frequency data set, the 250-metre buffer now also played a role.

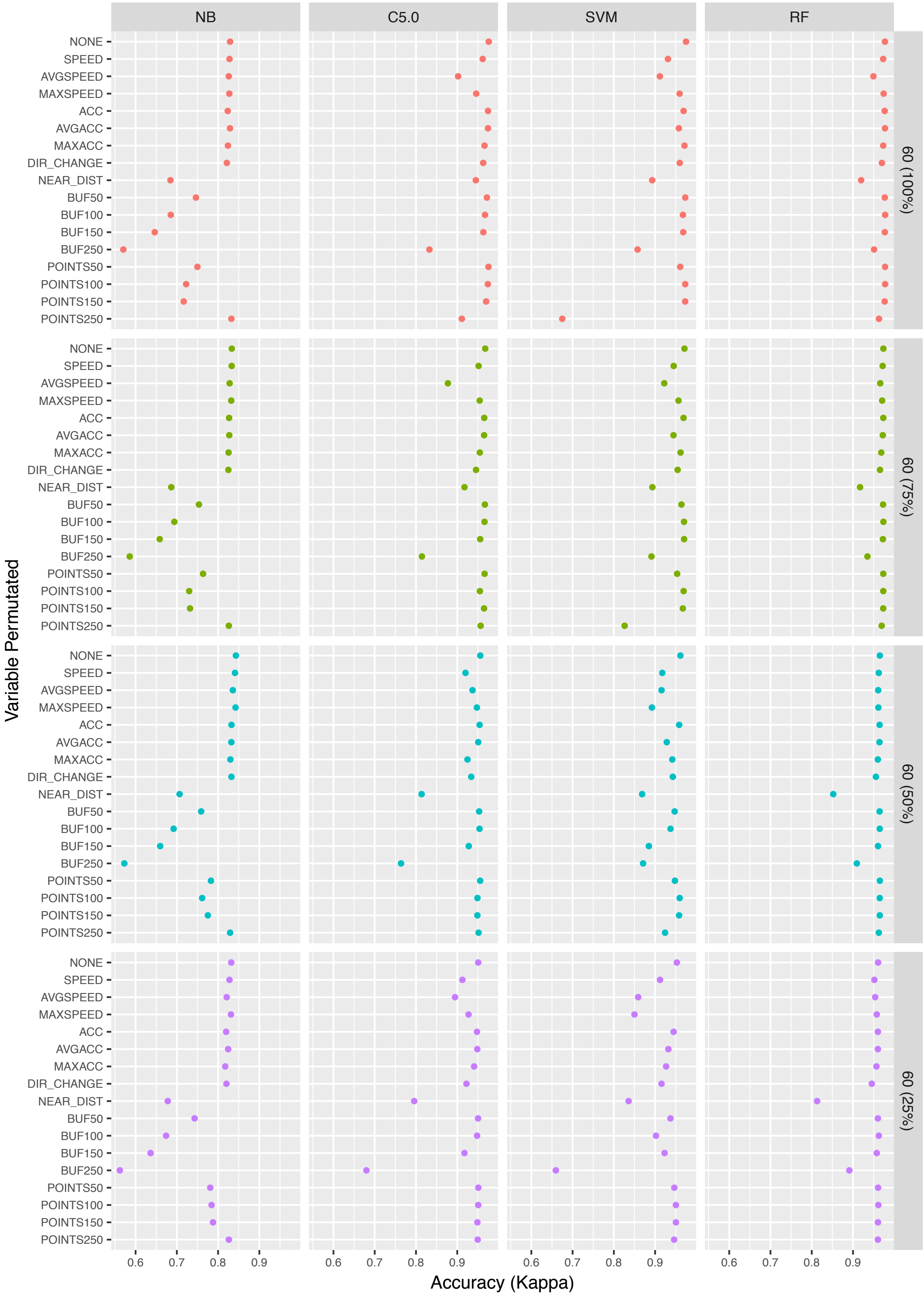
The differences in overall accuracy between the machine learning classifiers, as well as the differences between the importance of the variables both between and within the machine learning classifiers, suggest at least two things. First, overall the RF classifier is the most robust classifier. Second, additional spatial information is important to consider when training machine learning classifiers. Although it may be easier to only incorporate features that can be directly extracted from the GPS data (e.g. speed, acceleration, and direction change), the distance to the nearest road centre line, as well as the incorporation of a local density through a buffer or a time-constrained buffer, seem to play a significant role and should be considered for GPS data imputation methods. Notwithstanding this, the exact aim and purpose of the classification arguably plays a role in this decision. For instance, if the purpose of the classification is to give real-time information to the user, the cost of the classification, in terms of processing power and duration, may be decisive.

Figure 3.5 | Kappa values machine learning models. GPS data sets with 30-second measurement frequency with permutated variables.*



* See Appendix 3.D for the corresponding values.

Figure 3.6 | Kappa values machine learning models. GPS data sets with 60-second measurement frequency with permutated variables.*



* See Appendix 3.E for the corresponding values.

3.7 Conclusion and discussion

Where GPS technology can precisely register the spatiotemporal elements of activity-travel behaviour, travel characteristics need to be imputed from the data. However, especially rule-based methods that use dwell time for classification may disregard short activities. This chapter introduced a set of machine learning algorithms (naive Bayes, boosted C5.0, support vector machine, and random forest) to identify activity and travel episodes on a point-by-point basis. Another issue with GPS data imputation methods is that they often use attribute values to classify each data point individually. To ensure the underlying data structure of the GPS trajectories, local densities were incorporated into the models using multiple moving windows. The models were tested on a set of 200 artificially generated GPS activity-travel sequences with varying noise levels, and compared to three rule-based methods that served as a baseline.

Overall, the results showed that RF is the most accurate classifier. RF not only returns the highest kappa values for both the 30 and 60 second measurement frequency data sets and all sampled data sets, but it is also the most robust classifier. This robustness of the RF is not very surprising, because the trees in the forest are grown on different samples of the data, and at each node, three random variables are used in the split. In addition, whereas most rule-based classification techniques predominantly use speed and acceleration for classification, the incorporation of local densities through buffers and time-constrained buffers with different sizes proved to be important. In all the classifiers, at least one of the buffers or time-constrained buffers came forward as being of importance for classification.

It should be noted that these procedures were not tested on real GPS data. This implies that the reported quality of the classifiers for a real-life situation will strongly depend on how well the artificial GPS data simulate real GPS data. The better the artificial GPS data represent a real-life situation, the better the classifiers can be trained. While this is obviously a caveat, the fact that the parameters and noise levels of the artificial data can be precisely tuned opens doors for the exploration of real-life applications. Moreover, the availability of a ground truth is a major advantage of using artificially generated data. For future studies, a systematic comparison between models trained on artificial data and subsequently applied to real-life data could shed more light on this – although the availability of a ground truth remains crucial. Along the same lines, the concept of multiple moving spatiotemporal windows could also be applied to real-life data.

References

Ashbrook, D. & Starner, T. 2003. Using GPS to learn significant locations and predict movement across multiple users. *Personal and Ubiquitous Computing*. 7(5):275–286.

Auld, J., Williams, C., Mohammadian, A. & Nelson, P. 2009. An automated GPS-based prompted recall survey with learning algorithms. *Transportation Letters: The International Journal of Transportation Research*. 1(1):59–79.

Bohte, W. & Maat, K. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*. 17(3):285–297.

Breiman, L. 2001. Random forests. *Machine Learning*. 45(1):5–32.

Chai, Y., Chen, Z., Liu, Y., Tana & Ma, X. 2014. Space-time behavior survey for smart travel planning in Beijing, China. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection* *and analysis*. Hershey, Pennsylvania: IGI Global. 79–90.

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*. 16(6):293–305.

Chatterjee, K. & Bonsall, P. 2009. Special issue on evaluation of programmes promoting voluntary change in travel behavior. *Transport Policy*. 16(6):279–280.

Feng, T. & Timmermans, H.J.P. 2013. Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies*. 37:118–130.

Feng, T. & Timmermans, H.J.P. 2014. Multi-week travel surveys using GPS devices: Experiences in the Netherlands. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile Technologies for Activity-Travel Data Collection and Analysis*. Hershey, Pennsylvania: IGI Global. 104–118.

Feng, T. & Timmermans, H.J.P. 2016. Comparison of advanced imputation algorithms for detection of transportation mode and activity episode using GPS data. *Transportation Planning and Technology*. 39(2):180–194.

Geertman, S., De Jong, T. & Wessels, C. 2003. Flowmap: A support tool for strategic network analysis. In S. Geertman & J. Stillwell (eds.). (Advances in spatial science). *Planning support systems in practice*. Berlin, Heidelberg: Springer. 155–175.

Hagenauer, J. & Helbich, M. 2017. A comparative study of machine learning classifiers for modeling travel mode choice. *Expert Systems with Applications*. 78:273–282.

Jianchuan, X., Zhicai, J., Guangnian, X. & Xuemei, F. 2014. Smartphone-based travel survey: A pilot study in China. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 209–223.

Krygsman, S.C. & Nel, J.H. 2009. The use of global positioning devices in travel surveys - A developing country application. In *Proceedings of the 28th Southern African Transport Conference (SATC 2009)*. Pretoria: Southern African Transport Conference. 108–118.

Kuhn, M., Weston, S., Coulter, N. & Culp, M. 2015. *C50: C5.0 decision trees and rule-based models [R package version 0.1.0-24]*. [Online], Available: https://cran.r-project.org/package=C50.

Liao, L., Fox, D. & Kautz, H.A. 2005. Location-based activity recognition. In Y. Weiss, P.B. Schölkopf, & J.C. Platt (eds.). *Advances in neural information processing systems*. Vancouver: Annual Conference on Neural Information Processing Systems. [Online], Available: https://papers.nips.cc/book/advances-in-neural-information-processing-systems-18-2005 [2017, April 11].

Liaw, A. & Wiener, M. 2002. Classification and regression by randomForest. *R news*. 2(3):18–22.

Maat, K. & Timmermans, H.J.P. 2009. A causal model relating urban form with daily travel distance through activity/travel decisions. *Transportation Planning and Technology*. 32(2):115–134.

Meloni, I. & Sanjust, B. 2014. Using a GPS active logger to implement travel behavior change programs. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 325–340.

Meyer, D., Dimitriadou, E., Hornik, K., Weingessel, A. & Leisch, F. 2017. *e1071: Misc functions of the Department of Statistics [R package version 1.6-8]*. Vienna, Austria: Probability Theory Group (Formerly: E1071), TU Wien.

Miller, H.J. 2005. A measurement theory for time geography. *Geographical Analysis*. 37(1):17–45.

Nitsche, P., Widhalm, P., Breuss, S. & Maurer, P. 2012. A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys. In Vol. 48. *Procedia - Social and Behavioral Sciences*. Oxford: Elsevier. 1033–1046.

OpenStreetMap Contributors. 2016. *Planet Dump [Datafile from 26/07/2016 of BBBike extracts]*. [Online], Available: http://extract.bbbike.org/ [2016, July 26].

Prelipcean, A.C., Gidófalvi, G. & Susilo, Y. 2017. Transportation mode detection - An in-depth review of applicability and reliability. *Transport Reviews*. 37(4):442–464.

R Core Team. 2016. *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Richardson, A.J., Seethaler, R.K. & Harbutt, P.L. 2004. Design issues for before and after surveys of travel behaviour change. *Transport Engineering in Australia*. 9(2):103–118.

Richter, J., Friman, M. & Gärling, T. 2011. Soft transport policy measures: Gaps in knowledge. *International Journal of Sustainable Transportation*. 5(4):199–215.

Schoier, G. & Borruso, G. 2011. Individual movements and geographical data mining. Clustering algorithms for highlighting hotspots in personal navigation routes. In Vol. 6782. B. Murgante, O. Gervasi, A. Iglesias, D. Taniar, & B.O. Apduhan (eds.). (Lecture notes in computer science). *International conference on computational science and its applications*. Berlin, Heidelberg: Springer. 454–465.

Schuessler, N. & Axhausen, K.W. 2009. Processing raw data from Global Positioning Systems without additional information. *Transportation Research Record: Journal of the Transportation Research Board*. 2105:28–36.

Shafique, M. & Hato, E. 2016. Travel mode detection with varying smartphone data collection frequencies. *Sensors*. 16(5):716.

Shen, L. & Stopher, P.R. 2013. A process for trip purpose imputation from Global Positioning System data. *Transportation Research Part C: Emerging Technologies*. 36:261–267.

Shen, L. & Stopher, P.R. 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*. 34(3):316–334.

Shoval, N., Kwan, M.-P., Reinau, K.H. & Harder, H. 2014. The shoemaker's son always goes barefoot: Implementations of GPS and other tracking technologies for geographic research. *Geoforum*. 51:1–5.

Sterkenburg, R.P., Pierik, F.P. & De Vries, S. 2012. *Filtering GPS tracks: Cluster detection, cluster classification and transportation mode classification*. the Hague: the Netherlands Organisation for applied scientific research (TNO).

Stopher, P.R. & Greaves, S.P. 2007. Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice*. 41(5):367–381.

Stopher, P.R., Clifford, E., Swann, N. & Zhang, Y. 2009. Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy*. 16(6):315–324.

Thierry, B., Chaix, B. & Kestens, Y. 2013. Detecting activity locations from raw GPS data: A novel kernel-based algorithm. *International Journal of Health Geographics*. 12(1):14.

Wan, N. & Lin, G. 2013. Life-space characterization from cellular telephone collected GPS data. *Computers, Environment and Urban Systems*. 39:63–70.

Wan, N. & Lin, G. 2016. Classifying human activity patterns from smartphone collected GPS data: A fuzzy classification and aggregation approach. *Transactions in GIS*. 20(6):869–886.

Witten, I.H. & Frank, E. 2005. *Data Mining: Practical machine learning tools and techniques*. 2nd ed. San Francisco, CA: Elsevier.

Wolf, J. 2000. *Using GPS data loggers to replace travel diaries in the collection of travel data*. Published doctoral dissertation. Atlanta, Georgia: Georgia Institute of Technology.

Zhou, X., Yu, W. & Sullivan, W.C. 2016. Making pervasive sensing possible: Effective travel mode sensing based on smartphones. *Computers, Environment and Urban Systems*. 58:52–59.

Appendix 3.A | Artificial GPS track creation

The procedure below describes the steps that are involved in the generation of artificial GPS tracks. The procedure was adapted and extended from Thierry *et al.* (2013). The code is available on GitHub as a collection of Python scripts (https://github.com/jtvandijk/ artificial-gps-repo).

Activities

1. Selection of a random number of activity locations (minimum of two, maximum of nine) using points of interest (POI) within the Cape Town metropolitan region. POIs were extracted from OpenStreetMap (OSM) with the Java command line Osmosis for OS X (OpenStreetMap Contributors, 2016). A total of 5,378 POIs such as restaurants, cafés, bars, supermarkets, and schools that are within 100 metres from the road network were used for activity location selection.

2. Random assignment of a duration classification to each of the activities (short, medium, long). After randomly assigning a duration classification to each activity, the actual duration of the activity was simulated using a Poisson distribution. The seed value for λ was randomly drawn from a normal distribution with the following parameters:

   a) A mean of 8 minutes and a standard deviation of 3 minutes for activities with a short duration (e.g. picking someone up).
   b) A mean of 50 minutes and a standard deviation of 10 minutes for activities with a medium duration (e.g. buying groceries).
   c) A mean of 300 and a standard deviation of 60 minutes for activities with a long duration (e.g. work, school).

3. Generation of activity points (taking the measurement frequency into account) around the activity locations. GPS noise was simulated by shifting the coordinates from the original identified activity location. The shifts followed a bi-dimensional normal distribution centred around a mean of 0 with a standard deviation that was randomly drawn from a uniform distribution ranging from 10 to 200 metres.

Travel

1. Connection of the generated activity locations sequence (e.g. Activity 1 to Activity 2, Activity 2 to Activity 3, etc.) using shortest path analysis in Flowmap (Geertman, De Jong & Wessels, 2003).

2. Transformation of the resulting polylines into points and a random assignment of a transport mode. The spacing of the points was based on the measurement frequency and the speed

of the randomly assigned transport mode. The speeds for each point were recursively drawn from a normal distribution with the following parameters:

a) A mean of 80 km/h and a standard deviation of 10 km/h for a car driving fast (e.g. freeway).
b) A mean of 60 km/h and a standard deviation of 5 km/h for a car driving normally (e.g. urban setting).
c) A mean of 20 km/h and a standard deviation of 5 km/h for a car driving slowly (e.g. in a living area).
d) A mean of 15 km/h and a standard deviation of 3 km/h for cycling.
e) A mean of 5 km/h and a standard deviation of 2 km/h for walking.

3. Introduction of noise in GPS tracks by randomly shifting the trip points along their X and Y axes; the shifts followed a bi-dimensional normal distribution with a mean of 0 metres and a standard deviation of 10 metres.

Activity-travel GPS track

1. Merging of the activity points and the track points to get the artificial activity travel sequence.

2. Calculation of attribute values (e.g. speed, acceleration, distance to the nearest road centre line, etc.).

Appendix 3.B | Confusion matrices machine learning algorithms

Table 3.6 | Confusion matrices machine learning algorithms 30-second measurement frequencies

	100%	Move	Stay	**75%**	Move	Stay
NB	Move	30,880	2,360	Move	23,304	1,092
	Stay	764	36,194	Stay	581	27,959
SVM	Move	31,326	103	Move	23,569	109
	Stay	318	38,451	Stay	316	28,942
C5.0	Move	31,374	228	Move	23,585	207
	Stay	270	38,326	Stay	300	28,844
RF	Move	31,227	75	Move	23,513	64
	Stay	367	38,479	Stay	372	28,987

	50%	Move	Stay	**25%**	Move	Stay
NB	Move	15,426	693	Move	7,741	674
	Stay	356	18,804	Stay	205	9,028
SVM	Move	15,535	124	Move	7,752	109
	Stay	247	19,373	Stay	194	9,593
C5.0	Move	15,545	210	Move	7,757	145
	Stay	237	19,287	Stay	189	9,557
RF	Move	15,487	80	Move	7,743	79
	Stay	295	19,417	Stay	203	9,623

Table 3.7 | Confusion matrices machine learning algorithms 60-second measurement frequencies

	100%	Move	Stay	**75%**	Move	Stay
NB	Move	14,207	2,617	Move	10,744	1,831
	Stay	239	15,442	Stay	199	11,860
SVM	Move	14,207	129	Move	10,697	108
	Stay	239	17,930	Stay	246	13,583
C5.0	Move	14,238	178	Move	10.743	218
	Stay	208	17,881	Stay	200	13,473
RF	Move	14,191	83	Move	10,677	86
	Stay	255	17,976	Stay	266	13,605

	50%	Move	Stay	**25%**	Move	Stay
NB	Move	7,006	1,149	Move	3,594	575
	Stay	157	7,915	Stay	97	3,969
SVM	Move	6,946	88	Move	3,584	90
	Stay	217	8,976	Stay	107	4,454
C5.0	Move	6,968	145	Move	3,578	95
	Stay	195	8,919	Stay	113	4,449
RF	Move	6,947	65	Move	3,589	49
	Stay	216	8,999	Stay	102	4,495

Appendix 3.C | Sensitivity, specificity, positive predictive value, and negative predictive value for both the rule-based methods and the machine learning algorithms

Table 3.8 | Sensitivity, specificity, positive predictive value, and negative predictive value for both the rule-based methods and the machine learning algorithms with 30-second measurement frequency. *Move* as positive class.

		100%	**75%**	**50%**	**25%**
NB	Sensitivity	0.976	0.976	0.977	0.974
	Specificity	0.939	0.962	0.965	0.931
	Pos. pred. value	0.929	0.955	0.957	0.920
	Neg. pred. value	0.979	0.980	0.981	0.978
SVM	Sensitivity	0.990	0.987	0.984	0.976
	Specificity	0.997	0.996	0.994	0.989
	Pos. pred. value	0.997	0.995	0.992	0.986
	Neg. pred. value	0.992	0.989	0.987	0.980
C5.0	Sensitivity	0.992	0.987	0.985	0.976
	Specificity	0.994	0.993	0.989	0.985
	Pos. pred. value	0.993	0.991	0.987	0.982
	Neg. pred. value	0.993	0.990	0.988	0.981
RF	Sensitivity	0.988	0.984	0.981	0.975
	Specificity	0.998	0.998	0.996	0.992
	Pos. pred. value	0.998	0.997	0.995	0.990
	Neg. pred. value	0.991	0.987	0.985	0.979
RULE1	Sensitivity	0.379	0.379	0.364	0.381
	Specificity	0.684	0.714	0.753	0.809
	Pos. pred. value	0.496	0.521	0.544	0.621
	Neg. pred. value	0.573	0.583	0.594	0.615
RULE2	Sensitivity	0.999	0.995	0.989	0.979
	Specificity	0.427	0.441	0.479	0.575
	Pos. pred. value	0.589	0.594	0.606	0.654
	Neg. pred. value	0.999	0.991	0.981	0.971
RULE3	Sensitivity	0.997	0.986	0.967	0.937
	Specificity	0.575	0.618	0.684	0.769
	Pos. pred. value	0.658	0.680	0.712	0.768
	Neg. pred. value	0.995	0.982	0.963	0.937

Table 3.9 | Sensitivity, specificity, positive predictive value, and negative predictive value for both the rule-based methods and the machine learning algorithms with 60-second measurement frequency. *Move* as positive class.

		100%	75%	50%	25%
NB	Sensitivity	0.984	0.982	0.978	0.974
	Specificity	0.855	0.866	0.873	0.874
	Pos. pred. value	0.844	0.854	0.859	0.862
	Neg. pred. value	0.985	0.984	0.981	0.976
SVM	Sensitivity	0.984	0.978	0.970	0.971
	Specificity	0.993	0.992	0.990	0.980
	Pos. pred. value	0.991	0.990	0.988	0.976
	Neg. pred. value	0.987	0.092	0.976	0.977
C5.0	Sensitivity	0.986	0.982	0.973	0.969
	Specificity	0.990	0.984	0.984	0.979
	Pos. pred. value	0.988	0.980	0.980	0.974
	Neg. pred. value	0.989	0.985	0.979	0.975
RF	Sensitivity	0.982	0.976	0.970	0.972
	Specificity	0.995	0.994	0.993	0.989
	Pos. pred. value	0.994	0.992	0.991	0.988
	Neg. pred. value	0.986	0.981	0.977	0.978
RULE1	Sensitivity	0.408	0.399	0.404	0.530
	Specificity	0.735	0.765	0.797	0.822
	Pos. pred. value	0.552	0.576	0.612	0.708
	Neg. pred. value	0.608	0.614	0.629	0.683
RULE2	Sensitivity	0.970	0.954	0.945	0.938
	Specificity	0.381	0.427	0.480	0.604
	Pos. pred. value	0.556	0.571	0.590	0.658
	Neg. pred. value	0.940	0.921	0.917	0.923
RULE3	Sensitivity	0.956	0.927	0.896	0.872
	Specificity	0.676	0.718	0.735	0.782
	Pos. pred. value	0.702	0.725	0.728	0.764
	Neg. pred. value	0.950	0.925	0.900	0.883

Appendix 3.D | Kappa values activity recognition 30-second measurement frequency

Table 3.10 | Kappa values machine learning models. GPS data with 30-second measurement frequency (full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.911	0.988	0.986	0.987
DIR_CHANGE	0.905	0.980	0.979	0.984
SPEED	0.909	0.956	0.978	0.980
AVGSPEED	0.907	0.765	0.863	0.948
NEAR_DIST	0.795	0.931	0.962	0.950
MAXSPEED	0.902	0.979	0.965	0.986
ACC	0.878	0.980	0.982	0.986
AVGACC	0.910	0.980	0.984	0.986
MAXACC	0.887	0.975	0.972	0.983
BUF50	0.811	0.986	0.984	0.986
BUF100	0.748	0.981	0.984	0.987
BUF150	0.703	0.979	0.977	0.987
BUF250	0.620	0.976	0.906	0.984
POINTS50	0.826	0.969	0.981	0.986
POINTS100	0.790	0.985	0.974	0.986
POINTS150	0.772	0.981	0.976	0.981
POINTS250	0.899	0.793	0.891	0.968

Table 3.11 | Kappa values machine learning models. GPS data with 30-second measurement frequency (75 percent of the full data set) with permutated variables.

	NB	C5.0	SVM	RF
NONE	0.936	0.984	0.981	0.983
DIR_CHANGE	0.929	0.978	0.975	0.981
SPEED	0.934	0.962	0.976	0.980
AVGSPEED	0.931	0.819	0.927	0.956
NEAR_DIST	0.828	0.927	0.952	0.960
MAXSPEED	0.929	0.978	0.948	0.982
ACC	0.906	0.978	0.979	0.983
AVGACC	0.894	0.937	0.976	0.983
MAXACC	0.919	0.970	0.898	0.981
BUF50	0.828	0.977	0.978	0.984
BUF100	0.762	0.983	0.976	0.984
BUF150	0.716	0.979	0.976	0.984
BUF250	0.650	0.974	0.940	0.979
POINTS50	0.853	0.958	0.979	0.984
POINTS100	0.816	0.979	0.975	0.984
POINTS150	0.801	0.980	0.975	0.983
POINTS250	0.921	0.816	0.912	0.975

Table 3.12 | Kappa values machine learning models. GPS data with 30-second measurement frequency (50 percent of the full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.940	0.979	0.974	0.979
DIR_CHANGE	0.933	0.967	0.963	0.970
SPEED	0.935	0.964	0.967	0.975
AVGSPEED	0.933	0.919	0.935	0.965
NEAR_DIST	0.832	0.910	0.927	0.928
MAXSPEED	0.933	0.972	0.954	0.974
ACC	0.907	0.967	0.967	0.976
AVGACC	0.873	0.887	0.961	0.973
MAXACC	0.918	0.972	0.965	0.975
BUF50	0.824	0.962	0.963	0.977
BUF100	0.760	0.975	0.968	0.977
BUF150	0.711	0.976	0.922	0.978
BUF250	0.646	0.961	0.936	0.969
POINTS50	0.865	0.955	0.968	0.977
POINTS100	0.832	0.960	0.968	0.977
POINTS150	0.831	0.973	0.967	0.976
POINTS250	0.924	0.896	0.952	0.974

Table 3.13 | Kappa values machine learning models. GPS data with 30-second measurement frequency (25 percent of the full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.900	0.965	0.962	0.968
DIR_CHANGE	0.894	0.956	0.946	0.965
SPEED	0.899	0.948	0.937	0.970
AVGSPEED	0.895	0.940	0.953	0.968
NEAR_DIST	0.776	0.882	0.878	0.871
MAXSPEED	0.899	0.823	0.940	0.969
ACC	0.869	0.960	0.928	0.972
AVGACC	0.865	0.876	0.954	0.970
MAXACC	0.870	0.963	0.964	0.969
BUF50	0.800	0.946	0.964	0.973
BUF100	0.744	0.933	0.956	0.973
BUF150	0.701	0.856	0.895	0.962
BUF250	0.621	0.870	0.822	0.947
POINTS50	0.848	0.952	0.963	0.973
POINTS100	0.833	0.965	0.958	0.973
POINTS150	0.842	0.960	0.954	0.973
POINTS250	0.894	0.933	0.963	0.971

Appendix 3.E | Kappa values activity recognition 60-second measurement frequency

Table 3.14 | Kappa values machine learning models. GPS data with 60-second measurement frequency (full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.825	0.977	0.976	0.979
DIR_CHANGE	0.821	0.960	0.963	0.970
SPEED	0.828	0.931	0.962	0.973
AVGSPEED	0.826	0.911	0.902	0.949
NEAR_DIST	0.685	0.893	0.945	0.920
MAXSPEED	0.827	0.959	0.946	0.974
ACC	0.823	0.969	0.975	0.977
AVGACC	0.829	0.958	0.974	0.978
MAXACC	0.824	0.971	0.966	0.973
BUF50	0.746	0.973	0.972	0.977
BUF100	0.685	0.967	0.967	0.978
BUF150	0.646	0.968	0.963	0.977
BUF250	0.570	0.858	0.833	0.951
POINTS50	0.750	0.961	0.976	0.978
POINTS100	0.723	0.973	0.974	0.978
POINTS150	0.717	0.973	0.970	0.977
POINTS250	0.832	0.675	0.911	0.963

Table 3.15 | Kappa values machine learning models. GPS data with 60-second measurement frequency (75 percent of full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.836	0.971	0.966	0.971
DIR_CHANGE	0.825	0.955	0.946	0.965
SPEED	0.833	0.945	0.952	0.972
AVGSPEED	0.828	0.922	0.877	0.966
NEAR_DIST	0.687	0.893	0.918	0.917
MAXSPEED	0.832	0.957	0.954	0.971
ACC	0.827	0.969	0.965	0.973
AVGACC	0.827	0.944	0.965	0.972
MAXACC	0.825	0.962	0.955	0.969
BUF50	0.754	0.964	0.967	0.973
BUF100	0.694	0.970	0.966	0.974
BUF150	0.658	0.970	0.956	0.972
BUF250	0.586	0.891	0.814	0.935
POINTS50	0.763	0.953	0.966	0.973
POINTS100	0.730	0.969	0.955	0.973
POINTS150	0.732	0.967	0.965	0.973
POINTS250	0.826	0.826	0.957	0.970

Table 3.16 | Kappa values machine learning models. GPS data with 60-second measurement frequency (50 percent of full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.839	0.962	0.958	0.965
DIR_CHANGE	0.832	0.943	0.934	0.956
SPEED	0.841	0.917	0.920	0.962
AVGSPEED	0.836	0.916	0.937	0.961
NEAR_DIST	0.707	0.868	0.814	0.852
MAXSPEED	0.842	0.892	0.948	0.962
ACC	0.832	0.958	0.954	0.965
AVGACC	0.832	0.928	0.951	0.964
MAXACC	0.830	0.942	0.925	0.960
BUF50	0.759	0.947	0.956	0.964
BUF100	0.692	0.937	0.949	0.965
BUF150	0.660	0.885	0.949	0.961
BUF250	0.573	0.871	0.952	0.909
POINTS50	0.783	0.948	0.953	0.965
POINTS100	0.761	0.960	0.954	0.965
POINTS150	0.775	0.958	0.928	0.965
POINTS250	0.829	0.924	0.764	0.963

Table 3.17 | Kappa values machine learning models. GPS data with 60-second measurement frequency (25 percent of full data set) with permutated variables.

	NB	SVM	C5.0	RF
NONE	0.837	0.952	0.949	0.963
DIR_CHANGE	0.820	0.916	0.922	0.946
SPEED	0.828	0.912	0.913	0.952
AVGSPEED	0.821	0.859	0.895	0.954
NEAR_DIST	0.678	0.836	0.796	0.813
MAXSPEED	0.831	0.850	0.927	0.958
ACC	0.819	0.945	0.948	0.960
AVGACC	0.824	0.932	0.949	0.960
MAXACC	0.817	0.927	0.941	0.957
BUF50	0.743	0.937	0.950	0.960
BUF100	0.674	0.902	0.948	0.963
BUF150	0.636	0.923	0.917	0.958
BUF250	0.562	0.659	0.680	0.891
POINTS50	0.781	0.947	0.951	0.961
POINTS100	0.784	0.950	0.951	0.961
POINTS150	0.788	0.950	0.949	0.961
POINTS250	0.826	0.946	0.950	0.960

Chapter 4. Post-processing GPS tracks in reconstructing travelled routes in a GIS-environment: Network subset selection and attribute adjustment

Van Dijk, J.T. & De Jong, T.J., 2017, Post-processing GPS-tracks in reconstructing travelled routes in a GIS-environment. *Annals of GIS.* 23(3):203-217.

Abstract

The exact distance and routes travelled on an individual level are essential variables in determining the effectiveness of 'soft' transport demand management strategies. The ability to track individuals in great spatial detail by means of location-aware technologies such as GPS has opened avenues for gathering these data with great precision. Route reconstruction with positional data is typically done by a process referred to as map-matching, however, despite the large number of real-time map-matching algorithms developed, few studies have developed a post-processing map-matching algorithm in a Geographic Information System (GIS)-environment. This chapter presents two GIS-based map-matching methods that predominantly use a digital road network with speed and directionality attributes for route reconstruction of raw GPS trajectories. The methodologies were tested for a dataset in which actual routes travelled were known. Both explored procedures, the *connected subset* assignment procedure based on network subset selection and the *impedance reduction* assignment procedure based on attribute adjustment, provide accurate results. In addition, both procedures effectively deal with commonly GPS-induced problems such as measurement gaps and positional drift.

Keywords

Route reconstruction; Post-processing map-matching; GPS; Geographic Information Systems

4.1 Introduction

In recent years, there has been a rising interest in 'soft' transport demand management strategies (cf. Cairns, Sloman, Newson, Anable, Kirkbridge & Goodwin, 2008; Bonsall, 2009; Brög, Erl, Ker, Ryle & Wall, 2009; Chatterjee, 2009; Stopher, Clifford, Swann & Zhang, 2009; Zhang, Stopher & Halling, 2013). The objective of these strategies, often referred to as voluntary travel behaviour change (VTBC) interventions, is typically "to allow people to choose to change travel behaviour rather than to expect or force reactions in response to external stimuli or pressures" (Taylor & Ampt, 2003: 171), for example, with a public information campaign aimed at informing individuals about the negative impacts of private vehicle use and highlighting public transport alternatives. Important variables in assessing the effectiveness of VTBC interventions include the exact distances and routes travelled on an individual level – information that is often difficult to acquire with traditional data collection techniques.

Since the late 1990s, Global Positioning Systems (GPS) technology and, more recently, GPS-enabled smartphones have opened avenues for accurately reconstructing the actual distances and routes travelled. Typically this is done by assigning GPS measurement to a digital road network by a process referred to as map-matching (Quddus, Ochieng & Noland, 2007; Hashemi & Karimi, 2014). However, Dalumpines and Scott (2011: 102) note that:

> [a]lthough most map-matching approaches use geometric and topological analysis, two common built-in functions in most GIS [Geographic Information System] packages, very limited studies have attempted to develop a post-processing map-matching algorithm in a GIS platform.

Notwithstanding the limited deployment of GIS platforms in the development of post-processing map-matching algorithms, the inclusion of network geometry and topology is pivotal in correctly reconstructing routes, as GPS is affected by measurement gaps, cold starts and signal reflection (Schuessler & Axhausen, 2009a; International Transport Forum, 2015). Furthermore, when adequate GIS data are available, such as detailed land use and cadastral information, a GIS offers an attractive analytical environment to integrate route reconstruction with other analyses such as activity diary reconstruction and trip purpose imputation of raw GPS trajectories.

This chapter describes two GIS-based post-processing map-matching algorithms with modest data requirements (the methods require only a topologically correct road network with speed and directionality). The first proposed algorithm selects a subset of the network, whilst the second proposed algorithm adjusts the impedance of network segments in a shortest path route assignment. A shortest path is respectively fitted through the subset of the network and the network with adjusted impedance attribute data. In addition, whilst the majority of post-processing map-matching algorithms focus solely on home to work trips by car, the applicability of the proposed map-matching algorithms for walk and bike trips is briefly explored. Especially in the context of 'soft' transport demand management policies in which stimulating more sustainable modes of travel is often one of the objectives, the necessity to reconstruct the bicycle and walking routes taken by individuals becomes evident.

4.2 Literature review

Reconstructed actual routes travelled by individuals serve transportation researchers in a number of ways, for instance, by providing input to route choice modelling exercises, estimating travel times at different times of the day, and calibrating traffic models (Hashemi & Karimi, 2014), The exact distance travelled is also often an important variable in evaluating the effectiveness of VTBC interventions (Schuessler & Axhausen, 2009b; Stopher *et al.*, 2009; Bierlaire, Chen & Newman, 2013). Although data on routes travelled are difficult to acquire with traditional data collection techniques, the increasing availability and ubiquity of GPS technology has enabled the collection of large amounts of high-frequency locational data. However, these raw data need to be processed in a systematic way in order to accurately identify the routes that were taken by their users: a process which is often referred to as map-matching (Quddus *et al.*, 2007; Chen & Bierlaire, 2015). A distinction is usually made between real-time map-matching algorithms and post-processing map-matching algorithms. Whereas real-time map-matching procedures focus on identifying the road segment on which the user is travelling and the position of the user on that road segment, post-processing map-matching algorithms aim to reconstruct the actual route travelled for an entire trip after a large number of raw GPS points has been acquired (Hashemi & Karimi, 2014). In addition, in post-processing map-matching, the continuity of the path route is essential, which is not a condition in real-time map-matching. Implementations of post-processing map-matching therefore allow for the employment of more computationally intensive methods. Because real-time map-matching aims to find the actual *position* of the user in real time and post-processing map-matching aims to reconstruct the actual *path*, "real-time map-matching algorithms cannot directly be used in place of post-processing map-matching algorithms as they have to resolve different challenges" (Hashemi & Karimi, 2014: 154).

The translation of raw GPS trajectories into the best estimation of a route travelled by a user is not a trivial exercise (Krumm, Gruen & Delling, 2013). This can be largely attributed to two sources of error. First, GPS faces several limitations regarding its accuracy. The environment in which the user is situated can significantly affect the accuracy of the recorded data. Whilst in open areas, GPS can be accurate up to 5 metres, provided that there are at least four satellites in sight; but the accuracy may decrease to 50 metres as a consequence of signal blockage by, for instance, tall buildings leading to multipath signal reflection (Schuessler & Axhausen, 2009a; International Transport Forum, 2015). Second, the digital road network to which the GPS measurements have to be assigned is only a digital representation of the actual world. The quality of the road network is thus crucial in map-matching; missing or unconnected road segments can cause a map-matching algorithm to fail (Marchal, Hackney & Axhausen, 2005; Quddus *et al.*, 2007). The challenge in post-processing map-matching is, therefore, to find the actual route travelled in spite of these errors (Krumm, Letchner & Horvitz, 2006).

The most basic type of map-matching works by assigning each individual GPS measurement to the nearest road segment. The algorithm simply calculates, for each GPS location, the airline distance to the nearest road segment or nearest node, and the measurement is subsequently assigned to this road segment or node. The simplicity of this approach is at the same time its major limitation, because neither does it account for network topology nor does it account for inaccuracies in the measurements: the nearest link is not the correct link per se (Schuessler & Axhausen, 2009b). Accordingly, throughout the years, numerous map-matching techniques have been proposed that go beyond this basic method, although they have been predominantly developed in the domain of real-time map-matching (Quddus *et al.*, 2007; Hashemi & Karimi, 2014).

Whilst some of the existing map-matching techniques involve simple procedures, others are more advanced by employing, for instance, fuzzy logic, mathematical models (Bierlaire *et al.*, 2013) or Bayesian belief networks (Quddus, Noland & Ochieng, 2006; Quddus *et al.*, 2007; Blazquez & Miranda, 2014). Quddus *et al.* (2007) roughly categorise the methods into four types: geometric, topological, probabilistic, and other advanced techniques. However, these methods are often not suitable to employ in the context of post-processing GPS measurements because of the inherent differences in the problems that real-time and post-processing map-matching techniques are trying to address. Post-processing map-matching allows for more computationally intensive procedures, but they also have to perform faster than real-time, something which has been particularly emphasised by Schuessler and Axhausen (2009b). In addition, in post-processing map-matching algorithms the whole GPS trajectory of a trip should be taken into account in order to ensure connectivity of the matched route (Marchal *et al.*, 2005).

In the past decade, several map-matching techniques have been developed and implemented in the context of post-processing raw GPS trajectories. Early methods predominantly employed geometric procedures, ranging from a simple nearest node search to curve-to-curve matching (White, Bernstein & Kornhauser, 2000). However, as Schuessler and Axhausen (2009b: 3) note, "the main shortcoming of all geometric procedures is that they ignore the sequence of the GPS points over time as well as the connectivity of the network links". Topological procedures try to avoid this shortcoming by extending the geometric approaches, taking the sequence of GPS measurements and the connectivity of the network links into account (Chung & Shalaby, 2005; Velaga, Quddus & Bristow, 2009). More statistically advanced probabilistic methods have been proposed by Marchal *et al.* (2005) and Schuessler and Axhausen (2009b), who have developed algorithms where for each GPS measurement a set of candidate paths is selected using the multiple hypothesis technique. In this iterative procedure, a best candidate is selected from the set of candidate paths when the full sequence of GPS measurements has been processed (Schuessler & Axhausen, 2009b).

What all of the methods for post-processing GPS trajectories that have been mentioned so far have in common is their statistical or mathematical approach to solve the problem, often extended with geometrical and topological information: information that is inherent to the data model for transport networks in a GIS-environment (Papinski & Scott, 2011). As Dalumpines and Scott (2011: 105) argue, this is a missed opportunity, because:

> [t]his strength of GIS makes it an ideal platform in developing a post-processing map-matching algorithm that fully integrates network topology and attributes to match streams of GPS points to the road network.

This also draws attention to the quality of the available data; a digital road network with minimum attributes can already be utilised for a simple map-matching solution without predefined parameters by capitalising on this geometrical and topological information. Lastly, the fact that routes are inherently spatial manifestations on a road network also offers direct opportunities for both spatial and non-spatial data manipulation and visualisation (Papinski & Scott, 2011). In fact, after the map-matching procedure itself, supplementary analyses can be executed from within the GIS-environment such, as activity diary reconstruction and trip purpose imputation of raw GPS trajectories.

Although a few studies have developed a post-processing map-matching algorithm in a GIS-environment, they still adapt real-time map-matching procedures (Chung & Shalaby, 2005). Dalumpines and Scott (2011) have tried to fill this void by proposing the first post-processing map-matching algorithm fully integrated into a GIS-environment. They set out to convert the stream of GPS measurement into a polyline feature. Subsequently, they created a predefined buffer of roughly five to six times the horizontal accuracy of the measurement around the newly created polyline feature. In turn, this buffer was used to limit the network through which a shortest path could be fitted between the first measurement and the last measurement in the sequence.

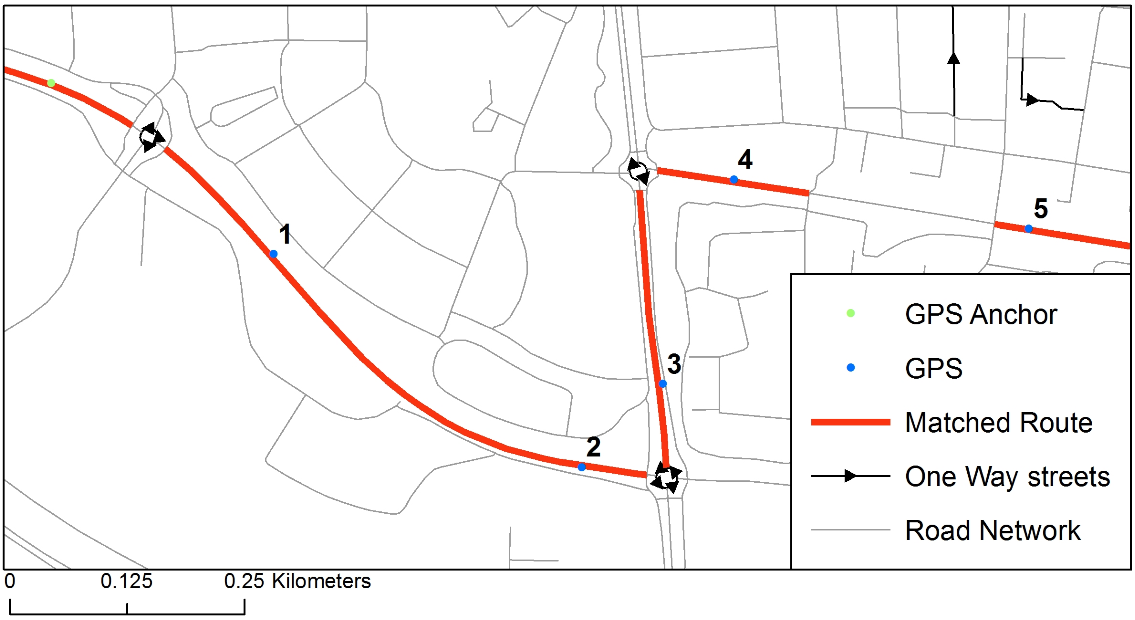
In 88 percent of the 99 routes Dalumpines and Scott (2011) analysed, this shortest path predicted a correct match with the actual route travelled. Whilst the buffer size with its predefined threshold distance should account for GPS inaccuracies and map inaccuracies, a limitation of the GIS-based buffer method is that the buffer size determines the success or failure of the proposed algorithm because the buffer restricts the shortest path to the network segments falling within the buffer. If the buffer size is too generous, a large number of irrelevant links can result in shorter, alternative, but incorrect, routes. If the buffer size is too small, the network might be too restrictive and completely prohibit the execution of the shortest path algorithm. In the following section, therefore, two purely GIS-based map-matching procedures are described and tested that circumvent the use of a buffer and do not use other predefined parameters.

4.3 Proposed GIS-based map-matching algorithms

A GIS is an extremely effective tool for the incorporation of a shortest path algorithm that can effectively solve for an optimal route based on a given impedance. However, the deployment of the shortest path between two activity locations, or the first and last measurement in a sequence of GPS measurements, is very unlikely to produce the actual route a user has travelled. In many occasions, people do not use the shortest path for reasons such as known congestions, safety, and simple preference. Whereas the shortest path on its own may not be suitable for reconstructing GPS routes, with an adjusted input the shortest path algorithm could still identify the actual route. A shortest path algorithm, for instance the one developed by Dijkstra (1959), requires two sources of input besides an origin and a destination: network segments and network attributes – both of which can be manipulated. Both of the methodologies suggested here work in a similar two-step approach, in which in the first step the network is either reduced or adjusted based on GPS measurements, and in the second step a shortest path between an origin and a destination is fitted through the selected or adjusted network. In both cases, the GIS extension Flowmap, specifically designed for handling spatial interaction and network data, is used for the execution of Dijkstra's (1959) shortest path algorithm (Geertman, De Jong & Wessels, 2003).

To illustrate the procedures of both proposed map-matching algorithms a set of six GPS points has been selected, as shown in Figure 4.1, of a short trip by car starting in an 'anchor' GPS point. Using a simple nearest link search, all GPS points correctly identify parts of the actual route travelled. However, apart from the road network segments between the first measurement after the anchor and the second measurement after the anchor, substantial gaps appear between the other GPS points, leading to an unconnected subnet or matched route. A relatively easy way to solve this is by increasing the measurement frequency; especially for in-vehicle GPS tracking devices, this is a viable solution. However, one should keep in mind that since the adoption of smartphones in GPS data collection, a higher measurement frequency is not always possible, because of the additional impact on the already short battery life of many devices. Furthermore, there are several one-directional road network segments present, of which those at the traffic circles seem to be important for an accurate route reconstruction. These particularities of the road network pose a simple map-matching problem in which network attribute information and directionality need be exploited to fit a correct route.

Figure 4.1 | Example of simple nearest link matching



4.3.1 Network subset selection: Connected subset

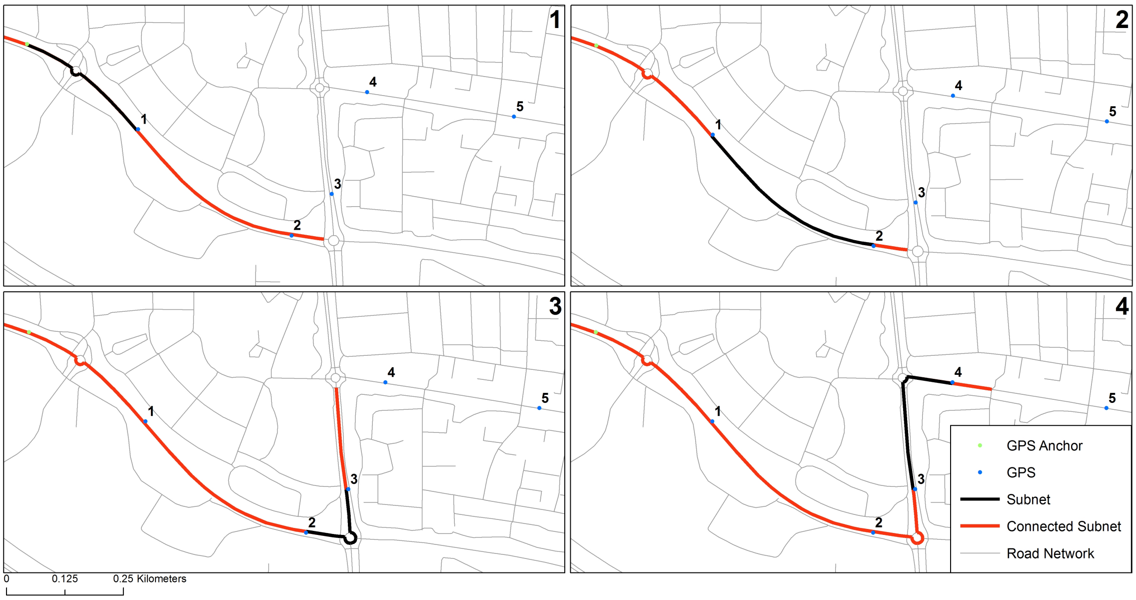
The first possible method suggested here is the *connected subset* assignment procedure, somewhat similar to the method used to make sense of locational data derived from cell phone records (Krygsman, Nel & de Jong, 2008) and to the real-time algorithm proposed by Miwa *et al.* (2012). Whereas Dalumpines and Scott (2011) use a buffer with a predefined threshold value to limit the network through which the shortest path algorithm finds a solution, the connected subset assignment takes a different approach by using the full network to connect consecutive GPS points by a series of 'mini' shortest path assignments, assuming that on such a short time interval the shortest path principle always applies as it is highly unlikely that detours are made between two GPS measurements. The combination of all these 'mini' shortest path assignments results in a fully connected subnetwork. In turn, a 'maxi' shortest path assignment is carried out through this subnetwork. The procedure involves the following steps:

1. Import a stream of GPS measurements (representing a single trip stage) into projected coordinates.

2. Connect all time-adjacent GPS measurements via the network by a series of 'mini' shortest path assignments, taking directionality and road type into account; when the mode of transport is known, a subnetwork is used for the initial matching process so that GPS measurements that were collected when the mode of transport was a car, can only be matched to road segments that allow for travel by car. The usage of a subnetwork reduces the chance of inducing a map-matching error because of GPS inaccuracies and map inaccuracies. As shown in Figure 4.2, the subnet (black) connects consecutive GPS measurements using a 'mini' shortest path assignment, and based on this subnet (red) the streets between consecutive measurements can be identified. The

outcome of this 'mini' assignment is a connected subnetwork derived from the union of the 'mini' shortest paths of each combination of consecutive GPS measurements.

Figure 4.2 | Creating a connected subnetwork using stepwise shortest path assignments



3. Connect the start and stop points of a trip by means of the shortest path, employing the subnetwork created in the previous step. The resulting route is the shortest path generated within the extent of the connected subnetwork created in the previous step, as illustrated in Figure 4.3; although in this case the subnetwork is relatively straightforward, subnetworks can get quite complicated, which necessitates the connection of the start and stop.

4.3.2 Attribute adjustment: Impedance reduction

The second method suggested here is the *impedance reduction* assignment procedure, somewhat similar to the voting-based algorithm proposed by Yuan *et al.* (2010). As opposed to the connected subset assignment procedure, the impedance reduction procedure reduces the impedance values of road segments based on the number of nearest GPS measurements. The result is that road segments with many hits will get a low impedance – establishing a 'waterbed' effect. In turn, a shortest path assignment is carried out through the full network with adjusted attributes, as opposed to the connected subset assignment procedure, in which the shortest path is fitted only through the connected subnetwork. The procedure involves the following steps:

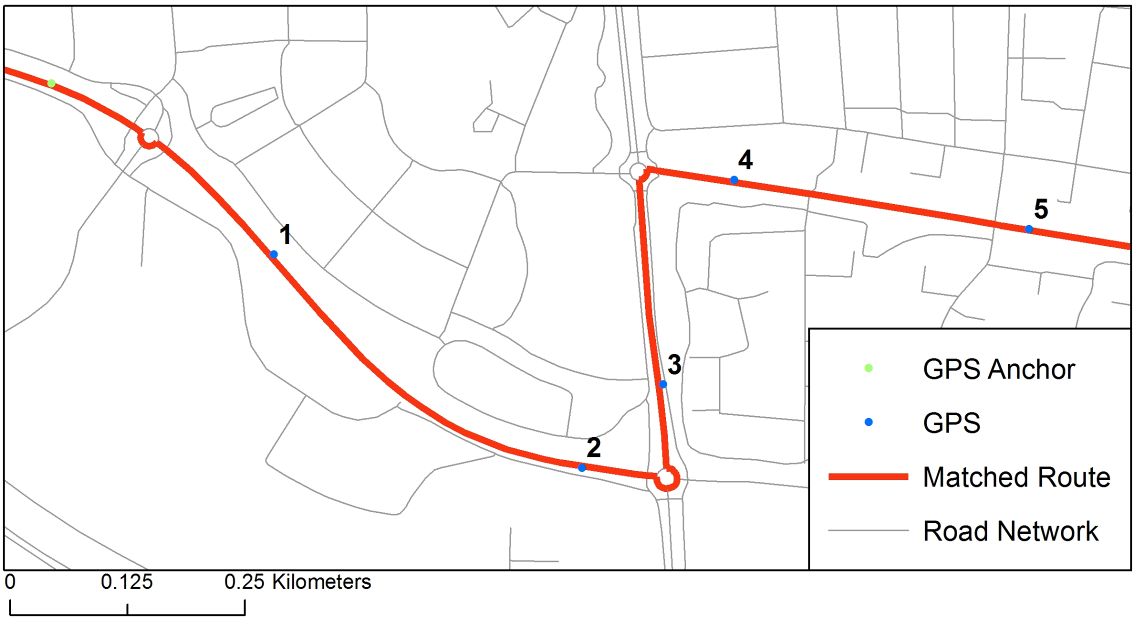
1. Import a stream of GPS measurements (representing a single trip stage) into projected coordinates.

2. Connect all the GPS measurements to the digital road network by means of an attribute-based near analysis using different subsets of the full digital road network: the car

network is used for car moves, the bicycle network is used for bicycle moves, the walking network is used for walk moves. In this way, road segments of, for instance, a highway, can only be assigned a nearest GPS measurement by qualifying modes of transport. Again, the usage of a subnetwork reduces the chance of inducing a map-matching error because of GPS inaccuracies and map inaccuracies.

3. Connect the start activity and stop activity of a trip by means of the shortest path, employing the adjusted network created in the previous step. The resulting route is the shortest path generated through the adjusted network; in our example resulting in the same solution as that presented in Figure 4.3 for the connected subset assignment procedure. Connect the start and stop activities of a trip by means of the shortest path employing the adjusted network created in the previous step. This way the shortest path assignment accounts for the network segments that were not included in the previous step of the matching procedure. The resulting route is the shortest path generated through the adjusted network; in our example resulting in the same solution as presented in Figure 4.3 for the connected subset assignment procedure.

Figure 4.3 | Correct solution with a shortest path through the connected subnetwork



4.4 Data input and pre-processing

A GPS data set was provided by Utrecht University, in which GPS measurements were available from colleague Maarten, who tracked all his movements in the period from December 03, 2014 to January 22, 2015, and from February 27, 2015, to March 20, 2015. All commuter trips had been done by car, but cycling and walking had also been used closer to home as alternative modes of transport. No public transport trips were recorded. In the first period, up to three different GPS loggers were deployed simultaneously; a Garmin 72H GPS tracking device, an Android G20 tablet with a MyTracks smartphone application, and a Trimble data logger. In the second period, two different applications were run on the same Android device. To that end, the

Trimble data logger was replaced with Tracklog, another smartphone application that also ran on the Android G20 tablet.[1] Well over 200,000 unique measurements that cover over 500 kilometres of unique road segments were registered over the period of data collection. Table 4.1 gives an overview of the full data set on which both methods were tested, including devices used, the total number of measurements, and the measurement frequency that was employed.

Table 4.1 | Devices and GPS measurements

Device	Number of measurements	Measurement frequency
Garmin 72H GPS tracking device	19,389	10 seconds
Android G20 tablet (MyTracks)	188,658	1 seconds
Android G20 tablet (TrackLog)	3,443	30 seconds
Trimble data logger	7,285	5 seconds

Table 4.2 | Trip diary definitions

Movement	Description
Stage / Move	a displacement with one travel mode.
Trip	the sequence of stages/moves between two subsequent activity locations
Stop	a non-movement during a stage
Activity Location	a location in space in which one interacts with for utilitarian or recreational purposes
Turning point	an arbitrary location that is part of a trip stage and is not an activity location, for instance, if one decides to turn around after taking a wrong turn
Transfer location	a location at the start or end of a move close to the actual activity location one is coming from or going to, for instance a parking lot close to work or close to a supermarket, or a location to switch travel mode initiating a next trip stage

4.4.1 GPS pre-processing

In preparation, all separate track recordings per device were merged in chronological order. This way, an automatic time relationship was created between the records that represent the switching off and switching on of the device, be it either on purpose or because of battery failure. After capturing the planar x- and y-coordinates, the measurements were transformed (in corroboration with Maarten) into a trip diary consisting of stages, trips, stops, activities, turning points, transfer locations, and travel modes, as shown in Table 4.2. It has to be noted that this corroboration was a necessity because the reconstruction of a trip diary based on GPS measurements without additional information is challenging (see for instance Schuessler &

[1] Tracklog is a smartphone application, running on the Android and iOS operating systems, developed at Stellenbosch University, South Africa, to track individuals by means of GPS to gather travel data on an individual level to assess the effectiveness of a VTBC intervention.

Axhausen, 2009a; Montini, Rieser-Schüssler, Horni & Axhausen, 2014; Shen & Stopher, 2014) and was outside the scope of the current study.

The pre-processing of the GPS measurements resulted in a trip diary including the actual mode of transport used for each trip stage. Table 4.3 gives an overview of all the devices with all the trip stages that are included in the analysis of both methods. Lastly, the actual routes travelled were added to a digital transport network. The digital network used was the official 2013 national road network from the Netherlands' Cadastre, Land Registry and Mapping Agency. Since 2012, the Basisregistratie Topografie has been freely available. This data set contains a detailed road network for all modes of transport. Unfortunately, speed and directionality were missing and had to be added for the relevant sections of the network, as these are a prerequisite for successfully executing a shortest path algorithm (Papinski & Scott, 2011).

Table 4.3 | Analysed trip stages

	Car	Bike	Walk	Total
Garmin	89	31	No Data	120
Android (MyTracks)	109	25	10	144
Trimble	23	4	No Data	27
Android (TrackLog)	52	24	8	84
Total	273	84	18	375

4.4.2 Route validation

To analyse the extent to which both post-processing map-matching methods were successful, all reconstructed routes needed to be validated.

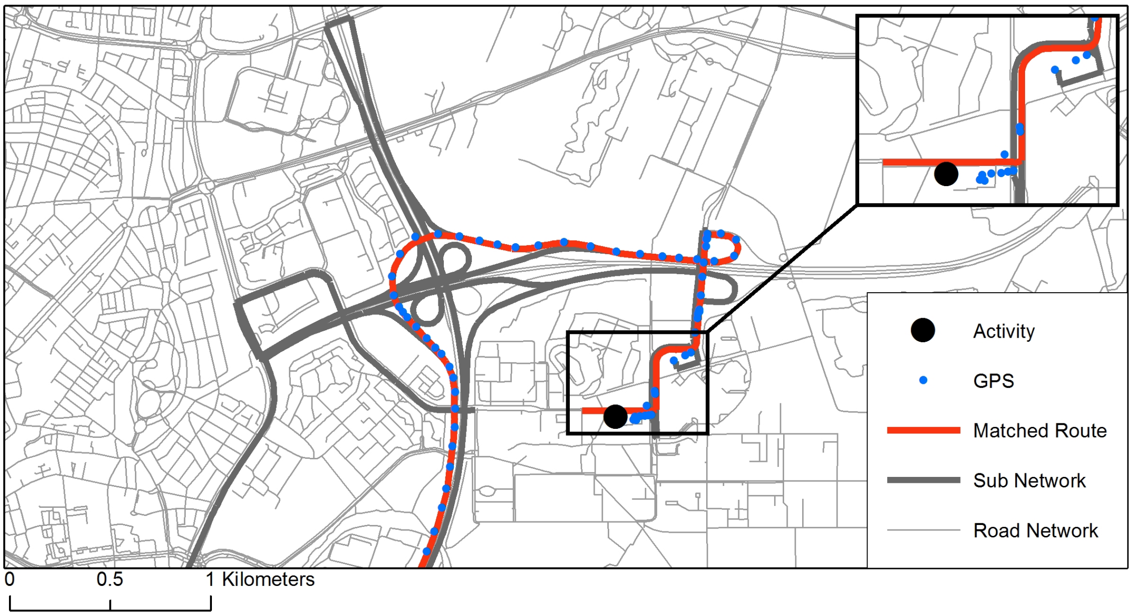
> Validation of a map-matching algorithm is essential to derive statistics on its performance in terms of correct link identification. Very few existing map-matching algorithms provide a meaningful validation technique (Quddus *et al.*, 2007: 322).

In a fashion comparable to Miwa *et al.* (2012), two indices were employed: the percentage of the road network links of the actual route that were identified correctly (% ACT) and the percentage of the road network links of the reconstructed route that were part of the actual route (% PRED). A score of 100 percent on both indices implies that the actual route travelled and the reconstructed route are identical. A score below 100 percent on ACT indicates that road segments of the actual route have been 'missed', whereas a score below 100 percent on PRED suggest that alternative, but incorrect, road segments have been incorporated in the reconstructed route, for example, because of GPS measurements 'hitting' an incorrect road segment.

4.5 Map-matching GPS measurements: Connected subnet

Before moving to the validation of the reconstructed trip stages for each device and each mode of travel, a graphical example of a successful execution of the procedure on a raw GPS trajectory is shown in Figure 4.4. It can be seen that quite a number of side roads are included in the subnetwork that was derived in the first step of the connected subset assignment. However, by executing the network subset selection through assigning 'mini' flows between consecutive GPS measurements, the connectivity of the subnetwork is warranted. Moreover, after exploiting the topology and geometry stored in the GIS network model, as well as the attribute information on directionality, Figure 4.4 shows that, despite the number of road segments that were included in the subnetwork and despite a noticeable gap in the raw GPS trajectory, the trip stage is correctly fitted through the complex connected subnetwork. This suggests that there are several GPS inaccuracies or map inaccuracies present in the example, but that after executing a series of 'mini' assignments between consecutive measurements the correct road segments are still included in the fully connected subnetwork.

Figure 4.4 | Example of a correct solution with a 'maxi' flow through the subnetwork



A summary of the results of the validation of the analysed routes with the connected subset assignment procedure is given in Table 4.4, and on average car moves and bike moves have high scores on both indices. Walk moves, on the other hand, did considerably less well. The fact that walking is less restricted to following the network and the low number of walk moves used in the analysis are largely responsible for these poor results. If we disregard the walk and bike moves for a moment and focus on the moves made by car, the results show index values close to 95 percent. However, there is a noticeable difference between the devices. The Android (MyTracks) had the highest scores. Not surprisingly, this device also had the highest measurement frequency.

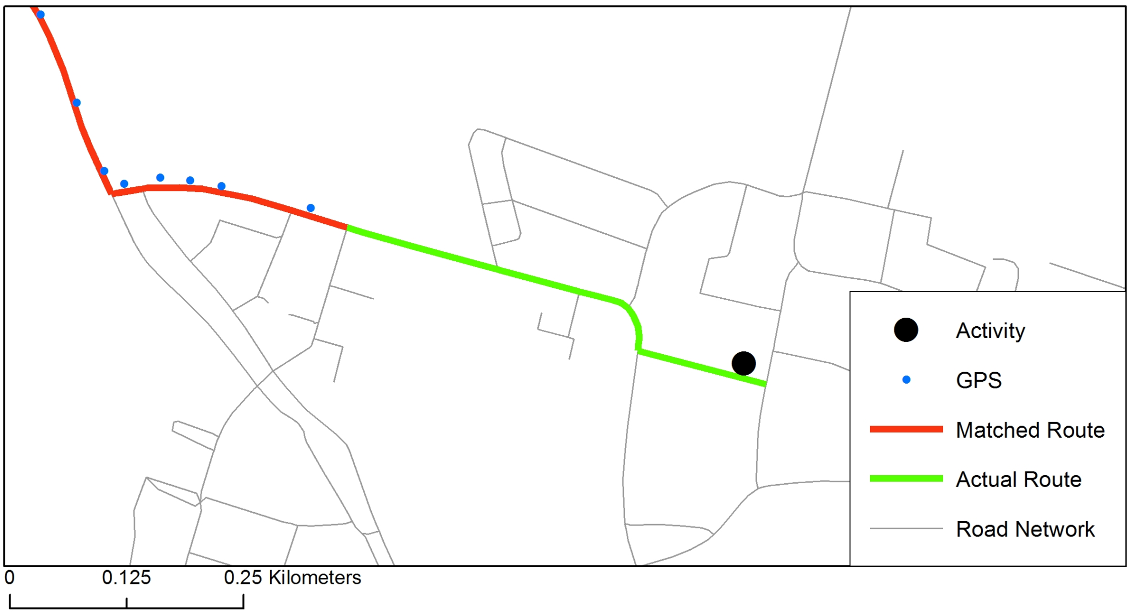
Notwithstanding the fact that the results of the procedure generally seem to be very good, a closer examination of the problems is necessary because the large number of moves, especially in the case of the Garmin, may camouflage some mistakes in the reconstructed routes. In general, the mistakes can roughly be divided into two categories: (1) fully GPS-induced mistakes that can only be solved by receiver technologies, and (2) procedural mistakes in which the procedure fails to correct for GPS errors. An example of a fully GPS-induced mistake is a cold start at the beginning of a trip stage, as illustrated in Figure 4.5. Because the road segments during the cold start are not included in the subnetwork, the 'maxi' shortest path assignment cannot exploit those segments and, therefore, does not find the complete actual route.

Table 4.4 | Percentages of accuracy: Connected subset

	Car		Bike		Walk[1]	
	% on ACT	% on PRED	% on ACT	% on PRED	% on ACT	% on PRED
Garmin	94.17	97.02	97.20	97.14	No Data	No Data
MyTracks	99.01	98.96	98.37	99.09	89.67	92.73
Trimble	89.92	98.50	100.00	100.00	No Data	No Data
TrackLog	92.69	93.94	96.26	97.79	88.75	100.00
Average	93.94	97.02	97.96	98.51	89.21	96.37

[1] Note that only a few walking trips have been recorded.

Figure 4.5 | Example of an incorrect solution with a 'maxi' flow through the subnetwork because of a cold start



Other than fully GPS-induced mistakes, procedural mistakes manifest themselves as a result of measurement errors of the GPS, which the procedure cannot account for. For instance, when as a consequence of inaccuracies in the GPS measurements the size of the subnetwork increases (Figure 4.6); due to a hit on a parallel road, the highway lane in the other direction will be incorporated into the final connected subnetwork because the 'mini' shortest path assignment made a successful attempt to connect the two consecutive points – although the correct road segment is also part of the mini subnetwork. Not only parallel routes can expand the subnetwork in unexpected ways, but also, as shown in Figure 4.7, multilevel exchanges can have this effect. The figure illustrates that the 'GPS from' point hits on the lower level of the multilevel exchange, whilst point 'GPS to' hits on the higher level, leading to unexpected road segments incorporated into the final subnetwork. Whereas these mistakes typically do not end up posing a problem, as already showed in Figure 4.4, the unfortunate combination of a few of these mistakes may result in a subnetwork that either forces the 'maxi' shortest path algorithm to take an incorrect route or creates a shorter circuit that differs from the actual route travelled. An example of this is given in Figure 4.8: because the matched route is incorporated into the subnetwork and allows for higher travel speeds, it is incorrectly identified as the travelled route. This problem is especially likely when a respondent's trajectory 'crosses' itself.

It is not only the extent of the subnetwork that can affect the quality of the reconstructed routes in the connected subset assignment procedure. The shortest path assignment itself can generate an incorrect route because of inaccuracies in the digital road network. Figure 4.9 illustrates such a procedural mistake. The actual route has a slightly higher impedance in terms of travel time than the matched route. However, because the shorter, but incorrect, route is incorporated in the connected subnetwork due to the road segments being 'hit' by a GPS measurement, the incorrect path is chosen over the actual path. Although the matched route is largely correctly reconstructed, the procedure in its current state cannot account for the errors in either measurement or road network in the situation presented. A similar situation is imaginable when the measurement frequency of the device is very low and inaccurate: the method depends on the quality of the selected subnetwork.

Figure 4.6 | Example of a 'mini' shortest path/matched route between two consecutive GPS measurements with matched network segments. GPS measurements matched on different sides of the freeway; all segments will be included in the subnetwork.

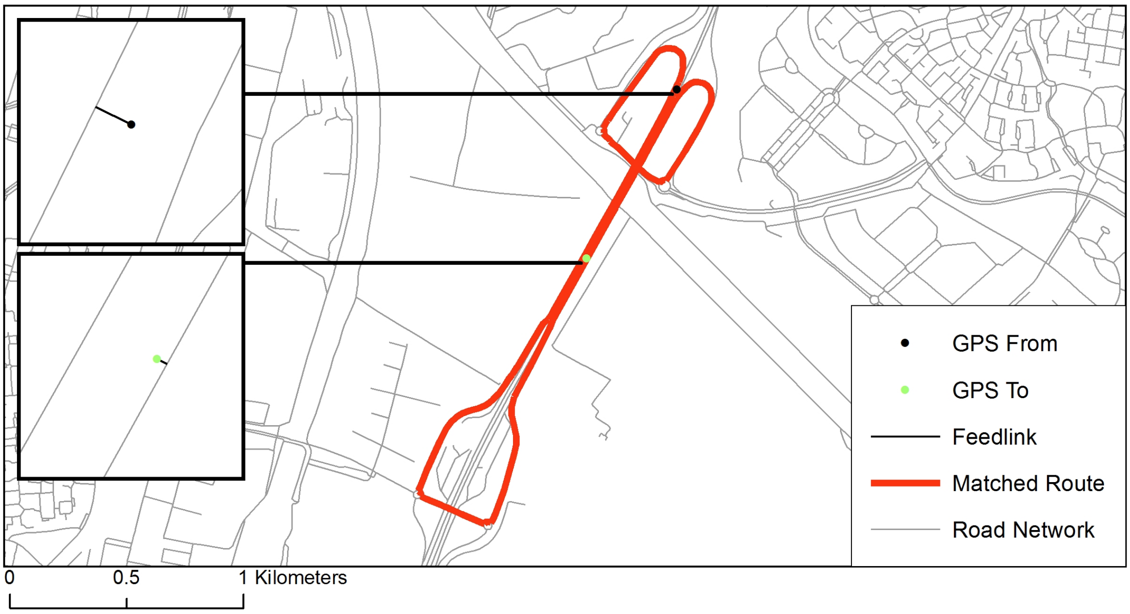


Figure 4.7 | Example of a 'mini' shortest path/matched route between two consecutive GPS measurements with matched network segments. GPS measurements matched on different parts of a multilevel interchange; all segments will be included in the subnetwork.

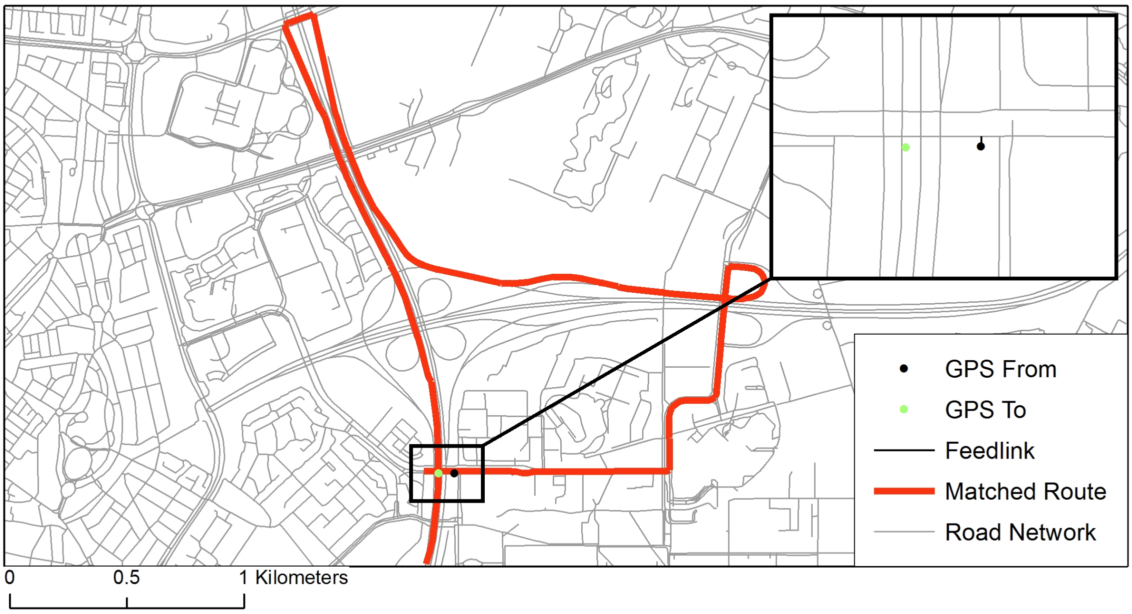


Figure 4.8 | Example of an incorrect solution with a 'maxi' flow through the subnetwork

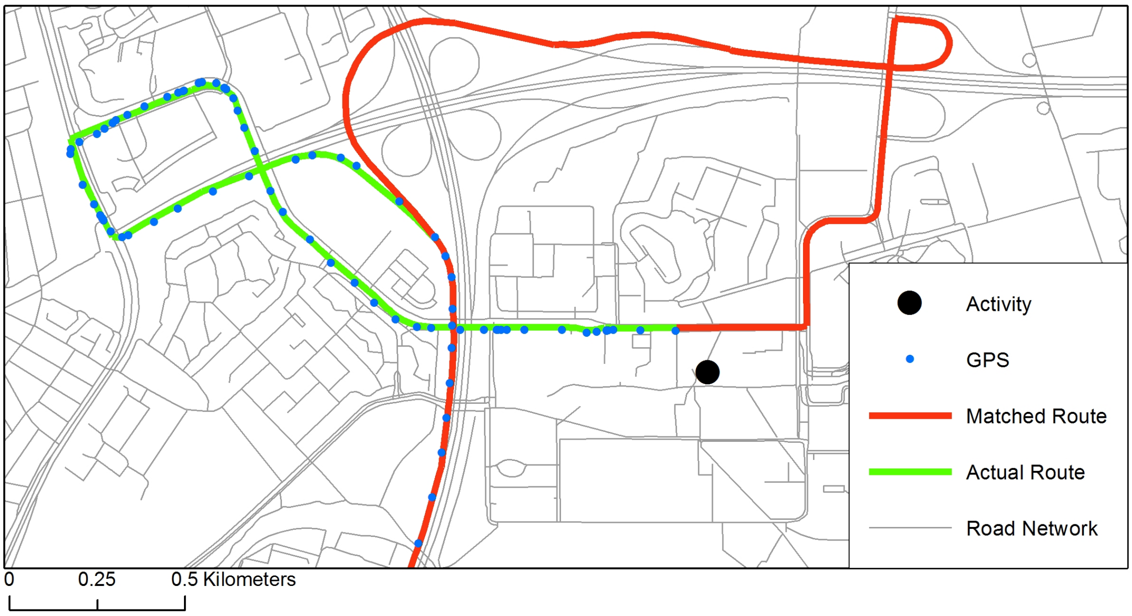


Figure 4.9 | Example of an incorrect solution with a 'maxi' flow through the subnetwork

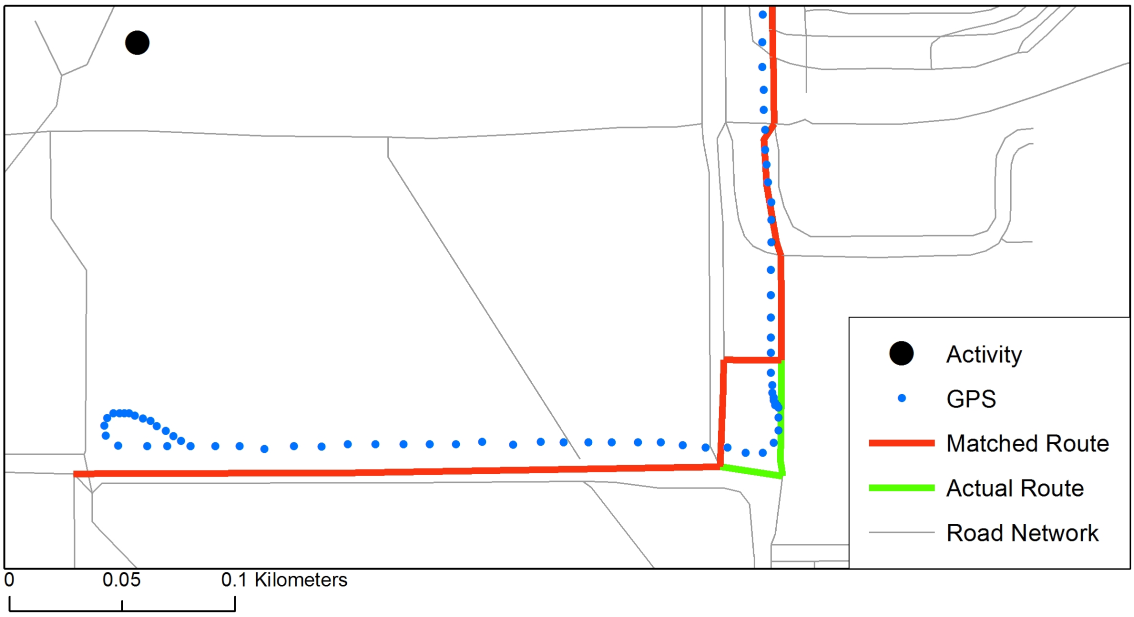
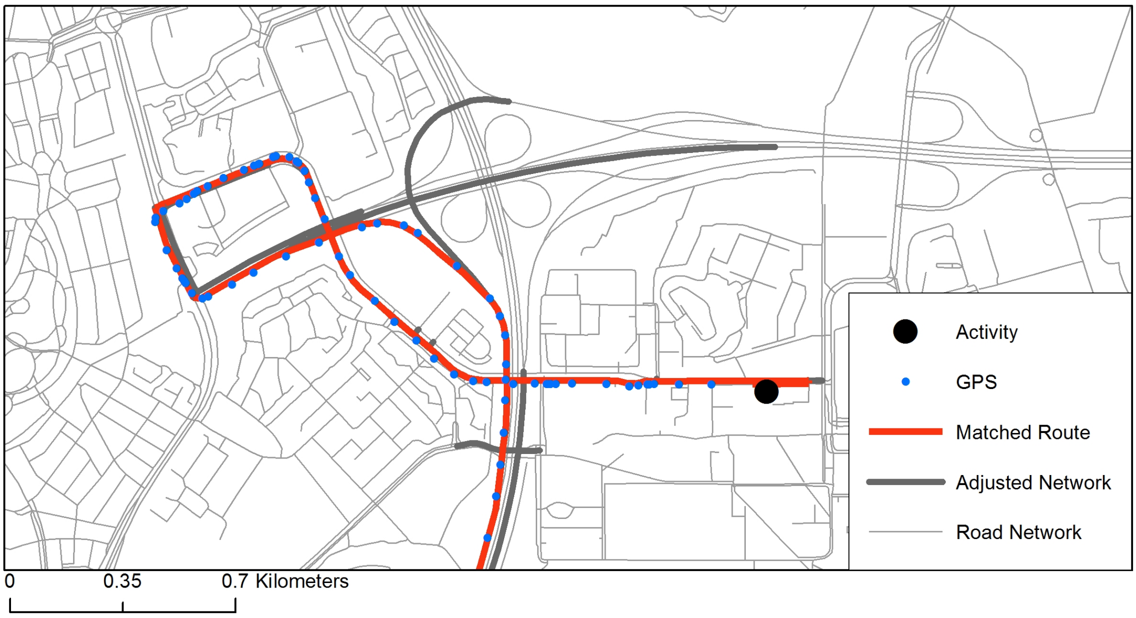


Figure 4.10 | Example of adjusted road segments by GPS hits



4.6 Map-matching GPS measurements: Impedance reduction

Before moving to the validation of the reconstructed trip stages for each device and each mode of travel, again a graphical example of a successful execution of the procedure is shown. Figure 4.10 shows a part of a raw GPS trajectory on top of the road network supplemented with the subnetwork of road segments that have undergone an impedance reduction as the first step of the impedance reduction assignment procedure. It can be seen that a few side roads are adjusted as well. In addition, some longer road segments are adjusted that extend well beyond the raw GPS trajectory. However, because the impedance is only adjusted for roads that are 'hit' by GPS measurements, the shortest path assignment does get drawn onto the actual route.

The results of the validation of the routes analyses with the impedance reduction assignment procedures are given in Table 4.5. The procedure is highly successful, as the individual scores per mode come close to a 100 percent score on both indices. In the case of MyTracks and Tracklog, even the scores for the walk mode equal 100 percent on both indices. It has to be noted that the walk moves are still hard to reconstruct because of the fact that they are not restricted to the network. This implies that for the actual routes, an approximation had to be used, i.e. the closest actual path over the network. A score of 100 percent on both indices for the walk mode thus means a perfect match between the reconstructed route and the approximation of the actual route taken. These results should be interpreted with caution. If we disregard the walk and bike moves for a moment again, and turn to the moves made by car, the indices show values well over 95 percent. However, there is a noticeable difference between the devices; especially for Tracklog, the scores on both indices are lower than the scores for the other devices. Again, the measurement frequency seems to affect the results when the number of hits does not adjust the impedance values enough to pull the matched route onto the actual route.

Table 4.5 | Percentages of accuracy: Impedance reduction

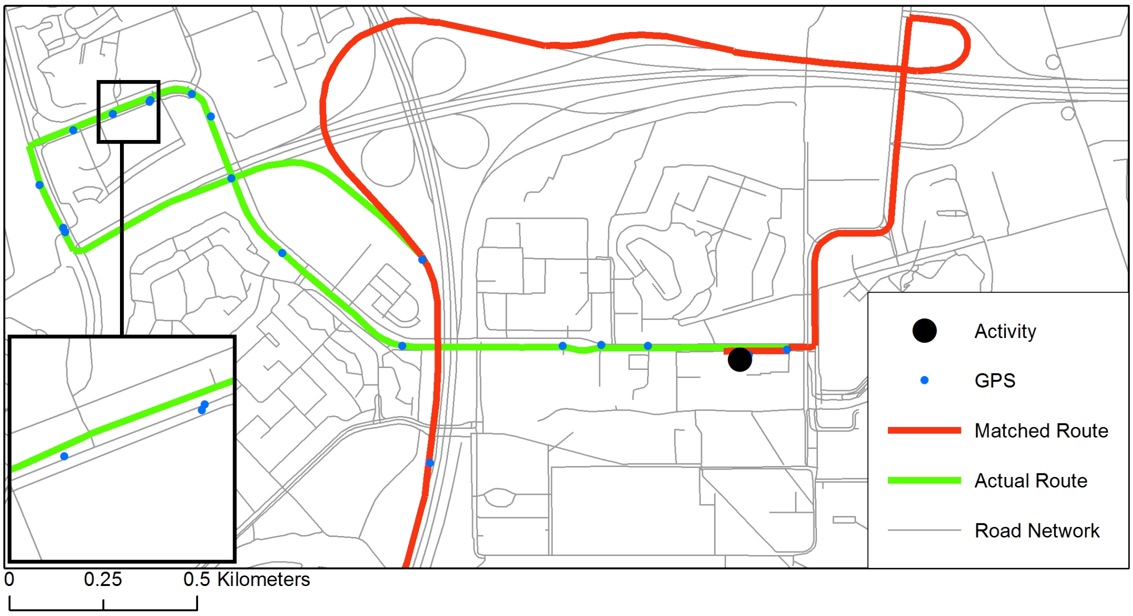
	Car		Bike		Walk[1]	
	% on ACT	% on PRED	% on ACT	% on PRED	% on ACT	% on PRED
Garmin	99.33	99.27	100.00	100.00	No Data	No Data
MyTracks	99.88	99.87	99.68	99.68	100.00	100.00
Trimble	97.34	97.72	95.51	95.51	No Data	No Data
TrackLog	94.26	95.62	96.20	96.86	100.00	100.00
Average	97.70	98.12	97.84	98.01	100.00	100.00

[1] Note that only a few walking trips have been recorded.

Although the impedance reduction algorithm seems to perform extremely well, not all indices count to a 100 percent. This again requires a closer look with respect to fully GPS-induced mistakes and procedural mistakes. Whilst the path over the complex intersection in Figure 4.10 is matched effortlessly to the actual path, the matched route as shown in Figure 4.11 is incorrect. In this case, the travel time for the shortest route, without a correction, is just over 190 seconds, whereas the travel time for the actual route is just over 253 seconds. After the impedance reduction, the travel time of the actual routes drops to 170 seconds. However, the actual shortest route also gets a minor reduction in impedance as a result of GPS hits, and decreases in travel time to 164 seconds.

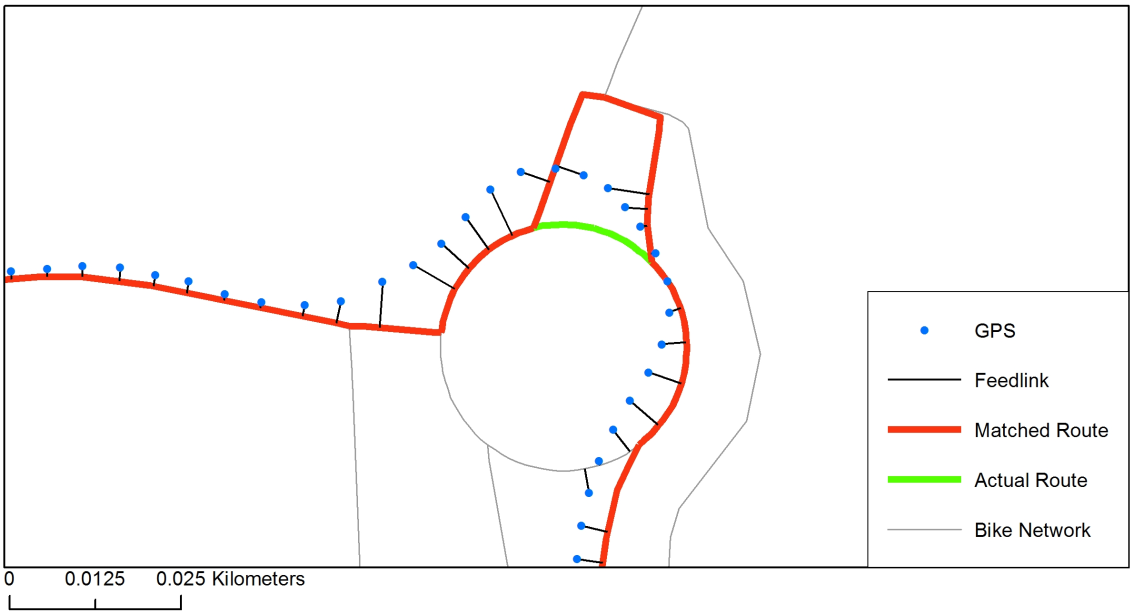
The low measurement frequency of Tracklog therefore does not compensate for the actual shorter route. This reveals that the impedance reduction procedure can lead to the same type of mistakes as the connected subset assignment procedure, especially if the measurement frequency is low. But it is not only a low measurement frequency that can affect the accuracy of the impedance reduction assignment procedure; a high measurement frequency can also generate an incorrect route as a consequence of inaccuracies in the GPS measurements or in the digital road network. Figure 4.12 illustrates such a problem: the measurements hit the road segments on different parts of the traffic circle as indicated by feed links. In turn, the matched route is being pulled off the actual route because the 'waterbed' effect draws the shortest path algorithm to the road segments with almost no impedance. Adjusting the impedance correction does not solve either of these problems. Where, for instance, increasing the impedance correction would solve the problem illustrated in Figure 4.11, it would exacerbate situations such as that in Figure 4.12.

Figure 4.11 | Example of an incorrect solution through the 'adjusted' network



	Impedance	Hits	Adjusted Impedance
Matched Route	190.8	4	164.6
Actual Route	253.3	15	170.4

Figure 4.12 | Example of an incorrect solution through the 'adjusted' subnetwork



4.7 Conclusion and discussion

Important variables in assessing the effectiveness of voluntary travel behaviour change (VTBC) interventions include the exact distances and routes travelled on an individual level. Whereas obtaining these data through paper-based survey instruments is extremely challenging, the increasing availability and ubiquity of locational data have, through map-matching procedures, considerably enhanced opportunities to accurately reconstruct travelled routes. This chapter, described two simple GIS-based post-processing map-matching methods. In addition, whilst post-processing map-matching algorithms have predominantly focused solely on home to work trips by car, the applicability of the proposed map-matching algorithms to walking and bike trips has been explored as well. Especially in the context of 'soft' transport demand management policies, in which stimulating more sustainable modes of travel is often one of the objectives, the necessity to reconstruct the bicycle and walking routes taken by individuals becomes evident.

Two procedures have been proposed that support route reconstruction in a GIS environment, using a digital road network and network attributes for route reconstruction of raw GPS trajectories without having to select predefined threshold values. In addition, the data model for transport networks in a GIS environment allows us to effortlessly take the directionality of the road-network into account. Whereas the first proposed algorithm selects a subset of the network, the second proposed algorithm adjusts the impedance attribute data. In turn, a shortest path is fitted through the subset of the network and the network with adjusted impedance attribute data, respectively. Both of the explored procedures lead to high scores on the both validation indices, especially for trips made by car.

Notwithstanding the good results, neither procedure can account for all the inaccuracies of the GPS measurements. In the case of the connected subset assignment procedure, this may surface as a consequence of a short circuit within the connected subnetwork, whilst in the impedance reduction assignment procedure, the impedance correction might not compensate for shorter, but incorrect, routes as a result of the measurement frequency. It has to be noted that in the case of the impedance reduction procedure, the major problems only come into existence when the measurement frequency is set too low; with the second highest measurement frequency of 10 seconds (by the Garmin), no large detour surfaced and the procedure could still be considered as robust. This signals that the deployment of the shortest path algorithm is not always applicable when the distance between two GPS measurements becomes too large. For future studies, a possible integration with the distance and time between two consecutive points as a proxy to assess which road segments are feasible to transverse could solve some of these issues; depending on the availability of a road network with detailed and accurate segment travel times. Another suggestion is to allow the initial matching to take more road segments into account by means of a distance-based ranking of nearby road segments, or by turning to more probabilistic methods, such as hidden Markov models, to identify candidate road segments (as opposed to the shortest path algorithm).

Both methods consist of a simple, reliable procedure to deal with some of the common challenges in post-processing map-matching using a minimum of input data. This effectiveness, combined with the fact that the operation can be executed within a GIS environment, make it easy to automate the whole procedure using, for instance, the Python scripting language, and extend it with other types of spatial analysis on individual spatial behaviour. For example, when adequate GIS data such as detailed land use and cadastral information are available, a GIS offers an attractive analytical environment to integrate route reconstruction with other analyses, such as activity diary reconstruction and trip purpose imputation. To conclude, it should be noted that this chapter was not per se intended to compete with existing state-of-the-art map-matching methods directly; it was rather intended to describe two relatively simple alternative methods which, by exploiting the capabilities of a GIS, can yield good results.

Acknowledgements

The financial assistance of the South African National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF. The sponsor did not play a role in executing this study. The authors also acknowledge Utrecht University for providing the GPS data sets used for the development and testing of the map-matching algorithms, as well as Maarten Zeylmans van Emmichoven for his enthusiasm.

References

Bierlaire, M., Chen, J. & Newman, J. 2013. A probabilistic map matching method for smartphone GPS data. *Transportation Research Part C: Emerging Technologies*. 26:78–98.

Blazquez, C.A. & Miranda, P.A. 2014. A real time topological map matching methodology for GPS/GIS-based travel behavior studies. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 152–170.

Bonsall, P. 2009. Do we know whether personal travel planning really works? *Transport Policy*. 16(6):306–314.

Brög, W., Erl, E., Ker, I., Ryle, J. & Wall, R. 2009. Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*. 16(6):281–292.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbridge, A. & Goodwin, P. 2008. Smarter choices: Assessing the potential to achieve traffic reductions using 'soft measures'. *Transport Reviews*. 28(5):593–618.

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England.

Transport Policy. 16(6):293–305.

Chen, J. & Bierlaire, M. 2015. Probabilistic multimodal map matching with rich smartphone data. *Journal of Intelligent Transportation Systems*. 19(2):134–148.

Chung, E.-H. & Shalaby, A. 2005. A trip reconstruction tool for GPS-based personal travel surveys. *Transportation Planning and Technology*. 28(5):381–401.

Dalumpines, R. & Scott, D. 2011. GIS-based map-matching: development and demonstration of a postprocessing map-matching algorithm for transportation research. In Vol. 1. S. Geertman, W. Reinhardt, & F. Toppen (eds.). (Lecture notes in geoinformation and cartography). *Advancing geoinformation science for a changing world*. Berlin, Heidelberg: Springer. 101–120.

Dijkstra, E.W. 1959. A note on two problems in connexion with graphs. *Numerische Mathematik*. 1(1):269–271.

Geertman, S., De Jong, T. & Wessels, C. 2003. Flowmap: A support tool for strategic network analysis. In S. Geertman & J. Stillwell (eds.). (Advances in spatial science). *Planning support systems in practice*. Berlin, Heidelberg: Springer. 155–175.

Hashemi, M. & Karimi, H.A. 2014. A critical review of real-time map-matching algorithms : Current issues and future directions. *Computers, Environment and Urban Systems*. 48:153–165.

International Transport Forum. 2015. *Big data and transport: Understanding and assessing options*. Paris: OECD / ITF. [Online], Available: http://www.internationaltransportforum.org/cpb/projects/mobility-data.html.

Krumm, J., Letchner, J. & Horvitz, E. 2006. Map matching with travel time constraints. In *Proceedings of the SAE 2007 World Congress*. Detroit: SAE International. 1–7.

Krumm, J., Gruen, R. & Delling, D. 2013. From destination prediction to route prediction. *Journal of Location Based Services*. 7(2):98–120.

Krygsman, S., Nel, J.H. & de Jong, T. 2008. The use of cellphone technology in activity and travel data collection in developing countries. In *Proceedings of the 8th International Conference on Transport Survey Methods*. Annecy, France.

Marchal, F., Hackney, J. & Axhausen, K. 2005. Efficient map matching of large Global Positioning System data sets: Tests on speed-monitoring experiment in Zürich. *Transportation Research Record*. 1935(1):93–100.

Miwa, T., Kiuchi, D., Yamamoto, T. & Morikawa, T. 2012. Development of map matching algorithm for low frequency probe data. *Transportation Research Part C: Emerging Technologies*. 22:132–145.

Montini, L., Rieser-Schüssler, N., Horni, A. & Axhausen, K.W. 2014. Trip purpose identification from GPS tracks. *Transportation Research Record: Journal of the Transportation Research Board*. 2405:16–23.

Papinski, D. & Scott, D.M. 2011. A GIS-based toolkit for route choice analysis. *Journal of Transport Geography*. 19(3):434–442.

Quddus, M., Noland, R. & Ochieng, W. 2006. A high accuracy fuzzy logic based map matching algorithm for road transport. *Journal of Intelligent Transportation Systems: Technology, Planning, and Operations*. 10(3):103–115.

Quddus, M. a., Ochieng, W.Y. & Noland, R.B. 2007. Current map-matching algorithms for transport applications: State-of-the art and future research directions. *Transportation Research Part C: Emerging Technologies*. 15(5):312–328.

Schuessler, N. & Axhausen, K.W. 2009a. Processing raw data from Global Positioning Systems without additional information. *Transportation Research Record: Journal of the Transportation Research Board*. 2105:28–36.

Schuessler, N. & Axhausen, K.W. 2009b. *Map-matching of GPS traces on high-resolution navigation networks using the Multiple Hypothesis Technique (MHT)*. Institute for Transport Planning and System (IVT), ETH Zurich, Zurich.

Shen, L. & Stopher, P.R. 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*. 34(3):316–334.

Stopher, P.R., Clifford, E., Swann, N. & Zhang, Y. 2009. Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy*. 16(6):315–324.

Taylor, M. & Ampt, E. 2003. Travelling smarter down under: Policies for voluntary travel behaviour change in Australia. *Transport Policy*. 10(3):165–177.

Velaga, N.R., Quddus, M. a. & Bristow, A.L. 2009. Developing an enhanced weight-based topological map-matching algorithm for intelligent transport systems. *Transportation Research Part C: Emerging Technologies*. 17(6):672–683.

White, C.E., Bernstein, D. & Kornhauser, A.L. 2000. Some map matching algorithms for personal navigation assistants. *Transportation Research Part C: Emerging Technologies*. 8(1–6):91–108.

Yuan, J., Zheng, Y., Zhang, C., Xie, X. & Sun, G.-Z. 2010. An interactive-voting based map matching algorithm. In *Proceedings of the 11th International Conference on Mobile Data Management*. Washington, DC, USA: IEEE Computer Society. 43–52. [Online], Available: http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=5489808 [2015, October 11].

Zhang, Y., Stopher, P. & Halling, B. 2013. Evaluation of south-Australia's TravelSmart project: Changes in community's attitudes to travel. *Transport Policy*. 26:15–22.

Part IV
Empirical results

Whereas the decades-long experience with household travel surveys typically depended precariously on the vagaries of human memory, agent's (...) movements and activities can now be tracked in great temporal and spatial detail.

Cottril *et al.* (2013: 59)

Chapter 5. Gathering individual travel data with GPS-enabled smartphones: A proof of concept study

Van Dijk, J.T. & Krygsman, S.C., 2015, Gathering individual travel data with GPS-enabled smartphones: A proof of concept study. In: *Proceedings of the 34th Southern African Transport Conference (SATC 2015).* Pretoria: Southern Africa Transport Conference. 448-460.

Abstract

Policies aimed at travel demand management (TDM) rely heavily on the gathering of accurate individual level activity and travel data to understand and unpack the demand for transport. Moreover, self-report based data collection methods face problems such as a high-respondent burden and inaccuracies in the number and duration of the reported trips. On that account, this chapter presents ongoing research to assess the reliability and feasibility of passively collecting high resolution spatiotemporal data on activity and travel behaviour using GPS-enabled smartphones. A small-scale pilot study was conducted in which respondents from the University of Stellenbosch, South Africa, were passively tracked for the course of two days by means of a purposefully designed smartphone application, termed Tracklog. The results of the small experiment indicate that while GPS technology in smartphones potentially holds a number of benefits for collecting activity and travel data, the technology is not without problems. This project has highlighted that these problems can be classified as: (1) user, (2) technology, and (3) methodology related problems. Notwithstanding these problems, the results indicate that gathering high resolution space-time data by means of GPS-enabled smartphones is feasible and that it opens doors to a range of possible applications that are unattainable by traditional survey methods.

Keywords

GPS; Smartphones; Spatiotemporal data collection; Geographic Information Systems

5.1 Introduction

Urban transport systems and road networks are under pressure as a result of a rapid increase in private vehicle ownership and increasingly complex and fragmented travel patterns (Axhausen, Zimmermann, Schönfelder, Rindsfüser & Haupt, 2002; Mokhtarian, Salomon & Handy, 2006; Järv, Ahas & Witlox, 2014). The highly visible externalities of these modern societal trends include daily congestion and traffic gridlock, but also environmental degradation. In 2003, road transport accounted for 18 percent of the worldwide CO_2 emissions. This dependency on road transport poses major challenges to the environment (Chapman, 2007; Graham-Rowe, Skippon, Gardner & Abraham, 2011; Finn, 2012). While developed countries are still responsible for 70 percent of the worldwide transport related greenhouse emissions, the relative contribution of upcoming economies such as China, India and South Africa is expected to grow significantly in the next decades (Hickman & Banister, 2010).

To address these challenges, policy makers and researchers have started to shift their attention to innovative measures to curb the demand for car use with travel demand management (TDM) policies (Kitamura, Fujii & Pas, 1997; Loukopoulos, Jakobsson, Gärling, Schneider & Fujii, 2004; Stradling & Anable, 2008). Also in South Africa the focus has shifted from supply-side to demand-side passenger transport planning (Behrens & Del Mistro, 2010). This is exemplified by the National Land Transport Act (Act 5 of 2009), in which it is stated that municipalities have to formulate and implement TDM techniques in their transport planning. Travel demand management, however, is not an undemanding task, as it:

> (…) relies on the agencies' ability to accurately predict future demand, suggest viable improvements and shift the demand away from the automobile to more sustainable transportation modes like walking, biking and transit (Jariyasunant, Carrel, Ekambaram, Gaker, Kote, Sengupta & Walker, 2011: 3).

In order to understand travel behaviour and accurately predict future demand, precise and reliable activity and travel data are essential (Shafique & Hato, 2015). However, obtaining disaggregated travel data through traditional methods such as paper-based survey instruments is a complex endeavour as a result of, amongst other reasons, a high respondent burden, the underreporting of trips, costly administrative processes and the variability of human travel behaviour (Behrens, 2004; Stopher, Clifford, Swann & Zhang, 2009; Jariyasunant *et al.*, 2011; Nitsche, Widhalm, Breuss, Brändle & Maurer, 2014). Although more cost-effective, computer-assisted methods do not solve all the limitations associated with traditional methods either, because all of these methods are contingent on the respondent's ability to accurately remember his or her movements and activities (Shafique & Hato, 2015). This makes administering and collecting travel surveys a challenging exercise, even more so for longitudinal studies (Behrens & Del Mistro, 2010).

Over the past decade, technological developments have advanced the state of the research with regard to travel surveys (Cottrill *et al.*, 2013). In particular, location-aware technologies (LAT) such as Global Positioning Systems (GPS) have greatly enhanced the opportunity to collect more accurate data on human spatiotemporal behaviour for longer periods of time (Shen & Stopher, 2014). These technologies are nowadays also available on most smartphones. The advancements in the technological realm in combination with the need for high-quality data on spatiotemporal human behaviour for transport management and planning, offers challenging new avenues for researchers and practitioners. This chapter describes ongoing efforts to assess the reliability and feasibility of smartphone tracking, in the context of South Africa, to suggest ways in which location-aware technology can be employed in collecting data on individual activity and travel behaviour.

5.2 Literature review

In order to capture, understand, predict and possibly manage travel behaviour, high-quality activity and travel data on an individual level is essential. Given the complexities associated with gathering disaggregated data, researchers have started to experiment with acquiring these data by means of new technologies such as mobile phones (cf. Asakura & Hato, 2004; Krygsman & Schmitz, 2005; Krygsman, Nel & de Jong, 2008), mobile phone call detail records (cf. Järv *et al.*, 2014), and Bluetooth technology (see for a recent overview of mobile technologies used in activity-travel data collection Rasouli & Timmermans, 2014). Increasingly, starting in the late 1990s, GPS technology and, more recently, GPS-equipped smartphone technology have been put forward as a solution in supplementing or (partly) substituting traditional travel data collection methods (see for instance Wolf, 2000; Bohte & Maat, 2009; Chen & Kwan, 2012; Nitsche, Widhalm, Breuss & Maurer, 2012; Nitsche *et al.*, 2014; Feng & Timmermans, 2013; Shen & Stopher, 2013, 2014; Shoval, Kwan, Reinau & Harder, 2014).

GPS technology provides opportunities to collect data that is otherwise difficult to acquire such as the exact start and end time of a trip, the chosen route and the distance travelled (Blazquez & Miranda, 2014). In addition, whereas paper-based travel surveys can be costly and put a significant cognitive burden on respondents (see for instance Behrens, 2004), GPS technology has the potential to reduce respondent burden whilst increasing the spatiotemporal accuracy of the collected data (Du & Aultman-Hall, 2007). As Cottril *et al.* (2013: 59) note:

> (...) whereas the decades-long experience with household travel surveys typically depended precariously on the vagaries of human memory, agents' (...) movements and activities can now be tracked in great temporal and spatial detail.

The increasing availability and ubiquity of smartphones make it even easier for researchers and policy makers to take advantage of GPS technology.

One of the major advantages of collecting locational data by means of smartphones when compared to dedicated GPS-devices and other mobile technologies, emanates from the fact that:

> [p]eople habitually carry their mobile phones with them much of the time as this pervasive technology offers its users to a means of constant and available communication as well as personal entertainment (Horanont, Phithakkitnukoon, Leong, Sekimoto & Shibasaki, 2013: 1)

Several studies have thus far begun to examine the use of smartphones in collecting data on activity and travel behaviour (e.g. Bierlaire, Chen & Newman, 2010; Nitsche *et al.*, 2012, 2014). Some researchers (for example Froehlich, Dillahunt, Klasnja, Mankoff, Consolvo, Harrison & Landay, 2009) provided their participants with a smartphone capable of tracking their movements, whilst others have experimented with delivering a smartphone application to participants who possess a smartphone (for instance Jianchuan, Zhicai, Guangnian & Xuemei, 2014).

Although there seems to be agreement that using mobile devices for the collection of activity-travel data is viable, a number of challenges are outlined for future studies. Not only is the privacy of the respondents often mentioned in this context, but also technological and methodological issues have been raised with regard to the level of battery consumption and the management of large data sets (Rasouli & Timmermans, 2014; Shen & Stopher, 2014). In addition, for GPS-based technologies to fully replace traditional travel surveys, analytical methods should be able to deal with imperfect locational information – especially because the signal of GPS receivers in smartphones can be blocked or weakened when the phone is carried in a purse or a pocket (Bierlaire, Chen & Newman, 2013).

5.3 Collecting and analysing GPS tracks

In August 2014, an experiment was conducted at Stellenbosch University which aimed to assess the reliability and feasibility of GPS-enabled smartphones in collecting data on transport and activity behaviour. An application named Tracklog was developed for the Android mobile operating system.[1] The choice of the Android operating system was based on the outcome of a survey conducted at Stellenbosch University in 2013, which indicated that the market share of Android was estimated to be 31% in 2013 and projected to increase to 47% in 2014 (Stellenbosch University, 2013). Upon activation, Tracklog allowed the recording of its user's positional information using the GPS location registered by the smartphone. This information included the geographical coordinates (*latitude, longitude*) and the time (*t*) of each measurement.

The recorded information was uploaded via the GSM network to a server from where the information could be downloaded by the researchers. A location measurement frequency of 30 seconds was set for the data collection, leading to the potential recording of 2,880 records per 24 hours. These records resulted in a database file of approximately 220 KB. Uploading these data over the cellular network did therefore not lead to high data costs, as on average the cost for 1 MB of data does not exceed R1. The uploaded information could also be accessed by the participants: a feedback website was created that allowed participants to log into a web interface where they could view their personal tracks projected onto a Google Maps background.

Participant recruitment was done by means of an e-mail invitation that was sent out to 100 individuals at Stellenbosch University, including post-graduate students, academics and support staff. The invitations yielded a non-representative sample of 15 individuals (6 females and 9 males). After accepting the invitation, the participants received a follow-up e-mail which included an informed consent form, the procedures of the experiment, and detailed instructions on how to install Tracklog. Because the project was reviewed by the Departmental Research Ethics Committee, participants were required to accept the terms and conditions of the informed consent form. Only when their consent was received, was he or she provided with a link to download Tracklog and requested to switch the application on for a period of two days. In order to remind the participants to switch on the application, every participant was sent an automated text message on both days of the experiment.

Data from the phones was automatically uploaded every hour to the servers from where the data could be downloaded in a database format (.csv). Using the Python scripting language, the spatiotemporal measurements of each individual were imported and projected in ArcGIS 10.2 to visualise the locational information. IBM SPSS Statistics version 22 was used to visualise the spatiotemporal trajectory.

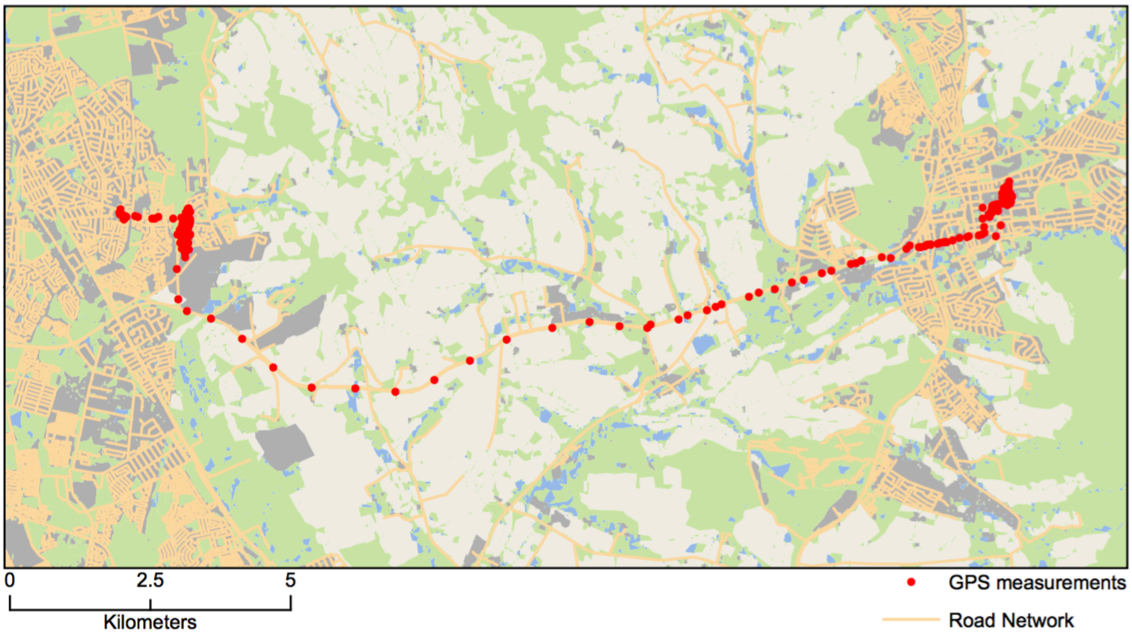
5.4 Reliability and feasibility of smartphone tracking

After the two days of tracking, the locational information was downloaded from the server and merged in chronological order. This way, an automatic time relationship was created between the records that represent the switching off and switching on of the device, be it either on purpose or as a result of other issues. The measurements were subsequently imported and

[1] At this stage Tracklog was only available for Android. In later stages of the project, the application was also made available for smartphones running iOS.

projected into a GIS environment (Hartebeesthoek LO19) – an example of which is given in Figure 5.1. It can be seen from the figure that there are both *clusters* of points, indicative of an activity, and individual *points*, indicative of travel. Taking the time of the occurring clusters into account, the clusters are indicative of a place of work in Stellenbosch (during office hours) and a place of residence in Kuilsrivier (outside office hours), respectively. In turn, the distance between two subsequent points indicative of travel, gives an indication of the speed with which was travelled.

Figure 5.1 | Output Tracklog – Projected GPS locations



Although elements of travel and activity patterns can be discerned from a simple visual inspection, a number of issues occurred during the tracking period. First, the automatic upload did not function as designed. Some phones experienced connectivity problems when switching from the Wi-Fi network to the cellular network and vice versa. Second, a number of participants reported that they could initially not log onto the Tracklog application, which withheld them from starting to record their locations. Third, it was found that the location (GPS) services of the smartphones of some users were not activated or that the smartphone's settings restricted background data transfer. This resulted in no locational information being recorded or uploaded at all. As a consequence of these technical issues, only 20 tracks (from 11 individuals) were suitable for further analysis.

For these 11 individuals, Table 5.1 presents a basic overview of the collected data. The last two columns of Table 5.1 refer to the number of records that were captured. In general the participants did not record their location for the full 48-hour period, but switched the application off during the evening. In order to give an indication of the quality of the data, the percentage of records that could have been recorded was calculated for each participant. The percentages are relative to the time the participant initiated the tracking and terminated the tracking, respectively – whether this was on purpose or as a consequence of a technical failure. A value of 100%

indicates that the phone continuously recorded locational information with a measurement frequency of 30 seconds between the starting and ending of Tracklog.

Table 5.1 | Overview of collected data

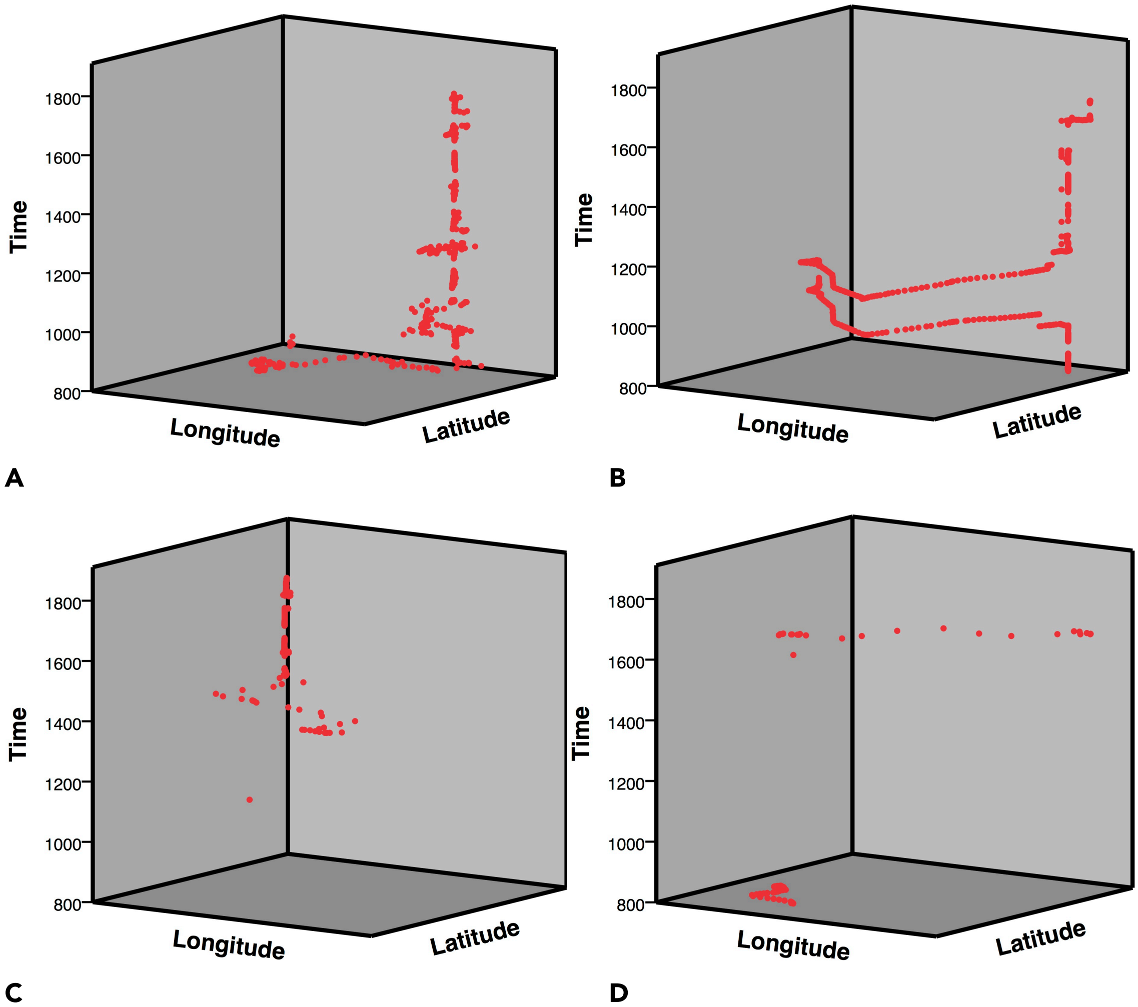
#	Phone Type	Android	Records Day 1		Records Day 2	
1	Samsung Galaxy S4 Mini	4.2.2	1259	95%	394	30%
2	Samsung Galaxy S4	4.2.2	No Data	No Data	460	87%
3	Motorola Moto E	4.4.4	1793	97%	1250	86%
4	Samsung Galaxy Pocket	2.3.6	653	37%	1007	47%
5	Samsung Galaxy S4	4.4.2	1320	125%	2572	124%
6	Samsung Galaxy S3 Mini	4.1.2	749	45%	170	15%
7	Sony Xperia Z1	4.4.4	463	69%	746	71%
8	Samsung Galaxy S4	4.2.2	No Data	No Data	827	62%
9	Samsung Galaxy S5	4.2.2	323	26%	76	9%
10	Samsung Galaxy S3	4.3.0	190	14%	50	6%
11	Samsung Galaxy S4	4.2.2	1268	65%	123	60%

In one instance in Table 5.1, the captured records exceed 100%. This implies that in this case the measurement frequency was lower than was intended. This may be attributed to the fact that Tracklog was set to synchronise its measurement frequency with the settings on the server. However, given the connectivity problems alluded to earlier, it is likely that Tracklog used a different measurement frequency instead. In other cases, it turned out that the measurement frequency was actually longer than 30 seconds. This most likely implies that when the smartphone did not have a GPS fix at the moment Tracklog requested it, the application recorded the location the moment the signal was available again. As noted by Blazquez and Miranda (2014: 154), "a validation of the GPS measurements is needed due to the data gaps caused by respondents forgetting their GPS device or satellite signal blockage".

The impact of a missing GPS fix should not be underestimated. Following Hägerstrand's conceptualisation of a space-time path (Hägerstrand, 1970), the daily life of an individual is embedded in a set of complex spatial and temporal arrangements that determine an individual's opportunities to engage in activities. This constitutes travel as a derived demand. As Pas (1980: 4) writes: "If all activities in which an individual wished to participate were located at the same place, that individual would be expected to undertake little or no travel at all." Whilst a number of missing data points might not be directly visible on a Cartesian plane, particularly when they occur at the same geographical location, they do become visible in a 3D representation of the participants' movements with time as a third dimension. The accuracy of the collected data therefore also depends on the temporal interval of the measurements. This is illustrated by Figure 5.2 in which the space-time paths, between 08h00 and 18h00, of four participants are visualised.

From the graphs in Figure 5.2 it becomes apparent that the temporal dimension of the data quality differs significantly between the four individuals. Whereas graphs A and B show a mostly uninterrupted path through time and space, graphs C and D show a dispersed set of recorded data points. When we turn to graph C, there is only one measurement around 11h30, and a continuous stream of measurements from 13h30 until 18h00. In graph D, on the other hand, most of the data points during the day are missing. This is problematic, because the activity locations the participant might have visited during those hours are not captured. In fact, if one would purely look at the geographical location, it seems that the individual did not engage in any travel during the day. If it were not for the time element, this measurement gap could go unnoticed. In turn, this could lead to incorrect conclusions with regard to his or her travel behaviour.

Figure 5.2 | Spatiotemporal representation of four GPS tracks (08h00 – 18h00)



Although not necessarily problematic with regard to determining an individual's activity pattern, graphs A and B also manifest some issues. In graph A, for instance, it is very clear that there is noise in the GPS measurements. This is a problem inherent to GPS technology: even when an individual is not moving, the GPS might record a change in location as a consequence of the configuration of the GPS satellites or a signal being reflected by an obstacle such as a building; a problem typically referred to as positional drift. Graph B also shows a number of gaps that can most likely be ascribed to signal blockage. This implies that the reliability of the captured locations should be critically reviewed and data heuristics should be employed to tease out the most likely activity patterns.

In the final part of the experiment, a small exit survey was distributed amongst the participants to capture their experience with the tracking process and the functioning of the application. The participants who responded to this exit survey (n = 7) were generally positive. One participant commented: "*I actually enjoyed taking part in the tracking and would certainly take part in any future studies if required.*" Privacy concerns did not seem to be an issue either. All of the participants (n = 5) who responded to this item, stated that they did not experience an intrusion of their privacy or a discomfort with regard to sharing their locational information.

All respondents (n = 7) did express a concern with regard to battery usage, and stated that they encountered a significant impact on their device's battery life; something often reported in the literature. This concern is not unfounded: in a battery consumption test, it was found that the GPS services that Tracklog requires indeed use considerable amounts of battery power. Using a Motorola Moto E as testing device, it was found that if the phone was put on standby for 24 hours, the battery drained to 71 percent of its capacity. If the phone was put on standby, but with Tracklog running, the battery drained to 33 percent of its capacity.[2] It should be noted that these numbers vary with the type of phone, user settings and usage characteristics. Notwithstanding this inconvenience, the majority (n = 6) stated that the Tracklog application did not interfere with the normal usage of their smartphone devices.

5.5 Conclusion and discussion: Lessons learnt

This work described ongoing efforts to assess the reliability and feasibility of smartphone tracking, in the context of South Africa, to suggest ways in which location-aware technology can be employed in collecting data on individual activity and travel behaviour. The results of this small-scale pilot study indicate that while GPS technology in smartphones potentially holds a number of benefits for collecting activity and travel data, the technology is not without problems. This project has highlighted that these problems can be classified as: (1) user, (2) technology, and (3) methodology related problems.

Despite the fact that the application that was used seemed to be quite well-received by the participants, a number of users struggled with installing and activating the application on their smartphone. Some users are thus more smartphone-literate than others, and care should be taken when designing future experiments. A related problem is that users tend to switch off Tracklog when they are home, which increases the chance that they will forget to switch it back on when they leave their homes again. The incorporation of, for instance, an 'active tracking' function could solve this problem by signalling the researchers that a user has forgotten to switch on their application. Another option currently being experimented with, is the usage of Near Field Communication (NFC) tags that automatically switch on the application. If one were to place an NFC tag on, for instance, a smartphone car mount, the application would not only switch on automatically, but it would also automatically record the exact starting time of a car trip.

While none of the participants in this study expressed concerns with regard to their privacy, this does not imply no person would have these concerns. In an attempt to mitigate these

---

[2] Motorola XT1021 running Android 4.4.4 with a Lithium-ion battery with a capacity of 1980 mAh. Wi-Fi disabled. No other functions of the phone were being used during the test.

concerns it is therefore recommended that a research project involving GPS technology should come with a clear description of what happens to the locational information being recorded and on who has access to this information. A concern that did emerge amongst the participants was related to the relatively high impact on their devices' battery levels. This should not just be regarded as a trivial complication; participants may actually drop out of the study to avoid their smartphone running out of battery. The battery consumption is, however, not solely a user related problem, but also technology related. Whereas it is hard to control the battery consumption of location services in general, the literature suggests the usage of additional sensors in order to capture location data more energy-efficiently – for instance, by only recording a location when the user is actually on the move.

Technological problems not only surfaced on the data collection side, but also during data analysis. In a small experiment such as this, with only 15,000 measurements, the number of measurements is manageable; but a large-scale study on the activity and travel behaviour of, for instance, 500 individuals for the course of five days with a measurement frequency of 30 seconds, could potentially lead to a data set comprising of 7.2 million measurements. When analysing such large numbers, one moves into the realm of big data analysis and thus significant requirements with regard to computing power. In addition, the automatisation of data cleaning processes, data heuristics and data analysis becomes imperative.

Aside from methodological problems with data analysis, methodological problems emerged furthermore from the measurement instrument itself. As only persons with a smartphone were eligible to participate, this leads to questions regarding the representativeness of the collected data. This is also related to different smartphone markets. In the context of this study, it meant that only individuals with Android-based operating systems qualified to participate. This not only requires additional attention to the research design, but might also call for the segmentation of the sample population, whereby different groups within a population are targeted with different data collection instruments. These issues should be addressed before smartphone technology can be used as a large-scale data collection technique.

Notwithstanding these problems, the collection of high-accuracy data using smartphones to address a number of the requirements imposed by activity-based analysis and spatiotemporal travel behaviour seems promising; both given the possibility of collecting data that would have been otherwise extremely difficult to acquire, and because of the spatiotemporal detail with which locational information is recorded. Sophisticated tracking systems may not be the perfect way of gathering truly reliable origin-destination data, if that is possible at all, but not only can they improve accuracy, they also open a whole new range of possible applications that were previously unattainable. One can think of accurately capturing both the financial and environmental costs of individual travel, precise route reconstruction that can be used in the calibration of traffic models, longitudinal studies with a decreased burden on the respondent, the measurement of behavioural changes that require a high level of spatial detail, a reduction in the reliance on retrospective surveys (and their associated inadequacies), increased access to hard-to-reach groups, and a cost reduction in large-scale travel surveys. Although GPS-equipped smartphones may not be a perfect instrument for measuring activity and travel behaviour yet, they are definitely part of the future of transport research.

Acknowledgements

The financial assistance of the South African National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF. The authors would also like to thank the two reviewers for their valuable comments.

References

Asakura, Y. & Hato, E. 2004. Tracking survey for individual travel behaviour using mobile communication instruments. *Transportation Research Part C: Emerging Technologies*. 12(3–4):273–291.

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

Behrens, R. 2004. Understanding travel needs of the poor: Towards improved travel analysis practices in South Africa. *Transport Reviews*. 24(3):317–336.

Behrens, R. & Del Mistro, R. 2010. Shocking habits: Methodological issues in analyzing changing personal travel behavior over time. *International Journal of Sustainable Transportation*. 4(5):253–271.

Bierlaire, M., Chen, J. & Newman, J. 2010. *Modeling route choice behavior from smartphone GPS data*. Lausanne: Transport and Mobility Laboratory. School of Architecture, Civel and Environmental Engineering. Ecole Polytechnique Fédérale de Lausanne.

Bierlaire, M., Chen, J. & Newman, J. 2013. A probabilistic map matching method for smartphone GPS data. *Transportation Research Part C: Emerging Technologies*. 26:78–98.

Blazquez, C.A. & Miranda, P.A. 2014. A real time topological map matching methodology for GPS/GIS-based travel behavior studies. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 152–170.

Bohte, W. & Maat, K. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*. 17(3):285–297.

Chapman, L. 2007. Transport and climate change: A review. *Journal of Transport Geography*. 15(5):354–367.

Chen, X. & Kwan, M.-P. 2012. Choice set formation with multiple flexible activities under space–time constraints. *International Journal of Geographical Information Science*. 26(5):941–961.

Cottrill, C.D., Pereira, F.C., Zhao, F., Dias, I., Lim, H.B., Ben-Akiva, M. & Zegras, P. 2013. Future Mobility Survey - Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board*. 2354:59–67.

Du, J. & Aultman-Hall, L. 2007. Increasing the accuracy of trip rate information from passive multi-day GPS travel datasets: Automatic trip end identification issues. *Transportation Research Part A: Policy and Practice*. 41(3):220–232.

Feng, T. & Timmermans, H.J.P. 2013. Transportation mode recognition using GPS and accelerometer data. *Transportation Research Part C: Emerging Technologies*. 37:118–130.

Finn, B. 2012. Towards large-scale flexible transport services: A practical perspective from the domain of paratransit. *Research in Transportation Business and Management*. 3:39–49.

Froehlich, J., Dillahunt, T., Klasnja, P., Mankoff, J., Consolvo, S., Harrison, B. & Landay, J.A. 2009. UbiGreen: Investigating a mobile tool for tracking and supporting green transportatoin habits. In *Proceedings of the 27th international conference on human factors in computing systems - CHI 09*. New York: ACM Press. 1043–1052.

Graham-Rowe, E., Skippon, S., Gardner, B. & Abraham, C. 2011. Can we reduce car use and, if so, how? A review of available evidence. *Transportation Research Part A: Policy and Practice*. 45(5):401–418.

Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional Science Association*. 24(1):6–21.

Hickman, R. & Banister, D. 2010. Low-carbon transport in a developed megalopolis: The case of London. In W. Rothengatter, Y. Hayashi, & S. Wolfgang (eds.). *Transport moving to climate intelligence: New chances for controlling climate impacts of transport after the economic crisis*. New York: Springer. 41–52.

Horanont, T., Phithakkitnukoon, S., Leong, T.W., Sekimoto, Y. & Shibasaki, R. 2013. Weather effects on the patterns of people's everyday activities: A study using GPS traces of mobile phone users. *PLoS ONE*.

8(12):1–14.

Jariyasunant, J., Carrel, A., Ekambaram, V., Gaker, D.J., Kote, T., Sengupta, R. & Walker, J.L. 2011. *The Quantified Traveler: Using personal travel data to promote sustainable transport behavior*. (UCTC-FR-2011-10). Berkeley: University of California Transportation Center. [Online], Available: http://www.uctc.net/research/papers/UCTC-FR-2011-10.pdf [2015, November 12].

Järv, O., Ahas, R. & Witlox, F. 2014. Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*. 38:122–135.

Jianchuan, X., Zhicai, J., Guangnian, X. & Xuemei, F. 2014. Smartphone-based travel survey: A pilot study in China. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 209–223.

Kitamura, R., Fujii, S. & Pas, E.I. 1997. Time-use data, analysis and modeling: Toward the next generation of transportation planning methodologies. *Transport Policy*. 4(4):225–235.

Krygsman, S.C. & Schmitz, P. 2005. The use of cellphone technology in activity and travel data collection. In *Proceedings of the 24th Southern African Transport Conference (SATC 2005)*. Pretoria: Southern African Transport Conference. 696–705.

Krygsman, S., Nel, J.H. & de Jong, T. 2008. The use of cellphone technology in activity and travel data collection in developing countries. In *Proceedings of the 8th International Conference on Transport Survey Methods*. Annecy, France.

Loukopoulos, P., Jakobsson, C., Gärling, T., Schneider, C.M. & Fujii, S. 2004. Car-user responses to travel demand management measures: Goal setting and choice of adaptation alternatives. *Transportation Research Part D: Transport and Environment*. 9(4):263–280.

Mokhtarian, P.L., Salomon, I. & Handy, S.L. 2006. The impacts of ICT on leisure activities and travel: A conceptual exploration. *Transportation*. 33(3):263–289.

Nitsche, P., Widhalm, P., Breuss, S. & Maurer, P. 2012. A strategy on how to utilize smartphones for automatically reconstructing trips in travel surveys. In Vol. 48. *Procedia - Social and Behavioral Sciences*. Oxford: Elsevier. 1033–1046.

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N. & Maurer, P. 2014. Supporting large-scale travel surveys with smartphones - A practical approach. *Transportation Research Part C: Emerging Technologies*. 43:212–221.

Pas, E.I. 1980. *Toward the understanding of urban travel behavior through the classification of daily urban travel/activity patterns*. Published doctoral dissertation. Evanston: Northwestern University.

Rasouli, S. & Timmermans, H.J.P. Eds. 2014. *Mobile technologies for activity-travel data collection and analysis*. 1st ed. Hershey, Pennsylvania: IGI Global.

Shafique, M.A. & Hato, E. 2015. A review of travel data collection methods. In *Proceedings of the International Conference on Civil Engineering and Applied Mechanics (ICCEAM 2015)*. Paris: World Academy of Science, Engineering and Technology. 1906–1909.

Shen, L. & Stopher, P.R. 2013. A process for trip purpose imputation from Global Positioning System data. *Transportation Research Part C: Emerging Technologies*. 36:261–267.

Shen, L. & Stopher, P.R. 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*. 34(3):316–334.

Shoval, N., Kwan, M.-P., Reinau, K.H. & Harder, H. 2014. The shoemaker's son always goes barefoot: Implementations of GPS and other tracking technologies for geographic research. *Geoforum*. 51:1–5.

Stellenbosch University. 2013. *Mobile technology survey revails new trends*. [Online], Available: http://blogs.sun.ac.za/it/2013/11/01/mobile-technology-survey-reveals-new-trends/ [2014, February 06].

Stopher, P.R., Clifford, E., Swann, N. & Zhang, Y. 2009. Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy*. 16(6):315–324.

Stradling, S. & Anable, J. 2008. Individual transport patterns. In 1st ed. R. Knowles, J. Shaw, & I. Dochterty (eds.). *Transport Geographies: Mobilities, flows and spaces*. Oxford: Blackwell Publishing Ltd. 179–195.

Wolf, J. 2000. *Using GPS data loggers to replace travel diaries in the collection of travel data*. Published doctoral dissertation. Atlanta, Georgia: Georgia Institute of Technology.

Chapter 6. Using GPS-based activity spaces and opportunity indicators for travel behaviour analysis

Van Dijk, J.T. & Krygsman, S.C., 2017, Using GPS-based activity spaces and opportunity indicators for travel behaviour analysis. *Journal of Urban Technology*. Accepted for publication.

Abstract

An activity space is a spatial expression of individual spatial behaviour that can play a role in visualizing and analysing travel behaviour. In this chapter, we use the data of a two-day tracking experiment to explore whether accessibility to opportunities as represented through activity spaces associate with different travel characteristics, including willingness to consider more sustainable modes of transport. Simultaneously, we draw attention to the question of how to represent activity spaces. Using data of 95 respondents, we introduce point-of-interest data as an indicator to represent the opportunities within individual activity spaces and we analyse its relationship to the respondents' self-reported willingness to consider more sustainable modes of travel. The results indicate that there is an association between higher scores on the activity space opportunity indicators and willingness to consider more sustainable modes of transport. Overall, the study shows the potential of using accessibility indicators derived from GPS-based activity spaces to gain insight into travel behaviour.

Keywords

Activity spaces; Points of Interest; Opportunity indicators; GPS; Smartphones

6.1 Introduction

Against the background of an unprecedented growth in private vehicle ownership and a decrease in vehicle occupancy in many cities around the world, the past decades have seen a growing academic and policy debate on how to encourage individuals to change to more sustainable ways of traveling. As such, in recent years, there has been a growing interest in 'soft' transport demand management strategies (see, for example, Bonsall, 2009; Brög, Erl, Ker, Ryle & Wall, 2009; Zhang, Stopher & Halling, 2013; Sanjust di Teulada, Meloni & Spissu, 2017). These are the strategies that aim to stimulate a behavioural change by, for example, informing people about the negative impacts of private vehicle use rather than by enforcing a behavioural change (Taylor & Ampt, 2003). These approaches are typically referred to as Voluntary Travel Behaviour Change (VTBC) interventions.

To date, several studies have attempted to evaluate the effectiveness of VTBC interventions; however, the results have been mixed (Bonsall, 2009; Brög *et al.*, 2009; Chatterjee, 2009; Zhang *et al.*, 2013). One possible reason for the uncertain outcomes of VTBC interventions and the limited success of TDM to alter unsustainable travel in general is that the spatial context in which travel behaviour takes place is not explicitly considered (Behrens & Del Mistro, 2010; Howarth & Polyviou, 2012). Yet, several studies have found that, for example, densities and accessibility to opportunities for activity participation have an important effect on travel demand (Meurs & Haaijer, 2001; Van Acker & Witlox, 2009). Another possible reason is that an understanding of the effect of a travel behaviour intervention requires an understanding of travel behaviour and travel patterns, even more so in the case of a VTBC intervention because its effectiveness is related to the complexity of the individual travel patterns it aims to alter (Sanjust di Teulada *et al.*, 2017).

One way to visualise and analyse travel behaviour within its spatial context is through the construction of activity spaces. An activity space is a spatial expression of individually revealed travel behaviour, and can be seen as an individual's spatial footprint (Rai, Balmer, Rieser, Vaze, Schönfelder & Axhausen, 2007; Li & Tong, 2016; Xu, Shaw, Zhao, Yin, Lu, Chen, Fang & Li, 2016). Over the years, activity spaces have been used to describe travel behaviour by moving beyond one-dimensional approaches, such as the ones focusing only on number of trips and trip distances (Patterson & Farber, 2015). At the same time, a variety of methods has been used to represent activity spaces, often using travel diaries as a source to identify the spatial dispersion of activity locations that an individual frequently visits. Yet, most of these representations seem to ignore the underlying spatial structure and as such overestimate the extent of individual travel (Li & Tong, 2016).

With the advancement of location-aware technologies, it has become easier to obtain accurate data on individual spatial behaviour, as well as on the geographical context in which this behaviour occurs (Cottrill, Pereira, Zhao, Dias, Lim, Ben-Akiva & Zegras, 2013; Prelipcean, 2016; Zou, Yu & Cao, 2016). Using data from a travel survey and a two-day tracking experiment, augmented with land use data and point-of-interest data, the present chapter therefore aims to explore whether accessibility to opportunities as represented through GPS-based activity spaces associate with different travel characteristics, including willingness to consider more sustainable modes of transport. In addition, it draws attention to the question of how to represent activity spaces. The remainder of this chapter starts with a brief discussion of the existing literature on activity spaces. This is followed by a description of the data collection, the data preparation, and the analysis of the GPS-based activity spaces.

6.2 Travel behaviour and activity spaces

Travel is often considered to be derived from the willingness or necessity to participate in activities. However, the daily life of an individual is embedded in a complex spatial and temporal context that governs the opportunities an individual has for activity participation (Ettema & Timmermans, 1997; Arentze & Timmermans, 2000). As for spatial context, it has long been known that there is a complex relationship between spatial structure, residential environment, and travel behaviour (Chen, Ma, Susilo, Liu & Wang, 2016). For instance, when it comes to daily activities, such as visiting the supermarket, there is a clear relationship between travel distance and the proportion of non-motorized trips (Meurs & Haaijer, 2001). As such, the spatial context of individual travel behaviour may be associated with different travel characteristics.

One way to visualise and analyse travel behaviour within its spatial context is to look at the spatial distribution and extent of the locations that an individual interacts with on a day-to-day basis. This is often referred to as an individual's activity space; although terminology and meaning have varied throughout the years, ranging from *actual action space* to *contact action space*. Activity spaces are consequently organized around an individual's home location, his/her work location, and/or other daily activity locations, as well as the travel between those locations

(Järv, Ahas & Witlox, 2014). Activity spaces are, thus, a spatial representation of individually revealed spatial behaviour, moderated by the spatial configuration of various types of opportunities. As such, activity spaces have been incorporated into a variety of applications across disciplines; for example, understanding travel behaviour (Dijst, 1999), studying the mobility behaviour of older adults (Hirsch, Winters, Clarke & McKay, 2014), analysing the interpersonal variability of travel behaviour (Järv *et al.*, 2014), assessing environmental exposure (Perchoux, Chaix, Cummins & Kestens, 2013), examining the accessibility of healthcare opportunities (Sherman, Spencer, Preisser, Gesler & Arcury, 2005), and estimating the effect of the built environment on walking trips (Tribby, Miller, Brown, Smith & Werner, 2017). See Patterson and Farber (2015) for an extensive review of the different applications of activity spaces.

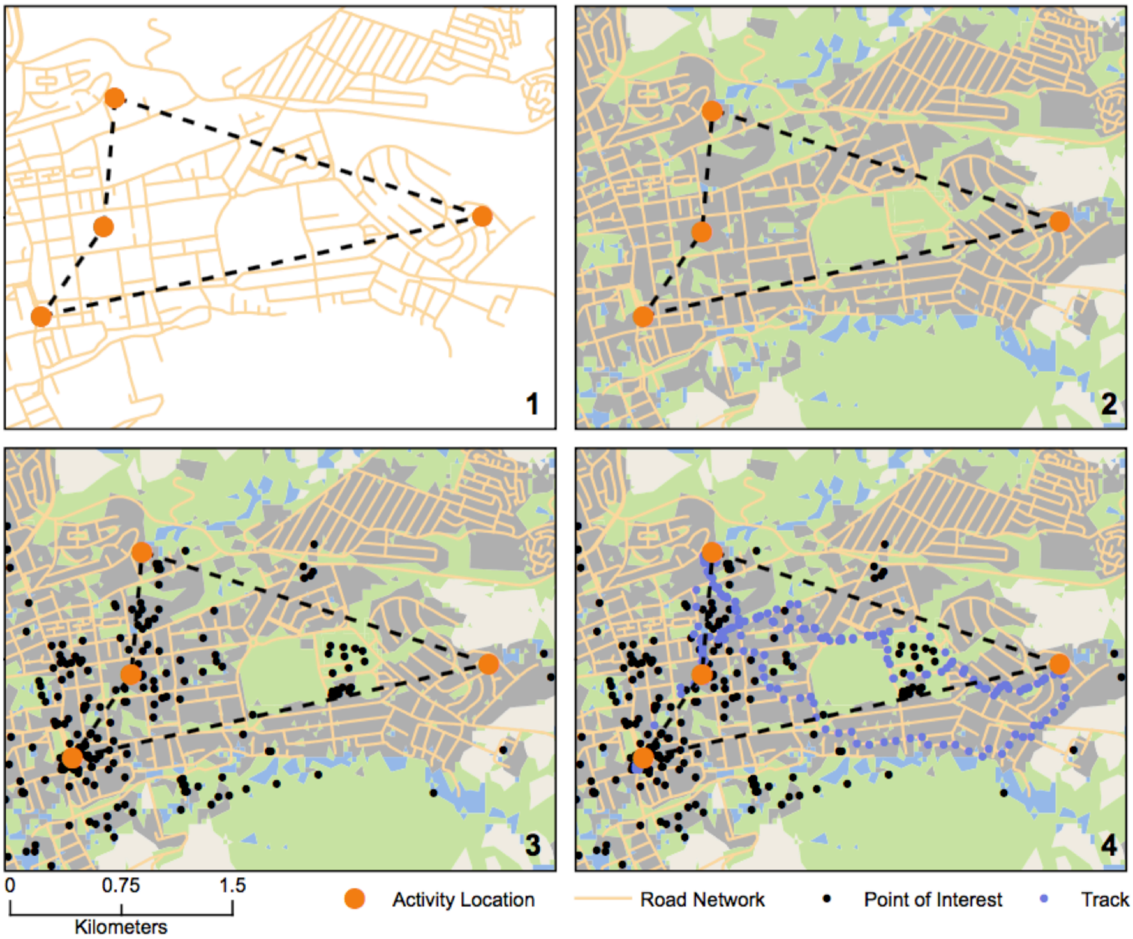
Whereas the idea of an activity space seems to be rather straightforward, studying it does pose some methodological challenges as its representation is a direct product of its operationalization. Several methods have been proposed in the literature. Examples of well-known operationalisations include the standard deviational ellipse (SDE) and the minimum convex polygon (MCP) of a set of points (Järv *et al.*, 2014; Patterson & Farber, 2015). The SDE was initially developed as a method to describe the distribution and orientation of a set of points. The MCP, on the other hand, most likely finds its origins in the ecological literature, and has been used to define animal home ranges as it represents the minimum bounding polygon that encompasses a set of points. In the transport literature, these sets of points are often the activity locations an individual has interacted with (cf. Schönfelder & Axhausen, 2003) or the locations of the furthest activities an individual has visited (cf. Li & Tong, 2016). Whereas both methods are relatively easy to understand and computationally inexpensive, they are likely to include areas with which individuals do not necessarily interact (Buliung, Roorda & Remmel, 2008; Patterson & Farber, 2015). From the perspective of using activity spaces to characterise travel behaviour, this may not be ideal. Nevertheless, the notion of the activity space operationalised as a standard deviational ellipse and a minimum convex polygon could still provide useful information because the area "represents the conceptual (perhaps minimal) area over which we know the individual is willing or able to engage in activities" (Newsome, Walcott & Smith, 1998: 362).

Besides the SDE and the MCP, kernel-density approaches and network-based approaches have been employed. Kernel densities transform a point pattern into a density surface that indicates potential for interaction. Additionally, it provides information on the intensity of the interaction. A disadvantage, however, is that the kernel function needs to be selected, and that the results strongly depend on the kernel bandwidth that needs to be estimated (Schönfelder & Axhausen, 2003). Alternatively, network-based approaches represent the area with which an individual is familiar with through his/her actual travel pattern. The easiest way to do this is by creating a buffer around the shortest paths that connect a sequence of activity locations (Patterson & Farber, 2015). However, rather than using shortest path analysis, location-based services such as GPS can also be used to identify the network segments to be incorporated into the network-based buffer (Tribby *et al.*, 2017). Similar as to the kernel-density approach, a disadvantage of the network-based approach is that the choice of parameters (i.e. in this case the size of the buffer) determines the activity space.

The different approaches not only pose difficulties with regard to the question of how to represent an activity space in the first place, but also with regard to the question of which indicators should be used to describe and compare activity spaces. Tribby *et al.* (2017), for example, use compactness as a means to compare network-based activity spaces to geometric representation of self-defined neighbourhoods. Alternatively, Newsome *et al.* (1998), who use an SDE, take the ratio between the minor and major axis of the ellipse and the size of the area of the ellipse to quantify the captured information. As such, the indicators that can be used to characterise an activity space are related to the representation of the activity space that has been employed. However, rather than looking only at the size or shape, in the context of travel behaviour analysis it may be more worthwhile to first consider the spatial contextualisation. Amongst other things, one may want to look into the opportunities for activity destinations that are present within an activity space, because this may give an indication of individual accessibility and availability of goods and services (Sherman *et al.*, 2005; Neutens, Schwanen & Witlox, 2011).

A graphical example of the importance of spatial contextualisation is given in Figure 6.1. The first part of the figure depicts four activity locations projected onto a road network that could, when connected, be part of an individual's activity space. When land use information is included in the second part of the figure, however, the question arises whether the land use classified as nature should be considered as part of the activity space or whether the urban form should be accounted for (Li & Tong, 2016). When points of interest are also added, as in the third part of the figure, a new spatial pattern emerges in which opportunities for activity destinations are particularly situated on the left-hand side. Whereas Li and Tong (2016), for instance, aptly draw attention to this limitation and argue for the incorporation of the urban form in the construction of activity spaces, the inclusion of the points of interest shows that urban form alone could also ignore important aspects of trip-making behaviour, as opportunities for activity destinations tend not to be homogenously distributed throughout the built environment. Lastly, when geo-locations, as acquired by GPS, are also incorporated as in the fourth part of the figure, it becomes evident that the individual's revealed travel behaviour is predominantly focused on three out of four of the initial activity locations. This example, therefore, suggests, that the combination of revealed spatial behaviour derived from GPS data and spatial contextualisation reveal important contextual information about an individual's travel pattern. Moreover, seemingly new technologies such as GPS directly convey information about the frequency and the spatial extent of those locations that an individual has interacted with (Järv *et al.*, 2014). As such, because location-aware technologies do not depend on the respondent's recollection of past activities and travel, they can provide a more accurate picture of spatial interaction with the built environment and individual accessibility to opportunities.

Figure 6.1 | An example of contextualised travel behaviour. (1) Activity locations with a road network; (2) Activity locations with a road network and land use indication; (3) Activity locations with a road network, land use indication and points of interest; (4) Activity locations with road network, land use indication and points of interest, overlaid with a GPS track.



6.3 Data collection

In September and October 2015, data collection for this study took place at Stellenbosch University, South Africa. The data collection was part of a research project on the travel patterns of staff and students on both the main campus and the Tygerberg satellite campus. Participants were recruited by means of an e-mail invitation that was sent to the total population of staff and students on both campuses; approximately 31,000 people. To stimulate participation, the student leaders of the university's residences were asked to remind their students of the importance of the study. The study was also promoted on the local student radio station. To stimulate staff participation, departmental heads were requested to help recruit staff members. The experiment itself consisted of two phases (in which a total number of three research instruments were used); a household transport questionnaire in phase one and a smartphone tracking application accompanied by a trip diary in phase two.

In the first phase of the experiment, the invitation was sent to the population of the smaller Tygerberg campus and a week later to the population of the main campus. In the invitation, it was explained that the research consisted of two phases which were both voluntary. The e-mail invitation contained a link which directed interested individuals to the online household transport questionnaire. After accepting the terms and conditions outlined in the ethical consent form, participants (*n = 1,130*) were requested to answer questions on their current travel behaviour. After data cleaning, 853 responses were suitable for analysis. In the second phase, participants who had filled out the questionnaire were contacted again with further instructions regarding the trip diary and the smartphone application. Participants were requested to complete the trip diary, which was available through an online survey tool, for two consecutive working days. Almost 25 per cent of the initial respondents accessed the online trip diary, but a smaller percentage actually completed it (*n = 200*). In addition, the quality of the trip diaries varied greatly, with many respondents completing only one trip and only a few respondents keeping an extensive diary recording every move during the course of the two days of the research project.

Respondents were asked to download a smartphone application named Tracklog for smartphones running iOS and Android from the respective App Stores (Apple App Store and Google Play Store). Upon activation, the application allowed the passive recording of its user's positional information using the GPS location registered by the smartphone. This information included the geographical coordinates and the time of each measurement. The recorded information was automatically uploaded every hour via the GSM network to a server from where the data could be extracted in a delimiter-separated file format (.csv). A location measurement frequency of 30 seconds was set for the data collection, leading to a potential recording of 2880 records per 24 hours. The uploaded information could also be accessed by the participants: a feedback website was created that allowed participants to log into a web interface where they could view their personal tracks projected onto a Google Maps background. Unfortunately, it was not possible for the users to annotate their data and the data did not contain travel specific information such as travel mode, trip purpose, and trip destination.

6.4 Travel behaviour, activity spaces, and opportunity densities

Before proceeding to the analysis of the activity spaces, it is important to briefly look at the overall characteristics of the staff and student population that responded to the household travel survey to contextualise the results of the activity space analyses. Accordingly, this section starts off by highlighting some of the results from the survey, followed by an exploration of the captured GPS tracks and the exploration of the GPS-based activity spaces.

6.4.1 Household travel survey

Of the 853 participants who responded to the survey, the majority were students (*n = 464*) versus a smaller number of staff members (*n = 389*). Also, the number of females (*n = 490*) was higher than the number of males *(n = 334*). More interestingly, in one of the questions, respondents were asked to indicate whether they have a vehicle that they use to travel to and from campus. The total number of responses to this question was 690, of which 551 answered 'yes' whilst only 139 respondents answered 'no'. As shown in Table 6.1, the main reasons for using a private vehicle are that it is perceived as the safest and the quickest option. Moreover, for many it is perceived to be the only option.

Table 6.1 | Reasons for using private transport, reasons that would encourage moving from private transport, and reasons that discourage moving away from private transport.[1]

Reasons to use private transport	Mode	Median
Private transport is the safest option (*n = 615)*	1	2
Private transport is the only option (*n = 636*)	1	1
Private transport is the most convenient option (*n = 627*)	1	3
Private transport is the quickest option (*n = 627*)	1	1
Private transport is the cheapest option (*n = 608*)	3	3
Private transport required for shopping (*n = 619*)	1	3
Private transport required for other people (*n = 616*)	5	4
Encouraging to move away from private transport		
Availability of reliable and affordable alternatives (*n = 624*)	1	2
Avoiding peak hour congestion (*n = 622*)	1	2
Reduction of carbon footprint (*n = 624*)	1	3
Reduction in transport costs (*n = 628*)	1	2
Increase in parking tariffs (*n = 604*)	3	3
Reductions in available parking (*n = 611*)	3	3
Discouraging to move away from private transport		
Public Transport is expensive (*n = 479*)	3	3
Safety of public transport (*n = 508*)	1	1
Unreliability of public transport (*n = 510*)	1	1
Waiting outside for public transport (*n = 497*)	1	3
Overcrowded conditions in public transport (*n = 500*)	1	1
Unfavourable weather conditions (*n = 499*)	1	2

[1] Mode and median of importance of reasons measured on a Likert-scale, with *1* indicating a very important reason and *5* indicating a reason of little importance. Please note that the number of valid responses (*n*) differs between the questions as a result of missing values.

Not only the reasons to use private transport, but also the reasons that would encourage the participants to move away from private transport and that discourage the participants to move away from private transport suggest that the safety and reliability of the transport mode are extremely important. In fact, these reasons seem to play a more important role than, for instance, one's reduction in carbon footprint or a reduction in one's transport costs. Concerns about the safety aspect of public transport in South Africa are consistent with other studies that have reported on the topic (cf. Walters, 2008; Lucas, 2012). Taken together, it is not surprising only 23 respondents indicated that they would consider walking (*n = 404*) and 75 respondents indicated that they would be willing to cycle (*n = 75*). On the other hand, when asked whether the respondent would consider carpooling (*n = 699*), 347 respondents answered 'yes' versus 322 respondents who answered 'no'; suggesting some possibilities for a willingness to change travel behaviour.

6.4.2 Smartphone data collection

A total of 176 participants registered onto the Tracklog system, but only 151 and 141 unique responses (tracks) were recorded on the two days of tracking, respectively. This implies that some users registered but did not actually initiate the tracking. After data collection, all tracks of the users who had also filled out the questionnaire were extracted from the server for further processing (*n = 292*). As GPS technology in smartphones also gives an indication of the positional error of the measurements, for each track, measurements with a horizontal accuracy of more than 50 metres were dismissed to remove possible unreliable measurements. However, when only a small number of satellites are in sight this horizontal accuracy cannot be estimated and returns a horizontal accuracy of 0. Therefore, the data cleaning procedure was extended to remove unrealistic measurements characterized by a speed of 150 kilometres per hour and higher. As this process resulted in several empty tracks (*n = 24*), the remaining tracks (*n = 261*) were imported (WGS84) and projected (Hartbeestehoek94 LO19) in ArcGIS 10.3 using the Python scripting language with the Pandas (McKinney, 2010) and ArcPy (ESRI, 2014) Python libraries.

After cleaning and preparing the tracking data, it was found that the cleaning process had a large impact on the number of measurements per track. As shown in Table 6.2, the full data set (*n = 285*) comprised on average of 594 measurements, whereas the remaining tracks (*n = 261*) comprised on average of 426 measurements. Whereas this large difference can be attributed to the cleaning process, it does raise the question of why so many measurements were filtered out. The most plausible cause for the unreliable measurements is the absence of a GPS fix when Tracklog requested the location – for instance, as a result of signal blockage inside a building. Another possibility is that the smartphone did not actually register the GPS location of the device. In some cases, it was found that Tracklog did not properly activate the GPS on smartphones running on iOS, and only registered the last known location of the phone, which, is under normal circumstances, refers to the location of the closest cell tower.

Table 6.2 | Descriptive statistics tracks before and after cleaning

	Average number of measurements	Minimum number of measurements	Maximum number of measurements	Std. Dev. number of measurements
Before cleaning (*n* = 285)	594	1	3093	612
After cleaning (*n* = 261)	426	1	3093	555

Not only the difference between the average number of measurements before and after cleaning but also the range of the number of measurements raises some questions. Whilst some tracks consist of only one measurement, others consist of more than 3,000 measurements. Regarding the almost empty tracks, two possible explanations come to mind. The first possible explanation is that the user actually decided to break off the tracking. The second possible explanation is that due to a technical failure, the application did not properly run in the background and was terminated due to memory starvation. Some smartphones running on the Android platform were especially affected by this malfunction. This strongly suggests that future versions of the application should incorporate a strategy for battery duty cycling to maintain a low footprint.

Overall, the quality of the tracks varied greatly. Figure 6.2 for instance, gives an example of a good quality track with a high number of measurements – on average, a measurement was captured every 27 seconds. It can be seen from the figure that there are both clusters of points, indicative of an activity, as well as individual points along road segments, indicative of travel. Figure 6.3, on the other hand, gives an example of a low-quality track with a low number of measurements – on average, only one measurement every 847 seconds. Because of the variations in the quality of tracks, it was decided that for the exploration of individual activity spaces, only tracks with at least 100 track points would qualify. When a respondent only initiated the tracking for only one day, the number of track points was evaluated directly against this criterion, whereas the tracks of the respondents who initiated the tracking for two days were firstly merged. This process resulted in 95 qualifying tracks being incorporated in the analysis.

Figure 6.2 | Example of a good-quality track – central Stellenbosch

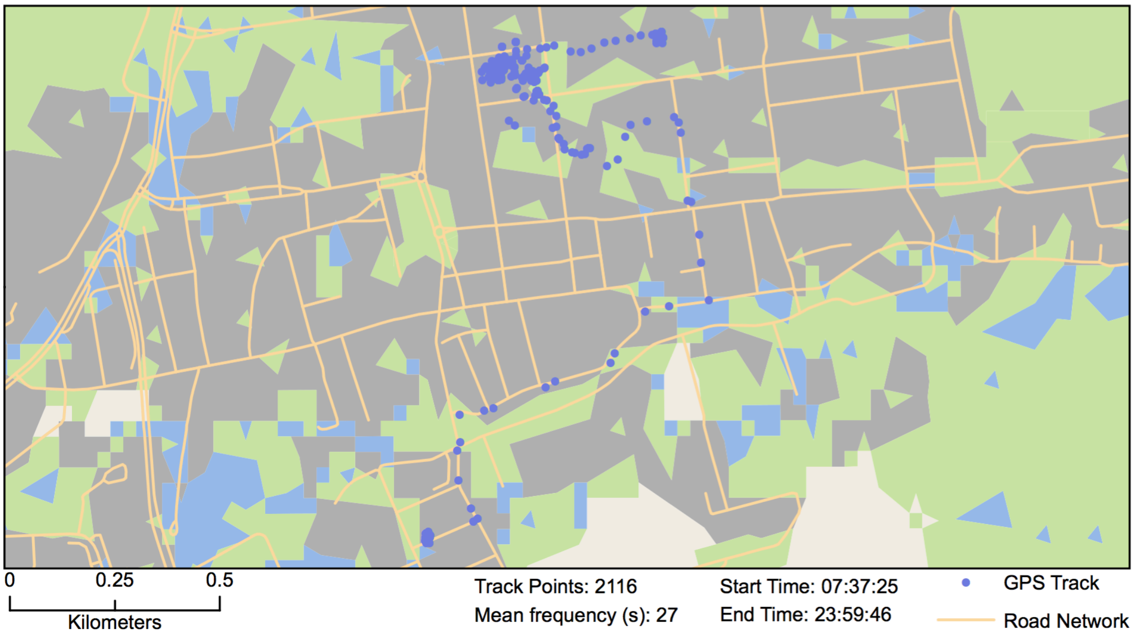
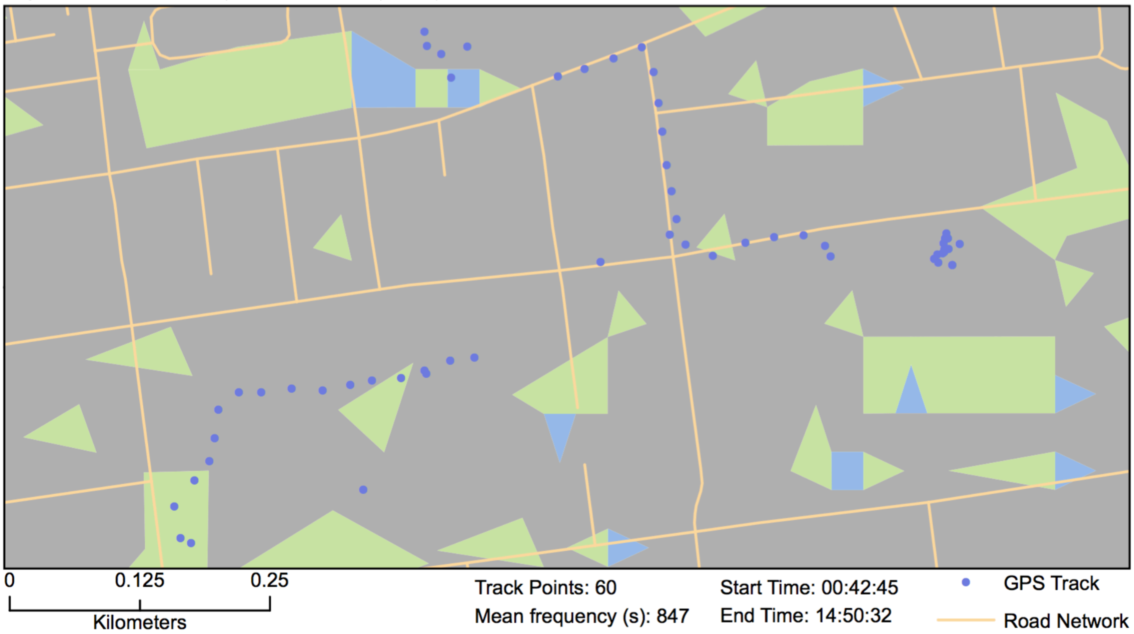


Figure 6.3 | Example of a low-quality track – central Stellenbosch



6.4.3 Activity spaces

As explained in the literature section, activity spaces pose some methodological challenges as their geometry is a product of how the activity space is measured, operationalised, and contextualised; which may, in turn, affect the results of any subsequent analysis. Because of this, it was decided to not just focus on one particular conceptualisation; in a similar fashion as Schönfelder and Axhausen (2003), Sherman *et al.* (2005), and Hirsch *et al.* (2014) have previously done. The following conceptualisations were chosen to be incorporated and directly applied to the qualifying tracks: minimum convex polygon (MCP), minimum convex polygon with a 200-metre buffer (MCP200), standard deviational ellipse (SDE), GPS track with a 200-metre buffer (BUFFER), and a characteristic-based hull with a 200-metre buffer (CHP) (see Downs & Horner, 2009). Table 6.3 summarizes the different conceptualisations and Figure 6.4 gives an example of each of them. As an addition to the basic conceptualisations, in accordance with Li and Tong (2016), all basic conceptualisations were intersected with land use classified as built-up to account for the structure of the urban form. The land use data were derived from the 2013-2014 national land-cover data set of South Africa (Department of Environmental Affairs, 2015; GeoTerraImage, 2015). This dataset, based on 30 by 30 metre raster cells derived from Landsat 8 imagery, was aggregated into four main land use categories, and the urban structure (i.e. built-up area) was extracted (Li & Tong, 2016). Figure 6.5 illustrates each of these basic conceptualisations with urban form correction.

Table 6.3 | Activity space representations

Representation	Explanation
MCP	Smallest possible polygon encompassing all points.
SDE	Directional distribution of a set of points at one standard deviation.
BUFFER	200-metre buffer around line feature of GPS track.
CHP	Delaunay triangulation on GPS track, subsequently removing the 5 percent polygons with the largest perimeter.

Figure 6.4 | Example of conceptualisations of activity spaces: (1) Minimum convex polygon and minimum convex polygon with a 200-metre buffer; (2) Standard deviational ellipse at one standard deviation; (3) 200-metre buffer around track; (4) Characteristic-based hull with a 200-metre buffer and 95% Delaunay triangulation.

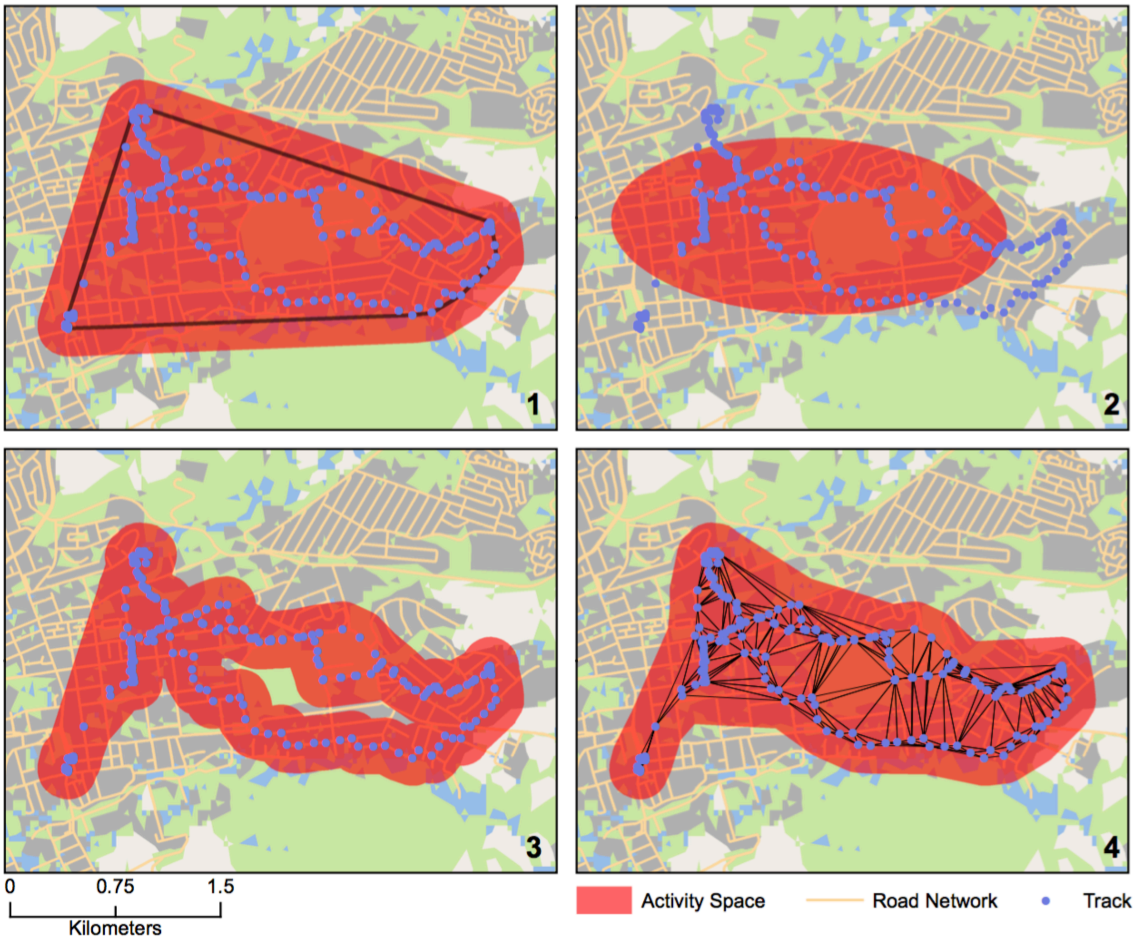


Figure 6.5 | Example of conceptualisations of activity spaces with urban form correction. (1) Minimum convex polygon and minimum convex polygon with a 200-metre buffer; (2) Standard deviational ellipse at 1 standard deviation; (3) 200-metre buffer around track; (4) Characteristic-based hull with a 200-metre buffer and 95% Delaunay triangulation.

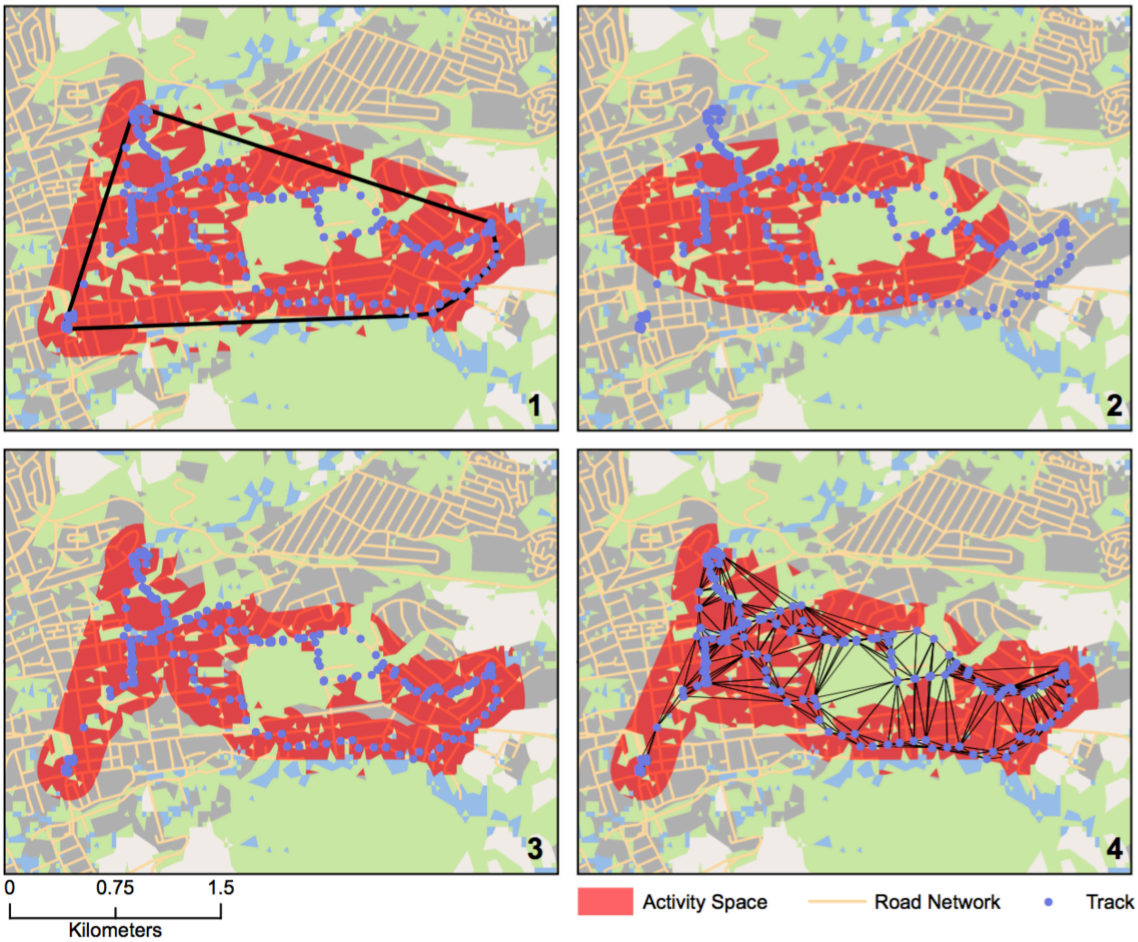
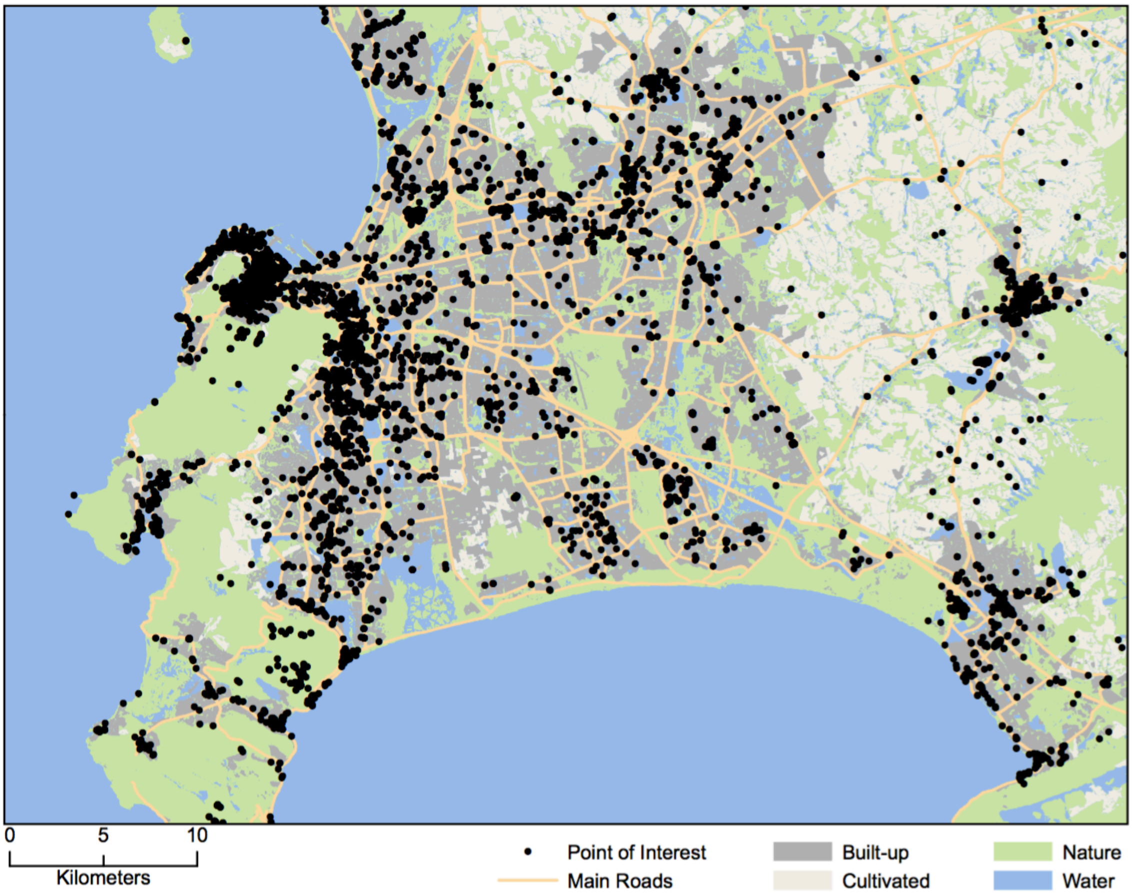


Figure 6.6 | Land use and selected Points of Interest Cape Town metropolitan region



Source: Department of Environmental Affairs (2015), GeoTerraImage (2015), OpenStreetMap Contributors (2016)

Whereas the urban form may be insightful, on its own it does not account for the distribution of activity destinations within the built-up area. To provide further spatial contextualisation for the activity spaces of the qualifying tracks, point-of-interest (POI) data were extracted from OpenStreetMap (OSM). OSM is one of the most popular and well-known platforms for geographic user-generated content. As opposed to proprietary data sources, OSM data is freely available under the Open Database License (OpenStreetMap Contributors, 2016). Whereas this data set is unlikely to include all POIs in the area, as OSM fully depends on donated and user-generated content, it is arguably the most complete POI data source for the region that is available to the public. Figure 6.6 shows the extracted points of interest, as well as the main land use categories within the research area of the wider Cape Town metropolitan region.

After downloading the full extract for the larger Cape Town and Stellenbosch area, relevant data were filtered through the Java command-line application Osmosis for OS X. The final data set ($n = 5,782$) consisted of POIs such as restaurants, bars, cafes, supermarkets, libraries, schools, shops, museums and hospitals. Less relevant amenities such as public toilets and private swimming pools were removed. To account for the diversity amongst the different POIs, they were categorized into 13 discrete amenity categories: government (e.g. fire station, police station), education (e.g. university), place of worship, groceries (e.g. supermarket, bakery), shop (e.g. clothing store, shoe store), trade (e.g. tailor, electrician), health (e.g. doctor, hospital), outgoing (e.g. pub, restaurant), sport, transport, other amenities (e.g. library, ATM, post office), other leisure (e.g. museum, hotel) and other. In turn, the points and the point diversity were translated into a set of activity space opportunity indicators that represent accessibility to opportunities, as shown in Table 6.4. Particular emphasis is on point clustering, the average number of neighbouring points within 100 metre, point density, and weighted point density. These indicators are subsequently used to see whether accessibility to opportunities as represented through activity spaces relate to different travel characteristics, including willingness to consider more sustainable modes of travel.

Table 6.4 | Activity space opportunity measures

Measure	Explanation
Points	Number of points
Point diversity	Ratio number of different POI categories out of maximum number of different POI categories
Point clustering	Ratio number of points POI category with highest number of points out of total number of points
Point neighbours	Average number of neighbouring points within 100 metre
Point density	Number of points / square kilometre
Weighted point density	Number of points weighted by point diversity / square kilometre

Against the background of the selected opportunity indicators, a first question to emerge is whether there is a relationship between these indicators and the distance between home and campus. To test for this relationship, Kendall's τ_b correlation was run in R (R Core Team, 2016) using the *Hmisc* software library (Harrell, 2016) between the self-reported distance from home to campus and the various opportunity measures. For verification purposes, a Spearman's Rank

Correlation was also executed. Table 6.5 shows the correlations for all types of activity spaces both with (U) and without the urban form correction. The correlations in the table show that in all cases there is a negative correlation between the distance and the opportunity indicator, suggesting that with increasing distances the value of the opportunity indicator decreases. However, there appears to be some variation. The point clustering and point neighbours indicators overall suggest a weaker association with the self-reported distance than the density and the weighted density do. In addition, for all opportunity indicators and all activity space representations, the association that includes the urban form correction is slightly weaker. This may suggest, as concluded by others, that the different representations of the activity spaces do not 'outperform' each other, but rather emphasize different aspects of individual accessibility (Sherman *et al.*, 2005).

Table 6.5 | Correlation of self-reported distance to campus with access to opportunities as represented through activity spaces [1]

	Clustering	Neighbours	Density	Weighted Density
MCP	-.396 (-.553)	-.408 (-.549)	-.572 (-.742)	-.549 (-.719)
MCP (U)	-.342 (-.501)	-.359 (-.488)	-.528 (-.687)	-.485 (-.637)
MCP200	-.374 (-.518)	-.399 (-.547)	-.599 (-.771)	-.606 (-.783)
MCP200 (U)	-.309 (-.436)	-.368 (-.507)	-.575 (-.740)	-.561 (-.732)
SDE	-.418 (-.572)	-.344 (-.467)	-.495 (-.653)	-.351 (-.447)
SDE (U)	-.340 (-.480)	-.290 (-.395)	-.386 (-.518)	-.273 (-.353)
BUFFER	-.265 (-.388)	-.357 (-.483)	-.560 (-.728)	-.551 (-.726)
BUF (U)	-.227 (-.335)	-.356 (-.472)	-.575 (-.740)	-.481 (-.635)
CHP	-.286 (-.408)	-.355 (-.478)	-.588 (-.765)	-.583 (-.767)
CHP (U)	-.265 (-.369)	-.343 (-.455)	-.522 (-.686)	-.498 (-.663)

[1] Correlations coefficients for Kendall's τ_b. Spearman's Rank Correlation in between brackets.

A second question to emerge is whether different groups can be identified within the types of activity spaces. Table 6.6 reports on the results of a series of Mann-Whitney U tests of three different groups: (1) males (n = *52*) versus females (n = *43*), (2) staff (n = *65*) versus students (n = *30*), and (3) using private transport to travel to campus (n = *84*) versus not using private transport to travel to campus (n = *5*). For the gender category (M/F), none of the tests are significant, except for MCP200 with urban correction. For the staff and student category (S/S), in some cases, there are some signs of differences. However, no consistent picture emerges across all activity space conceptualisations or throughout one of the opportunity measures. Lastly, the activity space conceptualisations also do not seem to significantly differ between respondents who use private transport to travel to and from campus and those who do not (CAR). It is apparent from this table that the opportunity indicators and the different activity spaces conceptualisations cannot conclusively characterize the groups.

Table 6.6 | Significant associations of characteristics (gender, staff member or student, private transport use or no private transport) with accessibility to opportunities as represented through activity spaces[1]

	Clustering			Neighbours			Density			Weighted Density		
	M/F	S/S	CAR	M/F	S/S	CAR	M/F	S/S	CAR	M/F	S/S	CAR
MCP	No	**	No	No	No	No	No	No	No	No	No	No
MCP (U)	No	No	*	No	*	No	No	**	No	No	**	No
MCP200	No	**	*	No	*	No	No	No	No	No	No	No
MCP200 (U)	No	*	**	No	No	No	*	**	No	No	**	No
SDE	No	*	No	No	No	No	No	No	No	No	No	No
SDE (U)	No	**	No	No	No	No	No	No	No	No	No	No
BUFFER	No	*	*	No	**	No	No	No	No	No	**	No
BUF (U)	No	No	*	No	No	No	No	No	No	No	No	No
CHP	No	*	No	No	**	No	No	No	No	No	No	No
CHP (U)	No	No	No	No	No	No	No	No	No	No	No	No

[1] Mann-Whitney U test on mean opportunity measure for different groups. * Significant at the 0.10 level (two-tailed). ** Significant at the 0.05 level (two-tailed).

So far, the results only cautiously suggest that the opportunity indicators are negatively associated with the distance to and from campus. However, this is not very surprising as larger distances would normally be associated with larger activity spaces and as such with, for instance, lower opportunity densities. A more promising question is whether there is a relationship between the opportunity indicators and the willingness to make a behavioural change. It can be hypothesized that respondents with higher scores on the opportunity measures have a higher degree of spatial freedom and, therefore, are more likely to express an interest in an alternative transport mode. To test this, three questions pertaining to the willingness of the respondent to consider carpooling, cycling, or walking to and from campus were selected from the survey.

Table 6.7 presents the results of the respondents who stated that they would be willing to consider carpooling to travel to and from campus ($n = 47$) against the respondents who stated that they would not be willing to do so ($n = 42$). The most obvious result to emerge from the table is that the average number of neighbours is significantly higher for the 'no' group across all activity space representations. Although not significant for all activity space conceptualisations, the means of the *point density* and the *weighted point density* are clearly higher for the respondents who answered 'no'. The activity spaces of respondents who said to be willing to carpool to work, for instance, seem to be associated with lower point densities than the activity spaces of the respondents who indicated not to be interested in carpooling.

Table 6.7 | Association of *willingness to consider carpooling to travel to and from campus* with access to opportunities as represented through activity spaces [1]

	Clustering			Neighbours			Density			Weighted Density		
	Carpool			Carpool			Carpool			Carpool		
	Yes	No	Sig.	Yes	No	Sig.	Yes	No	Sig.	Yes	No	Sig.
MCP	.289	.293	No	9.4	11.4	**	9.9	19.2	*	8.0	14.7	No
MCP (U)	.302	.264	No	10.2	12.4	**	15.7	27.9	**	13.4	21.9	**
MCP200	.251	.241	No	9.8	11.6	**	9.9	16.5	*	8.5	14.5	*
MCP200 (U)	.273	.261	No	10.8	12.7	**	16.2	25.5	**	14.4	22.9	**
SDE	.279	.316	No	8.6	10.7	*	12.3	27.6	No	9.0	15.2	No
SDE (U)	.313	.326	No	9.3	11.3	**	19.7	36.2	No	15.2	21.1	*
BUFFER	.272	.264	No	10.3	12.2	**	13.8	21.3	**	11.7	18.6	**
BUF (U)	.297	.295	No	11.0	13.2	**	21.6	31.7	**	18.5	27.8	**
CHP	.273	.266	No	10.5	12.2	**	15.9	22.2	No	13.7	18.7	No
CHP (U)	.293	.291	No	11.3	13.2	**	23.2	31.6	*	20.2	26.9	*

[1] Mann-Whitney U test on mean opportunity measure 'yes' (n = *42*) and 'no' (n = *47*) groups. * Significant at the 0.10 level (two-tailed). ** Significant at the 0.05 level (two-tailed).

Table 6.8 shows the results of the respondents who indicated that they are willing to consider cycling (n = *39*), not willing to consider cycling (n = *12*), or consider cycling not an option (n = *9*). Table 9 shows the results of the respondents who indicated that they are willing to consider walking (n = *5*), not willing to consider walking (n = *4*), or state that walking is not an option (n = *51*). When it comes to the willingness to consider cycling, the most interesting result is that both the *point density* and the *weighted point density* show a similar trend: the highest mean values are associated with those respondents who indicated that they would consider cycling, and the lowest mean values are associated with the respondents who said that cycling was not an option. These results are significant for all activity space conceptualisations, except for the SDE. A similar pattern emerges from Table 6.9 in the case of willingness to consider walking; however, the low number of cases in the 'yes' and 'no' category inhibit any further inferences. These results provide some evidence that the explanatory power of the opportunity indicators in the SDE is not as good as the explanatory power of the opportunity indicators of the other activity space conceptualisations. Moreover, the results provide some support for the notion that opportunity indicators could be useful in distinguishing between groups that might be willing to move to a more sustainable mode of transport.

Table 6.8 | Association of *willingness to consider cycling to travel to and from campus* with access to opportunities as represented through activity spaces [1]

	Clustering				Neighbours				Density				Weighted Density			
	Bike				Bike				Bike				Bike			
	Yes	No	N/O	Sig.	Yes	No	N/O	Sig.	Yes	No	N/O	Sig.	Yes	No	N/O	Sig.
MCP	.409	.310	.243	**	12.4	9.6	9.5	**	26.0	16.1	6.3	**	19.6	13.9	4.2	**
MCP (U)	.432	.321	.262	**	13.1	10.3	10.6	No	32.8	25.8	13.0	**	25.6	22.5	10.3	**
MCP200	.319	.219	.230	**	12.4	9.6	9.9	**	23.1	15.3	5.3	**	19.9	13.2	4.6	**
MCP200 (U)	.343	.243	.246	*	13.3	10.3	11.0	No	32.2	23.2	12.3	**	28.2	20.6	11.3	**
SDE	.403	.218	.249	**	10.8	8.9	8.9	No	33.1	17.2	15.0	*	18.3	14.2	6.3	No
SDE (U)	.384	.233	.276	No	10.8	9.3	9.8	No	42.6	25.1	23.8	No	24.6	20.4	13.4	No
BUFFER	.347	.225	.254	**	11.9	10.3	10.5	No	25.4	21.5	9.9	**	20.9	19.0	8.7	**
BUF (U)	.371	.243	.281	*	12.8	11.0	11.4	No	34.1	30.7	19.0	**	28.2	27.4	16.7	**
CHP	.338	.237	.256	**	12.4	10.2	10.6	No	27.4	21.8	10.5	**	22.7	18.3	8.8	**
CHP (U)	.360	.250	.276	*	13.3	10.9	11.6	No	35.6	29.6	18.6	**	29.7	25.5	16.4	**

[1] Kruskal-Wallis rank sum test on mean opportunity measure 'yes' (n = *39*), 'no' (n = *12*) and 'N/O' (No option) (n = 9) groups. * Significant at the 0.10 level (two-tailed). ** Significant at the 0.05 level (two-tailed).

Table 6.9 | Association of *willingness to consider walking to travel to and from campus* with access to opportunities as represented through activity spaces [1]

	Clustering				Neighbours				Density				Weighted Density			
	Walk				Walk				Walk				Walk			
	Yes	No	N/O	Sig.	Yes	No	N/O	Sig.	Yes	No	N/O	Sig.	Yes	No	N/O	Sig.
MCP	.568	.416	.252	**	11.2	11.6	9.8	No	25.0	27.8	9.1	No	15.8	23.7	6.7	No
MCP (U)	.566	.423	.273	No	11.4	12.3	10.9	No	32.8	37.3	15.9	No	21.8	31.9	13.0	No
MCP200	.423	.252	.232	*	10.0	11.4	10.3	No	22.4	25.3	7.9	**	18.0	21.5	7.0	**
MCP200 (U)	.474	.277	.248	**	10.5	12.2	11.3	No	32.4	32.3	15.3	**	27.0	21.7	14.1	*
SDE	.370	.238	.272	No	7.8	10.0	9.3	No	37.6	35.6	15.9	No	11.3	30.1	7.8	No
SDE (U)	.408	.252	.286	No	7.9	10.6	10.0	No	48.0	42.6	24.7	No	17.5	36.0	14.7	No
BUFFER	.423	.251	.258	No	10.2	11.8	10.7	No	27.8	30.4	12.1	**	22.5	25.4	10.7	**
BUF (U)	.476	.270	.281	No	10.7	12.6	11.6	No	38.9	38.3	21.2	**	32.1	32.1	18.6	No
CHP	.425	.275	.256	No	10.3	11.9	10.8	No	29.3	32.7	12.8	**	22.8	26.5	10.9	*
CHP (U)	.472	.293	.275	No	10.8	12.7	11.8	No	39.6	39.5	20.9	**	31.5	32.3	18.4	No

[1] Kruskal-Wallis rank sum test on mean opportunity measure 'yes' (n = *4*), 'no' (n = *5*) and 'N/O' (No option) (n = *51*) groups. * Significant at the 0.10 level (two-tailed). ** Significant at the 0.05 level (two-tailed).

6.5 Conclusion and discussion

The recent growth and application of location-aware technologies, such as GPS, have facilitated progress in accurately gathering spatial data on individual travel behaviour. The present chapter set out to explore whether accessibility to opportunities as represented through GPS-based activity spaces associate with different travel characteristics, including willingness to consider more sustainable modes of transport. In addition, it draws attention to the question of how to represent activity spaces. Activity spaces can be seen as an individual's spatial footprint of their day-to-day travel and activity behaviour and, thus, provide valuable information about the context in which this behaviour takes place. Based on the examination of 95 GPS tracks with different two-dimensional representations of activity spaces, however, there seems to be no evidence that one of the representations is better than the others; although it may be cautiously suggested that the explanatory power of the opportunity indicators in the SDE is not as good as that of the other activity space representations. This may be related to the fact that the SDE is by definition a statistical representation of the distribution and orientation of a set of points, and as such can exclude some areas in which an individual has actually interacted with. Whereas an SDE may still provide useful information, the advancement in the availability of geo-spatial technologies makes it easier to consider alternative representations that do include all visited locations. Furthermore, the results do indicate that there is some merit in considering opportunity indicators - not so much in terms of characterizing independent groups like gender, but more in terms of distinguishing between groups that may and may not be willing to consider a move towards a more sustainable mode of transport.

The possible relationship between the accessibility to opportunities as represented through an activity space and the willingness to consider a move towards a more sustainable mode of transport is important for two reasons. The first reason is that it suggests that higher values on the opportunity indicators, could be indicative of a higher degree of spatial freedom. These individuals may be more likely to respond to external stimuli and/or awareness programs when it comes to stimulating more sustainable modes of travel than individuals who have lower scores on the opportunity indicators. Some preliminary evidence for this is provided by the differences in point densities between respondents who said to be willing to cycle to work and who indicated not to be willing to cycle compared to the activity spaces of the respondents who said that cycling was not an option. The second reason is that it shows that GPS data and activity spaces, can be relatively easily be exploited and perhaps even be used in travel behaviour change schemes. For instance, activity spaces could be used for identifying target populations that exhibit the same type of spatial behaviour and have similar scores for their opportunity indicators.

While insightful, the major limitation of the current study is the period of data collection. Because activity-travel behaviour is extremely variable, the data collection period should be considerably longer to lead to more definitive conclusions. There are also some limitations when it comes to using activity spaces as part of travel behaviour analysis in general because activity spaces are naive towards space-time constraints. As an individual's life is governed by multiple constraints, such as coupling constraints with other household members, the actual bandwidth of opportunities available to an individual cannot be established by merely looking at their activity space. One possible way to consider this bandwidth of behavioural opportunities and allow for a differentiation between activity types is to turn to space-time accessibility (STA) measures. Because of these limitations, the results should be interpreted with caution and considered as only a preliminary exploration. Future work should explore whether the established relationships hold for larger populations, and whether an activity space typology can be derived and related to socioeconomic demographics. In addition, future work could explore which representation of an activity space 'better' explains the differences between different groups, or gives the best representation of activity-travel behaviour in general. The inclusion of trip purpose and travel mode information, for instance, could also improve the representations as the suitability of a representation may vary with its purpose.

Acknowledgements

The financial assistance of the South African National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF. We thank both anonymous reviewers for their comments and constructive criticism. We would also like to thank Ms Melanie Venter and Ms Jeanette Thiart for their contributions in the data collection and data preparation process. An early version of this paper was presented at Mobile Tartu 2016 in Tartu, Estonia.

References

Van Acker, V. & Witlox, F. 2009. Why land use patterns affect travel behaviour (or not). *Belgeo*. 1(1):5–26.

Arentze, T. & Timmermans, H.J.P. 2000. *ALBATROSS: A learning-based transportation oriented simulation system*. Eindhoven: European Institute of Retailing and Services Studies.

Behrens, R. & Del Mistro, R. 2010. Shocking habits: Methodological issues in analyzing changing personal travel behavior over time. *International Journal of Sustainable Transportation*. 4(5):253–271.

Bonsall, P. 2009. Do we know whether personal travel planning really works? *Transport Policy*. 16(6):306–314.

Brög, W., Erl, E., Ker, I., Ryle, J. & Wall, R. 2009. Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*. 16(6):281–292.

Buliung, R., Roorda, M. & Remmel, T. 2008. Exploring spatial variety in patterns of activity-travel behaviour: Initial results from the Toronto Travel-Activity Panel Survey (TTAPS). *Transportation*. 35(6):697–722.

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*. 16(6):293–305.

Chen, C., Ma, J., Susilo, Y., Liu, Y. & Wang, M. 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies*. 68:285–299.

Cottrill, C.D., Pereira, F.C., Zhao, F., Dias, I., Lim, H.B., Ben-Akiva, M. & Zegras, P. 2013. Future Mobility Survey - Experience in developing a smartphone-based travel survey in Singapore. *Transportation Research Record: Journal of the Transportation Research Board*. 2354:59–67.

Department of Environmental Affairs. 2015. *2013 / 2014 South African national land-cover dataset*. [Online], Available: https://www.environment.gov.za/mapsgraphics [2016, November 22].

Dijst, M. 1999. Two-earner families and their action spaces: A case study of two Dutch communities. *GeoJournal*. 48(3):195–206.

Downs, J.A. & Horner, M.W. 2009. A characteristic-hull based method for home range estimation. *Transactions in GIS*. 13(5–6):527–537.

ESRI. 2014. *ArcGIS Desktop: Release 10.3*. Redlands, CA: Environmental Systems Research Institute.

Ettema, D. & Timmermans, H.J.P. Eds. 1997. *Activity-based approaches to travel analysis*. Oxford: Pergamon.

GeoTerraImage. 2015. *2013 - 2014 South African national land-cover data. User report and metadata.* Pretoria, South Africa. [Online], Available: https://www.environment.gov.za/mapsgraphics.

Harrell, F. 2016. *Hmisc: Harrell Miscellaneous [R package version 4.0.0].* [Online], Available: https://cran.r-project.org/package=Hmisc.

Hirsch, J., Winters, M., Clarke, P. & McKay, H. 2014. Generating GPS activity spaces that shed light upon the mobility habits of older adults: A descriptive analysis. *International journal of health geographics*. 13(1):51.

Howarth, C.C. & Polyviou, P. 2012. Sustainable travel behaviour and the widespread impacts on the local economy. *Local Economy*. 27(7):764–781.

Järv, O., Ahas, R. & Witlox, F. 2014. Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*. 38:122–135.

Li, R. & Tong, D. 2016. Constructing human activity spaces: A new approach incorporating complex urban activity-travel. *Journal of Transport Geography*. 56:23–35.

Lucas, K. 2012. Transport and social exclusion: Where are we now? *Transport Policy*. 20:105–113.

McKinney, W. 2010. Data structures for statistical computing in Python. In *Proceedings of the 9th Python in Science Conference*. 51–56.

Meurs, H. & Haaijer, R. 2001. Spatial structure and mobility. *Transportation Research Part D: Transport and Environment*. 6(6):429–446.

Neutens, T., Schwanen, T. & Witlox, F. 2011. The prism of everyday life: Towards a new research agenda for time geography. *Transport Reviews*. 31(1):25–47.

Newsome, T.H., Walcott, W.A. & Smith, P.D. 1998. Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and space dimensions. *Transportation*. 25(4):357–377.

OpenStreetMap Contributors. 2016. *Planet Dump [Datafile from 26/07/2016 of BBBike extracts].* [Online], Available: http://extract.bbbike.org/ [2016, July 26].

Patterson, Z. & Farber, S. 2015. Potential path areas and activity spaces in application: A review. *Transport Reviews*. 35(6):679–700.

Perchoux, C., Chaix, B., Cummins, S. & Kestens, Y. 2013. Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place*. 21:86–93.

Prelipcean, A.C. 2016. *Capturing travel entities to facilitate travel behaviour analysis: A case study on generating travel diaries from trajectories fused with accelerometer readings*. Published licentiate thesis. Stockholm, Sweden: Royal Institute of Technology (KTH).

R Core Team. 2016. *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing.

Rai, R., Balmer, M., Rieser, M., Vaze, V., Schönfelder, S. & Axhausen, K. 2007. Capturing human activity spaces: New geometries. *Transportation Research Record: Journal of the Transportation Research Board*. 2021:70–80.

Sanjust di Teulada, B., Meloni, I. & Spissu, E. 2017. The influence of activity-travel patterns on the success of VTBC. *International Journal of Urban Sciences*. In press.

Schönfelder, S. & Axhausen, K.W. 2003. Activity spaces: Measures of social exclusion? *Transport Policy*. 10(4):273–286.

Sherman, J., Spencer, J., Preisser, J., Gesler, W. & Arcury, T. 2005. A suite of methods for representing activity space in a healthcare accessibility study. *International Journal of Health Geographics*. 4(1):24.

Taylor, M. & Ampt, E. 2003. Travelling smarter down under: Policies for voluntary travel behaviour change in Australia. *Transport Policy*. 10(3):165–177.

Tribby, C., Miller, H., Brown, B., Smith, K. & Werner, C. 2017. Geographic regions for assessing built environmental correlates with walking trips: A comparison using different metrics and model designs. *Health & Place*. 45(June 2016):1–9.

Walters, J. 2008. Overview of public transport policy developments in South Africa. *Research in Transportation Economics*. 22(1):98–108.

Xu, Y., Shaw, S., Zhao, Z., Yin, L., Lu, F., Chen, J., Fang, Z. & Li, Q. 2016. Another tale of two cities : Understanding human activity space using actively tracked cellphone location data. *Annals of the Association of American Geographers*. 106(2):489–502.

Zhang, Y., Stopher, P. & Halling, B. 2013. Evaluation of south-Australia's TravelSmart project: Changes in community's attitudes to travel. *Transport Policy*. 26:15–22.

Zou, Z., Yu, Z. & Cao, K. 2016. An innovative GPS trajectory data based model for geographic recommendation service. *Transactions in GIS*. In press.

Appendix 6.A – Tracklog screenshots

Figure 6.7 | Tracklog on iOS

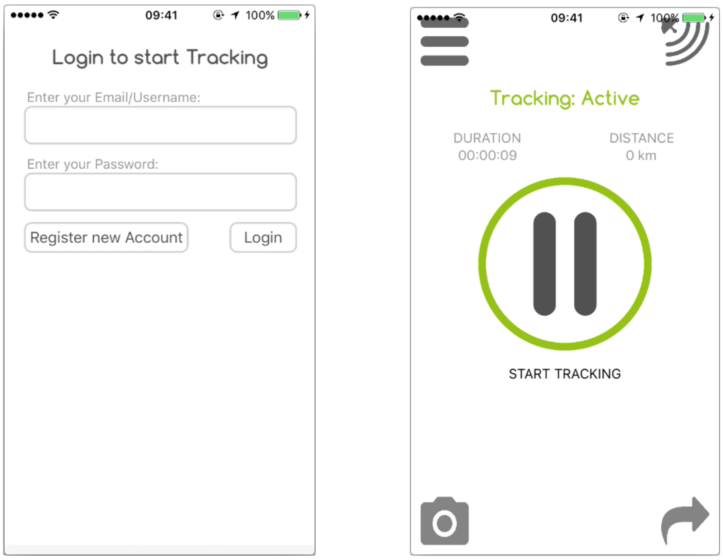
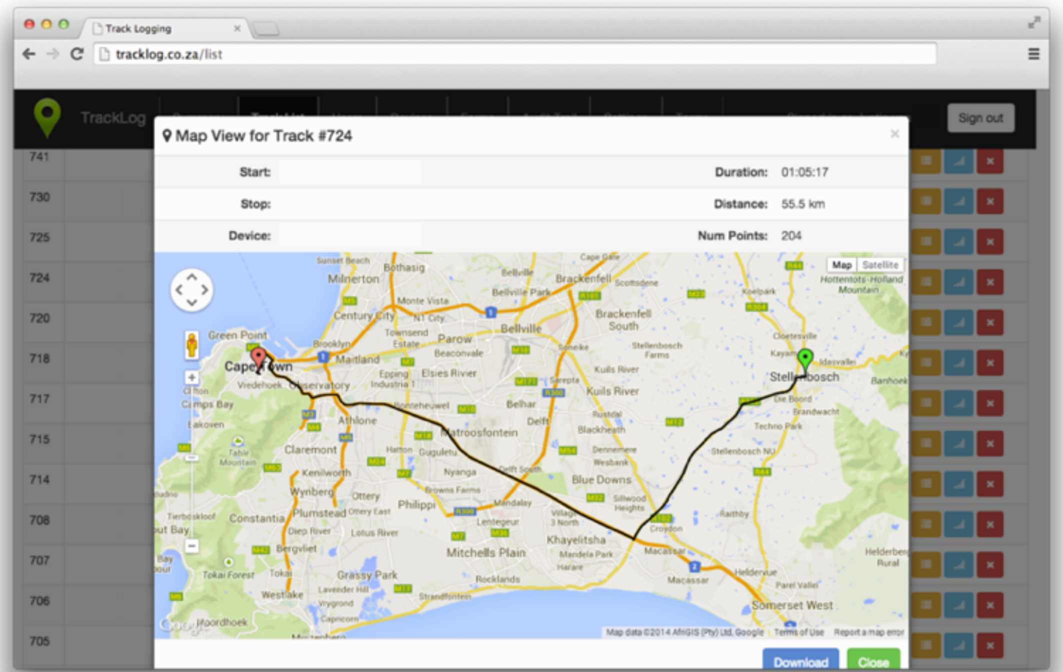


Figure 6.8 | Interface server back-end of Tracklog application



Chapter 7. Targeting voluntary travel behaviour change interventions by assessing spatiotemporal access to goods and services

Van Dijk, J.T., 2017, Targeting voluntary travel behaviour change interventions by assessing spatiotemporal access to goods and services. *Submitted for publication. Under review.*

Abstract

To stimulate people to use more sustainable modes of travel, it is essential to take their spatiotemporal opportunities to access goods and services into account. In this chapter, we use the data of a travel survey and a two-day tracking experiment, to explore how a spatiotemporal analysis of available opportunities could aid the design of a voluntary travel behaviour change intervention. For 636 respondents, individual levels of accessibility were calculated using space-time accessibility (STA) analysis. By relating access to opportunities to the respondents' self-reported willingness to consider cycling or walking to and from campus, the results indicate that there is an important difference between the respondents who are willing to consider alternative modes of travel, but are not able to, and the respondents who are willing and able to. For the design of a VTBC intervention, this is essential information because effort and resources should rather be targeted to the latter group.

Keywords

Travel behaviour change; Space-time accessibility; Potential path area; GPS; Smartphones

7.1 Introduction

Non-motorised transport (NMT) plays an important role in a sustainable transport system, because "non-motorised travel is seen as an inexpensive, efficient, and healthy mode of travel for covering short distance when compared to automobiles" (Lundberg & Weber, 2014: 165). As such, around the world, much research has been done on policies and schemes that try to stimulate people to move from their cars to more sustainable forms of transport, including NMT options and public transport alternatives. In the last decade, voluntary travel behaviour change (VTBC) interventions have been explored as a possible way of promoting sustainable transport. VTBC interventions aim to achieve a behavioural change by giving travellers feedback on their current behaviour and by informing travellers about possible travel alternatives (Taylor & Ampt, 2003). VTBC interventions come in many forms. Some examples include travel plans that encourage commuting employees not to use their private car; school travel plans that encourage parents not to bring their children to school by private car; ride-sharing schemes; and travel awareness campaigns (Bamberg & Möser, 2011). VTBC interventions thus tend to focus on breaking recurring travel patterns.

To date, several studies have attempted to evaluate the effectiveness of VTBC interventions. However, the outcomes have been mixed (Bonsall, 2009; Brög, Erl, Ker, Ryle & Wall, 2009; Chatterjee, 2009; Zhang, Stopher & Halling, 2013). One possible reason for this is that to understand individual responses to spatial policies an understanding of individual travel behaviour is required. Particularly, an individual's spatiotemporal opportunities to access goods, services, and travel alternatives play a central role in their response to a spatial policy (Kingham, Dickinson & Copsey, 2001). Yet, the consideration of these opportunities seems to have been largely neglected in the design of travel behaviour change interventions. In addition, one cannot look only at the overall number of opportunities in a neighbourhood or city, because "people move throughout the day, changing their levels of access to various goods and services as they participate in their daily activities" (Widener, Farber, Neutens & Horner, 2015: 72).

One way to account for differing levels of access to goods and services between individuals is to turn to space-time accessibility (STA) measures. STA measures do not consider a static urban environment, but rather take the individual as a unit of analysis. As such, these measures are capable of modelling an individual's spatiotemporal opportunities to access goods and services, given their personal space-time constraints. Although STA measures are computationally intensive and data-hungry, the increased capacity in computational power and fast-paced developments in the domain of GIS technology and location-aware technologies over the past decade have led many scholars to reconsider the use of STA measures in (geographical) research (cf. Kwan & Weber, 2003; Miller, 2007; Neutens, Schwanen & Witlox, 2011; Neutens, Delafontaine, Schwanen & van de Weghe, 2012).

Using data from a travel survey amongst staff and students of Stellenbosch University, augmented with two-day GPS tracking data, the present chapter explores how a spatiotemporal analysis of opportunities to access goods and services could be used to aid in the design of a VTBC intervention. In addition, GPS data are used to augment and contextualise the access to spatiotemporal opportunities. This chapter starts with a brief introduction to the existing literature on spatiotemporal accessibility, and its importance to studies concerned with travel behaviour change interventions. This is followed by a description of the data collection, data preparation, and STA analysis. The conclusion and discussion draw attention to the importance of considering spatiotemporal analysis in the design of a VTBC intervention.

7.2 Spatiotemporal accessibility

A key concept in geography and transport studies is accessibility. Accessibility is often defined as the "ease with which activities may be reached from a given location using a particular transportation system" (Morris, Dumble & Wigan, 1979: 91). Accessibility has often been measured on an aggregate level, such as a neighbourhood, census tract, or city. These aggregate accessibility measures, or place-based accessibility measures, typically consist of a reference location, a set of opportunities, and a measure of impedance between the reference location and the set of opportunities (Schwanen & De Jong, 2008). Two major critiques on place-based accessibility measures are that they do not capture the complexity of activity behaviour, and that they do not account for individual differences in space-time constraints and instead assign every individual the same level of accessibility.[1] As such, "place-based measures [are] insensitive to differences in accessibility between individuals and between sociodemographic segments within the population" (Neutens *et al.*, 2011: 29).

[1] An additional issue is the modifiable areal unit problem (MAUP), which entails that the outcomes of a geographical analysis ultimately depends on the unit of analysis that is used (Tribby *et al.*, 2016).

As opposed to place-based accessibility measures, person-based accessibility measures do account for individual differences, as they take the individual as a unit of analysis. The analysis of individual activities in space and time have their basis in time geography (Hägerstrand, 1970). In time geography, human spatial behaviour is understood to derive from the willingness or necessity of individuals to partake in activities. An important aspect of time geography is that it recognises that activities have both a spatial dimension (a location) and a temporal dimension (a duration) (Miller, 2007). This implies that individuals must carefully allocate their time to realise their preferred activity schedule; all of which depends on their available time for the activities and their available modes of transport. In addition, individuals have to account for at least three constraints on their behaviour. In his seminal work "What about people in regional science", Hägerstrand (1970) identified capability constraints, coupling constraints and authority constraints as the constraints governing spatiotemporal behaviour. Capability constraints refer to physiological and cognitive issues such as the need for sleep, shelter, and food. Coupling constraints play a role when individuals need or want to come together, and their individual space-time paths have to be synchronised. Authority constraints are those laws, rules, and norms that prohibit, inhibit, or steer activity participation, for example, the trading hours of a supermarket.

Because of the focus on the individual and the explicit attention to the context in which activity-travel behaviour is embedded, time geography is a powerful framework for understanding and analysing spatiotemporal behaviour (Kwan, 1999, 2013; Kwan & Weber, 2003; Widener *et al.*, 2015). Accessibility measures that originated from time geography are typically referred to as STA measures. These measures are based on a key time-geographical concept: the space-time prism. Given two activities, the space-time prism is the three-dimensional graphical representation of all possible paths that an individual could take between the end time of one activity and the start time of a second activity. The size and shape of the prism depend on the network distance between the two activities, the available time between the two activities, and the speed of the mode of transport being used (Neutens *et al.*, 2011). In turn, the two-dimensional derivative of the space-time prism, the potential path area (PPA), represents the area that an individual could potentially visit during these two activities (Schwanen & De Jong, 2008). The extent of the PPA, therefore, can be equated to an individual's opportunity to access activity locations or the potential for spatial interaction (Casas, Horner & Weber, 2009; Patterson & Farber, 2015).

PPA's can be considered as an important tool to quantitatively analyse prospective opportunities – opportunities that are crucial to consider when trying to stimulate travel behaviour change and achieve a modal shift (Dijst, 1995). From the perspective of stimulating more sustainable ways of travelling, the above discussion suggests that one's individual accessibility and constraints should be considered when designing, targeting, and evaluating a travel behaviour change intervention. Table 7.1, for instance, shows the different considerations that could play a role in the decision-making process of an individual before, during, and after a journey. In each of these stages, a possible travel behaviour change intervention is imaginable that could be part of a VTBC programme. Examples are suggesting an alternative travel mode (before a journey), proposing a less-congested route (during a journey), and providing information on CO^2 emissions (after a journey). However, if the spatiotemporal constraints of an individual do not allow for a certain travel behaviour change, that change is guaranteed to be unsuccessful.

Table 7.1 | Influencing the decision-making process of traveller

Before a trip	During a journey	After a journey
Choice of destination Necessity of the trip Choice of travel mode Choice of time	Choice of route Driving style	Feedback on costs of trip Information on alternatives

Source: Adapted from Howarth and Polyviou (2012: 768)

With STA analysis, it is possible to assess the feasibility of a particular activity programme (e.g. doing your shopping directly after work rather than making an additional trip) or the feasibility of using a different travel mode (e.g. suggesting walking or cycling as an alternative to the private car). Accordingly, over the years, many implementations of STA analysis (PPA analysis) have been implemented in various fields of research to understand spatial behaviour. (For a recent, extensive review on PPAs and activity spaces, see Patterson & Farber, 2015.) The consideration of spatiotemporal accessibility could, therefore, be brought into studies focusing on VTBC interventions. Moreover, technological advancements have now made it easier to collect accurate data on revealed individual spatiotemporal behaviour. Whereas researchers used to have to rely on self-reported activity and travel information gathered by pen-and-paper diaries or computer-administrated surveys, researchers can now capitalise on the opportunities provided by location-aware technologies such as GPS devices and smartphones (cf. Bohte & Maat, 2009; Nitsche, Widhalm, Breuss, Brändle & Maurer, 2014; Shoval, Kwan, Reinau & Harder, 2014).

7.3 Data collection and data preparation

To explore how a spatiotemporal analysis of opportunities to access goods and services could aid in the design of a VTBC intervention, a data set was used that was collected at Stellenbosch University, South Africa. With around 3,300 staff members and around 30,000 enrolled students in various degree programmes, Stellenbosch University is a key generator of travel demand in the area (Stellenbosch University, 2016). The data collection consisted of two phases in which a total number of three research instruments were used. In the first phase, a household travel survey was distributed, which yielded 853 valid responses. In the second phase, respondents were asked to track their movements with a GPS smartphone tracking application (Tracklog) and fill out an activity-trip diary for two consecutive days. The first day of tracking yielded 151 tracks, and on the second day of tracking, 141 tracks were recorded.[2]

[2] The data collection is described in more detail in Chapter 6.

7.4 Construction of potential path areas

A PPA model requires at least three inputs: a set of primary anchors, a time window, and a set of activity locations. Primary anchors are typically activity episodes that are mandatory or cannot be rescheduled, such as a job. The time window is the time between two consecutive primary anchors (i.e. the end time of the first primary anchor and the start time of the second primary anchor). The set of activity locations represents the choice universe an individual has, i.e. the locations with potential for spatial interaction. We used a PPA model with two primary anchors: home and work. The respondents' home and work locations were, wherever possible, geocoded from the household travel questionnaire they had completed. This resulted in 636 home and work location combinations. The choice universe was operationalised in two ways; with a hexagonal tessellation of the study area, and with point-of-interest (POI) data extracted from OpenStreetMap (OpenStreetMap Contributors, 2016).

The hexagonal choice universe was created by tessellating the larger Cape Town metropolitan area, in which Stellenbosch is situated, with hexagons of 0.1 square kilometres (each hexagon with an edge length of 200 metres). This yielded roughly 27,500 hexagons for the entire area that constituted the destination choice universe. To reduce the computational burden, it was decided not to increase the spatial resolution. The full tessellation was subsequently intersected with the built-up land use, enlarged with a 1000-metre buffer around the road network segments. The land use data were extracted from the national land-cover data set of South Africa (Department of Environmental Affairs, 2015; GeoTerraImage, 2015). The intersection of this buffer network with the tessellation provided the area for potential activity and travel engagement. Figure 7.1 shows the full tessellation and the area for possible spatial interaction.

Whereas the urban form may be insightful, on its own it does not account for the distribution of activity destinations within a built-up area, nor does it take temporal constraints on the activity location side into account. To limit the choice universe to actual activity locations, POI data were extracted from OpenStreetMap (OSM). OSM is a user-generated geographic data set that is available worldwide under the Open Database License, and is arguably the most complete data set of the area that is freely available (OpenStreetMap Contributors, 2016). All relevant POIs were downloaded for the large Cape Town metropolitan area (n = *5,782*) and subsequently categorised into 13 discrete amenity categories. Furthermore, typical trading hours were estimated for each amenity category. Table 7.2 shows the different categories, examples of the types of amenities that belong to each category, the number of extracted POIs for each category, and the assigned trading hours to be used in the analysis. Three amenity categories were not assigned any trading hours because one would normally not visit the activity locations within these categories daily (e.g. museum, hotel) or because one would not simply consider an alternative activity location than the usual activity location even if one is closer (e.g. church). These amenity categories were thus excluded from further analysis.

Figure 7.1 | Tessellated study areas with selected hexagons (inset: Stellenbosch)

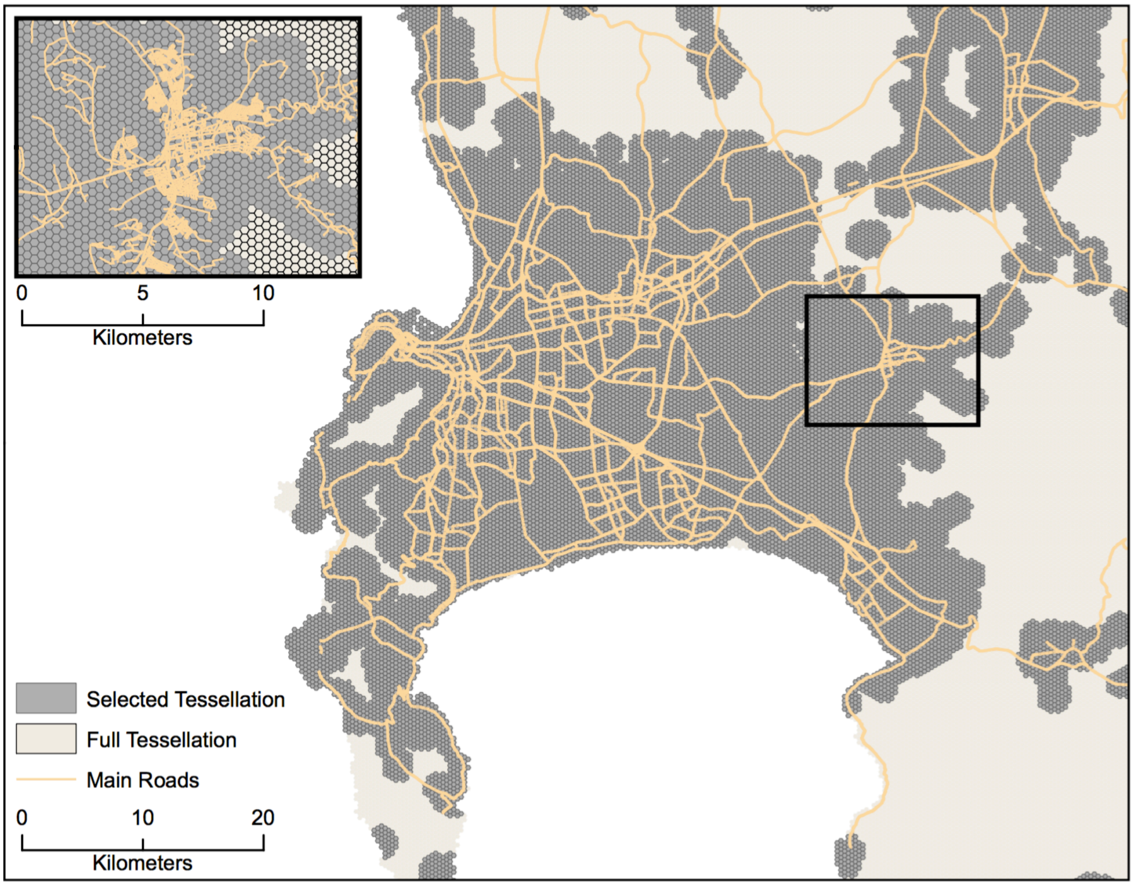


Table 7.2 | OpenStreetMap classification and assigned trading hours

Category	Examples	POIs	Trading hours
Government	Police station, town hall	141	08h00-16h00
Educational	School, university	750	08h00-16h30
Places of worship	Mosque, church	301	N/A
Groceries	Supermarket, butcher, bakery, mall	583	08h00-20h00
Shops	Department store, music store, sport clothing store	343	08h00-20h00
Special	Electrician, garden centre, mechanical workshop	182	08h00-17h00
Health	Doctor, hospital, clinic	117	08h00-17h00
Outgoing	Pub, restaurant, café, winery	963	12h00-24h00
Sport	Gym, sport fields	801	08h00-22h00
Transport	Train station, bus station	64	N/A
Other amenities	Library, post office, bank	701	08h00-16h30
Other leisure	Museum, hotel, guest house	570	N/A
Other	Other	356	N/A

Time windows were, where possible, directly derived from the household travel survey. A time window was established for those respondents who answered a question in the survey pertaining to the respondent's interest in a possible home-campus shuttle service. For the mornings, the time window for staff members was derived from the self-reported earliest feasible pick-up time from home, together with the time the respondent needed to be on campus. For the evenings, this was done by comparing the time the respondent is typically done on campus and the latest feasible drop-off time at home. For the students, the times between the earliest feasible and the latest feasible pick-up and drop-off times were used for the morning and evening time windows, respectively. For the respondents for whom a part, or all, of this information was not available, a morning time window of 60 minutes and an evening time window of 120 minutes were selected; similar to the time window that Widener *et al.* (2015) used, but with a more conservative morning time window estimate.

To allow for a network-based PPA, in which the space-time prism of a trip is constrained by network topology, speed, and connectivity (Tribby, Miller, Werner, Smith & Brown, 2016), the digital road network of the area (including average speed information and directionality for each road segment) was supplemented with a constant speed of 15 kilometres/hour and 4 kilometres/hour to facilitate a bicycle network and a walking network, respectively (Salonen, Broberg, Kyttä & Toivonen, 2014). This network was used to calculate the distances (measured in travel time) between the home location of every respondent and all destinations in both the hexagonal and POI choice universes. This process was repeated for all three travel modes, and saved in individual distance matrices. The same was done for the work location of every respondent. As such, every distance matrix contained all "the mode-specific, network-based travel times" (Schwanen & De Jong, 2008: 562) between the origin (i.e. home and work) and all destinations (i.e. activity hexagons and POIs).

After the primary anchors, the choice universes, the time windows, and the distance matrices were established, the actual calculation of the PPA consisted of two additional steps: the calculation of the future cone of the space-time prism – the hexagons or POIs that can be reached from the first primary anchor (given the travel times and time window), and the calculation of the past cone of the space-time prism – the hexagons or POIs from which the second primary anchor can be reached (given travel times and time window) (Schwanen & De Jong, 2008). In turn, the mode-specific PPA was computed by intersecting these two cones. The PPA thus shows the area with which individuals can interact, constrained by their time window, and if an interaction is possible, the time an individual can interact with that location.

7.5 Potential path area analysis on a hexagonal and POI choice universe

Once the PPA model was in place for all 626 cases, the potential path areas were calculated for different travel modes. Figure 7.2 gives an example of the hexagonal choice universe a respondent could reach by bike and by foot, given a reported time window of 60 minutes. In addition, it is shown how much time the respondent could spend at each location (i.e. a potential interaction) without being too late for the second primary anchor (in this case, work). If it is assumed that the respondent normally travels to work by car, the figure also suggests that this commuter would be able to substitute the car trip by another mode, such as cycling or walking. In this case, intermediate activity dwell time remains sufficient (indicated by green shaded hexagons) to allow for activity engagement. Figure 7.3, on the other hand, illustrates a different case. This respondent only reported having a time window of 30 minutes. If the respondent travels by car, s/he has some opportunities to engage in activities such as dropping off his/her children at school, but if s/he would travel by bike, the options to engage in additional activities become very limited. In addition, travelling by foot is for this respondent not possible: given the distance between home and work, as well as the time window, s/he would not be able to be in time for work, let alone engage in other activities.

The disadvantage of using the hexagonal choice universe is that a level of accessibility is assigned to all hexagons in the PPA, irrespective of whether there is actual potential for spatial interaction (i.e. an activity location). Figure 7.4 shows the PPA of the same individuals with the 60-minute time window, but the hexagonal choice universe is now replaced with the POI choice universe. For the different activities in the POI choice universe, the PPA now includes only actual activity locations. All coloured POIs correspond with the time that is available for interaction, whilst all black POIs are spatially inaccessible. If a respondent were to travel by bike, s/he would have access to quite a few activity locations in the city centre with which s/he could interact for more than 40 minutes. However, the trading hours of the POIs are not yet considered. To account for temporal accessibility, in Figure 8 the same situation is visualised, but now the trading hours (as defined in Table 7.2) of the POIs are considered. A comparison between Figure 7.4 and Figure 7.5 reveals that, especially in the city centre, several POIs turn out to be inaccessible. In Figures 7.6 and 7.7, on the other hand, the situation for the individual who reported to have a time window of 30 minutes is visualised. Again, the first figure only takes the spatial properties of the POI universe into account, whereas the second figure also accounts for the temporal properties. Figure 7.6 suggests that the respondent has quite some time in his/her morning commute to interact with a variety of locations for 20 to 30 minutes when using a bike. When walking to work, s/he still has several possibilities for shorter interactions of 10 to 20 minutes. However, Figure 7.7 indicates that none of the opportunities is temporally accessible (i.e. they are not open yet).

The comparison of the hexagonal activity destinations, POI activity destinations, and POI activity destinations that take trading hours into account for the two examples reveal that the individual levels of accessibility show a high degree of variability. Moreover, it draws attention to the importance of including a temporal element in accessibility analysis. For the design of a VTBC intervention, this information is crucial. When we look at the two cases that have been presented so far, it is very unlikely that trying to convince the second respondent to use the bicycle to get to work will be successful. For the first respondent, on the other hand, an intervention would be more likely to succeed from a spatiotemporal perspective. It is not only the larger time window that makes the first respondent a better candidate for an intervention; the start time of the second primary anchor (i.e. work) is also significant. Whereas the second respondent reports to have to start working at 07h30, the first respondent indicates s/he only has to be on campus at 09h00. Given the fact that most activity locations only open at 08h00, they are by definition not accessible to the second respondent without making him/her late for work.

Figure 7.2 | Network-based potential path area and possible time for activity engagement when travelling by bike and by foot (inset) with a morning time window of 60 minutes. Hexagonal choice universe.

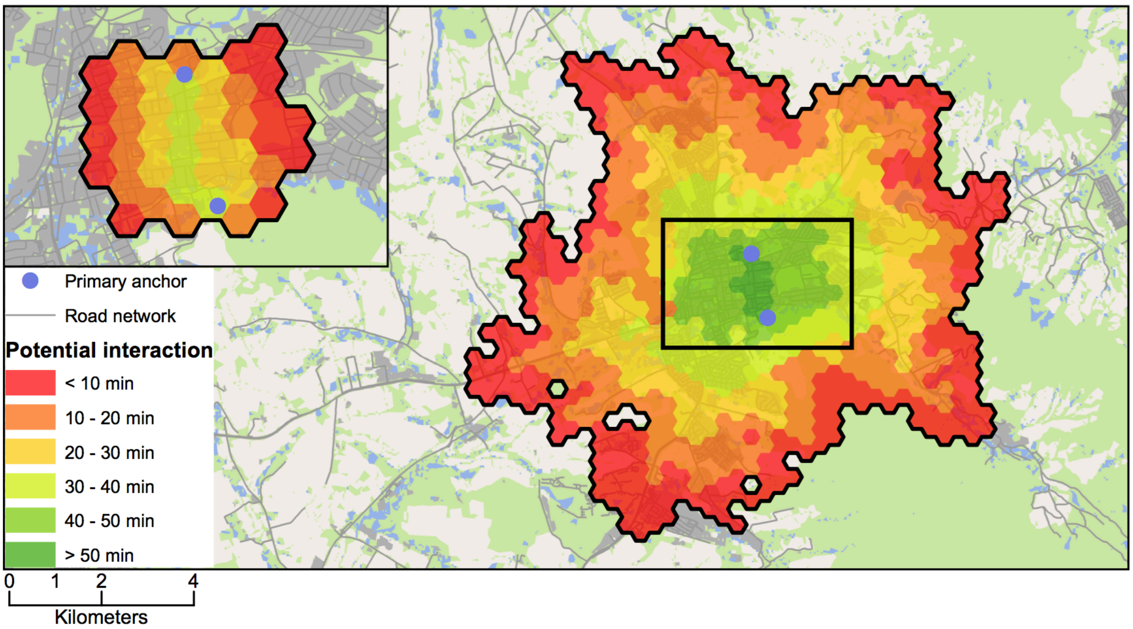


Figure 7.3 | Network-based potential path area and possible time for activity engagement when travelling by car and by bike (inset) with a morning time window of 30 minutes. Hexagonal choice universe.

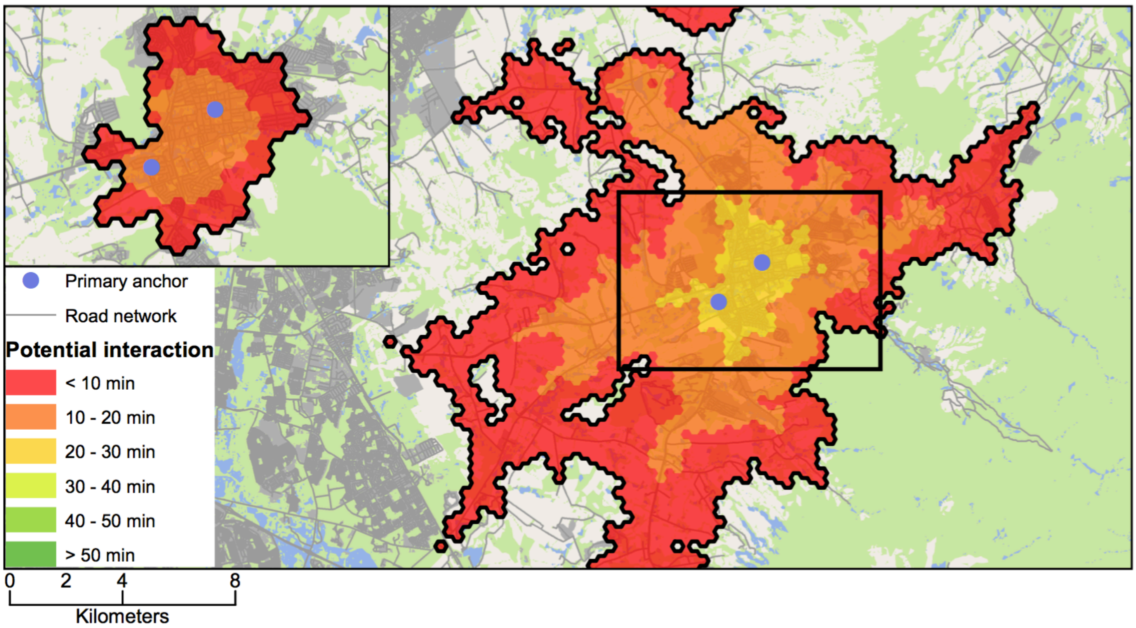


Figure 7.4 | Network-based potential path area and potential for spatial interaction when travelling by bike and by foot (inset) with a morning time window of 60 minutes. POI choice universe.

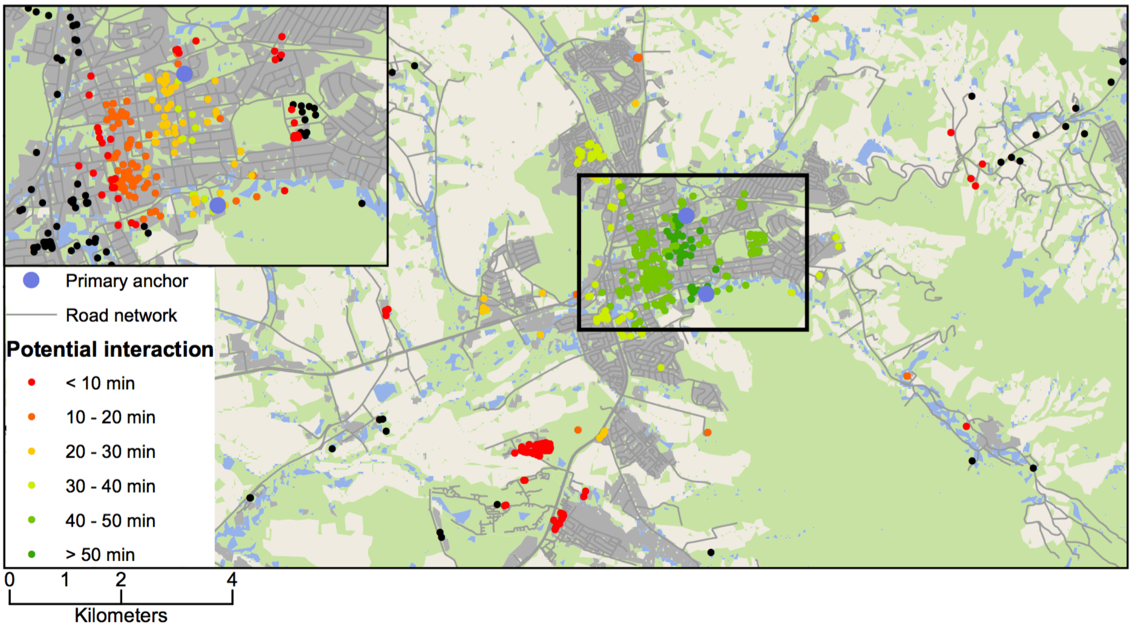


Figure 7.5 | Network-based potential path area and potential for spatial interaction when travelling by bike and by foot (inset) with a morning time window of 60 minutes. POI choice universe accounting for trading hours.

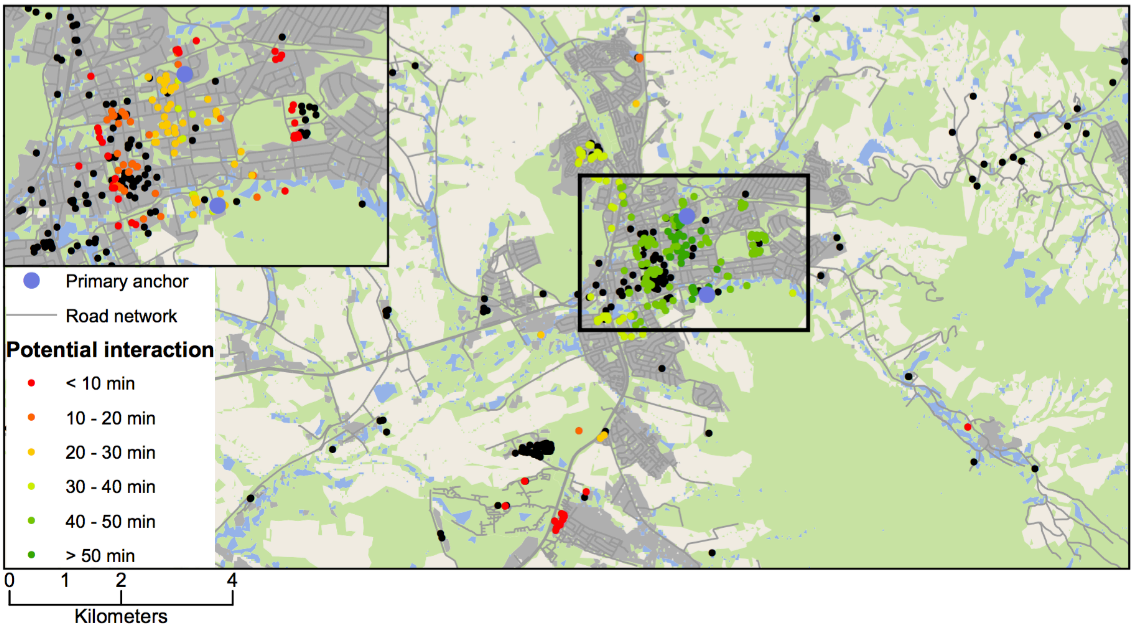


Figure 7.6 | Network-based potential path area and potential for spatial interaction when travelling by bike and by foot (inset) with a morning time window of 30 minutes. POI choice universe.

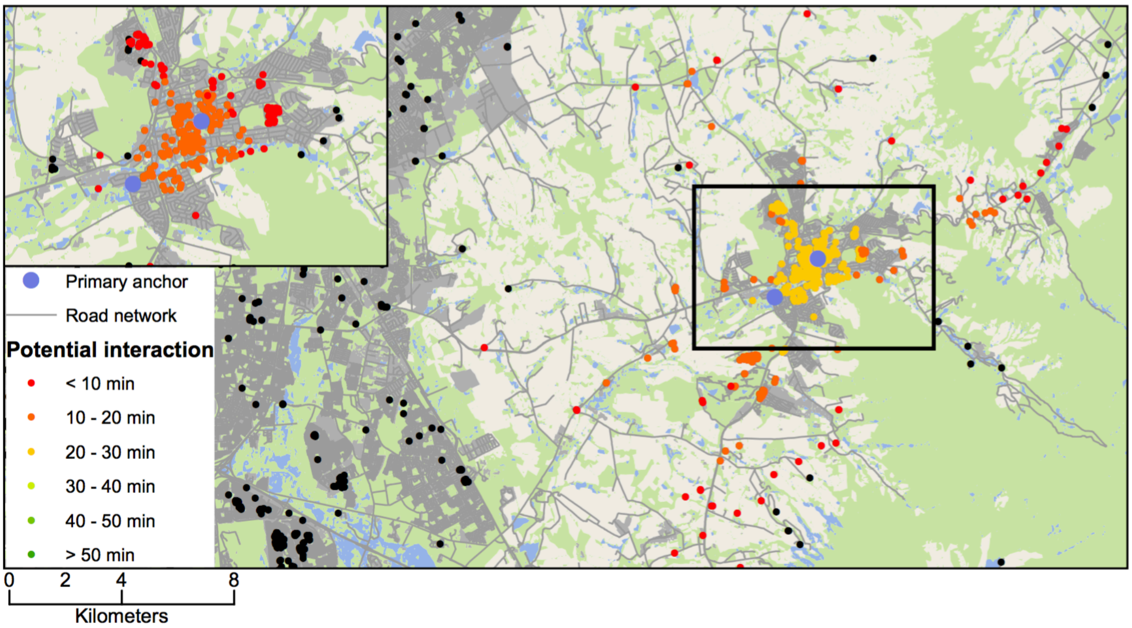
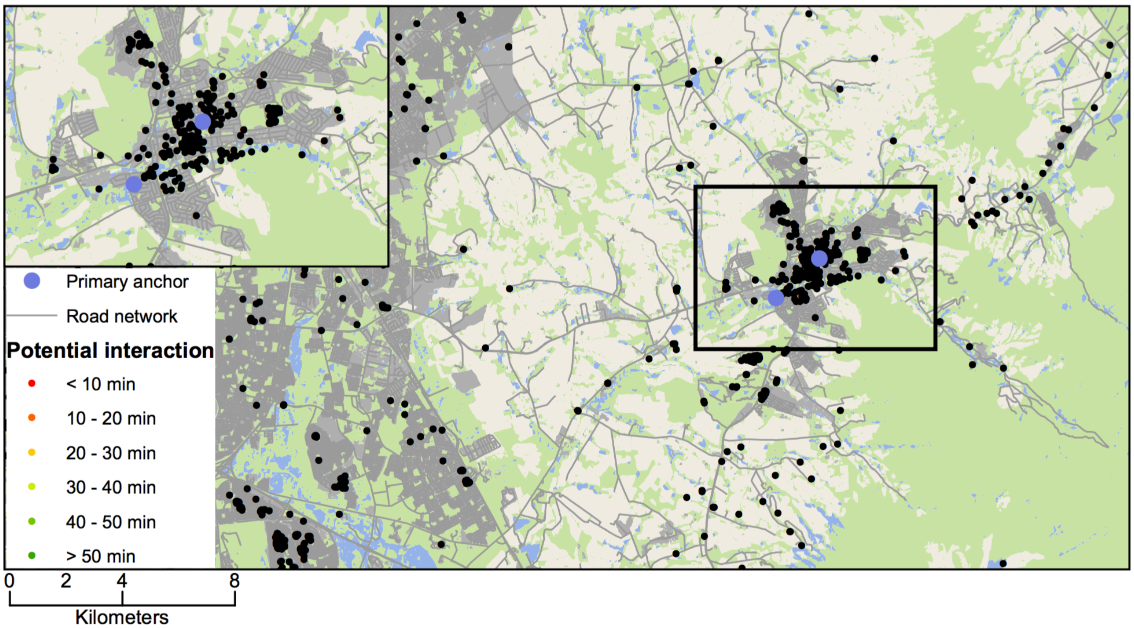


Figure 7.7 | Network-based potential path area and potential for spatial interaction when travelling by bike and by foot (inset) with a morning time window of 30 minutes. POI choice universe accounting for trading hours.



The importance of considering individual spatiotemporal accessibility to opportunities is further illustrated by Table 7.3. The table shows the average number of opportunities one can access in the morning (AM) and in the evening (PM) for the different sets of activity locations if a PPA for the mode of transport could be calculated. In addition, the average time for activity engagement is given. For the hexagonal choice universe (HEX), for example, 623 respondents have access to, on average, 11,271 locations in the morning. However, if the PPA is modelled with the bicycle as transport mode, it turns out that the constraints of only 437 respondents allow for spatial interaction. For the POI choice universe that takes trading hours into account (TIME), the numbers drop significantly. Only 123 respondents can engage in some form of activity in the morning when travelling by bike. Interestingly, the average number of accessible points in the evening commute is higher than the average number of accessible points in the morning commute across the board. This is not entirely surprising, because most people have more time available after work than before work. In addition, for the TIME choice universe, the trading hours play an important role, because many activity locations will still be open at the end of the day.

Table 7.3 | Average number of accessible opportunities and average available activity time within time window for available PPA

		PPA Car			PPA Bike			PPA Walk		
		Points	Time	*N*	Points	Time	*N*	Points	Time	*N*
HEX	AM	11,271	29	623	1,618	24	437	230	24	294
	PM	17,624	77	616	4,260	49	506	541	49	369
POI	AM	2,713	30	623	464	38	437	201	36	295
	PM	4,690	70	616	923	70	505	286	85	369
TIME	AM	2,533	55	145	493	51	123	163	70	90
	PM	2,647	68	615	576	64	503	166	76	367

Again, from the perspective of designing a VTBC intervention, the spatiotemporal accessibility measures suggest that interventions should be targeted towards individuals who have a higher level of spatiotemporal freedom. In fact, some people may express a willingness to change behaviour but are not actually able to (Kingham *et al.*, 2001). As such, two questions from the household travel questionnaire pertaining to the willingness of the respondent to consider cycling or walking to and from campus were selected and related to the availability of a PPA modelled on a commute by bicycle and a commute by foot. The results from the respondents who answered these questions are shown in Table 7.4. The columns refer to the availability of a PPA in the morning and in the evening, whereas the rows show how many respondents were willing to consider cycling or walking (Yes), how many respondents were not willing to consider cycling or walking (No), and how many respondents stated that cycling or walking is not an option (N/O).

Moving to the TIME choice universe, Table 7.4 suggests that a PPA could be modelled using a bicycle as mode of transport for 43 respondents, of which 19 respondents would be willing to consider cycling. On the other hand, for the 37 respondents who did indicate that they would be willing to consider cycling, the PPA set returned empty. These respondents may be willing to consider a more sustainable mode of transport, but may not be able to because of their spatiotemporal constraints. What is interesting is that 19 respondents also indicated that cycling is not an option, whereas from the person-based accessibility indicator (TIME), cycling should be an option for them. For the evening commute, there were even 93 respondents who stated that cycling is no option, whilst a PPA is in fact available. For the PPAs that have been modelled on accessibility by foot, on the other hand, the differences are less pronounced. Out of a total of 236 respondents who reported that walking is not an option, 225 and 181 respondents do indeed have no PPA available in the morning and in the evening, respectively.

Table 7.4 | Access to goods and services versus willingness to consider cycling and walking to travel to and from campus

		PPA Bike (AM)		PPA Bike (PM)		PPA Walk (AM)		PPA Walk (PM)	
		Yes	No	Yes	No	Yes	No	Yes	No
HEX	Yes	49	7	46	10	11	6	14	3
	No	27	17	34	10	9	22	15	16
	N/O	56	128	95	89	27	209	56	180
POI	Yes	49	7	46	10	11	6	14	3
	No	27	17	34	10	10	21	15	16
	N/O	56	128	94	90	27	209	56	180
TIME	Yes	19	37	46	10	6	11	14	3
	No	5	39	34	10	3	28	15	16
	N/O	19	165	93	91	11	225	55	181

Although there could be different reasons for the respondents to indicate that cycling or walking is not an option when a PPA is available (household commitments, for instance), it could be wondered whether the number of accessible activity locations could partly explain these differences. As such, the average number of opportunities for the three different groups (yes, no, no option), for the respondents that have a PPA available, were compared using a one-way ANOVA, as shown in Table 7.5. However, whereas in most cases the respondents who were not willing to consider cycling or walking had a smaller average of available opportunities, the tests did not return significant for any of the choice universes (HEX, POI, TIME), transport modes (bicycle, walking), or times of day (AM, PM). Despite the tests not being significant, the outcomes are still relevant. While the previous results indicated that spatiotemporal accessibility may be a necessary condition to consider in the design of a VTBC intervention, the outcomes of the tests clearly indicate that it is not a sufficient condition. In the context of a household, for instance, the juggling of responsibilities by parents could play a role. Another factor that is very likely to be important is safety. Females, for example, may be less willing to cycle or walk to and from campus at night.

Table 7.5 | Average number of available opportunities for spatial interaction versus willingness to consider cycling and walking to travel to and from campus[1]

		PPA Bike (AM)			PPA Bike (PM)			PPA Walk (AM)			PPA Walk (PM)		
		Avg.	Sd.	Sig.	Avg.	Sd.	Sig.	Avg.	Sd.	Sig.	Avg.	Sd.	Sig.
HEX	Yes	1,313	1,846		4,910	5,991		197	260		686	1,225	
	No	838	1,005	No	3,156	3,298	No	104	164	No	344	462	No
	N/O	2,489	4,231		3,712	5,219		918	4,100		953	2,803	
POI	Yes	385	258		1,046	1,373		207	87		262	127	
	No	319	116	No	646	602	No	144	86	No	237	102	No
	N/O	675	1,004		930	1,335		386	1,089		383	709	
TIME	Yes	317	267		664	940		164	43		176	103	
	No	228	131	No	422	500	No	150	85	No	132	78	No
	N/O	527	707		586	945		133	54		175	148	

[1] One-way ANOVA on potential for spatial interaction (in number of accessible opportunities) between 'yes', 'no', and 'N/O' (No option) groups. Outliers were removed. Shapiro-Wilk normality test is significant for all groups and as such the distribution of the residuals is not normal. Levene's test for homogeneity of variances indicates homoscedasticity for all groups. Kruskal-Wallis rank sum test also shows no significant results for any of the groups.

7.6 Potential path area analysis augmented with GPS

The previous section has shown how potential path area analysis can be used to identify the potential for spatial interaction based on geocoded data. Location-aware technologies, however, can help to further refine this choice set. Out of the 636 respondents who provided their home and work location in the questionnaire, 79 individuals also participated in the tracking experiment and had tracks of sufficient quality.[3] For these 79 individuals, an activity space was constructed. An activity space is a geometric representation of individually-revealed spatial behaviour, and can be seen as an individual's spatial footprint (Li & Tong, 2016; Xu, Shaw, Zhao, Yin, Lu, Chen, Fang & Li, 2016). The activity space conceptualisation that was chosen was the minimum convex polygon (MCP) with a 200-metre buffer.[4] In turn, this activity space was used to limit the activity locations within one's choice universe. This was done by intersecting the different mode-specific PPAs with the activity space.

Table 7.6 | Average number of accessible opportunities and average available activity time within time window for available PPA. Responses with associated GPS track.

		PPA Car			PPA Bike			PPA Walk		
		Points	Time	*N*	Points	Time	*N*	Points	Time	*N*
HEX	AM	10,938	22	77	1,442	20	46	944	20	25
	PM	16,045	59	76	3,598	37	60	927	43	35
POI	AM	2,736	23	77	464	35	46	386	26	26
	PM	4,268	60	76	791	56	59	406	73	35
TIME	AM	2,153	21	22	199	37	18	136	29	11
	PM	2,360	54	75	408	52	58	149	65	34
AS POI	AM	239	46	77	142	45	46	96	33	26
	PM	254	93	76	157	91	59	108	92	35
AS TIME	AM	133	43	22	86	43	17	68	37	11
	PM	149	81	75	92	78	58	62	79	34

Figures 7.8 and 7.9 show an example of the intersection of the GPS-based activity space with the POI destinations, and with the POI destinations in which trading hours are accounted for. The average number of accessible opportunities and average available activity time within the time window for the 79 respondents for the different choice universes are given in Table 7.6. It becomes clear from the figures and the table that the activity space effectively constrains the number of accessible POIs. The average number of POIs accessible by car (2,736) in the morning, for example, drops to 239 when intersected with the activity space (AS POI). When accounting for trading hours, the average number of accessible POIs within the activity space (AS TIME)

[3] Because of the variation of the quality of the tracking data, it was decided that for the exploration of potential path areas, only the tracks that consisted of at least 100 track points qualified. When a respondent initiated the tracking for only one day, the number of track points was evaluated directly against these criteria, whereas the tracks of the respondents who initiated the tracking for two days were first merged. The quality of the data is discussed at length in Chapter 6.

[4] See Chapter 6 for a more in-depth discussion on activity spaces and its different conceptualisations.

drops from 2,153 to 133. Although activity spaces are likely to include areas with which individuals do not necessarily interact, an activity space nevertheless provides useful information because the area "represents the conceptual (perhaps minimal) area over which we know the individual is willing or able to engage in activities" (Newsome, Walcott & Smith, 1998: 362).

Table 7.7 again relates the availability of a PPA modelled on a commute by bicycle and a commute by foot to the willingness of the respondent to consider cycling or walking. Once more, the columns refer to the availability of a PPA in the morning and in the evening, whereas the rows show how many respondents were willing to consider cycling or walking (Yes), how many respondents were not willing to consider cycling or walking (No), and how many respondents stated that cycling or walking is not an option (N/O). Interestingly, there were not many differences in the available PPAs between POI and AS POI, nor were there many differences in the available PPAs between TIME and AS Time. However, the combination of the points within the available PPAs that could potentially be used for activity engagement with the area of revealed spatial behaviour, may help in suggesting alternatives for activity destinations that could be relevant to the individual. Moreover, if multiple days of GPS data are available, this data could be used to help identify other primary anchors that could be used in modelling alternatives.

Table 7.7 | Access to goods and services versus willingness to consider cycling and walking to travel to and from campus. Responses with associated GPS track.

		PPA Bike (AM)		PPA Bike (PM)		PPA Walk (AM)		PPA Walk (PM)	
		Yes	No	Yes	No	Yes	No	Yes	No
HEX	Yes	11	1	11	1	2	1	3	0
	No	5	2	7	0	2	3	3	2
	N/O	6	25	15	16	4	38	9	33
POI	Yes	11	1	11	1	2	1	3	0
	No	5	2	7	0	3	2	3	2
	N/O	6	25	14	17	4	38	9	33
TIME	Yes	6	6	11	1	1	2	3	0
	No	0	7	7	0	0	5	3	2
	N/O	3	28	13	18	2	40	8	34
AS POI	Yes	11	1	11	1	2	1	3	0
	No	5	2	7	0	3	2	3	2
	N/O	6	25	14	17	4	38	9	33
AS TIME	Yes	6	6	11	1	1	2	3	0
	No	0	7	7	0	0	5	3	2
	N/O	2	29	13	18	2	40	8	34

Figure 7.8 | Network-based potential path area and possible time for activity engagement by bike and by foot (inset) with an evening time window of 120 minutes. POI choice universe with activity space correction.

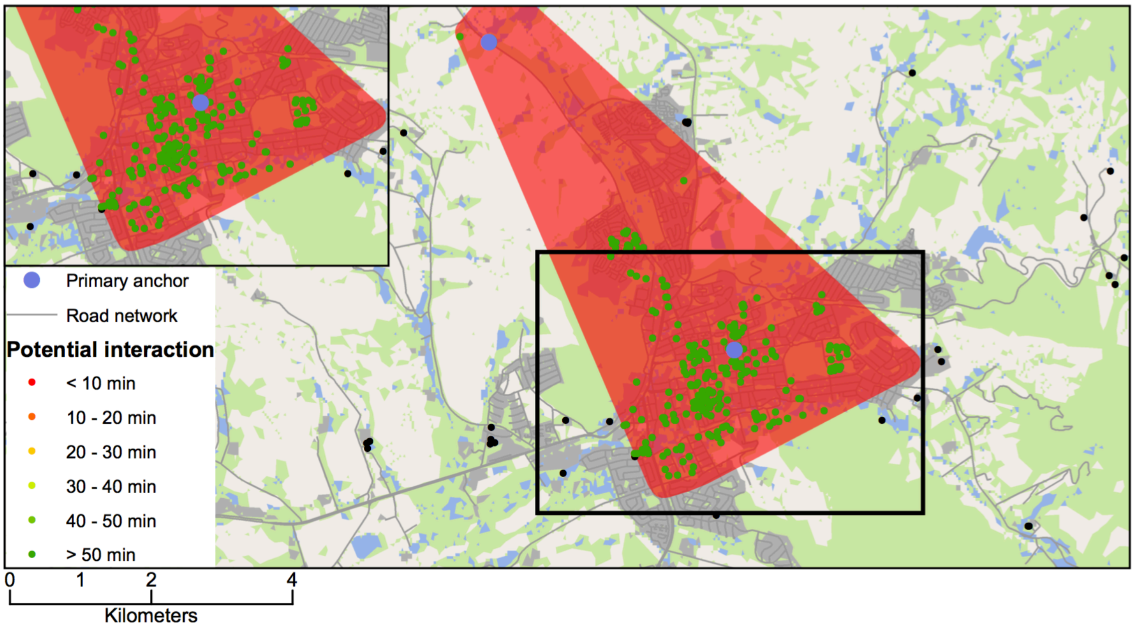
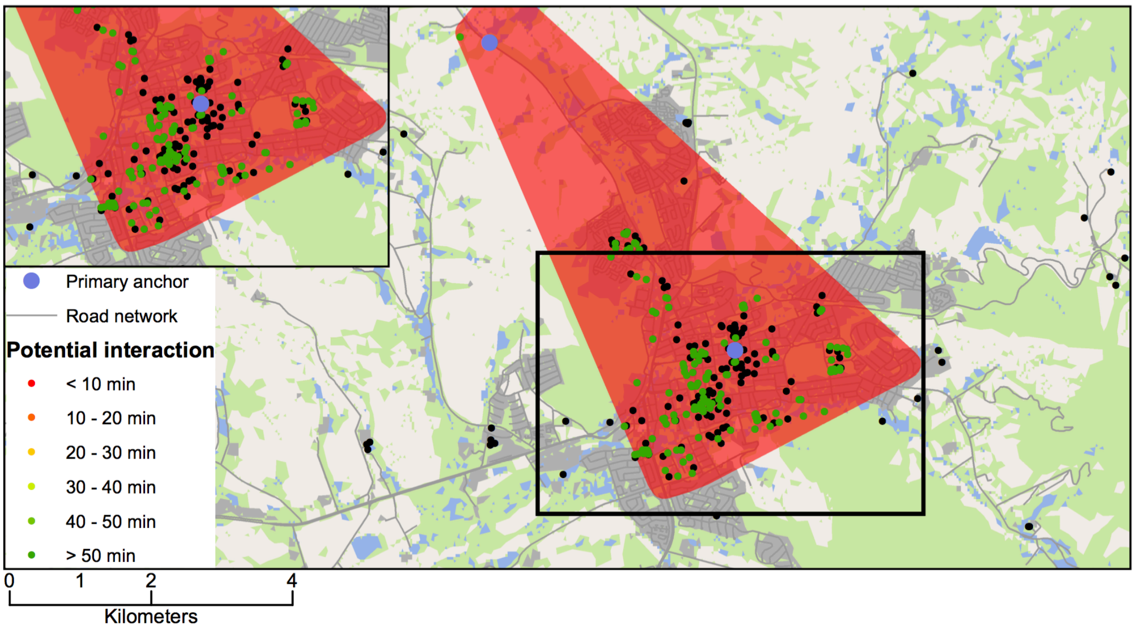


Figure 7.9 | Network based potential path area and possible time for activity engagement by car and by bike (inset) with an evening time window of 120 minutes. POI choice universe with activity space correction (accounting for trading hours).



7.7 Conclusions and discussion

This chapter set out to explore how a spatiotemporal analysis of opportunities to access goods and services could be used to aid in the design of a voluntary travel behaviour change (VTBC) intervention. Based on the examination of 636 combinations of geocoded home and work locations, accessibility indicators were calculated for different choice universes using space-time accessibility (STA) analysis. After relating access to opportunities to the respondents' self-reported willingness to consider cycling or walking to and from campus, the results indicated that there is an important difference between the respondents who are willing to consider alternative modes of travel but are not able to, and the respondents who are both willing and able to. In addition, it was shown how GPS data can be used to further augment and contextualise access to spatiotemporal opportunities by making use of individual activity spaces.

For the design of a VTBC intervention, these results are important for at least two reasons. The first is that the results suggest that STA analysis can be used to identify the individuals who are most likely to respond to an intervention. Interventions should be targeted at those individuals who have the opportunity to change their behaviour. The second reason is that with STA analysis, it is possible to assess the feasibility of a certain activity programme (e.g. doing shopping directly after work rather than making an additional trip) or the feasibility of using a different travel mode (e.g. walking or cycling as an alternative to driving). Whereas the activity programme in this chapter consisted of only home and work, the data can be used to make targeted suggestions for alternatives to activity locations and alternative modes of travel; even more so when the available choice universe can be constrained with an activity space.

While still insightful, a major limitation of this study was the period of data collection. Because activity-travel behaviour is extremely variable, day-to-day accessibility can also vary. A second limitation was the fact that a very simple activity programme with only two primary anchors and three available modes of travel was used for the calculation of the accessibility statistics. However, most people will have more than two primary anchors, and in some places public transport will be a viable travel alternative. Thirdly, although person-based accessibility measures can account for individual differences in accessibility, other individual characteristics were disregarded in this study. For a male individual, it may be acceptable to walk home in the evening, but females may have safety reservations.

Notwithstanding these limitations, the results clearly show the potential of space-time accessibility analysis for incorporation in the design of VTBC interventions. Accordingly, there are plenty of avenues for further research. A field experiment to identify potential candidates for an intervention STA analysis is recommended, for example. Yet more methodological approaches should also not be disregarded. For example, it could be explored how GPS data could be used to identify primary anchors and derive time windows, such as by GPS imputation of raw GPS trajectories, and subsequently to identify the activity locations that can be considered as primary anchors. Another possibility is to extend the analysis with a travelling salesman implementation to see what the most sustainable, and feasible, implementation of an activity programme could be.

Acknowledgements

The financial assistance of the South African National Research Foundation (NRF) towards this research is hereby acknowledged. Opinions expressed and conclusions arrived at, are those of the authors and are not necessarily to be attributed to the NRF. An early version of this paper was presented at Mobile Tartu 2016 in Tartu, Estonia.

References

Bamberg, S. & Möser, G. 2011. Please Mr. Brög, give us your data! Reply to the comment of Wall, Brög, Erl, Ryle, & Barta on our paper "The effectiveness of soft transport policy measures: A critical assessment and meta-analysis of empirical evidence". *Journal of Environmental Psychology*. 31(3):270–271.

Bohte, W. & Maat, K. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*. 17(3):285–297.

Bonsall, P. 2009. Do we know whether personal travel planning really works? *Transport Policy*. 16(6):306–314.

Brög, W., Erl, E., Ker, I., Ryle, J. & Wall, R. 2009. Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*. 16(6):281–292.

Casas, I., Horner, M.W. & Weber, J. 2009. A comparison of three methods for identifying transport-based exclusion: A case study of children's access to urban opportunities in Erie and Niagara counties, New York. *International Journal of Sustainable Transportation*. 3(4):227–245.

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*. 16(6):293–305.

Department of Environmental Affairs. 2015. *2013 / 2014 South African national land-cover dataset*. [Online], Available: https://www.environment.gov.za/mapsgraphics [2016, November 22].

Dijst, M. 1995. *Het elliptisch leven: Actieruimte als integrale maat voor bereik en mobiliteit - modelontwikkeling met als voorbeeld tweeverdieners met kinderen in Houten en Utrecht*. Published doctoral dissertation. Utrecht: Koninklijk Nederlands Aardrijkskundig Genootschap.

GeoTerraImage. 2015. *2013 - 2014 South African national land-cover data. User report and metadata*. Pretoria, South Africa. [Online], Available: https://www.environment.gov.za/mapsgraphics.

Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional Science Association*. 24(1):6–21.

Howarth, C.C. & Polyviou, P. 2012. Sustainable travel behaviour and the widespread impacts on the local economy. *Local Economy*. 27(7):764–781.

Kingham, S., Dickinson, J. & Copsey, S. 2001. Travelling to work: Will people move out of their cars. *Transport Policy*. 8(2):151–160.

Kwan, M.-P. 1999. Gender and individual access to urban opportunities: A study using space–time measures. *The Professional Geographer*. 51(2):211–227.

Kwan, M.-P. 2013. Beyond space (as we knew It): Toward temporally integrated geographies of segregation, health, and accessibility. *Annals of the Association of American Geographers*. 103(5):1078–1086.

Kwan, M.-P. & Weber, J. 2003. Individual accessibility revisited: Implications for geographical analysis in the twenty-first century. *Geographical Analysis*. 35(4):341–353.

Li, R. & Tong, D. 2016. Constructing human activity spaces: A new approach incorporating complex urban activity-travel. *Journal of Transport Geography*. 56:23–35.

Lundberg, B. & Weber, J. 2014. Non-motorized transport and university populations: An analysis of connectivity and network perceptions. *Journal of Transport Geography*. 39:165–178.

Miller, H. 2007. Place-based versus people-based Geographic Information Science. *Geography Compass*. 1(3):503–535.

Morris, J.M., Dumble, P.L. & Wigan, M.R. 1979. Accessibility indicators for transport planning. *Transportation Research Part A: General*. 13(2):91–109.

Neutens, T., Schwanen, T. & Witlox, F. 2011. The prism of everyday life: Towards a new research agenda for time geography. *Transport Reviews*. 31(1):25–47.

Neutens, T., Delafontaine, M., Schwanen, T. & van de Weghe, N. 2012. The relationship between opening hours and accessibility of public service delivery. *Journal of Transport Geography*. 25:128–140.

Newsome, T.H., Walcott, W.A. & Smith, P.D. 1998. Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and space dimensions. *Transportation*. 25(4):357–377.

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N. & Maurer, P. 2014. Supporting large-scale travel surveys with smartphones - A practical approach. *Transportation Research Part C: Emerging Technologies*. 43:212–221.

OpenStreetMap Contributors. 2016. *Planet Dump [Datafile from 26/07/2016 of BBBike extracts]*. [Online], Available: http://extract.bbbike.org/ [2016, July 26].

Patterson, Z. & Farber, S. 2015. Potential path areas and activity spaces in application: A review. *Transport Reviews*. 35(6):679–700.

Salonen, M., Broberg, A., Kyttä, M. & Toivonen, T. 2014. Do suburban residents prefer the fastest or low-carbon travel modes? Combining public participation GIS and multimodal travel time analysis for daily mobility research. *Applied Geography*. 53:438–448.

Schwanen, T. & De Jong, T. 2008. Exploring the juggling of responsibilities with space-time accessibility analysis. *Urban Geography*. 29(6):556–580.

Shoval, N., Kwan, M.-P., Reinau, K.H. & Harder, H. 2014. The shoemaker's son always goes barefoot: Implementations of GPS and other tracking technologies for geographic research. *Geoforum*. 51:1–5.

Stellenbosch University. 2016. *Statistical profile*. [Online], Available: http://www.sun.ac.za/english/statistical-profile-2014-test [2017, February 05].

Taylor, M. & Ampt, E. 2003. Travelling smarter down under: Policies for voluntary travel behaviour change in Australia. *Transport Policy*. 10(3):165–177.

Tribby, C.P., Miller, H.J., Werner, C.M., Smith, K.R. & Brown, B.B. 2016. Assessing built environment walkability using activity-space summary measures. *Journal of Transportation and Land Use*. 9(2):1–21.

Widener, M.J., Farber, S., Neutens, T. & Horner, M. 2015. Spatiotemporal accessibility to supermarkets using public transit: an interaction potential approach in Cincinnati, Ohio. *Journal of Transport Geography*. 42:72–83.

Xu, Y., Shaw, S., Zhao, Z., Yin, L., Lu, F., Chen, J., Fang, Z. & Li, Q. 2016. Another tale of two cities : Understanding human activity space using actively tracked cellphone location data. *Annals of the Association of American Geographers*. 106(2):489–502.

Zhang, Y., Stopher, P. & Halling, B. 2013. Evaluation of south-Australia's TravelSmart project: Changes in community's attitudes to travel. *Transport Policy*. 26:15–22.

Part V
Conclusion

Where geographical scholarship on transport has been plugged into broader disciplinary debates through engagement with questions around inequality, anthropogenic climate change and health, ongoing experimentation with big data and network analysis draws the field into interdisciplinary developments around data and methodology.

Schwanen (2017a: 355–356)

Chapter 8. Smartphones and transport research

The past decades have seen growing and ongoing academic and policy debate on how to encourage individuals to change to more sustainable ways of travelling. More recently, researchers have started to build on so-called location-aware technologies (LAT) such as Global Positioning Systems (GPS), looking for new methods to more accurately capture individual spatiotemporal travel patterns. This dissertation set out to expand on this research by using GPS-enabled smartphones for collecting and analysing individual travel data in South Africa. The aim of this research has been operationalised into four research questions. These research questions are addressed in the next sections. The theoretical contributions, policy implications, and the limitations are subsequently addressed. This is followed by some recommendations for future research and an overall conclusion.

8.1 Introduction

Especially among households in the higher-income quintiles, 'car as driver' and 'car as passenger' are the main modes of transport in South Africa (Statistics South Africa, 2013). In addition, non-motorised forms of transport such as cycling and walking are not considered favourably by many South Africans (Statistics South Africa, 2016). Particularly in urban areas, this is problematic because it leads to problems such as congestion, accidents, road decay, and reduced accessibility to goods and services. Moreover, the heavy reliance on road-based transport across the globe has profound environmental effects, including a decrease in local air quality, and worldwide emissions of harmful greenhouse gases (IPCC, 2014; Pojani & Stead, 2015). Accordingly, "[i]mplementing behavioural strategies aimed at reducing car use represents one of the most topical challenges for current transport research" (Meloni & Sanjust, 2014: 325).

Behavioural strategies aimed at reducing car use come in many forms, and are normally referred to as transport demand management strategies (TDM). TDM is an umbrella term for all interventions that try to alter travel behaviour in favour of more socially, environmentally, and economically sustainable alternatives (Taylor, 2007). Within the domain of TDM, a useful distinction can be made between 'hard strategies' and 'soft strategies'. Where hard strategies focus on restraining and managing car travel with regulations and economic disincentives, soft strategies aim to alter behaviour by means of information provision (Eriksson, Garvill & Nordlund, 2006). Soft strategies are often referred to as voluntary travel behaviour change (VTBC) programmes. In recent years, a large body of research has focused on the evaluation of these soft programmes (cf. Cairns, Sloman, Newson, Anable, Kirkbridge & Goodwin, 2008; Bonsall, 2009; Brög, Erl, Ker, Ryle & Wall, 2009; Chatterjee, 2009; Stopher, Clifford, Swann & Zhang, 2009; Zhang, Stopher & Halling, 2013; Meloni & Sanjust, 2014; Meloni, Sanjust Di Teulada & Spissu, 2016). However, despite many contributions, the current scientific debate on VTBC interventions contains several areas that require further research.

Firstly, most VTBC intervention evaluations rely heavily on self-reported data, and as such depend on the respondent's ability to accurately remember his or her daily movements and activity participation. However, precise measurement is required to determine the often small changes in travel behaviour that VTBC programmes reportedly achieve. Secondly, although technological developments in the field of location-aware technologies have greatly enhanced opportunities to collect accurate data on human spatiotemporal behaviour, analytical methods and data imputation methods are required. Thirdly, analytical tools that draw attention to individual accessibility and spatiotemporal access to transport and activity opportunities have not been employed in the design of VTBC strategies. The following aim was therefore formulated for this research:

> To evaluate the potential of location-aware technologies and the analysis of spatiotemporal behaviour for the design of a voluntary travel behaviour change (VTBC) intervention in South Africa.

8.2 Overview of the results

To address the aim of this research, four research questions were formulated in the introduction of this dissertation. This section summarises and interprets the findings from Chapters 2 through 7 that have addressed these questions.

> RQ1: What are the current insights and research gaps on the effectiveness of VTBC changes?

Chapter 2 assessed the effectiveness of VTBC programmes through a literature review. VTBC programmes have been implemented around the world, particularly in Australia, Japan, and the United Kingdom. Many of these programmes have yielded positive results. As such, there seems to be a fair amount of evidence in support of the idea that soft transport measures have the potential to reduce private vehicle usage and stimulate public transport ridership. However, there are several limitations to indicate that this conclusion may be too premature. For instance, the results of a number of studies suggest that behavioural effects after an intervention were not sustained (Taylor, 2007). In addition, it has been suggested that if soft measures are not implemented side by side with hard measures, behavioural change is unlikely (Gärling & Schuitema, 2007; Behrens, Adjei, Covary, Jobanputra, Wasswa & Zuidgeest, 2015). Furthermore, the success of an intervention may not solely depend on one's willingness to change, but also on available travel alternatives (Kingham, Dickinson & Copsey, 2001).

The literature review also revealed several methodological and contextual issues. Firstly, assessing the effectiveness of a VTBC programme requires travel behaviour data of both pre- and post-intervention. However, the cross-sectional research design that has been most frequently employed in the evaluation studies of VTBC programmes may not be suitable to properly identify the effects of such an intervention. Important reasons for this are that studies employing a cross-sectional research design disregard temporal changes and are susceptible to regression to the mean. In addition, because the expected changes are often relatively small, large sample sizes are necessary to accurately assess the effect of the intervention (Richardson, Seethaler & Harbutt, 2004; Stopher & Greaves, 2007; Stopher *et al.*, 2009).

Secondly, in many cases, VTBC programme evaluations have relied heavily on traditional activity-travel diaries or on prompted-recall methods. Yet, obtaining disaggregate multi-day travel data with these methods goes hand in hand with a high respondent burden and many inaccuracies in the collected data (Behrens, 2004; Stopher *et al.*, 2009; Jariyasunant, Carrel, Ekambaram, Gaker, Kote, Sengupta & Walker, 2011; Nitsche, Widhalm, Breuss, Brändle & Maurer, 2014; Prelipcean, 2016). This is problematic because for a precise evaluation of the effectiveness of a VTBC intervention, multi-day travel data are required that account for spatial and temporal differences, socio-demographic diversity, variety in travel modes, and differentiation in trip purposes (Richardson, 2003; Richardson *et al.*, 2004; Taylor, 2007; Bonsall, 2009; Nitsche *et al.*, 2014).

Thirdly, the majority of the evaluation studies have predominantly focused on reducing the number of car trips and the vehicle kilometres travelled; as such they do not properly account for the context in which travel behaviour occurs. Yet, several studies have found that, for example, densities and accessibility to opportunities for activity participation have an important effect on travel demand (Meurs & Haaijer, 2001; Van Acker & Witlox, 2009). Following Hägerstrand's (1970) time-geographical framework, it can be argued that "people's daily activity-travel is constrained by the spatiotemporal availability of alternatives for activity destinations" (Ren, Tong & Kwan, 2014: 330). The consideration of these opportunities seems to have been largely neglected in the design of travel behaviour change interventions.

> RQ2: To what extent are current GPS data processing techniques suitable for analysing the effectiveness of a VTBC intervention?

VTBC interventions have been found to be moderately successful in changing behaviour on individual and household level. Yet there also has been some controversy surrounding the effectiveness of such interventions over longer periods of time and whether the effects of an intervention can be sustained. One possible way to tackle this issue is by augmenting VTBC research with non-intrusive, location-aware technologies such as GPS technology, which can be used for longitudinal data collection (Bonsall, 2009; Brög *et al.*, 2009; Meloni & Sanjust, 2014). At the same time, this calls for analytical methods to process and analyse GPS data and extract information on activity and travel behaviour of individuals. With this third research question, possibilities for analysing raw GPS trajectories were investigated.

Where GPS technology can precisely register locational information, travel characteristics need to be imputed from the data (Shen & Stopher, 2014). As such, throughout the last decade, various methods have been developed for identifying trips, activities, and travel modes from raw GPS trajectories (Feng & Timmermans, 2016). However, rule-based methods that use dwell time for the classification of activity and trip episodes may disregard short activities. Chapter 3, therefore, described a set of machine learning algorithms to identify activity and travel episodes on a point-by-point basis. To account for the underlying spatial structure of the data, attribute information was supplemented with information derived from multiple moving spatial windows over preceding and succeeding points to estimate local point densities. Because a ground truth is essential to evaluating the accuracy of the algorithms, a set of 200 artificial GPS activity-travel sequences with varying noise levels was generated (Thierry, Chaix & Kestens, 2013). The results indicate that the random forest classifier in particular can lead to high accuracies. In addition, it was found that local densities are important variables in measuring the accuracy of the classifiers.

Other than activities, the exact distances and routes an individual has travelled are also essential variables in determining the effectiveness of soft transport demand management strategies. Furthermore, this information is required when one wants to provide the traveller with individualised feedback on costs and CO_2 emissions. For this reason, Chapter 4 presented two GIS-based map-matching methods that predominantly use a digital road network with speed and directionality attributes for route reconstruction of raw GPS trajectories. The methodologies were tested for a dataset in which actual routes travelled were known. Both explored procedures, the *connected subset* assignment procedure (based on network subset selection) and the *impedance reduction* assignment procedure (based on attribute adjustment), provide accurate results. In addition, both procedures effectively deal with commonly GPS-induced problems such as measurement gaps and positional drift. An important advantage of executing the algorithms in a GIS environment is that the whole procedure can be automated, for instance using Python, and allows for easy integration with other types of spatial analyses.

RQ3: To what extent are GPS-enabled smartphones suitable for acquiring high-resolution space-time data in South Africa?

In Chapter 5, research was presented which assessed the reliability and feasibility of passively collecting high-resolution locational data on activity and travel behaviour using GPS-enabled smartphones. A small-scale field test was conducted in which respondents from the University of Stellenbosch, South Africa, were passively tracked for two days using a purposely designed smartphone application. The results of the experiment indicated that while GPS technology in smartphones potentially holds several benefits for collecting activity and travel data, the technology is not without problems. A first problem is that some users tend to switch off the application when they arrive at home, which increases the chances that they forget to switch the application back on for their next trip. A second problem is the battery consumption of location-based services. If the GPS of a smartphone is continuously active, battery levels decrease significantly. This may also make people decide to switch off the tracking application. While this can potentially be solved by improving the battery duty cycling of the application, for instance with other sensors like the smartphone's accelerometer, this may lead to a decrease in data quality. A third problem is the representativeness of the sample. Because only users with qualifying smartphones could participate in the data collection, groups of potential participants may have been excluded.

Chapters 6 and 7 used data collected during a research project on the travel patterns of staff and students on both the main campus and the Tygerberg satellite campus of Stellenbosch University, South Africa. During this data collection, some additional problems than those identified in Chapter 5 came to light. Whereas 853 people responded to the household travel questionnaire, only 176 individuals registered onto the Tracklog system. In addition, only 151 and 141 unique responses (tracks) were recorded on the two respective days of tracking. This means that some users did register on the system but did not initiate the tracking. An additional issue was the quality of the data, which varied greatly between tracks. Notwithstanding these problems, the results indicated that gathering high-resolution space-time data by means of GPS-enabled smartphones is feasible and that it opens doors to a range of possible applications that are unattainable by traditional survey methods.

> RQ4: To what extent can space-time accessibility measures be integrated into modelling the feasibility of a VTBC intervention?

A major issue with the design of VTBC interventions is that many studies ignore the individual context in which spatiotemporal behaviour occurs (Behrens & Del Mistro, 2010; Howarth & Polyviou, 2012). Chapter 6 explored how accessibility to opportunities for activity participation within the spatial context of individual GPS-based activity spaces can help to distinguish between individuals according to their willingness to consider more sustainable travel modes. Activity spaces can be seen as an individual's spatial footprint of their day-to-day travel and activity behaviour and, thus, provide valuable information about the context in which this behaviour takes place (Newsome, Walcott & Smith, 1998; Perchoux, Chaix, Cummins & Kestens, 2013; Patterson & Farber, 2015). By augmenting the individual activity spaces with point-of-interest data to establish opportunity indicators, it was found that there is some evidence to suggest that there is a relationship between the number of opportunities within an individual's activity space and their willingness to consider walking or cycling to and from campus.

In Chapter 7, accessibility to opportunities was further explored. Based on the examination of 636 combinations of geocoded home and work locations, accessibility indicators were calculated with space-time accessibility (STA) analysis. STA measures are based on a key time-geographical concept: the space-time prism. Given two activities, the space-time prism is the three-dimensional graphical representation of all possible paths that an individual can take between the end time of the first activity and the start time of the second activity. The two-dimensional derivative of the space-time prism, the potential path area (PPA), represents the area that an individual can potentially visit during these two activities (Schwanen & De Jong, 2008). The extent of the PPA, therefore, can be equated to one's opportunity to access activity locations, or the potential for spatial interaction (Casas, Horner & Weber, 2009; Patterson & Farber, 2015). By relating access to opportunities to the respondents' self-reported willingness to consider cycling or walking to and from campus, the results indicate that there is an important difference between the respondents who are willing to consider alternative modes of travel but are not able to, and the respondents who are willing and able to. In addition, it was shown how GPS data could be used to further augment and contextualise access to spatiotemporal opportunities by making use of individual activity spaces.

8.3 Implications

As a whole, this dissertation supports the suggestions made in several studies that have called for GPS data collection methods to be integrated into the research on the effectiveness of voluntary travel behaviour change (VTBC) interventions (cf. Chatterjee, 2009; Chatterjee & Bonsall, 2009; Richter, Friman & Gärling, 2011; Chai, Chen, Liu, Tana & Ma, 2014; Meloni & Sanjust, 2014). However, the insights from the methodological contributions and empirical analyses presented in this dissertation have several implications for current methodological discussions on GPS data analysis vis-à-vis VTBC interventions, for theoretical discussions on the design and evaluation of VTBC interventions in general, as well as for policy makers and transport planners.

GPS data analysis

This dissertation has highlighted that research using GPS-enabled smartphones is not without challenges. These challenges include not only harnessing the tools to obtain geo-referenced data, but also acquiring new skills for cleaning and interpreting these data, and ultimately embedding these new data streams into a geographical-analytical framework. Accordingly, as argued in a recent contribution by Schwanen (2017a: 355–356), transport-geographical scholarship is drawn "into interdisciplinary developments around data and methodology". That these interdisciplinary developments are indeed challenging is illustrated by the striking absence of geographers in studies employing tracking technologies. Shoval *et al*. (2014), for example, found that only 13 percent of peer-reviewed research using these new technologies is published in geographic journals or involve geographers. Unlike the prominent role geographers have had in the development of Geographic Information Systems (GIS), and whereas spatial information is at the core of the geographic discipline, tracking technologies have been mostly employed by health and transport researchers (Birenboim & Shoval, 2016).

The importance of incorporating insights from different fields of study into GPS data imputation methods was illustrated by integrating multiple moving spatial windows into advanced machine learning algorithms. Although the idea of a moving spatial window to identify local point densities is not new (cf. Schuessler & Axhausen, 2009) and moving averages have been incorporated in machine learning algorithms before (cf. Feng & Timmermans, 2016; Shafique & Hato, 2016), the integration of these two concepts proved to be very effective in classifying GPS points into activity points and travel points. In addition, a major advantage of this method is that it does not require the specification of a minimum distance or time threshold that potentially signals activity engagement. As such, it reduces the chances that short activities are not recognised. Whilst for some implementations the recognition of short activities may not be essential, they could be route-determining. In the context of VTBC interventions, this is relevant information. If, for instance, someone briefly stops to take his/her children to day care on his/her way to work, this needs to be considered when designing a personalised intervention. If one were to suggest an alternative mode of transport that cannot accommodate this responsibility, it is very unlikely that the intervention will be successful. If, on the other hand, the short stop is of a more generic nature, a feasible alternative for an activity destination could be suggested.

Route reconstruction through map-matching is another important domain to which transport geographers could contribute. Most map-matching algorithms have been developed for real-time applications, but real-time and post-processing map-matching algorithms cannot be used interchangeably. Whereas real-time map-matching procedures aim to identify the position of the user on the road network, post-processing map-matching algorithms aim to reconstruct the actual route travelled for an entire trip (Hashemi & Karimi, 2014). In addition, in post-processing map-matching, the continuity of the path route is essential, which is not a condition in real-time map-matching. What most post-processing map-matching algorithms have in common is their statistical or mathematical approach to solving a problem, often extended with geometric and topological information (information that is inherent to the data model for transport networks in a GIS environment). Against this background, we have proposed two purely GIS-based post-processing map-matching algorithms, both of which yielded good scores on the employed validation indices.

Once the actual routes travelled are reconstructed, the information can be used to quantify metrics, such as vehicle kilometres travelled. If, for example, data on the type of car a road user drives are also available, this information can be relayed back to the individual, extended with an estimate of the vehicle operating costs. For a VTBC intervention, this is useful because "vehicle excise duty, insurance, servicing and depreciation are all fixed costs which, once paid, tend to be forgotten; thus, when comparing alternatives to car use for a specific journey many motorists only consider the cost of fuel" (Howarth & Polyviou, 2012: 765). In addition, unlike earlier implementations of post-processing map-matching, the applicability of the proposed map-matching algorithms to walking and bike trips was explored. In the context of VTBC interventions, the total distance travelled by bicycle or on foot can provide valuable information as well. For example, data on bicycle kilometres travelled could be used for goal-setting, such as encouraging users to reach a certain threshold distance travelled by bicycle.

Design of voluntary travel behaviour change interventions

Travel behaviour changes cannot be viewed as separate from the personal and spatiotemporal context in which travel behaviour is situated (Hägerstrand, 1970; Howarth & Polyviou, 2012). As such, in the introduction of this dissertation, attention was drawn to time geography, in which mobility is understood as being derived from time-use decisions that are constrained by different time regimes. This implies that for a VTBC intervention to be successful, there needs to be "compliance between the 'system' (…) and individual travel needs" (Howarth & Polyviou, 2012: 765). Against this background, activity spaces were used to explore the relationship between accessibility to opportunities and a willingness to consider alternative modes of transport. In addition, STA analysis was used to further investigate this relationship.

Analysis of the activity spaces showed that there is a relationship between opportunity indicators (i.e. points of interest) within one's activity space and the willingness to consider walking or cycling to campus (Chapter 6). This is important for the design of VTBC interventions for at least two reasons. The first reason is that it suggests that higher values on the opportunity indicators could be indicative of a higher degree of spatial freedom. These individuals may be more likely to respond to external stimuli and/or awareness programmes. The second reason is that it shows that GPS data and activity spaces can be exploited relatively easily, and potentially used in the design of travel behaviour change schemes. A possible implementation of activity spaces, for example, would be to identify target populations with similar levels of opportunity densities. Although activity spaces do not fully account for individual accessibility to opportunities, mainly because they are temporally naïve, they can still be used as a tool to identify potential candidates for a VTBC intervention.

Unlike activity spaces, STA analysis can take the temporal component of the urban context into account. For the design of a VTBC intervention, the results are again significant for at least two reasons. The first reason is that it suggests that accessibility is important to consider when designing a VTBC intervention. Similar to the activity spaces, STA analysis could be used to identify those individuals that are most likely to respond to an intervention. Interventions should rather be targeted to those individuals who have the opportunity to change their behaviour. The second reason is that with STA analysis, it is possible to assess the feasibility of a certain activity programme or the feasibility of using a different travel mode. Where the analysis in Chapter 7 only focused on home and work, the data can be used to make targeted suggestions for alternatives to activity locations and modes of travel. A combination of activity spaces with STA analysis could be more effective, as the available choice universe can be combined with revealed spatial behaviour.

Evaluation of voluntary travel behaviour change interventions

The fact that individuals are constrained (or enabled) by accessibility to opportunities raises the question of whether this information can be further exploited. An important discussion in travel demand management studies deals with the issue of self-selection bias (Cao, Mokhtarian & Handy, 2009; Van Wee, 2009). For example, studies have found evidence that people in suburban neighbourhoods have less sustainable travel habits than people in urban neighbourhoods. This raises the question of whether this effect can be attributed to the spatial environment, or rather that individuals "who prefer walking may consciously choose to live in neighbourhoods conducive to walking, and thus walk more" (Cao *et al.*, 2009: 360). Similarly, individuals who have a more positive attitude towards cycling could be more easily persuaded to cycle more (Gatersleben & Appleton, 2007). The same applies to VTBC interventions, people tend to self-select into participating in these programmes (Brög *et al.*, 2009; Stopher *et al.*, 2009).

Because of the issue of self-selection, it is difficult to measure the effectiveness of a broad implementation of a VTBC intervention. However, it could be argued that this is not a problem, but rather an opportunity. The analysis of individual GPS tracks, combined with data on the geographical context, shows that some individuals have more opportunities to make a change. Whether these individuals self-selected themselves to live in that geographical context or not stops mattering if the goal of a VTBC intervention is encouraging people to use more sustainable modes of travel. By augmenting research on the effectiveness of VTBC interventions with GPS data collection methods, it is then possible to identify these people more easily.

Policy and planning

The introduction to this research positioned transport as an important contributor to greenhouse gas emissions, and showed how VTBC interventions could be a cost-effective way to encourage people to travel more sustainably. The research presented in this dissertation has dealt primarily with identifying tools to be integrated into the design of these interventions. Therefore, the results have implications for policy and planning with regard to travel behaviour change initiatives. The main point here is that, as emphasised throughout the dissertation, spatiotemporal accessibility to opportunities and feasible alternatives play an important facilitating role in whether or not a travel behaviour change intervention is successful. No matter how good the information campaign or the company's transport policy, this means that policy makers are advised to consider identifying different groups with similar levels of accessibility and target the interventions to fit the spatial context.

It is not only the context that matters. Policy makers and transport planners should also pay close attention to socio-economic differences. Even more so in the context of South Africa, policy makers and transport planners should not aim for a broad implementation of a VTBC intervention as this is unlikely to be effective. Rather, the focus should be on the upper middle class who often drive private vehicles. Arguably, the most viable option is to identify large institutions that generate high levels of demand for transport, such as a university. The idea of differentiating between groups could work well. For example, students and staff members living close to campus could be targeted directly by encouraging them to walk or use bicycles. Although concerns about safety play a role in many South African cities and public transport systems (Walters, 2008, 2013), proper monitoring of high-traffic walking or cycling routes could support these initiatives.

Other than universities, large companies are potential candidates for designing an intervention. However, in such a case people may not live in the vicinity of the company premises. As such, this would require a different approach because an alternative mode of transport should be in place. If public transport is not an option, a company bus or shuttle service could serve as a high-occupancy alternative. Also, carpooling could be stimulated, for example by facilitating contact between individuals from the same company whose activity space is similar. Whereas one would most likely only reach the individuals who already have a positive attitude towards travelling more sustainably, one can provide the platform to actually help these individuals, in Prochaska's terms, moving from pre-contemplation to action (Prochaska, Velicer, Rossi, Goldstein, Marcus, Rakowski, Fiore, Harlow, Redding, Rosenbloom & Rossi, 1994).

8.4 Limitations

This research has provided important innovative insights into the collection and analysis of GPS data vis-à-vis the design of VTBC interventions. Nevertheless, there are some methodological and theoretical shortcomings to this research. A first limitation is found in the representativeness of the sample frame that was used in the empirical part of the research: a university's population is arguably very different from the general population. In addition, the use of smartphones for data collection may have biased the results towards younger respondents. It has been suggested that "the GPS survey is more suitable for the younger respondent, while traditional survey methods may be better for older respondents, because the younger respondents are more technology savvy" (Shen & Stopher, 2014: 318). Moreover, only people who have a qualifying smartphone could participate in the GPS data collection, potentially excluding certain respondents.

An issue related to the sampling frame was the sample size and the quality of the GPS data. Whereas a large number of staff members and students was targeted, the number of collected GPS tracks was relatively low. In addition, the quality of the some of the GPS tracks was not as good as was expected. Related to this is the period of data collection. Because activity-travel behaviour is extremely variable, the data collection period should be considerably longer to lead to more definitive conclusions. However, the fact that using GPS-enabled smartphones for data collection on travel behaviour was the first of its kind in South Africa does put this limitation into perspective.

Aside from the data collection, there were also two limitations on the side of the GPS processing. While data imputation to extract relevant information from raw GPS trajectories is crucial, a ground truth is required to assess the accuracy of these methods. Whereas for the developed route reconstruction methods the actual routes travelled were available, artificial GPS tracks had to be used for activity points recognition because of the many challenges involved in obtaining ground truth (Stopher, Shen, Liu & Ahmed, 2015; Feng & Timmermans, 2016). The creation of artificial data can be used as a starting point for methodological explorations, but it is most likely not sufficient because the noise and inaccuracies in real-life GPS-tracks are difficult to simulate. A second limitation relating to GPS processing was that, in this dissertation, attention was paid only to activity and travel point recognition and route reconstruction. Travel mode detection remained out of scope, whereas the travel mode used is actually relevant information in the context of designing a VTBC intervention.

The limitations on the side of the GPS processing side also have some impact on the overall result of this study. Whereas the random forest classifier provides good results to identify move points and stay points from a stream of raw GPS data and whereas the post-processing map-matching algorithms give good results for route reconstruction, the methods have not been applied on the collected Stellenbosch data and combined to provide a holistic view of travel behaviour. The main reason for this is that the current methods need to be further extended. For example, for the machine learning algorithms to be applied on the collected smartphone data, a training data set is required because machine learning algorithms are not directly transferable to a different context. Similarly, whilst the machine learning algorithms are effective in dividing a GPS trajectory into move points and stay points, a next step is required to identify the actual activity locations to determine the activity-travel sequence.

On the theoretical side, the use of time geography as a theoretical framework has also some disadvantages. The main issue here was that we only focused on individual accessibility, whereas, for instance, the juggling of household responsibilities was not accounted for (cf. Schwanen & De Jong, 2008). In addition, as put forward by Howarth and Polyviou (2012: 769), "[i]ssues such as habitual behaviour, reluctance to change, perceptions of travel mode, cost of alternative options, travel time, travel purpose, number of passengers, comfort, convenience and so forth all play a key role in influencing behaviour". As shown in Chapter 6, the safety aspect, for instance, seems to play an important role when considering alternative modes of transport. A VTBC may look like a possibility, but concerns about safety may deter a person from actually considering alternative options. This, once again, implies that individual spatiotemporal accessibility is a necessary, but not a sufficient condition for attaining a behavioural change.

A reflection on the position of this research within the domain of transport geography also highlights some limitations. For instance, research using quantitative methods to analyse GPS data falls within the new paradigm in human mobility research that capitalises on so-called 'big data' (Chen, Ma, Susilo, Liu & Wang, 2016; Kwan, 2016; Kwan & Schwanen, 2016). However, there is a danger of reductionism because "information on social identity (gender, race/ethnicity, etc.) or embodiment and lived experience is absent from most big datasets" (Schwanen, 2017a: 358). Similarly, "there is a historical hegemony of predominantly western worldviews, concepts, theories methods and research practices" (Schwanen, 2017b: 2) which needs to be questioned in the context of the Global South, ultimately leading to further decolonisation of the production of transport-geographical scholarship.[1]

8.5 Recommendations for future research

Through the usage of GPS-enabled smartphones and space-time accessibility analysis, this research has provided important innovative insights into the collection and analysis of GPS data vis-à-vis the design of VTBC interventions. Nevertheless, there are some important issues that have remained unexplored. This final section presents recommendations for future research that also could help address some of the limitations.

The conclusions of the literature review in Chapter 2 suggested two lines of research: research into using GPS-enabled smartphones to gather activity-travel data and methods to analyse these data, and research into the effectiveness of VTBC interventions. The work in this dissertation has preliminary dealt with the first line of research. Although there is still a lot of work to be done when it comes to analysing GPS data, "research into the effectiveness of TDM measures in breaking habits and fostering new behaviours is vital" (Gärling, Eek, Loukopoulos, Fujii, Johansson-Stenman, Kitamura, Pendyala & Vilhelmson, 2002: 68). Based on the various methodological critiques that surfaced in the literature review, it is suggested that VTBC interventions be analysed by means of field experiments that include a randomisation of participants to experimental and control groups. To account for the variability of travel behaviour, these should preferably be longitudinal studies or, at the very least, panel studies that cover multiple days of travel.

Building on better ways to assess the effectiveness of VTBC interventions, the individual context should be extended to incorporate the household context when modelling space-time accessibility. As mentioned in Section 8.4, individuals do not act in a vacuum, but are embedded in a social context, such as the household in which they live. The juggling of responsibilities between parents, for instance, is important here because "in car-dependent communities, a car trip shared between parent(s) and child(ren) is more than a utilitarian movement from A to B but also an expression of care and a moment of family time" (Schwanen, 2016: 132). GPS-enabled smartphones could help by including a social network element as part of the application's interface, which requests the user to supply information about with whom an activity was engaged in (cf. Yip, Forrest & Xian, 2015).

From the perspective of using GPS-enabled smartphones, a number of suggestions can also be made. Before one can design, administer, and evaluate a VTBC intervention through an application, significant work is required at the methodological side. The methods and tools that have been tested and described in the different chapters are necessary, but not sufficient. For example, in order to provide a comprehensive understanding of travel behaviour and

[1] It is worth noting Schwanen's (2017b) point in that in transport-geographical scholarship, a 'silent' reference is often made to the situation in the USA and/or Western Europe. For example, in relation to the perceived chaotic nature of minibus taxis, scholars sometimes seem to be inclined to represent this form of public transport "as second-best solutions operated by reckless, ill-disciplined drivers – framings of lack and absence compared to a silent referent [the USA and Western Europe] are never far away" (Schwanen, 2017b: 2).

opportunities for a VTBC intervention, travel modes should be imputed. Similarly, after the identification of move and stay points it will be necessary to identify the actual activity locations (including the type of activity location). These analyses can subsequently be extended by capitalising on the flexibility to incorporate a variety of feedback elements offered by smartphone applications. For example, precise data on routes travelled can be used to quantify metrics, such as vehicle kilometres travelled. If, as mentioned before, data on the type of car the user drives are also available, this information can be relayed back to the individual extended with an estimate of the vehicle operating costs. In addition, "change is more likely to occur and be sustained if peer behaviour is also evident (Howarth & Polyviou, 2012: 771)". A possible suggestion here is to develop an application which includes an element of gamification, so that users can 'compete' with each other for the highest sustainability score.

On the data analysis side, the usage of advanced machine learning classifiers should be further explored, for instance, to improve travel mode detection by building on Bayesian belief networks (Xiao, Juan & Zhang, 2015) or advancing artificial neural networks (Feng & Timmermans, 2016). Machine learning algorithms could potentially also be used in the context of identifying the individuals who are most likely to respond to a VTBC intervention. Without trying to fully 'predict' human mobility, it would be interesting to see whether people who have a positive attitude towards using more sustainable modes share similar characteristics in terms of levels of accessibility, time windows, socio-economic group, etc.

Future transport-geographical studies in which smartphones are used as a data collection tool should continue to explore cross-fertilisation between disciplines, including, but not limited to, transport geography, transport engineering, transport economics, psychology, computer science, economics, and data science. Travel behaviour (change) has many different aspects and, as such, demands an interdisciplinary approach to furthering our understanding and improving our policy and planning decisions. Yet "cleaning smartphone-collected data poses a challenge as the data volume is huge (…) and the techniques involved in cleaning such data are very different compared to cleaning ordinary survey data" (Yip *et al.*, 2015: 162). This requires a different approach from the research community, not just in terms of data collection, but also in its subsequent analyses, as this may require new skills.[2] Because, as mentioned previously, geographers are not leading in transport research using big data and tracking applications, it will be interesting to see to what extent geographers will engage in utilising these highly spatial data sources.

8.6 Conclusion

The main aim of this research was to evaluate the potential of incorporating location-aware technologies and the analysis of spatiotemporal behaviour for the design of a VTBC intervention in South Africa. Location-aware technologies, such as GPS-enabled smartphones, provide unprecedented opportunities to collect accurate data on individual spatial behaviour. Whereas smartphones are capable of both sourcing data and providing users with real-time information,

---

[2] The data analyses in this dissertation have been predominantly executed using Python (for spatial analysis) and R (for statistical analysis); as we did not have a coding background, these languages had to be learnt during the course of the project. However, for application development other skills are required.

feedback, and suggestions for travel alternatives, between sourcing the data and relaying feedback to individual commuters, significant research is required on how to obtain, clean, and interpret the data, as well as on how to account for individual spatiotemporal accessibility.

In this dissertation, it was established that there have been several attempts around the world to implement a VTBC intervention with various degrees of success. Based on an extensive literature study, two main approaches were identified to augment the current research on VTBC interventions. The first approach is to make use of field experiments that randomly divide participants into experimental and control groups. The second approach is to capitalise on the opportunities provided by location-aware technologies to gather more accurate data on individual spatial behaviour (Chapter 2). The main focus of this dissertation was on the second approach. However, these seemingly new technologies require analytical methods to transform the locational data into information on travel behaviour (i.e. impute travel characteristics). As such, methods derived from the domain of machine learning were explored to determine whether a participant is on the move or stationary (Chapter 3) with the random forest machine learning algorithm being the most effective. In addition, two novel GIS-based methods were proposed to use GPS data to reconstruct actual routes travelled by its user (Chapter 4). Both explored procedures, the *connected subset* assignment procedure based on network subset selection and the *impedance reduction* assignment procedure based on attribute adjustment, provide accurate results. In addition, both procedures effectively deal with commonly GPS-induced problems such as measurement gaps and positional drift.

In Chapter 5, attention was drawn to the actual usage of smartphones for travel data collection. Although promising, several issues were highlighted. For instance, user-related issues (e.g. user switching the application on and off), methodology-related issues (e.g. representativeness of the sample as the usage of smartphones forces the exclusion of participants that do not have these devices), and technology-related issues (e.g. impact of location-based services on the battery life of the device). In Chapter 6, the use of GPS-based activity spaces and opportunity indicators for travel behaviour analysis were explored. Based on the examination of GPS tracks with different two-dimensional operationalisations of activity spaces, it was found that the density of opportunities within an activity space is related to the size of the activity space: larger activity spaces have lower densities of opportunities than smaller activity spaces. This may suggest that individuals who have a low opportunity density are less likely to respond to external stimuli and/or awareness programmes than individuals who have a high opportunity density. The contribution of this analysis is that it lays a foundation for future work to analyse VTBC interventions and travel behaviour when the purpose of the trip is known. In Chapter 7, spatiotemporal access to goods and services, as measured through the concept of a potential path area, was accessed. By using different operationalisations of the choice universe (e.g. through a hexagonal choice universe or through points-of-interest data), it was investigated which participants have opportunities to make changes to their travel behaviour.

The findings of this dissertation indicate that there is much potential to incorporate location-aware technologies to augment the current research on VTBC interventions. Moreover, the results suggest that the incorporation of spatiotemporal measurements is crucial for the design and implementation of these interventions. Nevertheless, the findings also indicate that considerably more work needs to be done before a full VTBC intervention can be administered through a smartphone. For instance, the methods and analyses that have been described need to be extended with methodologies to, amongst other things, impute travel modes and identify activity locations from activity-travel sequences to provide a holistic view of travel behaviour and opportunities for a VTBC intervention.

References

Van Acker, V. & Witlox, F. 2009. Why land use patterns affect travel behaviour (or not). *Belgeo*. 1(1):5–26.

Behrens, R. 2004. Understanding travel needs of the poor: Towards improved travel analysis practices in South Africa. *Transport Reviews*. 24(3):317–336.

Behrens, R. & Del Mistro, R. 2010. Shocking habits: Methodological issues in analyzing changing personal travel behavior over time. *International Journal of Sustainable Transportation*. 4(5):253–271.

Behrens, R., Adjei, E., Covary, N., Jobanputra, R., Wasswa, B. & Zuidgeest, M. 2015. A travel behaviour change framework for the City of Cape Town. In *Proceedings of the 34th Southern African Transport Conference (SATC 2015)*. Pretoria: Southern African Transport Conference. 412–430.

Birenboim, A. & Shoval, N. 2016. Mobility research in the age of the smartphone. *Annals of the American Association of Geographers*. 106(2):1–9.

Bonsall, P. 2009. Do we know whether personal travel planning really works? *Transport Policy*. 16(6):306–314.

Brög, W., Erl, E., Ker, I., Ryle, J. & Wall, R. 2009. Evaluation of voluntary travel behaviour change: Experiences from three continents. *Transport Policy*. 16(6):281–292.

Cairns, S., Sloman, L., Newson, C., Anable, J., Kirkbridge, A. & Goodwin, P. 2008. Smarter choices: Assessing the potential to achieve traffic reductions using 'soft measures'. *Transport Reviews*. 28(5):593–618.

Cao, X., Mokhtarian, P.L. & Handy, S.L. 2009. Examining the impacts of residential self-selection on travel behaviour: A focus on empirical findings. *Transport Reviews*. 29(3):359–395.

Casas, I., Horner, M.W. & Weber, J. 2009. A comparison of three methods for identifying transport-based exclusion: A case study of children's access to urban opportunities in Erie and Niagara counties, New York. *International Journal of Sustainable Transportation*. 3(4):227–245.

Chai, Y., Chen, Z., Liu, Y., Tana & Ma, X. 2014. Space-time behavior survey for smart travel planning in Beijing, China. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 79–90.

Chatterjee, K. 2009. A comparative evaluation of large-scale personal travel planning projects in England. *Transport Policy*. 16(6):293–305.

Chatterjee, K. & Bonsall, P. 2009. Special issue on evaluation of programmes promoting voluntary change in travel behavior. *Transport Policy*. 16(6):279–280.

Chen, C., Ma, J., Susilo, Y., Liu, Y. & Wang, M. 2016. The promises of big data and small data for travel behavior (aka human mobility) analysis. *Transportation Research Part C: Emerging Technologies*. 68:285–299.

Eriksson, L., Garvill, J. & Nordlund, A.M. 2006. Acceptability of travel demand management measures: The importance of problem awareness, personal norm, freedom, and fairness. *Journal of Environmental Psychology*. 26(1):15–26.

Feng, T. & Timmermans, H.J.P. 2016. Comparison of advanced imputation algorithms for detection of transportation mode and activity episode using GPS data. *Transportation Planning and Technology*. 39(2):180–194.

Gärling, T. & Schuitema, G. 2007. Travel demand management targeting reduced private car use: Effectiveness, public acceptability and political feasibility. *Journal of Social Issues*. 63(1):139–153.

Gärling, T., Eek, D., Loukopoulos, P., Fujii, S., Johansson-Stenman, O., Kitamura, R., Pendyala, R. & Vilhelmson, B. 2002. A conceptual analysis of the impact of travel demand management on private car use. *Transport Policy*. 9(1):59–70.

Gatersleben, B. & Appleton, K.M. 2007. Contemplating cycling to work: Attitudes and perceptions in different stages of change. *Transportation Research Part A: Policy and Practice*. 41(4):302–312.

Hägerstrand, T. 1970. What about people in regional science? *Papers of the Regional Science Association*. 24(1):6–21.

Hashemi, M. & Karimi, H.A. 2014. A critical review of real-time map-matching algorithms : Current issues and future directions. *Computers, Environment and Urban Systems*. 48:153–165.

Howarth, C.C. & Polyviou, P. 2012. Sustainable travel behaviour and the widespread impacts on the local economy. *Local Economy*. 27(7):764–781.

IPCC. 2014. *Climate change 2014: Synthesis report. Contribution of working groups I, II and III to the fifth assessment report of the Intergovernmental Panel on Climate Change*. Core Writing Team, R.K. Pachauri, & L.A. Meyers (eds.). Geneva, Switzerland: IPCC.

Jariyasunant, J., Carrel, A., Ekambaram, V., Gaker, D.J., Kote, T., Sengupta, R. & Walker, J.L. 2011. *The Quantified Traveler: Using personal travel data to promote sustainable transport behavior*. (UCTC-FR-2011-10). Berkeley: University of California Transportation Center. [Online], Available: http://www.uctc.net/research/papers/UCTC-FR-2011-10.pdf [2015, November 12].

Kingham, S., Dickinson, J. & Copsey, S. 2001. Travelling to work: Will people move out of their cars. *Transport Policy*. 8(2):151–160.

Kwan, M.-P. 2016. Algorithmic geographies: Big data, agorithmic uncertainty, and the production of geographic knowledge. *Annals of the American Association of Geographers*. 106(2):274–282.

Kwan, M.-P. & Schwanen, T. 2016. Geographies of mobility. *Annals of the American Association of Geographers*. 106(2):243–256.

Meloni, I. & Sanjust, B. 2014. Using a GPS active logger to implement travel behavior change programs. In S. Rasouli & H.J.P. Timmermans (eds.). *Mobile technologies for activity-travel data collection and analysis*. Hershey, Pennsylvania: IGI Global. 325–340.

Meloni, I., Sanjust Di Teulada, B. & Spissu, E. 2016. Lessons learned from a personalized travel planning (PTP) research program to reduce car dependence. *Transportation*. 44(4):1–18.

Meurs, H. & Haaijer, R. 2001. Spatial structure and mobility. *Transportation Research Part D: Transport and Environment*. 6(6):429–446.

Newsome, T.H., Walcott, W.A. & Smith, P.D. 1998. Urban activity spaces: Illustrations and application of a conceptual model for integrating the time and space dimensions. *Transportation*. 25(4):357–377.

Nitsche, P., Widhalm, P., Breuss, S., Brändle, N. & Maurer, P. 2014. Supporting large-scale travel surveys with smartphones - A practical approach. *Transportation Research Part C: Emerging Technologies*. 43:212–221.

Patterson, Z. & Farber, S. 2015. Potential path areas and activity spaces in application: A review. *Transport Reviews*. 35(6):679–700.

Perchoux, C., Chaix, B., Cummins, S. & Kestens, Y. 2013. Conceptualization and measurement of environmental exposure in epidemiology: Accounting for activity space related to daily mobility. *Health & Place*. 21:86–93.

Pojani, D. & Stead, D. 2015. Sustainable urban transport in the developing world: Beyond megacities. *Sustainability*. 7(6):7784–7805.

Prelipcean, A.C. 2016. *Capturing travel entities to facilitate travel behaviour analysis: A case study on generating travel diaries from trajectories fused with accelerometer readings*. Published licentiate thesis. Stockholm, Sweden: Royal Institute of Technology (KTH).

Prochaska, J.O., Velicer, W.F., Rossi, J.S., Goldstein, M.G., Marcus, B.H., Rakowski, W., Fiore, C., Harlow, L.L., et al. 1994. Stages of change and decisional balance for 12 problem behaviors. *Health Psychology*. 13(1):39–46.

Ren, F., Tong, D. & Kwan, M.-P. 2014. Space–time measures of demand for service: Bridging location modelling and accessibility studies through a time-geographic framework. *Geografiska Annaler: Series B, Human Geography*. 96(4):329–344.

Richardson, A.J. 2003. Temporal variability of car use as an input to design of before and after surveys. *Transportation Research Record: Journal of the Transportation Research Board*. 1855:112–120.

Richardson, A.J., Seethaler, R.K. & Harbutt, P.L. 2004. Design issues for before and after surveys of travel behaviour change. *Transport Engineering in Australia*. 9(2):103–118.

Richter, J., Friman, M. & Gärling, T. 2011. Soft transport policy measures: Gaps in knowledge. *International Journal of Sustainable Transportation*. 5(4):199–215.

Schuessler, N. & Axhausen, K.W. 2009. Processing raw data from Global Positioning Systems without additional information. *Transportation Research Record: Journal of the Transportation Research Board*. 2105:28–36.

Schwanen, T. 2016. Geographies of transport I: Reinventing a field? *Progress in Human Geography*. 40(1):126–137.

Schwanen, T. 2017a. Geographies of transport II: Reconciling the general and the particular. *Progress in Human Geography*. 41(3):355–364.

Schwanen, T. 2017b. Geographies of transport III: New spatialiaties of knowledge production. *Progress in Human Geography*. In press.

Schwanen, T. & De Jong, T. 2008. Exploring the juggling of responsibilities with space-time accessibility analysis. *Urban Geography*. 29(6):556–580.

Shafique, M. & Hato, E. 2016. Travel mode detection with varying smartphone data collection frequencies. *Sensors*. 16(5):716.

Shen, L. & Stopher, P.R. 2014. Review of GPS travel survey and GPS data-processing methods. *Transport Reviews*. 34(3):316–334.

Shoval, N., Kwan, M.-P., Reinau, K.H. & Harder, H. 2014. The shoemaker's son always goes barefoot: Implementations of GPS and other tracking technologies for geographic research. *Geoforum*. 51:1–5.

Statistics South Africa. 2013. *Measuring household expenditure on public transport: In-depth analysis of the National Household Travel Survey 2013 data*. Pretoria: Statistics South Africa. [Online], Available: http://www.statssa.gov.za/publications/Report-03-20-11/Report-03-20-112013.pdf [2017, May 02].

Statistics South Africa. 2016. *Transport Series Volume I: Profile of non-motorised transport users: In-depth analysis of the National Household Travel Survey data*. Pretoria: Statistics South Africa. [Online], Available: http://www.statssa.gov.za/publications/Report-03-20-11/Report-03-20-112013.pdf [2017, May 02].

Stopher, P.R. & Greaves, S.P. 2007. Household travel surveys: Where are we going? *Transportation Research Part A: Policy and Practice*. 41(5):367–381.

Stopher, P.R., Clifford, E., Swann, N. & Zhang, Y. 2009. Evaluating voluntary travel behaviour change: Suggested guidelines and case studies. *Transport Policy*. 16(6):315–324.

Stopher, P.R., Shen, L., Liu, W. & Ahmed, A. 2015. The challenge of obtaining ground truth for GPS processing. *Transportation Research Procedia*. 11:206–217.

Taylor, M. 2007. Voluntary travel behavior change programs in Australia: The carrot rather than the stick in travel demand management. *International Journal of Sustainable Transportation*. 1(3):173–192.

Thierry, B., Chaix, B. & Kestens, Y. 2013. Detecting activity locations from raw GPS data: A novel kernel-based algorithm. *International Journal of Health Geographics*. 12(1):14.

Walters, J. 2008. Overview of public transport policy developments in South Africa. *Research in Transportation Economics*. 22(1):98–108.

Walters, J. 2013. Overview of public transport policy developments in South Africa. *Research in Transportation Economics*. 39(1):34–45.

Van Wee, B. 2009. Self-selection: A key to a better understanding of location choices, travel behaviour and transport externalities? *Transport Reviews*. 29(3):279–292.

Xiao, G., Juan, Z. & Zhang, C. 2015. Travel mode detection based on GPS track data and Bayesian networks. *Computers, Environment and Urban Systems*. 54:14–22.

Yip, N.M., Forrest, R. & Xian, S. 2015. Exploring segregation and mobilities: Application of an activity tracking app on mobile phone. *Cities*. 59:156–163.

Zhang, Y., Stopher, P. & Halling, B. 2013. Evaluation of south-Australia's TravelSmart project: Changes in community's attitudes to travel. *Transport Policy*. 26:15–22.