DISSERTATIONES
GEOGRAPHICAE
UNIVERSITATIS
TARTUENSIS
74

JANIKA RAUN

Mobile positioning data for tourism destination studies and statistics

DISSERTATIONES GEOGRAPHICAE UNIVERSITATIS TARTUENSIS
74

DISSERTATIONES GEOGRAPHICAE UNIVERSITATIS TARTUENSIS
74

JANIKA RAUN

Mobile positioning data for tourism destination studies and statistics

UNIVERSITY OF TARTU
Press

Department of Geography, Institute of Ecology and Earth Sciences, Faculty of Science and Technology, University of Tartu, Estonia

Dissertation was accepted for the commencement of the degree of *Doctor philosophiae* in human geography and regional planning at the University of Tartu on April 6th, 2020 by the Scientific Council of the Institute of Ecology and Earth Sciences, University of Tartu.

Supervisors: Dr. Anto Aasa
Senior Research Fellow in Human Geography
University of Tartu, Estonia

Professor Noam Shoval
Hebrew University of Jerusalem, Israel

Professor Rein Ahas
University of Tartu, Estonia

Opponent: Professor John Östh
Uppsala University, Sweden

Commencement: Senate Hall, University of Tartu, Ülikooli 18, on June 26th 2020 at 10:15.

Publication of this thesis is granted by the Institute of Ecology and Earth Sciences, University of Tartu.

ISSN 1406-1295
ISBN 978-9949-03-349-2 (print)
ISBN 978-9949-03-350-8 (pdf)

University of Tartu Press
www.tyk.ee

To the memory of Professor Rein Ahas who is much missed.

„Kõik PhD tööd on arenguetapiks. Aasta hiljem tundubki nigel. Ja see näitab, et oled arenenud. Selle etapi kaudu võid jõuda aga kaugele. Kui kaua selle üle mõtled, siis võib hiljem jõuda ☺„ –
Rein Ahas 29.03.2017

CONTENTS

LIST OF PUBLICATIONS

This dissertation is based on the four publications shown below. The first three have been published in international peer-reviewed journals, the fourth publication was accepted for publication in European Planning Studies on 8 April 2020. Publications are referred in the dissertation by their respective Roman numeral.

I. **Raun, J.,** Ahas, R., Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202–212. https://doi.org/10.1016/j.tourman.2016.06.006

II. Saluveer, E., **Raun, J.,** Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A., Silm, S. (2020). Methodological framework for producing national tourism statistics from mobile positioning data. *Annals of Tourism Research*, 81, 102895. https://doi.org/10.1016/j.annals.2020.102895

III. **Raun, J.,** Shoval, N., Tiru, M. (2020). Gateways for intra-national tourism flows: Measured using two types of tracking technologies. *International Journal of Tourism Cities*. https://doi.org/10.1108/IJTC-08-2019-0123

IV. Silm, S., Jauhiainen, J. S., **Raun, J.,** Tiru, M. (in press). Temporary population mobilities between Estonia and Finland based on mobile phone data and the emergence of a cross-border region. *European Planning Studies*.

Author's contribution to the articles: '*' denotes a minor contribution, '**'denotes a moderate contribution, '***' denotes a major contribution.

	I	II	III	IV
Original idea	**	**	**	*
Study design	**	**	***	*
Data processing and analysis	***	*	***	*
Interpretation of the results	***	**	***	**
Writing the manuscript	***	***	***	**

ABBREVIATIONS

CDR	Call Detail Record
DDR	Data Detail Record
GDPR	General Data Protection Regulation
GIS	Geographic Information System
GPS	Global Positioning System
ICTs	Information and Communication Technologies
IMSI	International Mobile Subscriber Identity
MMS	Multimedia Messaging Service
MNO	Mobile Network Operator
MPD	Mobile Positioning Data
RFID	Radio-Frequency Identification
SIM	Subscriber Identification Module
SMS	Short Message Service
UGC	User-Generated Content
Wi-Fi	Wireless networking technology

1. INTRODUCTION

Until the coronavirus outbreak in the beginning of 2020, the world has been facing an increasing number of tourists travelling worldwide, reaching up to 1.5 billion international tourist arrivals in 2019 and the prognosis done before the virus outbreak show that the growth is going to continue according to World Tourism Organization (2020). Thus, the contribution of tourism to the overall world economy has been steadily growing. Therefore, ever-increasing tourist numbers and expenditure on one side and possible unexpected negative impacts on the other side have, as a result, given rise to the need for a detailed and accurate overview of the volumes of tourism flows that our destinations are receiving. However, this need consists of two contradictions. First, there is a lack of consensus over the meaning and scope of the term 'tourism destination' itself. Second, there are relatively few data sets that could provide sufficient detailed information about the spatiotemporal characteristics of tourist visits to destinations on various geographical and temporal scales.

Despite the fact that destination is the fundamental concept in tourism and a key unit in tourism management and marketing, different definitions exist and there is a lack of homogeneity among them (Pearce, 2014). Older and more traditional understandings consider the destination as a geographical area with definite borders that is marketed as one complete unit, e.g., country, national park, island, city, old town, and can exist on several different geographical levels (Hall, 2005a; UNWTO, 2010). However, other approaches see destination as a socially constructed space without fixed geographical boundaries (Framke, 2002; Saarinen, 2004), that is in constant change and the meaning of it is formed in the minds of the tourists (Saraniemi & Kylänen, 2011). Therefore, due to the complexity of the term 'destination', various definitions exist that are rather fragmented and incomplete and derived from specific research interests; in some cases the term is even taken for granted and is not defined at all (Pearce, 2014).

Several authors who have tried to conceptualise the essence of tourism destination have also noted the need to analyse tourism destinations more thoroughly using different data sets to give meaning to the term using empirical analysis (Framke, 2002; Pearce, 2014; Pechlaner, Pichler, & Herntrei, 2012; Saarinen, 2004). One possibility could be to analyse destinations based on the movements of tourists and the places and routes they have passed through while on their tourism trips (Beritelli, Bieger, & Laesser, 2014; Pechlaner et al., 2012). Regardless, few studies have been made because there has been a lack of temporally and spatially precise data about the tourist visits made (Batista e Silva et al., 2018; Beritelli, Reinhold, & Laesser, 2020). Most commonly used traditional data sources, such as accommodation statistics, border-crossing statistics, visitor surveys, and travel agency data are, however, rather limited and do not take into account destinations at various geographical scales, or include the various forms of tourism and the increasing mobility of tourists. For example, accommodation statistics do not include information about same-day visitors, the

majority of 'Airbnb' guests, or people who are staying with friends or family. There is no border control between any of the European countries within the Schengen area. Furthermore, visitor surveys capture only a small number of people in specific settings and it is difficult to generalise the results to total populations. Therefore, it is important to find better data sources that could enable to analyse the movements of tourists at various geographical levels and temporal scales with sufficient detail, and as a result better understand the essence of tourism destinations.

Rapid advances in information- and communication technologies (ICTs) in recent decades have opened up several new possibilities for collecting data about human movement, including tourists. One example of this is the widespread distribution of mobile phones. In 2019, 67% of the world's population owned a mobile phone, and 65% of all connections were smartphones (GSMA, 2020). In Europe, North America, and China, that constitute the majority of the world's tourism international arrivals and receipts (World Tourism Organization, 2019), the subscriber penetration rates were 86%, 83%, and 82% respectively (GSMA, 2020). Therefore, vast majority of people are constantly carrying phones with them to everywhere they go, thereby providing an opportunity to collect detailed location information about the movements of people, including tourists (Shoval, 2007). According to a review by Shoval and Ahas (2016) the most popular tracking technology to be used in tourism research is GPS tracking, followed by mobile positioning data, Bluetooth tracking, geocoded social media, and photo database analysis. The number of studies using various tracking data sets is increasing, although the majority of them still have a fixed timeframe and a limited spatial extent, e.g., an attraction, a national park, or a city (Shoval & Ahas, 2016). Therefore, there is lack of research that analyses tourist movement on a wider geographical scale over long periods (such as months, or years).

An exception is passive mobile positioning data (MPD) – automatically collected log files by mobile network operators (MNO) for billing purposes – that enable to analyse tourist movement on various spatial scales, from local to global, and during different timeframes, from hours to years. MPD about tourist flows can therefore be used to analyse and understand the essence of tourism destinations. Although the need to study tourism destinations based on tourism flows has been pointed out before (Beritelli et al., 2014; Pechlaner et al., 2012), and even the possibility to use big data for that has been noted (Beritelli, Crescini, Reinhold, & Schanderl, 2019), few studies have been made. The novelty of this thesis lies in its attempt to bridge the gap between conceptual tourism destination research and empirical studies using big data for tourism movement analysis.

The general aim of the thesis is twofold:

1. to achieve a better understanding of how mobile positioning data can be used to produce tourism statistics that could be used for destination analysis;
2. to demonstrate the use and applicability of mobile positioning data in tourism destination analysis and conceptualising tourism destination.

This thesis focuses on incoming tourism to Estonia and uses inbound roaming data about the spatiotemporal movement of foreign visitors in different time periods in Estonia during 2011–2017. In order to achieve the general aim, the following tasks were set out:

a) to describe a method for extracting inbound tourism statistics from roaming data for an entire country, Estonia (**Article II**);
b) to evaluate the representativeness of tourism statistics derived from mobile data compared with accommodation statistics (**Article II**);
c) to analyse the geographical, temporal, and compositional dimensions of tourist destinations (the latter dimension being the country of origin of the tourist), based upon the spatiotemporal behaviour of foreign visitors in Estonia (**Articles I, III, IV**);
d) to deduce destination functions from the spatiotemporal behaviour of tourists (**Article III**).

2. THEORETICAL FRAMEWORK

2.1 Theoretical concepts of tourism destination

The World Tourism Organisation (UNWTO) defines tourism destination 'as the place visited that is central to the decision to take the trip. However, if no such place can be identified by the visitor, the main destination is defined as the place where he/she spent most of his/her time during the trip. Again, if no such place can be identified by the visitor, then the main destination is defined as the place that is the farthest from the place of usual residence' (UNWTO, 2010, p. 13). This definition remains the standard for collecting tourism statistics and is also the most widely used definition in tourism studies. However, this classic geographical approach is a simplification that does not allow to describe the complexity of destinations.

Therefore, the UNWTO's definition is a good example of a static definition that focuses on one aspect alone – place. This is the problem in most approaches that define tourism destinations. They tend to concentrate on and emphasise only some aspects whilst totally forgetting the other perspectives (Framke, 2002). However, there have been some authors who have tried to conceptualise the essence of tourism destination into one complete framework. See, for example, Lew's (1987) ideographic, organisational and cognitive perspectives; Framke's (2002) elaboration on business-related versus socio-cultural approach; Saraniemi and Kylänen's (2011) economic geography-, marketing and management,- and customer-oriented views on destination; or most recently, the review by Pearce (2014) who concluded that the most used concepts for depicting and analysing tourism destinations are industrial districts, clusters, networks, systems, social constructs, and the recurring elements in all of them are the geographical dimension, the mode of production and the dynamic dimension.

However, this static comprehensiveness has also been criticised by Beritelli et al., (2014), who deconstructs the essence of tourism destination and proposes an alternative dynamic viewpoint where 'a destination's space is the playground of different supply networks activated by visitor flows at different times of the year (week) and with different durations' (Beritelli et al., 2014, p. 406). Therefore, the physical areas that comprise destinations vary for different actors and tourists, and may overlap (Beritelli et al., 2014). This dynamic approach for depicting tourism destinations is also best in line with the current world situation in which the rapid changes in information and communication technologies are constantly changing the way we travel and communicate, as well as the way we carry out research (Law, Buhalis, & Cobanoglu, 2014; Law, Leung, Cheng, & Chan, 2020).

Even though the approach taken by Beritelli et al.'s (2014) is most fully in accordance with the current situation in terms of worldwide tourism, it was not founded with a focus on what kind of new digital data sources could be used for an empirical analysis of the concept, and neither does it take into account the changing role of tourism itself or overtourism problems. Therefore, in this thesis,

an additional theoretical framework will be proposed for tourism destinations, one that derives from the previous theoretical, methodological, and empirical work on the topic, and which focuses on the new technological developments that are making it possible to study the essence of tourists, their visits, and their destinations in much more detail than was previously possible. The proposed framework was first introduced and published in 2016 in **Article I**. This framework was mainly drawn up by keeping in mind the essence of mobile positioning data, but it also applies to other forms of location and social media data.

According to the framework, destination is the activity space of a tourist and it has five dimensions that can be measured with space-time tracking data (**Article I**). First, **geographical** dimension, where tourism is happening – where do tourists actually go and what places do they visit? Even though virtual travels exist today, people still prefer to physically go to places and see things with their own eyes. If the traditional approaches have mostly seen the geographical component of tourism destination as one continuous territory (Hall, 2005a; UNWTO, 2010) which could be marketed as one complete unit (Buhalis, 2000), then the possibility to use detailed, continuous location points from tracking data enable to provide a much more dynamic viewpoint of the geographical dimension of a destination. For example, detailed tracking data enable to distinguish smaller destination areas inside a bigger destination area and those destinations may overlap and still possess a different image. This existence of sub-destinations and overlapping among destinations has been stated in theoretical literature covering tourism destination planning perspectives (Beritelli et al., 2014; Dredge, 1999), but it has rarely been confirmed by means of detailed empirical analysis. Only recent developments in new digital tracking data sets have opened up the possibility of being able to carry out detailed analyses to this extent (Mckercher, Hardy, & Aryal, 2019).

Second, **temporal** dimension, when tourism is happening – when are tourists visiting and for how long do they stay? Tourist movements cannot be viewed without including the time factor. However, the traditional theories about tourism destinations have been relatively scarce in elaborating on the aspect of time, stating mostly that destinations can appear on various temporal scales whenever there is a visitor engaged in tourism activities (Framke, 2002; Saraniemi & Kylänen, 2011). Nevertheless, the recent advancements in and popularity around using tracking technologies have helped to rediscover the time geography concept that was developed by Hägerstrand (1970). The conceptual framework of time geography helps to describe tourist movement paths in time and space, and to understand the time-space constraints behind those movements. This theoretical knowledge, combined together with digital tracking data, can be used in tourism destination analysis to make it possible to understand tourist patterns on different temporal scales (Birenboim, Anton-Clavé, Russo, & Shoval, 2013; Grinberger, Shoval, & McKercher, 2014; Yun & Park, 2015).

Third, **compositional** dimension, who are visiting destinations? Even though tourists constitute the core of tourism, they have not been at the forefront of destination analysis. Compared to business and administrative actors they are less

frequently included in any such discussions (Pearce, 2014). The reason behind this lies in the lack of data, as so far detailed information about tourists themselves could be collected via surveys alone. However, the richness of various digital data sources enables to collect different information about users. For example, tourists who are active on social networking sites such as Twitter, Instagram, Flickr, or Weibo provide user info such as name, profile picture, short bio, or even home location. This information, combined together with the content and geographical location of the posts themselves, can be used to derive additional information such as age, gender, or place of residence (Toivonen et al., 2019). Therefore, segmenting visitors is somewhat different for each data set. For example, MPD enable to distinguish a user's country of origin, and, in addition, to use the spatio-temporal variables of the visits (e.g., time, duration, geographical locations, and the number of previous visits) to differentiate between tourists (Kuusik, Tiru, Ahas, & Varblane, 2011; Tiru, Kuusik, Lamp, & Ahas, 2010).

Fourth, **social** dimension, what is the destination all about? Traditional socio-cultural literature supports that destinations are socio-cultural structures without definite borders that are formed in the course of social practise (Framke, 2002; Saarinen, 2004; Saraniemi & Kylänen, 2011) and include also intangible aspects such as culture, image, and spatial narratives (Iwashita, 2003; Lichrou, O'Malley, & Patterson, 2008). Therefore, new digital data sets again open up new horizons for analysis, as tourists voluntarily share their emotions and opinions via text or photos, likes or comments, reviews or blogs on social media platforms with their friends and family or even wider audience (J. Li, Xu, Tang, Wang, & Li, 2018; Toivonen et al., 2019). This information can be used to gain deeper understanding of the destination and, for example, find out the core attractions and hot-spots of destinations (García-Palomares, Gutierrez, & Mínguez, 2015; Kádár, 2014), making it possible to identify connected places (Zoltan & McKercher, 2014).

Fifth, **dynamic** dimension, how is destination changing? This dimension is closely related to time and focuses on changes in tourism destinations over time. The most prominent theory behind it is Butler's (1980) 'Tourism Area Life Cycle' model which illustrates changes in the popularity of destinations over time as represented in the number of visitors. Thus, dynamic visitor numbers are the proof that destination areas evolve over time and, furthermore, social scientists also emphasise the creation and development of destinations (Saarinen, 2004; Saraniemi & Kylänen, 2011). Thus, understanding these changes is important, however, there is a limited number of data sets available with sufficient detail that could provide information over decades. New digital data can be of some help here. For example, bank transaction details are collected over long time periods and could shed some light to the expenditure patterns at destinations (Sobolevsky et al., 2014); or mobile positioning data collected by mobile network operators which include various information and could be used to analyse volumes and spatiotemporal patterns of visits over years (Eurostat, 2014a). Regardless, this has rarely been done due to privacy issues.

Thus, this five-dimensional framework is proposing a more dynamic approach, one in which destination is seen as the realised mobility space, or in other words,

the activity space of a tourist which can be derived from tracking data. It consists of the locations and routes travelled through (Golledge & Stimson, 1997) by a tourist over various time periods. Thus, it is possible to distinguish and analyse destinations based on who, when, and where is visiting. In addition, it is possible to analyse what tourists are doing and analyse changes over time. This approach is, most importantly, data driven. Therefore, there are several aspects of destination that this framework does not cover or where it is not applicable. However, destination research areas in which new digital tracking data sources, including user-generated data, could be of use is rapidly increasing. Therefore, it is important also to develop the theoretical conceptual underpinnings, because the theory that is available so far has been created from traditional data viewpoints alone.

2.2 Tourist movement analysis

When considering the fact that destination is formed by the movements of tourists, it is important to understand how tourism flows have been analysed and conceptualised. According to Leiper (1979), the tourism system consists of three geographical elements: origin, destination, and the linkage between the two. Global movements take place from origin to destination and local movements take place within a destination. Correspondingly, McKercher and Lew (2004) have defined the movement to destination as the 'transit component' and movements inside destination as the 'touring component'. In other words, also known as inter-destination movements and intra-destination movements. Inter-destination movements are mostly influenced by destination choice, which is dependent on many different factors, such as time availability, income, price level, human 'push' factors (personal motivations, prior visits, etc.), and physical 'pull' factors (geographical location, geomorphology, climate), and many others (Lau & McKercher, 2006; Oppermann, 1999). On the other hand, intra-destination movements are more heavily influenced by destination characteristics (e.g., locations of accommodation and attractions, the structure of the road network, transportation accessibility), and tourist characteristics (e.g., time budget, emotional attachment, tourist motivation and interests, knowledge and information of destination) (Lew & McKercher, 2006). For extensive reviews about factors that affect tourist movements see McKercher & Lew (2004) and McKercher & Zoltan (2014).

However, intra-destination movements cannot be analysed without defining the spatial extent of a destination. Therefore, it is most commonly predefined by the researchers as a geographical area. For example, local tourism destination – a concept of an area that can be consumed in a daytrip, as proposed by Lew and McKercher (2006). It was empirically tested using trip diaries and surveys in Hong Kong by McKercher and Lau (2008). The popularity of analysing intra-destination movements has grown rapidly in the recent decade due to the availability of new digital tracking data sets that provide sufficient spatiotemporal detail. Most prominent are intra-destination analysis in urban settings which use GPS data (Shoval & Ahas, 2016), referred to also as multi-attraction travel

(Caldeira & Kastenholz, 2017). Before these technical advancements took place, intra-destination movements received less attention due to the difficulties in obtaining and analysing the data. Specifically, traditional visitor statistics are collected at discrete points in the tourism system, and for this reason intra-national movement analysis are not possible.

Another concept that is strongly affected by the definition of destination is multi-destination travel. Quite often, in the case of multi-destination trips, destinations are defined as places in which a tourist spends an overnight (Ferrante, Abbruzzo, & De Cantis, 2017; Parroco, Vaccina, De Cantis, & Ferrante, 2012; Wu & Carson, 2008). Therefore, multi-destination trips are commonly analysed by taking into account the entire trip. However, this approach has its limitations, because places that have been visited during same-day trips or places in which no overnight was spent are, by definition, excluded (Wu & Carson, 2008). Nevertheless, there are no fixed spatial or temporal thresholds for distinguishing between intra-destination or multi-destination travel. As the essence of both of them is similar, only with the differences in temporal and spatial scales, thereby multi-destination research has faced similar problems in collecting sufficient data. Carrying out surveys or filling in trip diaries is time-consuming and expensive and the samples are small.

In this thesis, the focus is on one destination level – the country of Estonia – and all movements inside it during the entire visit. The concept behind tourist movements within a destination country follows the same principles as introduced for multi-destination trips: 'movements' between different locations and 'stays' at these locations (Oppermann, 1992a). The two components together can be generalised into a movement pattern (dynamic connectivity), which can be described by the intensity (the volume and frequency) of the movement (Bowden, 2003). Even though, country as a destination, is the common scale in tourism statistics data that is collected by national statistical bureaus, and is the main analysis unit used in tourism research, there are almost no sufficient data available about the movements inside a country. Despite the fact that, progress in ICTs has opened up new possibilities for analysing multi-destination trips or tourist movements on a larger scale (e.g., regional, national), there is still very little research done using tracking technologies on a national scale (Shoval & Ahas, 2016). Only some examples exist, such as using geolocated photos from Flickr in Austria (Önder, 2017), or passive mobile positioning data in Estonia (Ahas, Aasa, Roose, Mark, & Silm, 2008).

No matter which geographical scale one chooses to look at, the general structure behind tourist movements is the division of movement between different locations and stays at these locations. These movements can be expressed as lines, and stays at locations as points, fundamental concepts in geoinformatics. Furthermore, these movements can be routed to the actual road network. Therefore, GIS methods have been applied most to visualise tourist movement patterns (Grinberger et al., 2014; Lau & McKercher, 2006; van der Knaap, 1999). Likewise, movements can be expressed as linkages, and stays at locations as nodes, both serving as core structures in network analyses. Network analysis has been applied to tourism

movement analysis between destinations (nodes) e.g., in Sicily (Asero, Gozzo, & Tomaselli, 2015; D'Agata, Gozzo, & Tomaselli, 2013). More specifically, social network analyses have been applied to study, for example, overseas tourist movement patterns in Beijing during the Olympic Games (Leung et al., 2012) or cross-provincial tourism flows in China (Peng, Zhang, Liu, Lu, & Yang, 2016). In addition, the sequence alignment method has been used to analyse the sequential aspects of tourism movement (Bargeman, Joh, & Timmermans, 2002; Shoval & Isaacson, 2007a; Shoval, McKercher, Birenboim, & Ng, 2015).

2.3 Identifying destination functions from tourist movements

Tourist movement analysis is the key to identifying the role of each visited place within the trip as a whole. Thus tourism flows themselves are the core element in destination positioning research (Lew & McKercher, 2002; Liu et al., 2012) and in destination management (Beritelli et al., 2014). Tourism flows, in turn, are not random but are spatially and temporally differentiated (Wu & Carson, 2008), depending on the national and geographical background of the tourists (Bowden, 2003). Several studies have suggested that a hierarchical system of tourist flows exists, as some places attract the majority of tourists while the demand for others is weaker (Guedes & Jiménez, 2015; Murphy & Keller, 1990; Oppermann, 1992a).

Therefore, in order to understand the destination's role in the overall trip, different methodologies have been applied. Some of the earlier studies used various indices calculated based on visitor nights, to illustrate the function of destination. For example, Pearce and Elliott (1983) used a 'Trip Index' (the proportion of the total nights spent at a destination out of the total nights on the trip) to determine whether a place served as a main destination or rather as a stopover within a larger tour. Leiper (1989) came up with a similar descriptive measure, the 'Main Destination Ratio', defined as 'the percentage of arrivals by tourists in a given place for whom that place is the main or sole destination in a current trip' (p. 533) of the total arrivals. Besides analysing which places function as main or stopover destinations indices have been used to analyse the overall dispersal of the trip. For example, Oppermann (1992b) proposed the 'Travel Dispersal Index' which takes into account more than just the spatial dimension; it also considers the length of stay in the country, the number of overnight destinations, the number of different accommodation and transportation types, and travel organisation (whether a package tour or an individual trip). His results from Malaysia showed that tourists are mainly concentrated in the main cities of Kuala Lumpur and Penang, there by the regional economic impact is poor (Oppermann, 1992b). More recently Koo, Wu, and Dwyer (2012) supplemented the existing dispersal ratio approaches in Australia using the 'Individual Dispersal Propensity', which illustrates the dispersal tendency of an individual trip.

In addition to descriptive indices, conceptual models of travel itineraries have been used the most, to understand the relative position of a destination. The first attempts to model tourist itineraries were carried out in the early 1990s. Although, the initial conceptual models concentrated on general inter-destination movements from origin to destination and back (Flognfeldt, 1999; Lue, Crompton, & Fesenmaier, 1993; Mings & Mchugh, 1992; Oppermann, 1995). These were followed by Lew and McKercher's (2002) study conducted in Hong Kong which was one of the first to model the itinerary from the perspective of the relative location of a destination within the larger itinerary pattern. Later, McKercher and Lew (2004) argued that all previous patterns that had been developed could be aggregated into four broad itinerary types: a) single destination, with or without side trips; b) transit leg and circle tour at a destination; c) circle tour with or without multiple access, egress points; d) hub-and-spoke itinerary in which tourists base themselves in one place and take side trips to other places. Later on, in 2006 Lew and McKercher further developed their conceptualisation and concluded. that in the case of local destination areas (possible to make daytrips), tourist movement patterns can be characterised by territorial dimension, showing the spatial distance of the movement from the starting point, and linear dimension, indicating the complexity of the movement.

Previously described studies illustrate well how theoretical knowledge on tourist movement analysis has developed and evolved over the years. However, there is a lack of consensus when it comes to naming the destinations (location points, nodes) based on their relative position within the itinerary. Lohmann and Pearce (2010) criticise that concepts for various types of destination have been used rather loosely, and that more focus is given to tourist flows and itineraries. Therefore, several classifications and definitions now exist in parallel, depending on the context and discipline. Lew and McKercher (2002) identified five destination types based on the location within an itinerary: single (one main destination), gateway (first destination encountered), egress (last destination visited), touring, and hub (destination visited more than once during one trip); Pearce (2001b) identified between origin, destination, gateway, hub, multiple functions; Lohmann and Pearce (2010) distinguished between origin, destination, gateway, hub, stopover, and multiple functions. As can be seen from these classifications, destinations may have multiple and overlapping functions (Lohmann & Pearce, 2010; Pearce, 2001a, 2001b), which further complicates the creation of a clear classification system.

Of the destination functions that have been identified (excluding general origin and destination), the use of 'gateway' has been most prominent. Gateways are the entry and exit points to and from a destination (Burghardt, 1971; Pearce, 2001b). Therefore a gateway is always the first place encountered and it may be a destination itself (Lew & McKercher, 2002), but it does not have to be. Gateways can exist on various geographical scales, from entrances to national parks to regional and international gateways (Burghardt, 1971; Zurick, 1992). Regardless of scale, gateways have a mediating role between the outside world and the ancillary areas, controlling the flows of goods and people and linking networks to

one another (Burghardt, 1971). Quite often, capital cities are also primary national gateways, as the national and international transportation nodes (airports, seaports) are located there (Maitland, 2012). According to Wu and Carson (2008), cities that function as gateways are also more likely to become tourist hot-spots.

2.4 Traditional data sources for tourist movement analysis

Collecting statistics on tourism is 'essential to the measurement of the volume, scale, impact and value of tourism at different geographical scales from the global to the country level down to the individual destination' (Hall & Page, 2014, p.76). Therefore, it is a key source of information for destination managers, tourism planners, local administrators, and tourism researchers. The key to understanding the volumes and effects of tourism is to become aware of physical tourist flows (Eurostat, 2014a). The most traditional way to collect information about tourist volumes (arrivals, departures, visits, stays), expenditures (tourist spending), and tourist- and trip-related characteristics (motivation, interest, visited places etc.) is to carry out surveys completed by a sample of individuals or households. Surveys were seen as the main data source for domestic and outbound tourism statistics in the International Recommendations for Tourism Statistics 2008 (UNWTO, 2010) and in Eurostat's Methodological Manual for Tourism Statistics (Eurostat, 2014b).

In addition to surveys for residents that cover specific reference periods, border surveys and visitor surveys carried out at major tourism attractions or transportation hubs can also register information about trips being made by non-residents (Eurostat, 2014a). Surveys have a number of disadvantages. The main one being their high cost, resulting in generally small samples that do not represent the whole population, and consequent possible bias in the results (Beaman, Huan, & Beaman, 2004; Eurostat, 2014a, 2014b). Surveys are usually retrospective and the geographical scale is limited to a regional level (De Cantis, Parroco, Ferrante, & Vaccina, 2015). Another downside of surveys is the lack of clarity in definitions (Volo & Giambalvo, 2008), for example questions like what is a tourism trip or what is the main destination are left to decide on the subjective feeling of the respondent (Eurostat, 2014b) that may lead to divergent results. Thus those disadvantages may threaten the statistical comparability over longer periods or between different geographical areas.

Apart from surveys, national statistical institutes use data that have been reported by accommodation establishments to cover domestic and inbound tourism. The main reported variables are the number of tourists accommodated, and the number of overnight stays by type of accommodation and by country of residence (Eurostat, 2014a). The greatest limitation in the use of accommodation statistics is their failure to capture unobserved tourism, which has two components (De Cantis et al., 2015). First, nights spent in unofficial accommodation (e.g., summer houses, friends' and relatives' homes), also known as unmeasured tourism

(De Cantis et al., 2015). For example, a study made in Australia found that visiting friends and relatives travel constitutes 48% of Australia's total overnight tourism market, substantially more than previously assumed by the official data (Backer, 2012). Second, nights spent in official establishments, which are deliberately unreported for fiscal reasons (e.g., obtain tax advantages) (De Cantis et al., 2015; Guizzardi & Bernini, 2012). Another important set of problems arises from how the data are collected. Some tourists stay overnight at several different establishments during a trip and are counted at each one, constituting double counting (De Cantis et al., 2015). The bigger the territory under observation, the greater is the tendency to do a multi-destination trip and therefore the double counting increases (Parroco et al., 2012). Thus accommodation data do not enable to analyse tourist movements, just volumes.

A number of potential problems, from item definition to sample bias, therefore exist. Neither official surveys nor traditional statistics on accommodation reveal the spatiotemporal patterns of tourists at a sufficiently fine resolution (Batista e Silva et al., 2018) to allow the analysis of tourists' movements over the entire duration of their visits. Therefore, to fill the gap where national and centrally organised data collection systems do not provide sufficient information about the movement patterns of tourists, tourism researchers have used additional methods to collect data about tourist movements. The most well-known are direct observation and self-completed trip diaries and maps that may be combined with pre or follow-up interviews, known also as time-space budget technique (McKercher & Lau, 2009; Shoval & Isaacson, 2007b). Former is done either unobtrusively or with the permission from the tourist by following the individual and marking down the movement trajectory of the visitor and the time spent in important tourism attractions. However, this has been used rarely, as the data collection is resource and time-consuming and, in consequence, the samples are relatively small. Furthermore, it also raises many ethical questions (McKercher & Lau, 2009).

The latter method, time-space budget technique, is more common in tourism studies. It focuses on the frequency, sequencing, and duration of tourist activities over a given period e.g., from one day up to a week (Shoval & Isaacson, 2007b). This is mostly based on recall time-space diaries in the form of a questionnaire or interview, or on real time diaries or maps that are filled in by the tourists themselves. In the first case, people may not remember all the details about their movement sequences and timings. The other approach, however, is a burden to the tourists as they have to keep a detailed track of the visited places while touring. Using maps is good for tracking the exact routes taken, but it has an inherent scale problem, resulting in loss of fine detail (McKercher & Lew, 2004). Therefore, in both cases, the quality of the results is dependent on the participant's memory, enthusiasm, and will to actively and truly take part.

2.5 New emerging data sources for tourist movement analysis

The essence of tourism and travel behaviours has changed in recent decades due to rapid advances in ICTs. The number of novel ICT-based data sources has likewise increased, giving researchers new ways of analysing tourism. Contrary to traditional data, big data is 'generated continuously, seeking to be exhaustive and fine-grained in scope, and flexible and scalable in its production' (Kitchin, 2014, p.2). High-powered computation and new analytical techniques have also made it possible to overcome the challenges of big data over the recent decades, such as volume, velocity, high relationality, exhaustivity and variety, timeliness and dynamism, messiness and uncertainty (Kitchin, 2014). Many of these new big data sources measure human activity or movement; hence, they can be used in tourism research and statistics (Demunter, 2017).

In a comprehensive review of studies using big data in tourism research, Li,Xu, Tang, Wang, and Li (2018) recognise three primary sources for generating big data, namely tourists as users, devices or sensors used by tourists, and various operations related to tourists. Hereafter, first two data sources are described in more detail as they enable to do tourism movement analysis on various geographical and temporal scales. Tourism-related operations such as web searching (e.g., Google, Baidu), webpage visiting (e.g., Wikipedia page views), online booking and purchasing (e.g., Booking.com) on the other hand are mostly used for predicting tourist flows and demand estimation, not for analysing the actual visits or calculating the actual number of visits, and they are least used in tourism research (Demunter, 2017; J. Li et al., 2018).

Li et al., (2018) found in their review that most of the research on big data related to tourism is derived from user-generated content (UGC) such as online photos and texts. The main sources for photos are photo-sharing platforms (e.g., Flickr, Panoramio, Instagram) and textual data originate mainly from online reviews (e.g., Tripadvisor, Booking, Expedia, Yelp, etc.) and posts from Twitter and its Chinese alternative Sina Weibo (J. Li et al., 2018). The value of UGC for tourism research and especially tourist movement analysis lies in the metadata embedded in photos or texts, which includes information about the time and location when the photo or text was produced. This spatiotemporal information enables to analyse tourist flows and volumes at various spatial and temporal scales.

Because UGC are low/no-cost and easy to access several studies using different data sets have been made. For example, geo-located photos from Flicker have been used for tourist movement analysis on global (Vu, Li, Law, & Zhang, 2018), regional (Girardin, Fiore, Ratti, & Blat, 2008; Stienmetz & Fesenmaier, 2019), and city level (Kádár, 2014; Vu, Li, Law, & Ye, 2015). In addition, photos have been used for tourism demand analysis (Önder, Koerbitz, & Hubmann-Haidvogel, 2016) and multi-destination trip classifications (Önder, 2017) on country level. Additionally, geo-located tweets have been used to understand tourist mobility patterns on a global (Hawelka et al., 2014), European (Provenzano,

Hawelka, & Baggio, 2018) and a regional (Chua, Servillo, Marcheggiani, & Moere, 2016) level. Although UGC are more increasingly being used in tourism research and tourist flow analysis, it is rarely included in the production of official tourism statistics. However, correlations with official tourism statistics have been shown. For example, Kádár (2014) used geotagged photos from Flickr and found a good correlation with official statistics on tourist arrivals, bed nights and singular attraction attendance in Vienna, Prague, and Budapest. Tenkanen et al., (2017) used data from Instagram, Twitter, and Flickr and found that social media-based monthly visitation patterns match relatively well with official visitor counts for national parks in Finland and South Africa.

According to the review by Li et al., (2018), the second most widely used big data source in tourism research comes from devices or sensors that have been used by tourists. In the era of the 'Internet of Things', devices with various sensors (e.g., smartphones) track tourists' movements and provide high-quality spatio-temporal data, such as GPS data, mobile positioning data, Bluetooth data, radio-frequency identification (RFID) data, and Wi-Fi data (J. Li et al., 2018; Shoval & Ahas, 2016). Furthermore, data from traffic sensors, street cameras or weather stations could be used for analysis or correlating the tourist numbers, but these sources are not further elaborated here. Of the various types of tracking data available, GPS data have been the most widely used in tourism research, followed by mobile positioning data (Shoval & Ahas, 2016). Bluetooth, RFID and Wi-Fi data have been used less, however distributing an ad-hoc sensor network over a specific study area has vast potential for registering small-scale movements including indoor movements (Tsai & Chung, 2012; Versichele et al., 2014).

Following the division by Shoval and Ahas (2016), tourism research using tracking technologies have gone through three major stages. There are no clear temporal boundaries between the stages, because the progress of using tracking data sets in tourism research varies for each specific data set, however the general characteristics of the stages are the same for all tracking data sets. The very first studies were made to understand how these new digital tracking data sets could be used to study human mobility in general (Ahas & Mark, 2005; Asakura & Hato, 2001; Isaacson & Shoval, 2006) and soon after, specific studies with the focus on tourists followed (Ahas, Aasa, Mark, Pae, & Kull, 2007; McKercher & Lau, 2009; Shoval, 2008). The main aim of those studies was to test the potential of the data and the feasibility of the methods. Early experimental studies were followed by numerous descriptive studies that focused on revealing actual visitor patterns at destinations and therefore hoping to uncover new aspects of tourism (Ahas et al., 2008; Shoval & Isaacson, 2007b). The success and continuity of this phase is mainly due to the widespread distribution of smartphones with its various inbuilt sensors that enable to do research more easily (Birenboim & Shoval, 2015). Therefore, the descriptive phase is still ongoing but is increasingly being taken over by studies that analyse more complex and fundamental questions on tourism and consumer behaviour (Shoval & Ahas, 2016).

To name some examples, GPS and mobile positioning data have been used to understand tourists' time–space resource allocation decisions (Grinberger &

Shoval, 2018), and to comprehend how different factors such as weather (Caldeira & Kastenholz, 2018; McKercher, Shoval, Park, & Kahani, 2015), prior destination experience (Caldeira & Kastenholz, 2017), hotel location (Shoval, McKercher, Ng, & Birenboim, 2011), or giving out incentives (Shoval, Kahani, De Cantis, & Ferrante, 2020) affect tourist movement behaviour. The insights gained from using tracking technologies for tourism movement analysis have been suggested to use as an input to improve marketing decisions (Kuusik et al., 2011; Mckercher et al., 2019). Furthermore, tracking data have been combined with real-time tourist emotions and experiences (Birenboim, 2016; Birenboim, Reinau, Shoval, & Harder, 2015; Shoval, Schvimer, & Tamir, 2018a, 2018b; Zakrisson & Zillinger, 2012).

However, as the majority of tourism research using tracking technologies have used GPS data, the sample sizes are relatively small, the research area is usually confined to a limited area, and the timescale is fixed to a certain period. For example, GPS studies made in theme park (Birenboim et al., 2013), island (Y. Li, Yang, Shen, & Wu, 2019) or a city (Grinberger & Shoval, 2018; Mckercher et al., 2019; Shoval, 2008) lasting from one day up to two weeks. On the other hand, cellular network based mobile positioning data can be used to cover vast areas and long periods of time. The essence of mobile positioning data and its use in tourism research is described in more detail in the next chapter.

2.6 Using MPD in tourism studies and tourism statistics

The first studies to use cellular network-based data from mobile phones can be dated back to the 2000s. They mostly show the potential of the data for use in studies on travel behaviour (Asakura & Hato, 2001, 2004), the planning and administration of public life (Ahas & Mark, 2005) and urban analysis (Ratti, Pulselli, Williams, & Frenchman, 2006). However, since 2007, location data from mobile phones have been actively used in tourism studies as well (Ahas, Aasa, Mark, et al., 2007; Ahas et al., 2008; Ahas, Aasa, Silm, & Tiru, 2007; McKercher & Lau, 2009), but the majority of studies using mobile positioning data are still focused on methodological questions and the use of data in human mobility, transportation and urban analysis research (Wang, He, & Leung, 2018). The main reasons behind it are difficulties in data access and the lack of analytical knowledge and computation power among tourism researchers to cope with the vast amounts of data that is produced in mobile networks.

MPD can be obtained via ‘active’ or ‘passive’ means. Active positioning implies the tracking of mobile phones using special queries with permission from the phone user; this is seldom used due to privacy concerns and relatively small samples (Ahas et al., 2008; McKercher & Lau, 2009). To note, sometimes also tracking via an installed application on smartphones (e.g., applications that use GPS data) can be referred to as active positioning. The majority of mobile phone tracking studies which have used cellular network data in tourism research have used the alternative, passive MPD (Shoval & Ahas, 2016). Passive MPD consist

of the location coordinates of mobile phones in a cellular network, which are collected and stored automatically by the mobile network operator (MNO) for the purposes of customer billing, network maintenance, and performance monitoring (Ahas, Aasa, Mark, et al., 2007; Ahas et al., 2008).

The most common passive MPD set is the call detail record (CDR), which consists of data entries of active phone use, such as incoming and outgoing calls and sent messages (SMS, MMS). Less used are data detail records (DDR), internet protocol data records (IPDR), and probed data from signalling information such as location update or cell handover. The temporal preciseness of the data improves with every mentioned data type, specifically the more data records there are, the more detailed the data are. The spatial preciseness of the data is mainly dependent on the distribution of mobile network cells, which in turn are determined by the population density and pattern (Ahas, Aasa, Mark, et al., 2007). Hence, the data are geographically more accurate in densely populated urban areas and near major roads but less accurate in rural areas. Moreover, the more frequent in time are the data records, the more continuous is the track in space and the smaller is the randomness in the data.

Because passive MPD are collected automatically from everyone who is using a mobile phone, it is possible to detect various mobility types, from everyday commuters to people making a holiday trip once a year. Thus, MPD are a valuable source to analyse the entire extent of human mobility in three dimensions time, space, and number of trips, a model proposed by Hall (2005b) to describe different forms of temporary mobility, including tourism. Passive MPD can be used to measure the volume (arrivals and departures) of tourists and to analyse tourist- and trip-related characteristics (country of origin, time, length and spatial extent of visits) in all three forms of tourism: domestic, inbound, and outbound. In turn, the good coverage and granularity of MPD would enable to start a deeper theoretical discussion on how to delineate tourism from other forms of mobility, but it is not the focus of this dissertation. The methodological steps on how to identify tourist visits from MPD are described in Chapter 3.1 and more thoroughly in **Article II**.

After testing the feasibility of passive MPD and its analysis methods in tourism research in the early 2000s, several studies have been made. CDR data gained from MPD have been used, for example, to analyse domestic tourists' seasonal moves to second homes in Estonia (Silm & Ahas, 2010), and to identify and construct spatial patterns of domestic tourist flows between 32 cities in France (Vanhoof et al., 2017). Inbound and outbound tourism trips can be identified thanks to roaming contracts between MNOs, which allow the use of mobile phones in countries other than those in which they are registered (Ahas, Aasa, Mark, et al., 2007; Ahas et al., 2008). In the case of inbound tourism, CDRs have been used for example, to analyse seasonality in foreign tourists' space consumption (Ahas, Aasa, Mark, et al., 2007), segment repeat visitors (Kuusik et al., 2011), measure destination loyalty (Tiru et al., 2010), evaluate travel distances of events visitors and regular visitors (Nilbe, Ahas, & Silm, 2014), identify tourist destinations based on visitor flows (**Article I**), find generalised movement patterns

of tourists (Baggio & Scaglione, 2018), identify destination functions (**Article III**). There have been fewer studies of outbound tourism, given the fact that the spatial accuracy of MPD for outbound trips is often not disaggregated further than the country visited. However, outbound trips have been used to compare ethnic (Silm & Ahas, 2014) and generational differences in those who travel abroad (Masso, Silm, & Ahas, 2019), and to segment travellers into different visitation groups such as tourists, cross-border workers, foreign workers, and transnationals (Ahas, Silm, & Tiru, 2017).

The uniqueness of mobile positioning lies in the fact that it is the only new digital data source used in tourism research that has vast potential of being used in official national tourism statistics and has been already used for that reason. So far, however, there are only two countries in the world using MPD to produce national tourism statistics: Estonia (Eesti Pank, 2020) and Indonesia (Lestari, Esko, Sarpono, Saluveer, & Rufiadi, 2018). Regardless of the potential of using MPD in tourism statistics, only a few wide-ranging examples are available, although the number is growing. In Estonia, the database of CDRs covers the entire country and one full economic cycle (inbound data have been collected since 2004 and outbound data since 2008), and tourism statistics have been published since 2008 (Ahas et al., 2008). In Indonesia, mobile signalling data combined with cross-border surveys have been used since 2016 (Lestari et al., 2018). In addition, there have been some small-scale pilot studies, for example in the Netherlands (Heerschap, Ortega, Priem, & Offermans, 2014), France (Gitton, 2016), Italy (Dattilo & Sabato, 2017), Finland (Nurmi, 2018), and other countries, the results of which have mainly been presented at international conferences.

The potential of using MPD in tourism statistics has been noted at the European and international level as well. In 2012 Eurostat initiated a study of the use of MPD in tourism statistics, resulting in a comprehensive report 'Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics' (Eurostat, 2014a). Furthermore, the United Nations has a Global Working Group on Big Data for Official Statistics, which has been working on the creation of a 'Handbook on the use of Mobile Phone data for Official Statistics' (United Nations, 2019). Both documents outline the major weaknesses and advantages of MPD. The former includes the complexity of access, the lack of qualitative information (e.g., trip purpose, transport mode, preferences etc.), difficulties in quality assessment, over and under coverage problems and methodological issues of statistical production from MPD. The latter being the timeliness, automatic production, cost-effectiveness, good consistency and resolution of data in time and space.

3. METHODOLOGY

This thesis is based on four original research articles that implement MPD in tourism destination research. The method for extracting cross-border tourism statistics from roaming data for an entire country, Estonia is described in **Article II** and summarised in the next chapter (3.1) of this thesis. An important part of the second paper is the evaluation of the representativeness of results derived from MPD compared with accommodation statistics. The methodology behind the comparison is discussed in Chapter 3.2 and the results are presented in Chapter 4.1. The main focus of the first article (**Article I**) is on the dimensions of tourism destination which can be measured using tracking data, in the current thesis, MPD. The findings from the first article are complemented with the spatiotemporal movement pattern analysis of visitors from Finland in **Article IV**. The methodology for analysing the dimensions of tourism destination is described in Chapter 3.4 and the results are presented in Chapters 4.2, 4.3, and 4.4. Destination functions derived from movement data, such as gateways, are identified and analysed in the third paper (**Article III**). The methodology behind identifying destination functions is described in Chapter 3.4, and the results are in Chapter 4.5. An overview of the articles is given in Table 1.

Table 1. An overview of study area, period, data and methods being used in Articles I, II, III, and IV.

	Article I	**Article II**	**Article III**	**Article IV**
Study area	Estonia	Estonia	Estonia, Israel	Estonia, Finland
Study period	2011–2013	2015–2017	2015–2017[a]	2014–2016
Forms of tourism included	inbound	inbound, outbound	inbound	inbound, outbound
Data	MPD	MPD, accommodation data	MPD, GPS	MPD
MPD type	CDR	CDR, DDR	CDR, DDR	CDR, DDR
No of MNOs	1	2	2	2
Analysis methods	Binary logistic regression	Linear regression	Descriptive statistics, GIS analysis	Descriptive statistics, GIS analysis
Main focus of the article for destination analysis	Geographical, temporal, compositional dimensions of destination	Feasibility of the data for tourism destination analysis	Identifying destination functions	Geographical and temporal dimensions of destination by composition

[a] MPD were collected in 2016 and GPS data in 2015–2017.

3.1 Framework for generating tourism statistics from MPD for destination analysis

Mobile Network Operators in Estonia collect three types of data: a) domestic operator customer data in Estonia (domestic), b) foreign roaming service users' data in Estonia (inbound), c) outbound roaming data for operators of clients travelling abroad (outbound). It is in accordance with the three forms of tourism. The main focus of this thesis is on inbound data. The data used in this thesis is collected from one or two largest MNOs in Estonia, with a combined market share of 73% in 2017 (Tehnilise Järelevalve Amet, 2017). Both CDR and DDR data are used that include information about outgoing and incoming call activities (phone calls and SMS messages) and using the internet and data services. The use of mobile phones is widespread in Estonia, according to the Eurobarometer survey carried out in 2017, a total of 96% of households in Estonia have at least one mobile phone (European Commission, 2018). The methodological steps how to generate meaningful tourism statistics for destination analysis from MDP are described as follows (Figure 1). Full methodological details can found in **Article II**.

Raw data description. Initial data processing is done in the MNO's systems. Specifically, MNOs provide a geographical reference table for the call events using the geographical location coordinates of antennae. MNOs extract the mobile country code for outbound and inbound data from the first three digits of the international mobile subscriber identity (IMSI) and then plot it to ISO 3166-1 alpha-2 letter country codes. MNOs generate pseudonymous identification codes for each phone, format the data, prepare it for transmission and transfer the data to processing machines by secure data exchange tunnels with file encryption.

After the initial processing, each record in the CDR database includes the following information: a) the randomly generated identification number (ID) of the phone used, b) the time of the call activity, c) the mobile country code for roaming data, and d) the antenna identification (cell ID) for inbound and domestic data (Table 2). The ID is pseudonymous and constant for each individual phone user for the whole period represented in the database (if the contract of the phone user did not change) and across the different databases (domestic and outbound). The ID ensures anonymity of the data for researchers and cannot be associated with a specific phone number or individual. The mobile country code of the SIM card is seen as the country of origin of the tourist.

Table 2. Examples of roaming data.

ID	Time	Country code	Cell ID
244217726502772	31.07.2017 11:41:27	LV	KCPBI14
240156976964286	31.07.2017 17:37:33	FI	KCPBI14
250072138923688	31.07.2017 18:02:03	RU	LDUCA14
262127856866986	31.07.2017 21:29:53	FI	TELBI13
350193355690996	01.08.2017 00:11:56	DE	LDUCA14

Data cleaning. In order the get correct visitation information from MPD it is important to clean the data from errors and carry out initial data quality control. This process comprises: a) removing duplicates (e.g., same information collected several times); b) checking the quality of the data and antenna coordinates (e.g., antennae coordinates should be within expected country boundaries); c) detecting missing data and re-requesting it from the MNOs (e.g., missing data for some area, period, antennae, or a particular type is missing); d) removing technical subscribers and machines distinguished by a corresponding attribute; and e) removing data of 'travel SIM cards' (e.g., Estonian SIM cards sold in foreign countries and only used there).

Visit detection. In this thesis the term 'visit' is used when there is location information available on the use of foreign mobile phones in Estonia. Although, in the international recommendations for tourism statistics visit is a sub-unit of trip. Specifically, 'trip refers to the travel by a person from the time of departure from his usual residence until he/she returns' and it 'is made up of visits to different places' (UNWTO, 2010, p.17). Because there are no data available from the user on other countries visited besides Estonia, which would constitute a round trip from home to home, term visit is used instead. Visits are identified as temporally close location events (call activities) (Tiru et al., 2010). For inbound visits, a new visit is registered when the maximum time period between call activities exceeds 155 hours (6.5 days). E.g., when there is a data gap for >6.5 days, then most probably the person has left the country and there is a new visit. Identified visits may include cases where the border is not actually crossed, such as sailors using Estonian roaming services who do not physically land or occasional roaming events near borders. Those visits are excluded from further analysis.

Identifying visitor groups. Visitor groups are defined based on the duration and number of visits made. Four types of visitor can be distinguished: a) transit visitors; b) migrant workers (including students); c) cross-border commuters (e.g., ship and aircrews, truck drivers, etc.); and d) tourists (defined as all other visitors). Tourism visits are found by eliminating the first three visitor groups where the number and duration of visits is not typical to traditional tourism. Specifically, a transit visit lasts less than four hours and is made in pre-defined transit corridors

in Estonia with the of aim reaching another country. Migrant workers and cross-border commuters spend more than 183 days of the year in a foreign country. The former, make less than two visits a month in at least seven months of a twelve-month period. The latter make more than two visits a month to neighbouring countries, such as Finland, Russia, Latvia, Lithuania, and Sweden, in at least seven months of a twelve-month period. All other visits are tourism visits, which can be divided based on duration, either as same-day or multi-day visits. The process of tourist visit detection from MPD by removing non tourism-related data is shown in Figure 1.

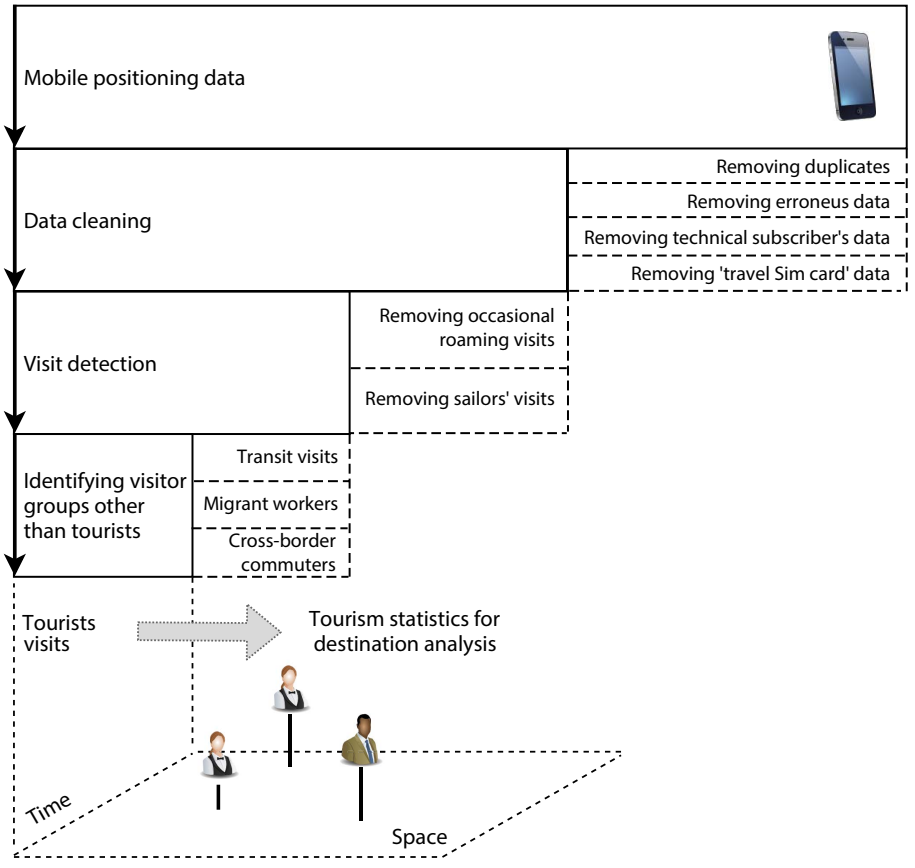

Figure 1. The process of tourist visit detection from MPD by removing non-tourism related data and extracting statistics for tourism destination analysis. Source: **Article II**, Figure 1 (modified).

Generating statistics for destination analysis. The attributes found so far are the number of visitors and visits. In addition, each visit has an undefined number of geographical location points, based on what it is possible to identify visited places. Furthermore, each visit has a start and an end time, based on what it is possible to calculate the duration of visits. For each inbound visitor the country of origin is known and the number of previous visits during the data collection

period can be calculated. Aforementioned variables can be used to sum up the number of visits and visitors, for example, by the spatial extent of the visit, time of the visit or by country of origin. Thus, tourism destination can be described and analysed using the characteristics of tourism flows as follows: geographical (visited places such as municipalities, counties, cities), temporal (time of the visit, duration of the visit), compositional (tourist's country of origin).

3.2 Evaluating the representativeness of MPD

In **Article II** the monthly indicators from MPD are compared with accommodation data by country of origin. For inbound tourism, data from Statistics Estonia (2019) is used covering the monthly number of nights spent at accommodation establishments by country of residence. This includes data from collective tourist accommodation establishments, such as hotels, hostels, motels, guesthouses, holiday camps, flats, cottages, bed-and-breakfasts. The advantage of mobile data is that it allows us to identify same-day visits. However, same-day visits are excluded from the representativeness analysis because there are no other data sets available to compare the numbers gained from MPD.

For comparison, the Pearson correlation coefficient was applied between the accommodation data and the number of nights derived from MPD. In addition, a simple linear regression model was applied where the number of nights spent by month in the accommodation data was used as the dependent variable and monthly numbers from MPD were used as the independent variable. One general model was calculated for all the nights spent during the study period for all the countries, and additional six models for the countries with the most visitors to Estonia were done. This included Finland, Latvia, Russia, Germany, Sweden, Norway. Because of the seasonal character of the inbound tourism to Estonia, where most of the nights are spent in high season (summer months), all models were seasonally adjusted using the mean monthly residual.

3.3 Analysing destination dimensions

MPD are a valuable source to study tourism destination, because the data cover the entire country over long time periods. Thus, destination does not have to be a pre-defined unit, but it can be distinguished based on the results from the spatio-temporal tourist flow analysis. Therefore, **geographical** (where do the visitors actually go), **temporal** (when do they go and for how long do they stay) and **compositional** (who are visiting) dimensions can be analysed.

In the first article (**Article I**) the respective variables included in the analysis are the number of different counties visited (n=15), visitation time and duration, and country of origin. The visitation time is determined by the time of the first call activity of the visit. It is divided into four categories according to the season: spring (March to May), summer (June to August), autumn (September to

November), and winter (December to February) visits (**Articles I, III**). The duration of the visit is calculated as the number of days between first and last call activity. Same day trips are lasting for one day (**Articles I, III**). Country of origin is derived from the mobile country code used in roaming data.

Binary logistic regression is carried out to predict the probability of visiting two specific destination areas in Estonia, Saaremaa and Tartumaa, instead of visiting other places in Estonia (**Article I**). At first separate models for each aforementioned variable are calculated and then full models including all the variables are done. In addition, the probability to visit other countries in Estonia is also evaluated. Hereinafter the names Saaremaa and Tartumaa are used interchangeably with Saare County and Tartu County, respectively.

The analysis in the first article is complemented with the analysis done in the fourth article (**Article IV**), where the geographical and temporal dimensions are assessed from the point of view of one country of origin only – Finland. The regional distribution of tourist visits from Finland to Estonia is analysed on the municipality level. In addition, the duration of visits, the distribution of the start and end times of the visits between days of the week, and the seasonal distribution of visits made from Finland to Estonia are evaluated.

3.4 Identifying destination functions

Tourist movement analysis can be used to identify destination functions. In this thesis the focus is on one function only – gateway – the main entry and exit point to and from a destination. In the third article (**Article III**) gateways are identified for the whole destination country Estonia using the location of 'first call activity' of the visit. Gateways are distinguished on the municipality level by the overall share of first call activities made in each municipality (n=213) in whole Estonia. The locations of first call activity correlate with the locations of last call activity, that strengthen the validity of the method being used for identifying gateways.

3.5 The ethics of data usage

Using MPD in tourism research and producing statistics entails several concerns such as privacy, ethics and surveillance fear among mobile phone users. This dissertation is conducted in compliance with European Union and Estonian legislation. The collection, storage, processing and analytical procedures for the data comply with European Parliament and Council of the European Union requirements regarding the protection of natural persons with regard to the processing of personal data and on the free movement of such data (REGULATION (EU) 2016/679) (European Parliament and Council of the European Union, 2016), and the processing of personal data and the protection of privacy in the electronic communications sector (DIRECTIVE 2002/58/EC) (European Parliament and Council of the European Union, 2002).

4. RESULTS

4.1 MPD as a valuable data set for tourism destination analysis

The results from the second article (**Article II**) confirm that MPD are a valuable data source for producing inbound tourism statistics. In general, the monthly number of nights spent in Estonia derived from MPD is strongly correlated with the number of nights spent according to official accommodation statistics (Pearson R is 0.96; $P < 0.0001$; $n = 36$ months). Both data sets show the seasonal rhythm in the number of nights spent by incoming tourists (**Article II**, Figure 3). According to the MPD the number of nights spent is higher in July and December, and lower in other months. The results from the linear regression model show that the monthly numbers generated from MPD explain 92% of the variation in the number of nights spent according to accommodation data (**Article II**, Table 2). Nevertheless, there are differences among countries. For example, in some countries (e.g., Russia, Germany) MPD do not capture all visits compared to the numbers from accommodation data, but in some cases (e.g., Norway, Latvia) MPD show much higher values than accommodation data (Figure 2).

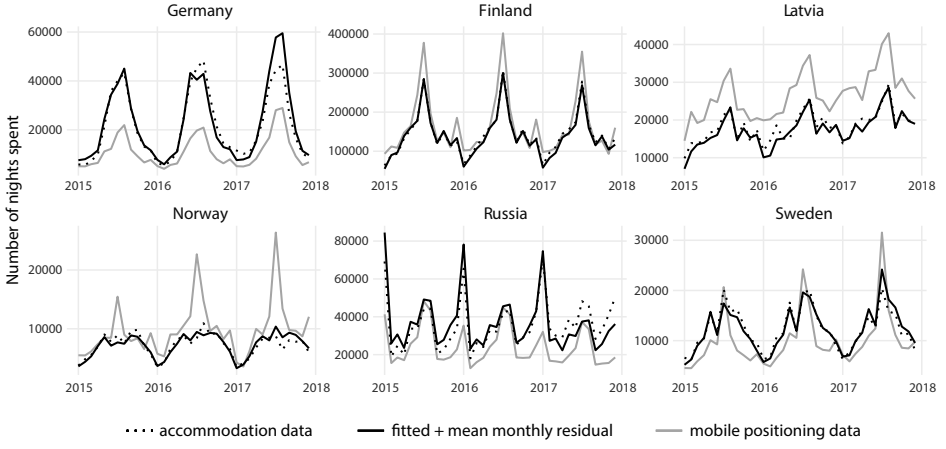

Figure 2. The monthly number of nights spent according to accommodation and MPD-based statistics in the six countries together with the fitted linear model. Source: **Article II**, Figure 4 (modified).

Previous comparison focused on overnight visits, but in addition MPD enable the detection of same-day visits also known as excursionists. For example, in the data set used in the second article 54.9% of all inbound tourist visits were actually same-day visits (**Article II**). Unfortunately, no thorough analysis to evaluate the representativeness of same-day visits derived from MPD have not been done in this thesis, because there are no comparable base line data available. Same-day visits are not included in accommodation statistics and are often underestimated by surveys (Eurostat, 2014a, 2014b). However, the inclusive character of MPD enable to detect various forms of mobility, e.g., from transit visitors passing through to regular second home visitors. The distribution of different visitor types

based on frequency (number of visits per year) and duration (number of days per year) are shown in **Article II**, Figure 2 and **Article IV**, Figure 2.

Thus, the use of MPD enable to broaden the horizons of classical conceptualisation of tourism that is especially important for tourism destination analysis. Because it is possible to analyse the movement of individual visitors throughout their entire visit in Estonia, irrespective of whether they spend a night here or not, or where do they spend it (e.g., staying at friend's place or in a hotel). Therefore, the geographical, temporal, and compositional aspects of tourism destinations are evaluated as follows in the next sections.

4.2 Geographical dimension of destinations

Detailed information about the actual visited places is of utmost importance for tourism destination analysis. The spatiotemporal analysis of foreign tourist's visits to Estonia enable to identify the most visited places and routes. The results show that Tallinn is overwhelmingly the most visited municipality in Estonia. For example, in 2016, a total of 70% of all inbound tourism visits included a stay in Tallinn (**Article III**). Tallinn is followed by Viimsi, Pärnu, Tartu, and Narva municipalities (Figure 3). In addition, the main transportation corridors can be identified, such as the route between Tallinn and Riga which serves to explain the abundance of visits to municipalities such as Häädemeeste, Tahkuranna and Märjamaa.

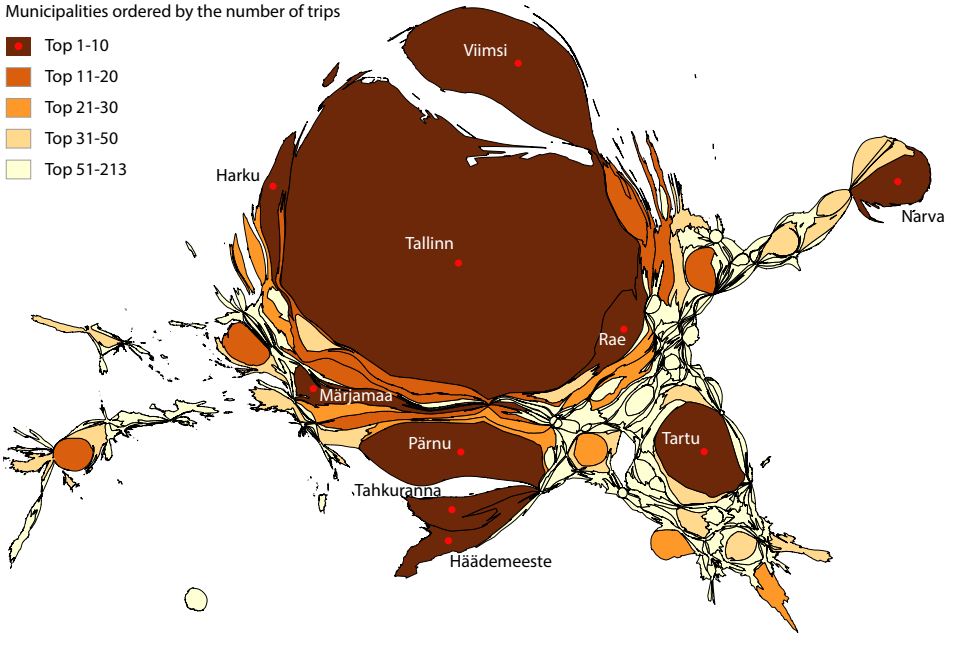

Figure 3. Most visited municipalities in Estonia in 2016. The size of each municipality is distorted in proportion to the number of visits it welcomed. Source: **Article III**, Figure 2.

The preference to visit Tallinn and Harju County alone is especially visible among Finnish tourists. During the period 2014–2016 a total of 73.5% of their visits were made solely to Tallinn and Harju County (**Article IV**, Figure 6). Even if the visits made only to Tallinn and Harju County are excluded, Harju County and, particularly, Tallinn still dominate among all other visits made to Estonia. For example, in the period from 2011 to 2013 approximately 40% of all visits (excluding visits to Tallinn and Harju County alone) still included a stay in Tallinn (**Article I**).

MPD enable to analyse visitation patterns to smaller destination areas. The results from the second article confirm that Saare and Tartu Counties function as two different destinations (**Article II**). Both counties have distinct images and different groundings for marketing. Saaremaa is mainly known for its spas, nature and recreation, while Tartumaa is famous for its university and is therefore more oriented to cultural and conference tourism. These differences are also reflected by the visitation patterns to these destinations (Figure 4). The main visited counties when visiting Saaremaa are Harjumaa (39.5%), Läänemaa (24.3%) and Pärnumaa (16.5%), while for Tartu County the other most visited counties are Harjumaa (30.5%), Jõgevamaa (11.5%) and Valgamaa (11.4%) (**Article I**).

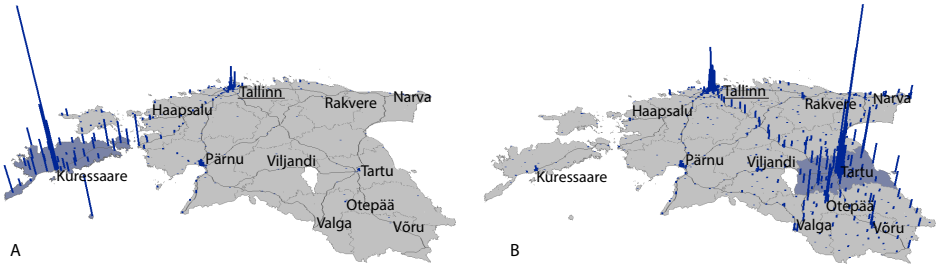

Figure 4. The distribution of call events in antennae locations when visiting (A) Saare County (blue), (B) Tartu County (blue) 2011–2013. Source: **Article I**, Figure 2 (modified).

4.3 Temporal dimension of destinations

The spatial movements of tourists cannot be analysed without the temporal dimension. According to time geography, a concept developed by Hägerstrand (1970), the movements of individuals should be also projected with the temporal parameters in addition to geographical coordinates. As was seen from the comparison between MPD and accommodation statistics, inbound tourism to Estonia is seasonal, with high peaks in the summer months (**Article II**). More than one third of visits to the country are made in summer (**Articles I, III**). In some rare cases there are also winter peaks (e.g., Russian tourists spending their New Year in Estonia). However, when analysing seasonality in smaller destination areas in Estonia, differences are significant. For example, tourism in Saaremaa is more seasonal than tourism in Tartumaa (Figure 5). Therefore, the odds of

visiting Saare County instead of other places in Estonia are higher in summer months, but for visiting Tartu County, the odds are higher in the off-season (**Article I,** Table 3). Nevertheless, the majority of visits to Tartumaa are still made in summer months, but the proportion is smaller compared to other case study areas.

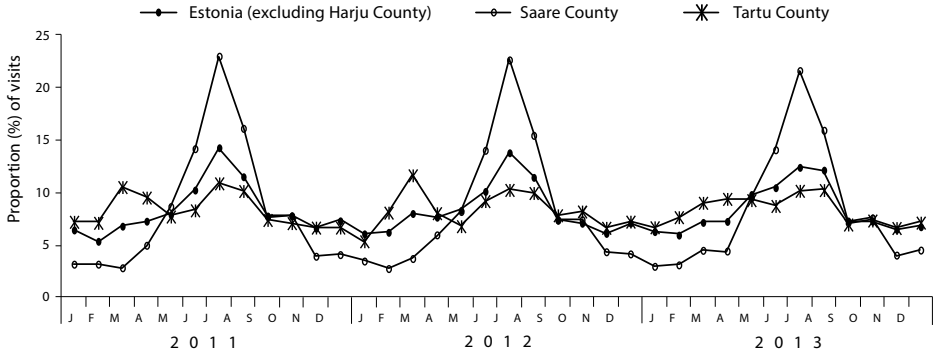

Figure 5. The monthly proportion of visits made to three case study areas in 2011–2013. Source: **Article I**, Figure 3 (modified).

In addition to seasonal pattern, differences in the number of visits received on different week days can be also distinguished. Because Estonia is a popular week-end holiday destination for neighbouring countries such as Finland, Latvia and Russia, the majority of visits to Estonia are made for weekends. For example, the majority of visits made from Finland to Estonia start on Fridays or Saturdays (Figure 6) and end on Saturdays or Sundays (**Article IV**).

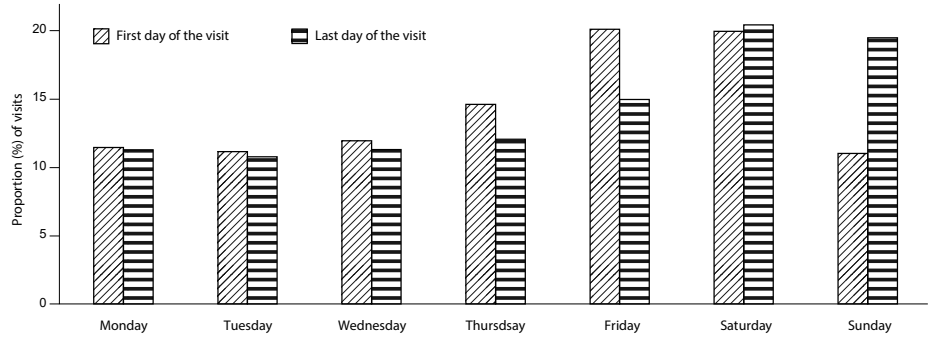

Figure 6. The distribution of start and end times for the visits from Finland to Estonia during 2014–2016. Source: **Article IV**, Figure 5 (modified).

Besides the time of the visit, MPD enable to analyse the duration of visits. In 2016 the mean duration of an inbound visit to Estonia was 2.16 days (**Article III**). Again this number varies between destinations. Shortest visits are made to Tallinn only, 1.48 days in 2016 (**Article III**), and the farther the destination from main entry point Tallinn the longer are the visits. For example, the mean duration of visits made to Saaremaa in 2011–2013 was 3.4 days (**Article I**). Results from

binary logistic regression also confirm that visits longer than three days are twice as likely to take place in Saare County than in Estonia as a whole (**Article I**, Table 3).

4.4 Compositional dimension of destinations

In this thesis the compositional dimension of the destination is based on the country of origin of the tourists who visit the destinations. Visitor's country of origin is derived from the mobile country code of the SIM card. The majority (ca 70%) of total inbound visitors to Estonia are from three neighbouring countries Finland, Latvia and Russia (**Articles I, II, III**). However again, there are differences among smaller destination areas inside Estonia. For example, Saare County is mostly visited by people from Finland, but the majority of visits to Tartu Country are made by visitors from Latvia (**Article I**). This fact is also confirmed with the results from binary logistic regression analysis, where the odds of visiting Tartu County instead of the whole of Estonia are significantly higher among Latvian (1.53) than Finnish visitors (**Article I**, Table 3). The share of visitors from Russia is smaller in Saaremaa than in Tartumaa or on average in Estonia. These spatial preferences among visitors from different countries can be partially explained by the relative location of the country of origin compared to Estonia. Therefore, visitors from Finland mostly stay in Tallinn and surroundings, because the harbour in Tallinn is the main entry point for them, yet visitors from Latvia are more oriented to the southern counties of Estonia. Difference spatiotemporal behaviour among different countries also appear on the temporal scale, e.g., German tourists are more summer oriented and tourists from Russia are less sensitive to season (**Article I**, Figure 4).

4.5 Destination functions

Tourist movement analysis enable to identify different functions of visited places based on their relative location in the visitation pattern. In this thesis the main focus is on gateway as the first place encountered by tourist at a destination. The geographical level used for destination function analysis in this thesis is entire country, Estonia. Gateways are found by the location of first call activity. In 2016 57% of the first call activities were in Tallinn (**Article III**). Together with neighbouring municipalities such as Viimsi, Harku and Rae the share of northern gateway reaches up to 66% (Figure 7). The antennae in Viimsi and Harku municipalities register the first call activities made on ships before reaching the port and the antennae in Rae municipality register the first call activities made after landing in Lennart Meri international airport. Therefore, as the international airport and harbour are located in the capital, Tallinn, it also functions as the main gateway to the country. Unfortunately, 78% of those visits with a first call activity in Tallinn never leave the city (**Article III**).

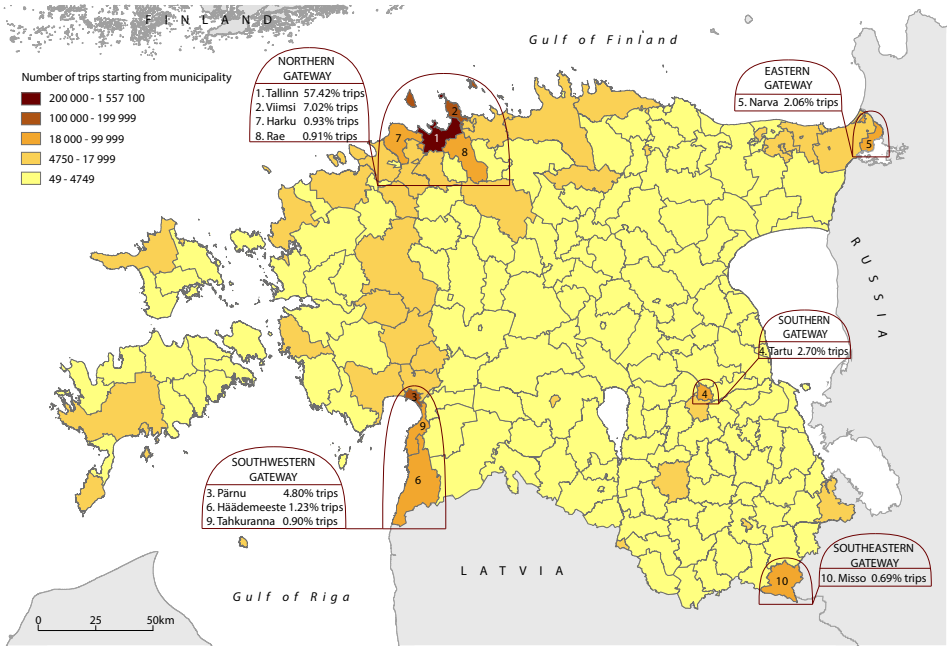

Figure 7. The location of the main entry points to Estonia identified by the share of total number of first call activities made in each municipality. Source: **Article III**, Figure 1.

Besides Tallinn and its surrounding municipalities smaller gateways also exist, but the share of first call activities in those places is remarkably lower. Second biggest gateway can be named as the southwestern gateway and it consists of the visits made to Häädemeeste, Tahkuranna and Pärnu municipalities (Figure 7). Therefore, it mostly functions as a gateway for tourists from Latvia and Lithuania travelling by car or by bus. In addition, the second biggest city in Estonia – Tartu – functions as a gateway thanks to the daily international flight connection to Helsinki. Furthermore, smaller gateways such as the city of Narva on the eastern border of Estonia and Misso municipality in southeast of Estonia that borders both Latvia and Russia exist (**Article III**). It is important to note, that when comparing the location of gateways with the most visited municipalities (Figure 3), differences are minimal. Therefore, quite often the gateway also functions as a destination itself and tourists do no travel further away from the entry point.

5. DISCUSSION

5.1 The use and applicability of MPD in tourism destination analysis

Destination has quite often been taken for granted in tourism research. It is assumed that destination is a fixed geographical area where tourism takes place and it offers services and attractions to visitors. This traditional geographical point of view is for example a widespread approach in statistics collected on tourism (Eurostat, 2014b; UNWTO, 2010), marketing (Buhalis, 2000) and a common ground in various research undertaken about tourism destinations. However, the geographical approach has been criticised by socio-cultural authors because of its static nature that do not take into account the dynamic character of destinations that changes over time and can be perceived differently for each visitor (Framke, 2002; Saarinen, 2004; Saraniemi & Kylänen, 2011). Nevertheless, neither destination as a geographic area nor as a socio-cultural concept only, are helpful in destination management and marketing. Thus, the latest discussions on the essence of tourism destinations derive mostly from the tourism planning and management literature where several authors have developed conceptual spatial models. This is nothing new, e.g., see the overview by Dredge (1999), but the article from Beritelli et al., (2014) has drawn the focus to tourism flows which constitute the destination and should be the core unit used for destination management.

This is the point to which current thesis contributes. Specifically, the model proposed by Beritelli and Laesser (2017) that distinguishes visitor flows based on five questions: Who? Where? When? What? Why? Although, the authors pointed out that big data could be useful for identifying visitor flows, to date the data and information about visitor flows in their case studies is primarily provided by local actors (e.g., entrepreneurs, guides) only (Beritelli et al., 2019). This thesis has proposed a similar theoretical framework for identifying tourism destinations, based on the visitor flows, but it is fully derived keeping in mind the essence of digital tracking data sets, specifically mobile positioning data. According to the proposed framework (**Article I**), destination has five measurable dimensions that can be analysed using data about tourism flows: geographical (where?), temporal (when?), compositional (who?), social (what?) and dynamic (how?). In this thesis, the first three (where, when and who) are analysed empirically. However, the social and dynamic dimension could easily be analysed using additional data (e.g., social media data) and longer timeframes (e.g., years).

The results from the analysis of **geographical** dimension confirm the previous knowledge, that destinations may overlap and exist on different geographical levels (Buhalis, 2000; Framke, 2002; Saarinen, 2004). For example, a relatively small country in size and population such as Estonia, has smaller distinguishable destination areas inside it. In this thesis the comparison is made between two counties in Estonia Saaremaa and Tartumaa. Both destinations have distinct visitation patterns that do not overlap, except in the gateway of Harju County,

where most of the visits to Estonia start (**Article I**). When analysing all the visits made to Estonia in one year, the situation is strongly influenced by the same-day visits to Tallinn, made mostly by visitors from Finland (**Article III**). Therefore, in order to gain a more detailed picture, studies using samples of visits to smaller areas are necessary. The maximum spatial accuracy of MPD is at the network cell level, thus it is not possible to define the geographical boundaries of a destination in terms of metres, but it is enough to distinguish between city neighbourhoods or villages. This is already detailed enough for the destination management and planning perspectives.

From the time geography perspective, a concept developed by Hägerstrand (1970), the geographical movements of tourists cannot be viewed without the **temporal** parameters. MPD enable to analyse when tourism is happening at destinations and for how long. Inbound tourism to Estonia is seasonal, with high peaks in the summer months (**Articles, I, II, III**) as shown by (Ahas, Aasa, Mark, et al., 2007). However, this seasonality varies for smaller destination areas in Estonia and between different visitor groups. For example, tourism in Tartu County is less seasonal than it is in Saare County, mainly due to the differences in tourism supply (**Article I**). Saaremaa is oriented towards holidays and recreation with more offerings in the warmer summer season, while Tartumaa and, especially, Tartu city attracts visitors all year round as a cultural and educational centre. Thus this temporal dimension clearly shows when a particular place actively functions as a destination, in other words when destination is activated by visitor flows (Beritelli et al., 2014). Second important temporal aspect is the duration of the visit. There is no doubt that tourist movement is influenced by time availability and the increase in the length of stay also raises the probability to visit more places. Analysis made in the first and third article also affirm that. Specifically, shortest visits (1.5 days) are made to Tallinn only (**Article III**) and when the distance of a destination from the main gateway increases, the average duration of visit also increases. Thus, the mean average duration of visits to Tartumaa is 2.9 days and to Saaremaa 3.4 days (**Article I**).

Third destination dimension analysed in this thesis is **compositional** – who are the people visiting. MPD enable to distinguish between domestic and foreign visitors. In this thesis the focus is on inbound visitors and their country of origin derived from the SIM card. Thus, it is possible to analyse which places are popular among visitors from different countries. Overall, the cumulative visitor numbers are strongly affected by the distance decay principle. Therefore, more visitors to Estonia come from nearby countries such as Finland, Latvia and Russia (**Articles I, III**). These findings are in line with the previous research describing the effect of distance (Nilbe et al., 2014). Compositional dimension can be analysed also from the geographical and temporal perspective. Thus, it is possible to compare movement patterns of visitors from different countries. For example, visitors from Finland tend to stay in Tallinn and its surroundings (**Article IV**), while visitors from Latvia mostly stay in the southern places of Estonia such as Pärnu, Tartu and Saaremaa (**Article I**). From temporal perspective visitors from different countries also have different seasonal patterns; visitors from long-haul countries

tend to be more summer-oriented (e.g., Germany), while visitors from neighbouring countries are relatively insensitive to season (e.g., Russia) (**Article I**). Thus, the geographical background of tourists strongly influences the tourism flow patterns inside destination country, as found in previous studies made in Malaysia (Oppermann, 1992a), Barbados (Reid & Reid, 1997), USA (Hwang, Gretzel, & Fesenmaier, 2006) and China (Bowden, 2003).

As stated before, **social** and dynamic dimensions are not empirically studied in this thesis because the current MPD sets used do not enable directly these kind of analysis. However, a different MPD set about calling networks (who is calling whom) could reveal the local social network and connected places (see e.g., Puura, Silm, & Ahas, 2017). Thus, it would be possible to compare the connected places for local residents, domestic visitors and inbound tourists. In addition to mobile positioning based movement data about connected places, additional data sources could be used to more fully understand tourist preferences and experiences. For example, content analysis using different social media data such as text and photos that have geographical and temporal attributes (Toivonen et al., 2019) can provide more information about visited places, activities undertaken, and opinions expressed. When it comes to analysing the **dynamic** dimension and understanding changes in tourism destinations over time as has been discussed by Butler (1980) and Saarinen (2004), longitudinal data is needed. MPD can be used here, as inbound roaming data in Estonia have been collected since 2004. Additional data from social media (e.g., Twitter or Flicker) could also be a potential source for more longitudinal analysis.

In addition to answering simple questions such as where tourism is happening, when it is happening and who are the tourists; the role of each visited place within the whole visit inside a destination country can be identified and analysed using MPD. In this thesis, the focus is on gateway, defined as the first place tourist encounters when visiting a destination country (Burghardt, 1971; Pearce, 2001b). For Estonia, the capital Tallinn **function**s as the main gateway to country; 57% of all the inbound visits made in 2016 had a first call activity in Tallinn. This is easily understood because the main transportation nodes (airport, harbour) are located there. Together with neighbouring municipalities the share for northern gateway Tallinn increases to 66% of all visits (**Article III**). The problem arises because 78% of the visits that start from Tallinn never leave the city. Thus, Tallinn functions both as a gateway and as a main destination. However, a place having multiple functions (Beritelli & Laesser, 2017; Lohmann & Pearce, 2010; Pearce, 2001a) is not a problem itself, but the fact that tourists are concentrated into one place resulting on the one hand in economic success but on the other in social and environmental problems which may lead to overtourism (Koens, Postma, & Papp, 2018).

The detailed knowledge gained from the analysis of MPD in regard to the geographical, temporal, and compositional dimensions of tourism destination and its functions is, above all, important for better **tourism destination management and planning**. The need to have a more dynamic and detailed conceptualisation of tourism destinations derives directly from the planning and management

interest. The most recent approach is the flow-based destination marketing and management (Beritelli et al., 2014; Reinhold, Laesser, & Beritelli, 2019), discussed in the beginning of this chapter, where the main unit of analysis are the strategic visitor flows at the destination. Thus, analysing MPD about the movements of tourists within a destination country is a valuable input information for this kind of tourism destination management and marketing approach. Based on the results gained from MPD analysis, it is possible to show the actual visitor movement patterns and distinguish destinations quantitatively, based on the number of visitors a place receives in time and space. Thus, the results from MPD illustrate which places are preferred and visited together or which places are not connected by visitor flows. This helps to overcome the traditional approach in which destinations are managed by following administrative borders (Dredge, 1999), even though these borders may not have any meaning to the tourists themselves. This has still been widely applied, because there have not been enough sufficient detailed data to act otherwise. However, there are already examples where the knowledge gained from new data sources is already in everyday use. For example, the Estonian Tourist Board, uses a monitoring tool of visitor flows in Estonia based on MPD.

Detailed knowledge about the movement of visitors in time and space is especially important in managing the effects of tourism (Lew & McKercher, 2006). This becomes relevant when some places are receiving more visitors at certain times than that place can actually sustain. It may lead to negative social, cultural, and environmental impacts (Koens et al., 2018). The situation in Estonia has not yet been seen as overtourism per se, because the overall numbers are still manageable, but the current results from identifying destination functions (**Article III**) revealed the relative substantial importance of Tallinn for the entire country. Thus, already some arrangements could be made to alleviate the pressure on the capital and distribute tourists to other places in Estonia as well. For example, day trip itineraries from Tallinn could be promoted, because a greater dispersal of tourists is first and foremost important for regional destinations (Koo et al., 2012; Oppermann, 1995). In order to contribute to the dispersal of tourists, changes have to be made also in terms of planning transportation and infrastructure and in designing marketing campaigns (Lew & McKercher, 2006). Knowledge about tourism flows gained from MPD analysis enable to optimise transportation connection, open up new ones where needed, and market connected places or routes to visitors.

According to the flow-based view by Beritelli et al., (2014), destination is activated by the visitors flows. Rapid developments in ICT have opened up several new possibilities to measure and analyse visitor flows but current studies have mostly been descriptive, and not taking one step further to also conceptualise destinations based on the results from new data sources. For example, Chua et al., (2016) have described the temporal, spatial, and demographic aspects of tourism flows in southern Italy using data from Twitter, but they do not focus on the essence of tourism destinations behind the tourism flows. On the other hand, theoretical studies which describe the essence of tourism destinations rarely

use empirical data, the models and conceptualisations that have been developed by authors are based on their impressions, observations, and intuitive thoughts, as well as considerable reading (Butler, 2015). The **novelty** of this thesis lies in the combination of those two sides. Therefore, MPD are used to conceptualise tourism destinations based on the actual movement patterns of tourists, while adding a level of empirical knowledge to the theoretical discussion.

5.2 An evaluation of using MPD in tourism statistics and destination research

In this thesis, passive MPD are used to calculate tourism statistics. The results are then used for tourism destination analysis. In order to evaluate the representativeness of MPD, the results gained from MPD are compared with the monthly number of nights spent at accommodation establishments in Estonia by country of residence from the same time period. The results confirm that MPD are strongly correlated with official accommodation data for inbound visits (**Article II**). However, there are some general differences and variations can be seen between countries. For example, MPD showed higher values in the high season, such as in July and December, and relative underestimates in autumn and spring. It may be speculated that some visitors during the high season are staying with friends and family and not in official accommodation establishments as shown in Australia (Backer, 2012). In addition, Statistics Estonia only partially collects information about people who are staying in holiday rentals that are being offered through Airbnb. Furthermore, some of the data may deliberately be left unreported for fiscal reasons as has been found to be the case in Italy (De Cantis et al., 2015; Guizzardi & Bernini, 2012).

When focusing on the differences in the top six countries, again over- and underestimation examples can be found. Latvia, Norway, and Finland represent a situation in which there are more nights being spent in Estonia according to MPD than can be seen from official statistics. It may be theorised that a substantial number of people from Estonia have moved to Finland and Norway, and when they visit Estonia, they no longer use their Estonian SIM card and are staying with friends and relatives. More detailed analysis of the mobility of transnationals between Estonia and Finland (Ahas et al., 2017) has been done in **Article IV**. For Latvia it may be speculated that a lot of them are working in Estonia; however, according to the visitor group methodology applied, they still belong to the tourist group (**Article II**), and therefore they are short-time workers who do not stay at official accommodation establishments in Estonia. Russia and Germany are the examples where MPD underestimate the number of nights spent in Estonia. This might be explained by high roaming prices and preferences in phone use, such as no need to make a phone call or to send an SMS, or a desire not to be interrupted.

The more often a phone is used the more data are produced and the more accurate are the results derived from MPD. The use of mobile phones depends on

various factors such as roaming prices (Eurostat, 2014a), personality traits (Butt & Phillips, 2008) and sociodemographic and -economic characteristics (e.g., gender, age, income) (Ahas et al., 2008; Wei & Lo, 2006). Differences in mobile phone use in turn may cause several under- and over coverage issues in overall number of nights spent and in the spatiotemporal coverage of the visits. For example, the share of same-day visits may be overestimated and the number of visited places may be underestimated. However, it is important to emphasise that the overall number of nights spent derived from MPD are in line with results gained from accommodation data as proven in **Article II**. Possible differences in the spatiotemporal extent of the visits between different data sets are also minimal as shown before (Ahas et al., 2008) and confirmed again in this thesis where 82% of inbound visitors stayed in Tallinn based on MPD and 77% according to accommodation data (**Article III**). Unfortunately, no comparison is made to evaluate the share of major gateways derived from MPD (**Article III**). This could be done using data from ports, airports and border control (only for the border with Russia, because Schengen rules apply to other border crossings), however it was not done because the data would represent all travellers and not only tourists, and such data are not available on a monthly basis.

Differences in phone use together with difficulties in access to data and lack of qualitative information are the three major **disadvantages of MPD**. Access to MPD from MNOs has so far been the biggest obstacle in the use of MPD when it comes to producing official statistics or carrying out scientific research. Problems arise from (inter)national regulatory limitations and legislation. MNOs are unwilling to provide data due to the necessity for maintaining business confidentiality, or due to privacy issues, fear of being tracked, and a general disapproval by society of such methods. Despite this, there are two countries in the world where MPD are an official source for tourism statistics, Estonia (Eesti Pank, 2020) and Indonesia (Lestari et al., 2018). In addition, there are several countries in which pilot studies for producing statistics have been conducted (United Nations, 2019), and even more scientific articles where MPD have been used for tourism studies (Shoval & Ahas, 2016). However, as access is restricted there is no unified methodology available and no cross-domain analysis, and therefore the reproducibility and overall transparency of the research is somewhat limited. Another major shortcoming of MPD is the lack of qualitative information, such as means of transport, type of accommodation, expenditure, trip purpose, tourist preferences etc. This kind of information can be collected by carrying out additional surveys among visitors.

Despite of the limitations described, MPD have many **advantages** over traditional data sets used in tourism research and statistics. One of the biggest is the ability to evaluate mobility indicators for a much larger sample, both automatically and cost-effectively. In the case of Estonia, the MNO market is quite equally shared between three operators, so that using data from one MNO represents information about one third of people living in Estonia and this can be generalised across the entire population. Thus, compared to traditional surveys, the cost and burden of data collection does not depend on the number of observations or

people in the sample. Surveys, on the other hand, entail high costs just because of that, thus the samples are relatively small, that in turn may lead to possible bias in the results (Beaman et al., 2004; Eurostat, 2014a, 2014b). Another major advantage is that MPD have good consistency and resolution of data in time and space that enable to distinguish between various mobility types based on people's longitudinal spatiotemporal behaviour (**Articles II, IV**). Good consistency and resolution over time means that MPD can be collected almost in real-time and for longer periods. Thus, the timeliness of MPD helps to grasp increasing tourist numbers and provide quick indicators in near real-time; however, the faster the process the higher are the costs for maintenance, software and hardware. Good consistency and resolution over time enable to follow tourists throughout their visits in Estonia. This is not possible with accommodation or border statistics, where tourists are registered at discrete locations that may result in the double counting effect (De Cantis et al., 2015). Other traditional methods such as surveys or trip diaries are mostly retrospective and the quality depends on the memory of the tourists. Compared to other digital big data sets the advantages of MPD are again their large sample and good contingency over space and time, while difficulties in access consist of the only major disadvantage.

However, MPD do not substitute traditional data sets, rather complement them. For more widespread use of MPD in tourism statistics three key aspects need to be improved: develop better regulation and legislation to access the data; find additional baseline data to compare the results derived from MPD, and create unified and transparent methodology for extracting indicators generated from MPD (**Article II**). In addition, several stakeholders are looking for even more detailed data in time and space on attraction level that would mean to take into use different MPD sets than used in this thesis, e.g., signalling data that are collected more frequently than CDR or DDR data. All of these developments have to be in line with privacy conditions that are set out by the GDPR and national legislation. Despite the difficulties and the need for future improvements, this thesis and other preceding studies which have used MPD-based statistics in tourism destination analysis have shown the vast potential of the data and its ability to complement conceptual studies that are carried out in regard to tourism destinations, while also providing guidance when it comes to making marketing and management related decisions.

To conclude, when compared with traditional accommodation, border, and visitor statistics, and survey data from questionnaires or trip diaries, inbound passive MPD enable to study tourism destinations more precisely and effectively because the spatial and temporal accuracy of the data are better; the tracking periods are longer; tracking allows to follow tourists throughout their visit; and digital data collection and processing are easy and timeliness. Good coverage of the data enables to do destination analysis based on the actual movements of tourists and to distinguish destinations based on that. Thus it provides an opportunity to add spatial and temporal scale to the so far magical 'place' defined as tourism destination by the World Tourism Organisation (UNWTO, 2010). However, this thesis focuses only on inbound tourism to Estonia, thus it is not

possible to say if and which other countries foreign tourists visited besides Estonia and how did the visit to Estonia situate in the whole trip. Nevertheless, MPD enable to capture the diverse, multi-faceted, highly dynamic social phenomenon (Beritelli & Laesser, 2017; Larsen, Urry, & Axhausen, 2007) that tourism is, in a manner that was nearly impossible before.

6. CONCLUSIONS

This thesis has opened up a discussion to find a new point of view for conceptualising tourism destinations in light of recent rapid developments in ICT-based space-time tracking technologies. The focus of this thesis is on inbound tourism to Estonia, based on the roaming activities of foreign visitors in Estonian mobile networks. This thesis contributes to academic tourism research and private and public sector tourism-related activities in three major areas which are described below.

Theoretical and conceptual contribution. In the age of digital tracking data, destinations in this thesis are conceptualised as the activity spaces of visitors that can be analysed using data about visitor movements. A theoretical framework is proposed in which destinations consist of five measurable dimensions: geographical, temporal, compositional, social, and dynamic, out of which the first three are analysed empirically in this thesis. The results from a tourist flow analysis based on mobile positioning data reveal that destinations exist on different geographical levels, that they may overlap, and that smaller destination areas such as Saaremaa and Tartumaa can be distinguished inside Estonia. Therefore, it is possible to identify places that are visited together and which function as a destination for visitors. Adding the temporal dimension enables to see and analyse changes in visitation numbers over time. Thus, it becomes possible to analyse, for example, the seasonal or weekly rhythms of destinations and the duration of visits to those destinations. The compositional dimension adds a human touch to the concept by analysing who are those people who are actually visiting the places. From mobile positioning data it is possible to use the country of origin of the visitor. Including more relevant data (e.g., social media data, longitudinal MPD) would also enable to analyse the social and dynamic dimensions of tourism destination. This kind of data driven approach which focuses on tourism flows together with a clear theoretical grounding is necessary to overcome the divide between the theory-oriented and data-oriented works that currently dominate in tourism research.

Methodological contribution. In this thesis passive mobile positioning data are used to analyse the spatiotemporal behaviour of foreign visitors in Estonia. Data are collected automatically in the mobile network operator's systems for billing purposes, and consist of the time and location of the call activity and the country of origin of the person. There are only two countries in the world that use mobile positioning data for producing official tourism statistics – Estonia and Indonesia. In addition, pilot studies describing the feasibility of the method have been published in scientific journals or presented at international conferences. Nevertheless, no unified methodology exists and there is a lack of transparency regarding the data access and the methodological steps done in order to gain meaningful indicators from mobile data. Therefore, this thesis describes a method for extracting

inbound tourism statistics from roaming data for an entire country, Estonia. The representativeness of tourism statistics derived from mobile data are compared with accommodation statistics. The results show a good correlation between the monthly number of nights spent in Estonia according to two data sets. Thus, statistics derived from MPD can therefore be are a valuable source for detailed destination analysis, because of the representativeness and good consistency in space and time. In addition, MPD also enable to capture those visitors who are not registered in the official accommodation statistics, such as, people who stay with their friends and family, or who use Airbnb, or who are visiting for one day only.

Practical contribution to tourism planning. A detailed spatiotemporal analysis of tourist movements in Estonia, enables to add a geographical and temporal scale to what so far has been a rather ambiguous term – destination – which has too often been taken for granted. This is valuable input for tourism destination researchers, planners, and management organisations whose decisions have suffered from a lack of accurate data. MPD enable to identify connected places at different times and to amend marketing and management strategies accordingly. In addition, based on tourist movements, it is possible to identify different destination functions. This thesis focused on main gateways and showed the major role that Tallinn as the main entry and exit point to Estonia has over the entire country. Therefore, in order to distribute the benefits gained from tourist visits more evenly over the entire country, additional measures are needed. For example, improving the capacity of smaller gateways such as regional airports or creating convenient day trip itineraries from Tallinn. Up to date, reliable and granular data is of utmost importance when it comes to dealing with the social, economic, and environmental impacts of tourism flows to ensure the sustainability of destinations. The Estonian Tourism Board has been using a web-based destination monitoring tool based on passive MPD and the methodology presented in this thesis.

Future perspectives. While writing this chapter already three months have passed since the last MPD set was forwarded to scientists in Estonia (at the end of 2019). Thus with every day that passes the data gap is increasing and, at some point, it will not be possible to retrospectively gain the data from the operators because the data are regularly deleted. However, at the same time new horizons are opening up. Specifically, because of the rapid spread of coronavirus across the world, which means that governments are looking for answers in terms of how to act in this new situation. The Estonian Government Emergency Committee has focused on this by ordering Statistics Estonia to prepare an analytical output that uses anonymous data from electronic communications, with the purpose of preventing the spread of the coronavirus (COVID-19). Let's hope that something good may come out of this really sad and difficult situation, and that the use of anonymised mobile positioning data at a general level in regard to the movements of people, including tourists, will again be possible one day. Smart decisions based on the knowledge gained from data analysis are needed every day, and not only during a crisis.

ACKNOWLEDGEMENTS

I am truly grateful to all the people with whom I have discussed and shared ideas on how new digital big data sets could be used more often in tourism research, on how to carry out statistical analysis when everything gives significant results, on how to continue to look for relevant literature when hope has faded, on how to prioritise and find time for 'PhD work', and how to carry on working, even when it seems that everything new and cool has already been done by someone else.

Out of all those people, my three supervisors have guided me the most. I am first and foremost grateful to Professor Rein Ahas, who saw the potential in me and opened the door for me into the interesting academic world in which I have been some years now. Your good sense of humour and cheerfulness helped me through the tough times. I'll raise a glass to you from wherever you are watching! Rein's friendly and open character is also a big reason why I got an amazing experience to be co-supervised by Professor Noam Shoval from the Hebrew University of Jerusalem. Thank you, Noam, for the wise suggestions and feedback throughout my studies. Anto, thank you for your unruffled attitude and for being there when I needed to cross the finish line.

I am very happy that I was not alone during my PhD journey, and that I had an amazing bunch of great academic people beside me. Mobility Lab: Age, Siiri, Veronika, Anniki, Laura, Kristi, Pilleriine, Kristiina, Daiga, Arvi, Ingrid, Indrek, Mirjam. Positium: Erki, Margus, Kaisa, Egle. Mobility Lab's international friends: Menno-Jan, Frank, Jukka, Tuuli, Hendrikki, Vuokko, Olle, Kerli. And all other colleagues from the University of Tartu who have helped me or inspired me during my studies. Thanks to all of you, the journey in pursuing a PhD degree has been a truly amazing one.

During my PhD studies I have received generous support from the European Commission European Regional Development Fund in the form of smart specialisation scholarship for PhD students, a Dora Plus short study visits scholarship, and the Kristjan Jaak scholarship for study periods abroad to attend conferences, PhD workshops, and seminars. Thanks to the Erasmus+ Mobility agreement between the University of Tartu and the Hebrew University of Jerusalem I have had the opportunity to visit my supervisor Professor Shoval several times in Jerusalem.

This dissertation was supported by the Israel Science Foundation (Grant No. 1240/15), Estonian Research Council Grant PRG306, Grant IUT2-17 and the Estonian Science Infrastructure Road Map project 'Infotechnological Mobility Observatory (IMO)'; ESPON programme project BIG DATA & EGC (EE/SO2/055/2018); Academy of Finland Research 'Urbanization, Mobilities and Immigration (URMI)'; ERDF through STACC; and EU Regional Development Foundation projects TERIKVANT (3.2.0802.11e0043) and 'Real-time Location-based Big Data Algorithms' (3.2.1201.13e0009).

Last but not least I want to thank my family and friends who believe in me and have always been there for me. Thank you!

Janika Raun
Home office, Tartu, 2020

SUMMARY IN ESTONIAN

Mobiilpositsioneerimise andmete kasutamine turismiuuringutes ja -statistikas

Teoreetiline sissejuhatus Sihtkoht on keskne mõiste turismis ja põhiline üksus sihtkoha arenduses ja turunduses. Traditsioonilised käsitlused vaatlevad sihtkohta geograafiliselt ühtse ja piiritletud alana, mida turistid külastavad ja mida turundatakse tervikuna (Hall, 2005a; UNWTO, 2010). Sotsiaalkultuurilised käsitlused iseloomustavad aga sihtkohta kui sotsiaalselt konstrueeritud ruumi (Framke, 2002; Saarinen, 2004), mis on pidevas muutumises ja mille geograafilisi piire on keeruline määrata ning mis eksisteerib pigem turistide meeltes (Saraniemi & Kylänen, 2011). Mitmed autorid, kes on püüdnud sihtkoha olemust käsitleda, on tunnistanud vajadust empiiriliste uuringute järele (Framke, 2002; Pearce, 2014; Pechlaner *et al*., 2012; Saarinen, 2004). Üheks võimaluseks sihtkohti uurida on teha seda turistide tegeliku liikumise, külastatud paikade ja läbitud teede põhiselt (Beritelli *et al*., 2014; Pechlaner *et al*., 2012). Paraku ei ole aga selliseid uuringuid seni väga palju tehtud, kuna on puudunud ajaliselt ja ruumiliselt piisavalt täpsed külastusandmed (Batista e Silva *et al*., 2018).

Traditsioonilised andmeallikad – majutusstatistika, piiriületusandmed, küsitlused elanike ja turistide seas, reisipäevikud – ei taga detailseks analüüsiks vajalikku ruumilist ja ajalist täpsust ning neil on ka mitmeid teisi puudujääke. Näiteks ei kajastu majutusandmetes ühepäevakülastajad ja mitteametlikes majutusasutustes (nt sõbra, sugulase või Airbnb üüripinnal) ööbijad – olukord, mida tuntakse ka nähtamatu turismina (De Cantis *et al*., 2015). Ühtlasi võivad turistid ühe reisi jooksul ööbida mitmes erinevas kohas, tekitades topeltloendamise probleemi (De Cantis *et al*., 2015). Seega ei ole majutusandmete põhjal võimalik analüüsida turisti liikumist kogu külastuse vältel. Schengeni viisaruumis puudub Euroopa riikide vahel ühtne piirikontroll ja piiriületajate registreerimine. Küsitluste peamiseks puuduseks on kõrge hind, kuna läbiviimine on tööjõumahukas, mistõttu ei saa andmeid koguda suure valimi kohta ning teha üldistust kogu rahvastikule (Eurostat, 2014b, 2014a). Küsitlused ja reisipäevikud on üldjuhul tagasivaatavad, kuid inimesed sageli ei mäleta enam täpseid külastatud kohti ja nii kannatab andmete kvaliteet (De Cantis *et al*., 2015; McKercher & Lau, 2009; Shoval & Isaacson, 2007b). Seega on vaja uusi ja paremaid andmeallikaid, mis võimaldaksid sihtkohtade olemust turistide liikumiste põhjal ajas ja ruumis täpsemini mõõta.

Info- ja kommunikatsioonitehnoloogia kiire areng on loonud mitmeid uusi võimalusi inimeste paiknemise ja liikumise uurimiseks. Mobiil- ja nutitelefonide laialdase leviku tõttu kannab üha suurem osa ühiskonnast pidevalt endaga kaasas seadet, mille abil on võimalik teada saada väga täpseid asukohaandmeid inimeste, sh ka turistide paiknemise ja liikumise kohta (Shoval, 2007). Levinumateks digitaalseteks asukohaandmeteks turismiuuringutes on GPS-i, mobiilpositsioneerimise, Bluetooth'i ja geolokeeritud sotsiaalmeedia postituste ja fotode andmed (Shoval & Ahas, 2016). Enamik asukohaandmeid kasutavad uuringud on aga

ajaliselt ja ruumiliselt piiritletud. Näiteks on enim kasutatavate GPS andmete põhjal turistide liikumist uuritud teemapargis (Birenboim *et al*., 2013), saarel (Y. Li *et al*., 2019) linnades (Grinberger & Shoval, 2018; Mckercher *et al*., 2019; Shoval, 2008) ning nende kestus on üldjuhul mõni päev kuni nädal. Seega puuduvad suuremat geograafilist piirkonda või eri ruumitasemeid hõlmavad pikemaajalised uuringud. Erandiks on passiivse mobiilpositsioneerimise andmed, mis salvestuvad automaatselt mobiilsideoperaatori logides ja võimaldavad turistide kõnetoimingute asukohtade põhjal uurida liikumisi erinevatel geograafilistel tasemetel, kohalikust kuni globaalseni, pika perioodi (nt aastate) vältel.

Seega on mobiilpositsioneerimise andmete abil võimalik analüüsida sihtkohti tegelike külastusandmete alusel. Sellist turismivoogude põhist lähenemist on varasemalt rõhutanud ka sihtkoha turunduse ja juhtimise teemadega tegelevad autorid (Beritelli *et al*., 2014; Dredge, 1999; Pechlaner *et al*., 2012). Seejuures on välja toodud ka digitaalsete suurandmete kasutuselevõtu vajadust (Beritelli *et al*., 2019), paraku on aga selliseid uuringuid vähe. Suurandmeid kasutavad turismiuuringud kalduvad olema andmetest orienteeritud ning ei arvesta laiema teoreetilise taustaga ning illustreerivad pigem andmete erinevaid kasutusvõimalusi. Teoreetilises uuringutes seevastu on harva empiiriliste andmete analüüsil tuginevaid selgitusi, need tuginevad pigem autorite arvamustele, vaatlustele, kogemustele ja lugemusele (Butler, 2015). Selle doktoritöö uudsus seisneb just nende kaks poole – andmete ja teooria – omavahelises sidumises. Digitaalsed asukohaandmed turistide liikumistest võimaldavad anda empiirilise sisu muidu küllaltki abstraktsele ja iseenesest võetavale mõistele nagu sihtkoht.

Doktoritöö keskendub Eestisse sisenevale turismile perioodil 2011–2017. Välisturistide liikumist analüüsitakse passiivse mobiilpositsioneerimise rändlusteenuse andmetel Eestis. Doktoritööl on kaks peamist eesmärki:

1. Mõista, kuidas mobiilpositsioneerimise andmeid on võimalik kasutada turismistatistika tootmiseks, et neid oleks võimalik kasutada sihtkoha analüüsides.
2. Näidata, kuidas mobiilpositsioneerimise andmetest toodetud turismistatistikat on võimalik kasutada detailsemaks sihtkoha analüüsiks ja mõtestamiseks.

Peamiste eesmärkide täitmiseks seati neli alameesmärki:

a) Kirjeldada turismistatistika tootmise metoodikat passiivse mobiilpositioniseerimise andmetest, kasutades selleks Eestisse sisenevate välisturistide rändlusteenuse andmeid.
b) Hinnata mobiilpositsioneerimise andmetest toodetud statistika esinduslikkust, võrreldes saadud tulemusi ametliku majutusstatistikaga.
c) Analüüsida mobiilpositsioneerimise andmetest tuletatud külastuste põhjal sihtkohtade geograafilist ulatust, ajalist rütmi ja turistide koosseisu.
d) Tuvastada sihtkoha funktsioon külastajate liikumise alusel.

Andmed ja metoodika. Doktoritöös kasutatakse passiivse mobiilpositsioneerimise andmeid, mis tekivad kõnetoimingu tegemisel või andmeside kasutamisel. Need salvestuvad automaatselt mobiilioperaatori süsteemides ning sisaldavad

tegevuse aega ja kohta mobiiliantenni täpsusega. Andmed salvestuvad nii Eesti elanike liikumise kohta Eestis (siseturism), Eesti elanike välisreiside kohta (väljaminev turism), kui välisriikide elanike viibimisest Eestis (sissetulev turism). Seega on mobiiliandmed kasulikud erinevate turismitüüpide uurimiseks. See doktoritöö keskendub sissetulevale turismile, mille puhul on võimalik lisaks külastuse ajale ja kohale eristada ka riik, kus on registreeritud külastaja telefoni SIM-kaart, ehk on võimalik uurida turistide liikumisi päritoluriikide lõikes.

Enne külastusandmete põhiste analüüside tegemist tuleb läbida mitmed olulised sammud. Esmane andmete töötlus toimub mobiilioperaatorite süsteemides, mille käigus saab iga kõnetoimingu sooritaja unikaalse pseudonüümse identifitseeritava numbri, mis on muutumatu kogu andmeperioodi vältel ning mida ei saa seostada ühegi konkreetse telefoninumbri ega inimesega. Seejärel edastatakse krüpteeritud andmed turvalisi kanaleid pidi uurijatele. Esmased andmed sisaldavad siiski veel mitmeid vigu, mis tuleb andmete puhastamise ja töötlemise käigus kõrvaldada. Näiteks eemaldatakse duplikaadid, tehnosüsteemide (nt SIM-kaardiga väravad) andmed ja muud vigased kirjed ning kui avastatakse puuduvaid andmeid (nt konkreetse perioodi või ala kohta) päritakse need operaatoritelt uuesti. Pärast andmete korrastamist eristatakse Eestis tehtud ajaliselt lähestikuste kõnetoimingute alusel külastused (Tiru *et al*., 2010). Kui kahe järjestikuse kõnetoimingu vaheline aeg on üle 155 tunni (6,5 päeva), eeldatakse, et tegu on uue külastusega. Leitud külastuste hulgas võib aga olla selliseid juhtumeid, kus tegelikult ei ole Eestis käidud. Näiteks meremeeste kõned laevadelt, mille registreerivad Eesti rannikul olevad mobiilimastid või ka juhuslik rändlus Venemaa ja Läti piiril, kus riiki ei ole sisenetud, aga Eesti mastid registreerivad kõne. Mõlemat tüüpi n-ö valekülastused eemaldatakse külastuste andmestikust.

Külastuste tuvastamisele järgneb nende analüüs kestuse ja külastuse sageduse alusel, et välistada andmestikust need külastused, mida tõenäoliselt ei tehtud turismi eesmärkidel. Näiteks esmalt eemaldatakse kõik külastused, mis asuvad tuntud transiidikoridorides (nt Ikla-Pärnu-Tallinn) ja kestavad vähem kui neli tundi. Lisaks transiitkülastustele eemaldatakse andmestikust külastused, mille kogu kestus Eestis on pikem kui pool aastat (>183 päeva) (Eurostat, 2014b), ehk tegu võib olla inimese elukohaga. Sellesse gruppi kuuluvad valdav osa pendelrändajate ja pikaajaliste töötajate külastusest. Kõik külastused, mis alles jäävad on kirjeldatud metoodika kohaselt turismikülastused, mille põhjal on võimalik sihtkohti täpsemalt analüüsida.

Doktoritöös analüüsitakse Eesti mobiilsideoperaatorite riigisisese rändlusteenuse kõnetoimingute andmeid eri perioodidelt vahemikus 2011–2017. Andmete esinduslikkust hinnatakse võrdluses ametliku majutusstatistikaga (Statistics Estonia, 2019). Passiivse mobiilpositsioneerimise andmetest tuletatud külastusandmed võimaldavad uurida turistide külastuste geograafilist ulatust ja ajalist rütmi (külastuse aeg ja kestus) päritoluriikide lõikes. Lisaks saab külastusandmete põhjal eristada erinevaid sihtkoha funktsioone. Töös on keskendutud peamisele sisenemisvärvale, mis on leitud külastuse esimese kõnetoimingu asukoha alusel omavalitsuse täpsusega.

Tulemused ja arutelu. Mobiilpositsioneerimise andmetest tuletatud külastusandmete põhjal leitud ööbimiste hulk kuus korreleerub tugevalt vastava näitajaga ametlikus majutusstatistikas. Mobiiliandmed ülehindavad tulemusi turismi kõrghooajal juulis ja detsembris ning pigem alahindavad madalhooajal. Mobiiliandmete esinduslikkus varieerub ka riigiti. Näiteks Saksamaa ja Venemaa turistide puhul näitavad majutusandmed kõrgemaid tulemusi kui mobiiliandmed. Samas võrreldes mobiiliandmetega, on Soome, Läti ja Norra turistide külastused majutusandmetes oluliselt alahinnatud. Siinkohal on oluline rõhutada, et majutusandmetes ei kajastu ööbimine sõprade ja sugulaste juures ega suur osa Airbnb üüripindadel ööbinutest. Samas mobiiliandmete põhised tulemused on otseses seoses telefoni kasutusega. Kui inimene näiteks kõrgete rändlushindade tõttu oma telefoni Eestis olles ei kasuta, siis ta selles andmestikus ka ei ole.

Külastuste geograafilise paiknemise alusel on võimalik välja selgitada enim külastatavad omavalitused. Eesti vaieldamatult enim külastatud omavalitsus on Tallinn, kuhu jõudis 2016. aastal 80% kõikidest väliskülastajatest. Tallinnale järgnevad Viimsi vald, Pärnu, Tartu ja Narva. Lisaks eristuvad külastuste ruumilise paiknemise alusel ka peamised liikumisteed, näiteks Tallinn-Ikla-Riia, kuhu jäävad Märjamaa, Tahkuranna ja Häädemeeste vallad. Lisaks on mobiiliandmete alusel võimalik analüüsida väiksemate sihtkohtade külastusi. Selles töös vaadeldi lähemalt Saaremaa ja Tartumaa külastusi, millest selgus, et mõlemad maakonnad toimivad eraldiseisvate sihtkohtadena. Maakondade ruumilised külastusmustrid erinevad ka teiste külastatud piirkondade osas. Ainsaks ühiseks kohaks nende maakondade külastamisele on Tallinn, kust paljud Eesti külastused alguse saavad. Paraku ei ole ka mobiiliandmete analüüsi põhjal võimalik tõmmata selgeid piirjooni sihtkohale, kuna andmete ruumiline täpsus sõltub mobiilimastide paiknemise tihedusest. Seega andmed on täpsemad linnapiirkonnas, kus on maste pindalaühiku kohta enam, aga seegi ei anna võimalusi eristada kõrvuti paiknevate atraktsioonide külastust.

Lähtuvalt Hägerstrandi (1970) ajageograafia põhimõtetest ei saa külastuste ruumilist paiknemist analüüsida aja dimensioonita. Sarnaselt majutusstatistikast teadaolevale kinnitavad ka mobiiliandmed Eestis turismi sesoonset iseloomu, kus *ca* 35% külastustest tehakse kolmel suvekuul. See probleem muutub aga oluliseks just väiksemate sihtkohtade puhul. Näiteks Saaremaa puhul on vastav näitaja 53%, seevastu külastused Tartumaale on küllaltki ühtlaselt kogu aasta peale jaotunud. Selline hooajaline erinevus külastustes kinnitab selgelt, kui oluline on turismiplaneerijate ja -korraldajate jaoks lähtuda sihtkoha mõtestamisel sellest, millal eri piirkonnad üldse tänu külastajate olemasolule sihtkohana aktiviseeruvad (Beritelli *et al*., 2014). Lisaks külastuse hulga muutusele ajas, saab mobiiliandmete põhjal uurida külastuse kestust. Mida rohkem on turistil aega, seda tõenäolisemalt külastab ta reisi jooksul ka erinevaid paiku. Seda kinnitavad ka selle töö tulemused, kus lühimad külastused tehtu ainult Tallinnasse (1,5 päeva) ja pikimad Saaremaale (3,5 päeva).

Lisaks külastuste ajalisele ja ruumilisele ulatusele on selles töös uuritud ka erinevusi turistide külastusmustrites päritoluriigiti. Turisti kodumaa on tuletatud telefoni SIM-kaardi registreerimise riigi järgi. Eesti väliskülastajatest *ca* 70% on

pärit naaberriikidest: Soomest, Lätist ja Venemaalt. See tõendab selgelt kauguse summutavat mõju, mis on kinnitust leidnud ka varasemates töödes (Nilbe *et al*., 2014). Külastuste ruumiline ulatus ja ajaline rütm erineb samuti päritoluriigiti. Eesti asukoha tõttu on lätlaste külastused enam koondunud Lõuna-Eesti maakondadesse ja saartele, soomlaste külastused on aga valdavalt Tallinna kesksed. Ajaliselt on kaugemate riikide turistid (nt Saksamaa) enam suvele orienteeritud kui lähinaabritest (nt Venemaa) turistid. Seega kinnitavad ka selle doktoritöö tulemused varasemat teadmist, et turistide geograafilisel päritolul on oluline mõju külastusmustritele (Bowden, 2003; Hwang *et al*., 2006; Oppermann, 1992a; Reid & Reid, 1997).

Turistide külastuste aegruumilise analüüsi põhjal on võimalik eristada sihtkoha funktsioone koha suhtelise asukoha kaudu, võrreldes seda teiste külastatud kohtadega (Lew & McKercher, 2002; Liu *et al*., 2012). Selles töös keskenduti sisenemisväravale kui esimesele kohale, mida turist riiki sisenedes läbib. Külastuste esimeste kõnetoimingute analüüsist selgus, et 66% Eesti külastustest saab alguse Tallinnast või selle lähivaldadest, nagu Viimsi, Harku ja Rae. See on lihtsasti selgitatav, sest Tallinnas asuvad rahvusvaheline lennujaam, sadam ning rongi- ja bussiühendused. Teised väiksemad sisenemisvärvad – Ikla, Narva ja Luhamaa piiripunkt – jäävad juba oluliselt Tallinnale alla. Eraldi väärib märkimist Tartu kui ainukene koht lisaks Tallinnale, kuhu käib hetkel regulaarne rahvusvaheline lennuliin Helsingist. Laiem probleem sisenemisvärava analüüsil seisneb aga selles, et ligi 80% Tallinnast algavatest külastustest, ei jõuagi sealt kaugemale, ehk pealinn toimib lisaks värvale ka külastajate peamise sihtkohana riigis. See, et kohal on mitu funktsiooni (Beritelli & Laesser, 2017; Lohmann & Pearce, 2010; Pearce, 2001a) ei ole iseenesest probleem vaid tõsiasi, et turistid on ruumis koondunud ühele küllaltki piiratud alale. Tallinna puhul on selleks kindlasti vanalinn, mis võib suvistel tipptundidel täituda tuhandete kruiisituristidega. See omakorda võib tekitada mitmeid sotsiaalmajanduslikke ja keskkonna probleeme, viies halvimatel juhtudel liigturismini (Koens *et al*., 2018).

Kokkuvõte. Empiiriliselt mõõdetavad parameetrid sihtkohtade külastatavuse kohta võimaldavad sihtkohti eristada, võrrelda ja grupeerida ning sihtkoha olemust ajas ja ruumis mõõta. Täpne teadmine sellest, kes, millal ja milliseid sihtkohti külastab, on vajalik eelkõige turismi arendustegevuste kavandamisel, koostöö korraldamisel, taristu planeerimisel ning turundustegevuse suunamisel. Paraku ei ole sellised mobiiltelefonide kasutusel põhinevad ajaliselt ja ruumiliselt täpsed andmed veel kõigis riikides lihtsasti kättesaadavad. Maailmas on ainult kaks riiki, Eesti ja Indoneesia, kus ametlik turismistatistika põhineb mobiilpositsioneerimise andmetel. Seega on potentsiaali selliste andmete laialdasemaks kasutuseks turismiuuringutes, kuid esmalt on vaja lahendada seadusandlikud ja regulatiivsed piirangud andmete kättesaadavuses. Seejuures on oluline silmas pidada isikuandmete kaitse üldmääruse (GDPR) peamisi printsiipe nagu andmete ja riskide minimeerimine ehk siis eelkõige privaatsuse tagamine. Kättesaadavuse probleemi kõrvale jättes, on mobiilpositsioneerimise andmete näol tegu väärtusliku andmeallikaga, mille peamisteks tugevusteks on selle automaatne kogumine, kulutõhusus

ja väga hea ajaline ja ruumiline katvus. Selle doktoritöö tulemused kinnitasid, et mobiiliandmed on võrdväärsed ja kohati paremadki ametlikust majutusstatistikast ning võimaldavad analüüsida leitud külastuste põhjal turismisihtkoha dünaamilist olemust, andes seeläbi empiirilise sisendi ka teoreetilisteks sihtkoha aruteludeks.

REFERENCES

Ahas, R., Aasa, A., Mark, Ü., Pae, T., & Kull, A. (2007). Seasonal tourism spaces in Estonia: Case study with mobile positioning data. *Tourism Management*, *28*(3), 898–910. https://doi.org/10.1016/j.tourman.2006.05.010

Ahas, R., Aasa, A., Roose, A., Mark, Ü., & Silm, S. (2008). Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. *Tourism Management*, *29*, 469–486. https://doi.org/doi.org/10.1016/j.tourman.2007.05.014

Ahas, R., Aasa, A., Silm, S., & Tiru, M. (2007). Mobile Positioning Data in Tourism Studies and Monitoring: Case Study in Tartu, Estonia. *Information and Communication Technologies in Tourism 2007. Proceedings of the International Conference in Ljubljana, Slovenia, 2007.*, 119–128. https://doi.org/10.1007/978-3-211-69566-1

Ahas, R., & Mark, Ü. (2005). Location based services—new challenges for planning and public administration? *Futures*, *37*(6), 547–561. https://doi.org/10.1016/j.futures.2004.10.012

Ahas, R., Silm, S., & Tiru, M. (2017). Eestist lähtuv hargmaisus mobiilside rändlusandmete alusel. In T. Tammaru (Ed.), *Eesti inimarengu aruanne 2016/2017 Eesti rändeajastul* (pp. 82–90). Tallinn: Eesti Koostöö Kogu.

Asakura, Y., & Hato, E. (2001). Behavioral monitoring of public transport users through a mobile communication system. *Journal of Advanced Transportation*, *35*(3), 289–304. https://doi.org/10.1002/atr.5670350307

Asakura, Y., & Hato, E. (2004). Tracking survey for individual travel behaviour using mobile communication instruments. *Transportation Research Part C: Emerging Technologies*, *12*, 273–291. https://doi.org/10.1016/j.trc.2004.07.010

Asero, V., Gozzo, S., & Tomaselli, V. (2015). Building tourism networks through tourist mobility. *Journal of Travel Research*, *55*(6), 751–763. https://doi.org/10.1177/0047287515569777

Backer, E. (2012). VFR travel: It is underestimated. *Tourism Management*, *33*(1), 74–79. https://doi.org/10.1016/j.tourman.2011.01.027

Baggio, R., & Scaglione, M. (2018). Strategic visitor flows and destination management organization. *Information Technology and Tourism*, *18*(1–4), 29–42. https://doi.org/10.1007/s40558-017-0096-1

Bargeman, B., Joh, C.-H., & Timmermans, H. (2002). Vacation behavior using a sequence alignment method. *Annals of Tourism Research*, *29*(2), 320–337. https://doi.org/10.1016/S0160-7383(01)00065-2

Batista e Silva, F., Marín Herrera, M. A., Rosina, K., Ribeiro Barranco, R., Freire, S., & Schiavina, M. (2018). Analysing spatiotemporal patterns of tourism in Europe at high-resolution with conventional and big data sources. *Tourism Management*, *68*, 101–115. https://doi.org/10.1016/J.TOURMAN.2018.02.020

Beaman, J. G., Huan, T.-C., & Beaman, J. P. (2004). Tourism Surveys: Sample Size, Accuracy, Reliability, and Acceptable Error. *Journal of Travel Research*, *43*, 67–74. https://doi.org/10.1177/0047287504265514

Beritelli, P., Bieger, T., & Laesser, C. (2014). The new frontiers of destination management: Applying variable geometry as a function-based approach. *Journal of Travel Research*, *53*(4), 403–417. https://doi.org/10.1177/0047287513506298

Beritelli, P., Crescini, G., Reinhold, S., & Schanderl, V. (2019). How Flow-Based Destination Management Blends Theory and Method for Practical Impact. In N. Kozak & M. Kozak (Eds.), *Tourist Destination Management: Instruments, Products, and Case Studies* (pp. 289–310). https://doi.org/10.1007/978-3-030-16981-7_17

Beritelli, P., & Laesser, C. (2017). The Dynamics of Destinations and Tourism Development. In D. R. Fesenmaier & Z. Xiang (Eds.), *Design Science in Tourism: Foundations of Destination Management* (pp. 195–214). https://doi.org/10.1007/978-3-319-42773-7_13

Beritelli, P., Reinhold, S., & Laesser, C. (2020). Visitor flows, trajectories and corridors: Planning and designing places from the traveler's point of view. *Annals of Tourism Research*, *82*, 102936. https://doi.org/10.1016/j.annals.2020.102936

Birenboim, A. (2016). New approaches to the study of tourist experiences in time and space. *Tourism Geographies*, *18*(1), 9–17. https://doi.org/10.1080/14616688.2015.1122078

Birenboim, A., Anton-Clavé, S., Russo, A. P., & Shoval, N. (2013). Temporal activity patterns of theme park visitors. *Tourism Geographies*, *15*(4), 601–619. https://doi.org/10.1080/14616688.2012.762540

Birenboim, A., Reinau, K. H., Shoval, N., & Harder, H. (2015). High-Resolution Measurement and Analysis of Visitor Experiences in Time and Space: The Case of Aalborg Zoo in Denmark. *The Professional Geographer*, *67*(4), 620–629. https://doi.org/10.1080/00330124.2015.1032874

Birenboim, A., & Shoval, N. (2015). Mobility research in the age of the smartphone. *Annals of the Association of American Geographers*, *106*(2), 283–291. https://doi.org/10.1080/00045608.2015.1100058

Bowden, J. (2003). A cross-national analysis of international tourist flows in China. *Tourism Geographies*, *5*(3), 257–279. https://doi.org/10.1080/14616680309711

Buhalis, D. (2000). Marketing the competitive destination of the future. *Tourism Management*, *21*(1), 97–116. https://doi.org/10.1016/S0261-5177(99)00095-3

Burghardt, A. F. (1971). A hypothesis about gateway cities. *Annals of the Association of American Geographers*, *61*(2), 269–285. https://doi.org/10.1111/j.1467-8306.1971.tb00782.x

Butler, R. W. (1980). The Concept of a Tourist Area Cycle of Evolution: Implications for Management of Resources. *The Canadian Geographer/Le Géographe Canadien*, *24*(1), 5–12. https://doi.org/10.1111/j.1541-0064.1980.tb00970.x

Butler, R. W. (2015). The evolution of tourism and tourism research. *Tourism Recreation Research*, *40*(1), 16–27. https://doi.org/10.1080/02508281.2015.1007632

Butt, S., & Phillips, J. G. (2008). Personality and self reported mobile phone use. *Computers in Human Behavior*, *24*(2), 346–360. https://doi.org/10.1016/J.CHB.2007.01.019

Caldeira, A. M., & Kastenholz, E. (2017). Tourists' spatial behaviour in urban destinations: The effect of prior destination experience. *Journal of Vacation Marketing*, *24*(3), 247–260. https://doi.org/10.1177/1356766717706102

Caldeira, A. M., & Kastenholz, E. (2018). It's so hot: predicting climate change effects on urban tourists' time-space experience. *Journal of Sustainable Tourism*. https://doi.org/10.1080/09669582.2018.1478840

Chua, A., Servillo, L., Marcheggiani, E., & Moere, A. Vande. (2016). Mapping Cilento: Using geotagged social media data to characterize tourist flows in southern Italy. *Tourism Management*, *57*, 295–310. https://doi.org/10.1016/j.tourman.2016.06.013

D'Agata, R., Gozzo, S., & Tomaselli, V. (2013). Network analysis approach to map tourism mobility. *Quality and Quantity*, *47*(6), 3167–3184. https://doi.org/10.1007/s11135-012-9710-7

Dattilo, B., & Sabato, M. (2017). Travelling SIM and Trips: an approach to make mobile phone data usable in tourism statistics. *New Techniques and Technologies for Statistics*. https://doi.org/10.2901/EUROSTAT.C2017.001

De Cantis, S., Parroco, A. M., Ferrante, M., & Vaccina, F. (2015). Unobserved tourism. *Annals of Tourism Research*, *50*, 1–18. https://doi.org/10.1016/j.annals.2014.10.002

Demunter, C. (2017). *Tourism statistics: Early adopters of big data*. https://doi.org/10.2785/762729

Dredge, D. (1999). Destination place planning and design. *Annals of Tourism Research*, *26*(4), 772–791. https://doi.org/10.1016/S0160-7383(99)00007-9

Eesti Pank. (2020). International travel statistics. Retrieved June 12, 2019, from http://statistika.eestipank.ee/#/en/p/MAKSEBIL_JA_INVPOS/1410

European Commission. (2018). *Special Eurobarometer 462 E-Communications and Digital Single Market*. https://doi.org/10.2759/043384

European Parliament and Council of the European Union. (2002). Directive concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications). *Official Journal of the European Communities*, *L 201*, 37–47.

European Parliament and Council of the European Union. (2016). Regulation on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (Data Protection Directive). *Official Journal of the European Union*, *L 119*, 1–88.

Eurostat. (2014a). *Feasibility Study on the Use of Mobile Positioning Data for Tourism Statistics – Consolidated Report*. https://doi.org/10.2785/55051

Eurostat. (2014b). *Methodological manual for tourism statistics: version 3.1*. https://doi.org/10.2785/67001

Ferrante, M., Abbruzzo, A., & De Cantis, S. (2017). Graphical models for estimating network determinants of multi-destination trips in Sicily. *Tourism Management Perspectives*, *22*, 109–119. https://doi.org/10.1016/j.tmp.2017.03.004

Flognfeldt, T. (1999). Traveler geographic origin and market segmentation: The multi trips destination case. *Journal of Travel & Tourism Marketing*, *8*(1), 111–124. https://doi.org/10.1300/J073v08n01_07

Framke, W. (2002). The Destination as a Concept: A Discussion of the Business-related Perspective versus the Socio-cultural Approach in Tourism Theory. *Scandinavian Journal of Hospitality and Tourism*, *2*(2), 92–108. https://doi.org/10.1080/15022250216287

García-Palomares, J. C., Gutierrez, J., & Mínguez, C. (2015). Identification of tourist hot spots based on social networks: A comparative analysis of European metropolises using photo-sharing services and GIS. *Applied Geography*, *63*, 408–417. https://doi.org/10.1016/j.apgeog.2015.08.002

Girardin, F., Fiore, F. D., Ratti, C., & Blat, J. (2008). Leveraging explicitly disclosed location information to understand tourist dynamics: A case study. *Journal of Location Based Services*, *2*(1), 41–56. https://doi.org/10.1080/17489720802261138

Gitton, F. (2016). Testing the use of cell phone data for inbound tourism statistics for France. *14 Th Global Forum on Tourism Statistics*. Venice.

Golledge, R. G., & Stimson, R. J. (1997). *Spatial behavior: a geographic perspective*. The Guilford Press.

Grinberger, A. Y., & Shoval, N. (2018). Spatiotemporal contingencies in tourists' intradiurnal mobility patterns. *Journal of Travel Research*, *58*(3), 512–530. https://doi.org/10.1177/0047287518757372

Grinberger, A. Y., Shoval, N., & McKercher, B. (2014). Typologies of tourists' time–space consumption: A new approach using GPS data and GIS tools. *Tourism Geographies*, *16*(1), 105–123. https://doi.org/10.1080/14616688.2013.869249

GSMA. (2020). *The Mobile Economy 2020*. Retrieved from www.gsmaintelligence.com

Guedes, A. S., & Jiménez, M. I. M. (2015). Spatial patterns of cultural tourism in Portugal. *Tourism Management Perspectives*, *16*, 107–115. https://doi.org/10.1016/J.TMP.2015.07.010

Guizzardi, A., & Bernini, C. (2012). Measuring underreporting in accommodation statistics: evidence from Italy. *Current Issues in Tourism*, *15*(6), 597–602. https://doi.org/10.1080/13683500.2012.667071

Hägerstrand, T. (1970). What about people in Regional Science? *Papers of the Regional Science Association*, Vol. 24, pp. 6–21. https://doi.org/10.1007/BF01936872

Hall, C. M. (2005a). *Tourism: Rethinking the Social Science of Mobility*. Essex: Pearson Education Limited.

Hall, C. M. (2005b). Reconsidering the geography of tourism and contemporary mobility. *Geographical Research*, *43*(2), 125–139. https://doi.org/10.1111/j.1745-5871.2005.00308.x

Hall, C. M., & Page, S. (2014). *The geography of tourism and recreation: environment, place and space* (4th ed.). New York: Routledge.

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, *41*(3), 260–271. https://doi.org/10.1080/15230406.2014.890072

Heerschap, N., Ortega, S., Priem, A., & Offermans, M. (2014). Innovation of tourism statistics through the use of new big data sources. *12th Global Forum on Tourism Statistics*. Prague.

Hwang, Y.-H., Gretzel, U., & Fesenmaier, D. R. (2006). Multicity trip patterns. Tourists to the United States. *Annals of Tourism Research*, *33*(4), 1057–1078. https://doi.org/10.1016/j.annals.2006.04.004

Isaacson, M., & Shoval, N. (2006). Application of Tracking Technologies to the Study of Pedestrian Spatial Behavior∗. *The Professional Geographer*, *58*(2), 172–183. https://doi.org/10.1111/j.1467-9272.2006.00524.x

Iwashita, C. (2003). Media construction of Britain as a destination for Japanese tourists : Social constructionism and tourism. *Tourism and Hospitality Research*, *4*(4), 331–340. https://doi.org/10.1177/1468797606071477

Kádár, B. (2014). Measuring tourist activities in cities using geotagged photography. *Tourism Geographies*, *16*(1), 88–104. https://doi.org/10.1080/14616688.2013.868029

Kitchin, R. (2014). Big Data, new epistemologies and paradigm shifts. *Big Data & Society*, *1*(1), 205395171452848. https://doi.org/10.1177/2053951714528481
Koens, K., Postma, A., & Papp, B. (2018). Is overtourism overused? Understanding the impact of tourism in a city context. *Sustainability*, *10*(12), 4384. https://doi.org/10.3390/su10124384
Koo, T. T. R., Wu, C. L., & Dwyer, L. (2012). Dispersal of visitors within destinations: Descriptive measures and underlying drivers. *Tourism Management*, *33*(5), 1209–1219. https://doi.org/10.1016/j.tourman.2011.11.010
Kuusik, A., Tiru, M., Ahas, R., & Varblane, U. (2011). Innovation in destination marketing. *Baltic Journal of Management*, *6*(3), 378–399. https://doi.org/10.1108/17465261111168000
Larsen, J., Urry, J., & Axhausen, K. W. (2007). Networks and tourism. Mobile Social Life. *Annals of Tourism Research*, *34*(1), 244–262. https://doi.org/10.1016/j.annals.2006.08.002
Lau, G., & McKercher, B. (2006). Understanding tourist movement patterns in a destination: A GIS approach. *Tourism and Hospitality Research*, *7*, 39–49. https://doi.org/10.1057/palgrave.thr.6050027
Law, R., Buhalis, D., & Cobanoglu, C. (2014). Progress on information and communication technologies in hospitality and tourism. *International Journal of Contemporary Hospitality Management*, *26*(5), 6. https://doi.org/10.1108/IJCHM-08-2013-0367
Law, R., Leung, D., Cheng, I., & Chan, C. (2020). Progression and development of information and communication technology research in hospitality and tourism A state-of-the-art review. *International Journal of Contemporary Hospitality Management*, *32*(2), 511–534. https://doi.org/10.1108/IJCHM-07-2018-0586
Leiper, N. (1979). The framework of tourism: Towards a definition of tourism, tourist, and the tourist industry. *Annals of Tourism Research*, *6*(4), 390–407. https://doi.org/10.1016/0160-7383(79)90003-3
Leiper, N. (1989). Main destination ratios: Analyses of tourist flows. *Annals of Tourism Research*, *16*(4), 530–541. https://doi.org/10.1016/0160-7383(89)90007-8
Lestari, T. K., Esko, S., Sarpono, Saluveer, E., & Rufiadi, R. (2018). Indonesia's Experience of using Signaling Mobile Positioning Data for Official Tourism Statistics. *15th Global Forum on Tourism Statistics*. Cusco.
Leung, X. Y., Wang, F., Wu, B., Bai, B., Stahura, K. A., & Xie, Z. (2012). A social network analysis of overseas tourist movement patterns in Beijing: The impact of the Olympic Games. *International Journal of Tourism Research*, *14*(5), 469–484. https://doi.org/10.1002/jtr.876
Lew, A. A. (1987). A framework of tourist attraction research. *Annals of Tourism Research*, *14*(4), 553–575. https://doi.org/10.1016/0160-7383(87)90071-5
Lew, A. A., & McKercher, B. (2002). Trip destinations, gateways and itineraries: The example of Hong Kong. *Tourism Management*, *23*(6), 609–621. https://doi.org/10.1016/S0261-5177(02)00026-2
Lew, A. A., & McKercher, B. (2006). Modeling tourist movements: A local destination analysis. *Annals of Tourism Research*, *33*(2), 403–423. https://doi.org/10.1016/j.annals.2005.12.002
Li, J., Xu, L., Tang, L., Wang, S., & Li, L. (2018). Big data in tourism research: A literature review. *Tourism Management*, *68*, 301–323. https://doi.org/10.1016/J.TOURMAN.2018.03.009

Li, Y., Yang, L., Shen, H., & Wu, Z. (2019). Modeling intra-destination travel behavior of tourists through spatio-temporal analysis. *Journal of Destination Marketing & Management*, *11*, 260–269. https://doi.org/10.1016/j.jdmm.2018.05.002

Lichrou, M., O'Malley, L., & Patterson, M. (2008). Place product or place narrative(s)? Perspectives in the Marketing of Tourism Destinations. *Journal of Strategic Marketing*, *16*(1), 27–39. https://doi.org/10.1080/09652540701794429

Liu, F., Zhang, J., Zhang, J., Chen, D., Liu, Z., & Lu, S. (2012). Roles and functions of tourism destinations in tourism region of South Anhui: A tourist flow network perspective. *Chinese Geographical Science*, *22*(6), 755–764. https://doi.org/10.1007/s11769-012-0557-6

Lohmann, G., & Pearce, D. G. (2010). Conceptualizing and operationalizing nodal tourism functions. *Journal of Transport Geography*, *18*(2), 266–275. https://doi.org/10.1016/j.jtrangeo.2009.05.003

Lue, C., Crompton, J. L., & Fesenmaier, D. R. (1993). Conceptualization of multi-destination pleasure trips. *Annals of Tourism Research*, *20*(2), 289–301. https://doi.org/10.1016/0160-7383(93)90056-9

Maitland, R. (2012). Capitalness is contingent: Tourism and national capitals in a globalised world. *Current Issues in Tourism*, *15*(1–2), 3–17. https://doi.org/10.1080/13683500.2011.634891

Masso, A., Silm, S., & Ahas, R. (2019). Generational differences in spatial mobility: A study with mobile phone data. *Population, Space and Place*, *25*(2), e2210. https://doi.org/10.1002/psp.2210

Mckercher, B., Hardy, A., & Aryal, J. (2019). Using tracking technology to improve marketing: insights from a historic town in Tasmania, Australia. *Journal of Travel & Tourism Marketing*, *36*(7), 823–834. https://doi.org/10.1080/10548408.2019.1580243

McKercher, B., & Lau, G. (2008). Movement patterns of tourists within a destination. *Tourism Geographies*, *10*(3), 355–374. https://doi.org/10.1080/14616680802236352

McKercher, B., & Lau, G. (2009). Methodological Considerations when Mapping Tourist Movements in a Destination. *Tourism Analysis*, *14*(4), 443–455. https://doi.org/10.3727/108354209X12596287114138

McKercher, B., & Lew, A. A. (2004). Tourist flows and the spatial distribution of tourists. In A. A. Lew, M. Hall, & A. M. William (Eds.), *A companion to tourism* (pp. 36–48). https://doi.org/10.1002/9780470752272.ch3

McKercher, B., Shoval, N., Park, E., & Kahani, A. (2015). The [Limited] Impact of Weather on Tourist Behavior in an Urban Destination. *Journal of Travel Research*, *54*(4), 442–455. https://doi.org/10.1177/0047287514522880

McKercher, B., & Zoltan, J. (2014). Tourist flows and spatial behavior. In A. A. Lew, C. M. Hall, & A. M. Williams (Eds.), *The Wiley Blackwell Companion to Tourism* (pp. 33–44). https://doi.org/10.1002/9781118474648.ch2

Mings, R. C., & Mchugh, K. E. (1992). The spatial configuration of travel to Yellowstone National Park. *Journal of Travel Research*, *30*(4), 38–46. https://doi.org/10.1177/004728759203000406

Murphy, P. E., & Keller, C. P. (1990). Destination travel patterns: An examination and modeling of tourist patterns on Vancouver Island, British Columbia. *Leisure Sciences*, *12*(1), 49–65. https://doi.org/10.1080/01490409009513089

Nilbe, K., Ahas, R., & Silm, S. (2014). Evaluating the Travel Distances of Events Visitors and Regular Visitors Using Mobile Positioning Data: The Case of Estonia. *Journal of Urban Technology*, *21*(2), 91–107. https://doi.org/10.1080/10630732.2014.888218

Nurmi, O. (2018). Improving the accuracy of outbound tourism statistics with mobile positioning data. *15th Global Forum on Tourism Statistics*. Cusco.

Önder, I. (2017). Classifying multi-destination trips in Austria with big data. *Tourism Management Perspectives*, *21*, 54–58. https://doi.org/10.1016/j.tmp.2016.11.002

Önder, I., Koerbitz, W., & Hubmann-Haidvogel, A. (2016). Tracing Tourists by Their Digital Footprints. *Journal of Travel Research*, *55*(5), 566–573. https://doi.org/10.1177/0047287514563985

Oppermann, M. (1992a). Intranational tourist flows in Malaysia. *Annals of Tourism Research*, *19*(3), 482–500. https://doi.org/10.1016/0160-7383(92)90132-9

Oppermann, M. (1992b). Travel dispersal index. *Journal of Tourism Studies*, *3*(1), 44–49.

Oppermann, M. (1995). A model of travel itineraries. *Journal of Travel Research*, *33*(4), 57–61. https://doi.org/10.1177/004728759503300409

Oppermann, M. (1999). Predicting destination choice - A discussion of destination loyalty. *Journal of Vacation Marketing*, *5*(1), 51–65. https://doi.org/10.1177/135676679900500105

Parroco, A. M., Vaccina, F., De Cantis, S., & Ferrante, M. (2012). Multi-destination trips and tourism statistics: Empirical evidences in Sicily. *Economics: The Open-Access, Open-Assessment E-Journal*, *6*(2012–44), 1–27. https://doi.org/10.5018/economics-ejournal.ja.2012-44

Pearce, D. G. (2001a). An integrative framework for urban tourism research. *Annals of Tourism Research*, *28*(4), 926–946. https://doi.org/10.1016/S0160-7383(00)00082-7

Pearce, D. G. (2001b). Towards a regional analysis of tourism in Southeast Asia. In P. Teo, T. C. Chang, & K. C. Ho (Eds.), *Interconnected worlds: Tourism in Southeast Asia* (pp. 27–43). Oxford: Pergamon.

Pearce, D. G. (2014). Toward an Integrative Conceptual Framework of Destinations. *Journal of Travel Research*, *53*(2), 141–153. https://doi.org/10.1177/0047287513491334

Pearce, D. G., & Elliott, J. M. C. (1983). The trip index. *Journal of Travel Research*, *22*(1), 6–9. https://doi.org/10.1177/004728758302200102

Pechlaner, H., Pichler, S., & Herntrei, M. (2012). From mobility space towards experience space: Implications for the competitiveness of destinations. *Tourism Review*, *67*(2), 34–44. https://doi.org/10.1108/16605371211236150

Peng, H., Zhang, J., Liu, Z., Lu, L., & Yang, L. (2016). Network analysis of tourist flows: A cross-provincial boundary perspective. *Tourism Geographies*, *18*(5), 561–586. https://doi.org/10.1080/14616688.2016.1221443

Provenzano, D., Hawelka, B., & Baggio, R. (2018). The mobility network of European tourists: A longitudinal study and a comparison with geo-located Twitter data. *Tourism Review*, *73*(1), 28–43. https://doi.org/10.1108/TR-03-2017-0052

Puura, A., Silm, S., & Ahas, R. (2017). The Relationship between Social Networks and Spatial Mobility: A Mobile-Phone-Based Study in Estonia. *Journal of Urban Technology*, *25*(2), 7–25. https://doi.org/10.1080/10630732.2017.1406253

Ratti, C., Pulselli, R. M., Williams, S., & Frenchman, D. (2006). Mobile Landscapes: Using Location Data from Cell-Phones for Urban Analysis. *Environment and Planning B: Planning and Design*, *33*, 727–748. https://doi.org/10.1068/b32047

Reid, L. J., & Reid, S. D. (1997). Traveler Geographic Origin and Market Segmentation for Small Island Nations: The Barbados Case. *Journal of Travel & Tourism Marketing*, *6*(3–4), 5–21. https://doi.org/10.1300/J073v06n03_02

Reinhold, S., Laesser, C., & Beritelli, P. (2019). Flow-based destination management and marketing: a perspective article. *Tourism Review*, *75*(1), 174–178. https://doi.org/10.1108/TR-05-2019-0193

Saarinen, J. (2004). "Destinations in change": The transformation process of tourist destinations. *Tourist Studies*, *4*(2), 161–179. https://doi.org/10.1177/1468797604054381

Saraniemi, S., & Kylänen, M. (2011). Problematizing the Concept of Tourism Destination: An Analysis of Different Theoretical Approaches. *Journal of Travel Research*, *50*(2), 133–143. https://doi.org/10.1177/0047287510362775

Shoval, N. (2007). Sensing Human Society. *Environment and Planning B: Planning and Design*, *34*(2), 191–195. https://doi.org/10.1068/b3402com

Shoval, N. (2008). Tracking technologies and urban analysis. *Cities*, *25*(1), 21–28. https://doi.org/10.1016/j.cities.2007.07.005

Shoval, N., & Ahas, R. (2016). The use of tracking technologies in tourism research: A review of the first decade. *Tourism Geographies*, *18*(5), 587–606. https://doi.org/10.1080/14616688.2016.1214977

Shoval, N., & Isaacson, M. (2007a). Sequence Alignment as a Method for Human Activity Analysis in Space and Time. *Annals of the Association of American Geographers*, *97*(2), 282–297. https://doi.org/10.1111/j.1467-8306.2007.00536.x

Shoval, N., & Isaacson, M. (2007b). Tracking tourists in the digital age. *Annals of Tourism Research*, *34*(1), 141–159. https://doi.org/10.1016/j.annals.2006.07.007

Shoval, N., Kahani, A., De Cantis, S., & Ferrante, M. (2020). Impact of incentives on tourist activity in space-time. *Annals of Tourism Research*, *80*, 102846. https://doi.org/10.1016/j.annals.2019.102846

Shoval, N., McKercher, B., Birenboim, A., & Ng, E. (2015). The application of a sequence alignment method to the creation of typologies of tourist activity in time and space. *Environment and Planning B: Planning and Design*, *42*(1), 76–94. https://doi.org/10.1068/b38065

Shoval, N., McKercher, B., Ng, E., & Birenboim, A. (2011). Hotel location and tourist activity in cities. *Annals of Tourism Research*, *38*(4), 1594–1612. https://doi.org/10.1016/j.annals.2011.02.007

Shoval, N., Schvimer, Y., & Tamir, M. (2018a). Real-Time Measurement of Tourists' Objective and Subjective Emotions in Time and Space. *Journal of Travel Research*, *57*(1), 3–16. https://doi.org/10.1177/0047287517691155

Shoval, N., Schvimer, Y., & Tamir, M. (2018b). Tracking technologies and urban analysis: Adding the emotional dimension. *Cities*, *72*(1240), 34–42. https://doi.org/10.1016/j.cities.2017.08.005

Silm, S., & Ahas, R. (2010). The seasonal variability of population in Estonian municipalities. *Environment and Planning A*, *42*(10), 2527–2546. https://doi.org/10.1068/a43139

Silm, S., & Ahas, R. (2014). Ethnic Differences in Activity Spaces: A Study of Out-of-Home Nonemployment Activities with Mobile Phone Data. *Annals of the Association of American Geographers*, *104*(3), 542–559. https://doi.org/10.1080/00045608.2014.892362

Sobolevsky, S., Sitko, I., Grauwin, S., Combes, R. T. Des, Hawelka, B., Arias, J. M., & Ratti, C. (2014). Mining Urban Performance: Scale-Independent Classification of Cities Based on Individual Economic Transactions. *ArXiv*, *abs/1405.4*. https://doi.org/10.1140/epjds31.

Statistics Estonia. (2019). Statistical database. Retrieved February 13, 2019, from http://pub.stat.ee/px-web.2001/dialog/statfile1.asp

Stienmetz, J. L., & Fesenmaier, D. R. (2019). Destination Value Systems: Modeling Visitor Flow Structure and Economic Impact. *Journal of Travel Research*, *58*(8), 1249–1261. https://doi.org/10.1177/0047287518815985

Tehnilise Järelevalve Amet. (2017). *Elektroonilise side ülevaade-III kvartal 2017*.

Tenkanen, H., Di Minin, E., Heikinheimo, V., Hausmann, A., Herbst, M., Kajala, L., & Toivonen, T. (2017). Instagram, Flickr, or Twitter: Assessing the usability of social media data for visitor monitoring in protected areas. *Scientific Reports*, *7*(1), 17615. https://doi.org/10.1038/s41598-017-18007-4

Tiru, M., Kuusik, A., Lamp, M.-L., & Ahas, R. (2010). LBS in marketing and tourism management: Measuring destination loyalty with mobile positioning data. *Journal of Location Based Services*, *4*(2), 120–140. https://doi.org/10.1080/17489725.2010.508752

Toivonen, T., Heikinheimo, V., Fink, C., Hausmann, A., Hiippala, T., Järv, O., … Di Minin, E. (2019). Social media data for conservation science: A methodological overview. *Biological Conservation*, *233*, 298–315. https://doi.org/10.1016/J.BIOCON.2019.01.023

Tsai, C.-Y., & Chung, S.-H. (2012). A personalized route recommendation service for theme parks using RFID information and tourist behavior. *Decision Support Systems*, *52*, 514–527. https://doi.org/10.1016/j.dss.2011.10.013

United Nations. (2019). *Handbook on the use of Mobile Phone data for Official Statistics*. Retrieved from https://unstats.un.org/bigdata/taskteams/mobilephone/

UNWTO. (2010). International Recommendations for Tourism Statistics 2008. In *Series M No. 83/Rev.1*. https://doi.org/10.18111/9789284419876

van der Knaap, W. G. M. (1999). Research report: GIS-oriented analysis of tourist time-space patterns to support sustainable tourism development. *Tourism Geographies*, *1*(1), 56–69. https://doi.org/10.1080/14616689908721294

Vanhoof, M., Hendrickx, L., Aare, P., Verstraeten, G., Ploetz, T., & Smoreda, Z. (2017). Exploring the use of mobile phone data for domestic tourism trip analysis. *Netcom, Networks and Communications Studies*, *31*(3/4), 335–372. https://doi.org/10.4000/netcom.2742

Versichele, M., de Groote, L., Claeys Bouuaert, M., Neutens, T., Moerman, I., & Van de Weghe, N. (2014). Pattern mining in tourist attraction visits through association rule learning on Bluetooth tracking data: A case study of Ghent, Belgium. *Tourism Management*, *44*, 67–81. https://doi.org/10.1016/j.tourman.2014.02.009

Volo, S., & Giambalvo, O. (2008). Tourism Statistics: Methodological Imperatives and Difficulties: The Case of Residential Tourism in Island Communities. *Current Issues in Tourism*, *11*(4), 369–380. https://doi.org/10.1080/13683500802140398

Vu, H. Q., Li, G., Law, R., & Ye, B. H. (2015). Exploring the travel behaviors of inbound tourists to Hong Kong using geotagged photos. *Tourism Management*, *46*, 222–232. https://doi.org/10.1016/j.tourman.2014.07.003

Vu, H. Q., Li, G., Law, R., & Zhang, Y. (2018). Travel Diaries Analysis by Sequential Rule Mining. *Journal of Travel Research*, *57*(3), 399–413. https://doi.org/10.1177/0047287517692446

Wang, Z., He, S. Y., & Leung, Y. (2018). Applying mobile phone data to travel behaviour research: A literature review. *Travel Behaviour and Society*, *11*, 141–155. https://doi.org/10.1016/J.TBS.2017.02.005

Wei, R., & Lo, V.-H. H. (2006). Staying connected while on the move: Cell phone use and social connectedness. *New Media & Society*, *8*(1), 53–72. https://doi.org/10.1177/1461444806059870

World Tourism Organization. (2019). International Tourism Highlights, 2019 Edition. In *International Tourism Highlights, 2019 Edition*. https://doi.org/10.18111/9789284421152

World Tourism Organization. (2020). World Tourism Barometer and Statistical Annex, January 2020. In *UNWTO World Tourism Barometer (English version)* (Vol. 18). https://doi.org/https://doi.org/10.18111/wtobarometereng

Wu, C.-L., & Carson, D. (2008). Spatial and temporal tourist dispersal analysis in multiple destination travel. *Journal of Travel Research*, *46*(3), 311–317. https://doi.org/10.1177/0047287506304046

Yun, H. J., & Park, M. H. (2015). Time–Space Movement of Festival Visitors in Rural Areas Using a Smart Phone Application. *Asia Pacific Journal of Tourism Research*, *20*(11), 1246–1265. https://doi.org/10.1080/10941665.2014.976581

Zakrisson, I., & Zillinger, M. (2012). Emotions in motion: tourist experiences in time and space. *Current Issues in Tourism*, *15*(6), 505–523. https://doi.org/10.1080/13683500.2011.615391

Zoltan, J., & McKercher, B. (2014). Analysing intra-destination movements and activity participation of tourists through destination card consumption. *Tourism Geographies*, *17*(1), 19–35. https://doi.org/10.1080/14616688.2014.927523

Zurick, D. N. (1992). Adventure travel and sustainable tourism in the peripheral economy of Nepal. *Annals of the Association of American Geographers*, *82*(4), 608–628. https://doi.org/10.1111/j.1467-8306.1992.tb01720.x

PUBLICATIONS

CURRICULUM VITAE

Name: Janika Raun
Date of birth: 01.08.1990
E-mail: janika.raun@ut.ee
Address: Department of Geography, Institute of Ecology and Earth Sciences, University of Tartu, Tartu Ülikool, 46 Vanemuise st, 51003, Tartu, Estonia

Education:

2015–2020	University of Tartu, PhD in Human Geography
2015	University of Twente, ITC | Faculty of Geo-Information Science and Earth Observation, Erasmus traineeship
2012–2015	University of Tartu, MSc in Human Geography and Regional Planning, *cum laude*
2009–2012	University of Tartu, BSc in Geography
2006–2009	Saaremaa Co-Educational Gymnasium, gold medal
1997–2006	Lümanda Primary School

Continuing education at University of Tartu:

2019	Scholarship of Teaching (P2AV.TK.733)
2018	Introduction to Programming (MTAT.TK.012)
2018	Teaching Science and Technology in Higher Education (K1PR.TK.091)
2018	Scholarship of Teaching (P2AV.TK.722)
2018	Moodle-Based Learning-Centred Materials (K1PR.TK.089)
2017	Group Supervision of Students Research Papers (K1PR.TK.075)
2016	Seminars for Supervisors of Students Research Papers (P2AV.TK.816)
2016	About Programming (MTAT.TK.006)
2015	Learning and Teaching in Higher Education (HTHT.TK.144)

Work experience:

2016–…	University of Tartu, Junior Research Fellow in Human Geography
2014–2016	University of Tartu, Department of Geography, specialist
2012–2013	University of Tartu, Department of Geography, assistant

Research interests:
Tourism geography, consumer behaviour, destination functions, tourism impacts, destination marketing, urban mobility, tracking technologies, GIS analysis.

Publications:

Silm, S., Jauhiainen, J. S., **Raun, J.,** Tiru, M. (in press) Temporary population mobilities between Estonia and Finland based on mobile phone data and the emergence of a cross-border region. *European Planning Studies*..

Raun, J., Shoval, N., Tiru, M. (2020). Gateways for intra-national tourism flows: Measured using two types of tracking technologies. *International Journal of Tourism Cities*.

Saluveer, E., **Raun, J.,** Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A., Silm, S. (2020). Methodological framework for producing national tourism statistics from mobile positioning data. *Annals of Tourism Research*, 81, 102895.

Raun, J., Aasa, A., Saluveer, E., Tiru, M., Silm, S. (2019). Measuring tourism destinations using mobile positioning data. In Pae, T., Mander, Ü., (Eds.), Publicationes Instituti Geographici Universitatis Tartuensis. Uurimusi eesti-keelse geograafia 100. Aastapäeval. (pp.186−203). Tartu: University of Tartu Press.

Raun, J., Ahas, R., Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202–212.

Other scientific activities and awards:

2019–2021	University of Tartu Scholarship of Teaching and Learning 'Hea õpetamise grant'
2015	III prize in the national student research competition for MSc thesis
2012–…	Member of the organising committee of Mobile Tartu conference

Supervised master's theses:

Kristjan Erik Loik, 2020, (sup) Janika Raun, Välisturistide korduvkülastused Eestis (Foreign repeat visitors in Estonia), University of Tartu.

Merli Ilves, 2020, (sup) Janika Raun, Suvekodude külastamine ja seda mõjutavad tegurid Eestis (Mobility patterns to second homes and its influencing factors in Estonia), University of Tartu.

Helen Kruut, 2016, (sup) Heli Tooman; Janika Raun, Raviturismi arendusvõimalused Saaremaa näitel (Developing Medical Tourism in Saaremaa), University of Tartu.

Conference presentations:

Raun, J. (2020, January). *Measuring tourism destinations with tracking data*. Oral presentation at the The 27th Annual International eTourism Conference (ENTER) PhD Workshop, Surrey, Great Britain.

Saluveer, E., Raun, J., Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A., Silm, S. (2019, September). *Generating national tourism statistics from mobile positioning data.* Oral presentation at the NECTAR Cluster 5 Conference 'Sustainable Tourism in the Digital World', Visby, Sweden.

Raun, J., Shoval, N. (2019, June) *The role of major gateways on national tourism flows.* Oral presentation at the 15th biannual NECTAR conference Towards Human Scale C ities – Open and Happy, Helsinki, Finland.

Raun, J., Shoval, N. (2019, April). *The role of major gateways on national tourism flows: A tale of two countries.* Oral presentation at the Association of American Geographers (AAG) Annual Meeting, Washington, US.

Raun, J. (2018, May). *The role of a major gateway in national tourism.* Oral presentation at Workshop for invited participants on: Challenges for European City Tourism, Jerusalem, Israel.

Raun, J., Ahas, R. (2016, November). *Defining usual environment with mobile positioning data.* Oral presentation at Global Tourism Forum, Venice, Italy.

Raun, J. (2016, September). *Measuring tourism destinations using mobile tracking data.* Oral presentation at NECTAR Cluster 5 Leisure, Recreation and Tourism Launch workshop 'Tourism and Transport: exploration of inter-dependencies', Lugano, Switzerland.

Raun, J. (2016, March). *Analysing the Influence of Weather Conditions on Every-day Activities in Estonia.* Oral presentation at the Association of American Geographers (AAG) Annual Meeting, San Francisco, US.

Raun, J., Ahas, R., Tiru, M. (2015, December). *Defining usual environment with mobile positioning data*. Oral presentation at Consumer Behaviour in Tourism Symposium (CBTS), Munich, Germany.

Raun, J., Ahas, R., Tiru, M. (2015, June). *Defining usual environment with mobile positioning data.* Oral presentation at Nordic Geographers Meeting (NGM), Tallinn, Estonia.

Raun, J. (2015, April). *Defining usual environment with mobile positioning data.* Oral presentation at the Association of American Geographers (AAG) Annual Meeting, Chicago, USA.

Raun, J., Ahas, R. (2014, December). *Distinguishing tourism destinations with behavioural data.* Oral presentation at conference Consumer Behaviour in Tourism Symposium (CBTS), Brunico; Italy.

Raun, J., Ahas, R. (2014, April). *Distinguishing Estonian tourism destinations with space-time tracking data.* Oral presentation at the Association of American Geographers (AAG) Annual Meeting, Tampa, US.

Raun, J., Ahas, R. (2013, October). *Distinguishing tourism destinations with behavioural data.* Oral presentation at the conference Mobile Ghent, Ghent, Belgium.

Raun, J., Ahas, R. (2013, March). *Distinguishing tourism destinations with behavioural data.* Oral presentation at the annual conference New Techniques and Technologies for Statistics (NTTS), Brussels, Belgium.

ELULOOKIRJELDUS

Nimi: Janika Raun
Sünniaeg: 01.08.1990
E-post: janika.raun@ut.ee
Aadress: Tartu Ülikool geograafia osakond, Vanemuise 46, Tartu, 51003, Eesti

Haridus:

2015–2020	Tartu Ülikool, PhD inimgeograafia erialal
2015	Twente Ülikool, Erasmuse praktika
2012–2015	Tartu Ülikool, MSc geograafias inimgeograafia ja regionaalplaneerimise suunal, *cum laude*
2009–2012	Tartu Ülikool, BSc geograafias
2006–2009	Saaremaa Ühisgümnaasium, kuldmedal
1997–2006	Lümanda Põhikool

Täiendõpe Tartu Ülikoolis:

2019	Grandiakadeemia 2019: oma õpetamise uurimine: ideed ja meetodid (P2AV.TK.733)
2018	Programmeerimise alused (MTAT.TK.012)
2018	Loodus- ja täppisteaduste õpetamine ülikoolis (K1PR.TK.091)
2018	Ettevalmistus oma õpetamise arendamiseks ja uurimiseks (P2AV.TK.722)
2018	Õppimise toetamine õppematerjalidega Moodle'is (K1PR.TK.089)
2017	Üliõpilaste uurimistööde rühmajuhendamine (K1PR.TK.075)
2016	Üliõpilaste uurimistööde juhendajate koolitus (P2AV.TK.816)
2016	Programmeerimisest maalähedaselt (MTAT.TK.006)
2015	Õppimine ja õpetamine kõrgkoolis (HTHT.TK.144)

Töökogemus:

2016–……	Tartu Ülikool, inimgeograafia nooremteadur
2014–2016	Tartu Ülikool, geograafia osakond, spetsialist
2012–2013	Tartu Ülikool, geograafia osakond, assistent

Uurimisvaldkonnad:
Turismigeograafia, turistide liikumine sihtkohas, sihtkoha funktsioonid, turismi mõjud, tegevusruumiuuringud, mobiilpositsioneerimise ja GPS uuringute metoodika arendamine, linnageograafia, mobiilsuslahendused, GIS, kaardid.

Publikatsioonid:

Silm, S., Jauhiainen, J. S., **Raun, J.,** Tiru, M. (trükis) Temporary population mobilities between Estonia and Finland based on mobile phone data and the emergence of a cross-border region. *European Planning Studies*..

Raun, J., Shoval, N., Tiru, M. (2020). Gateways for intra-national tourism flows: Measured using two types of tracking technologies. *International Journal of Tourism Cities*.

Saluveer, E., **Raun, J.,** Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A., Silm, S. (2020). Methodological framework for producing national tourism statistics from mobile positioning data. *Annals of Tourism Research*, 81, 102895.

Raun, J., Aasa, A., Saluveer, E., Tiru, M., Silm, S. (2019). Turismisihtkoha ajalise ja ruumilise ulatuse mõõtmine mobiilpositsioneerimise andmetega. Pae, T., Mander, Ü., (Toim.), Publicationes Instituti Geographici Universitatis Tartuensis. Uurimusi eestikeelse geograafia 100. aastapäeval. (186−203). Tartu: Tartu Ülikooli kirjastus.

Raun, J., Ahas, R., Tiru, M. (2016). Measuring tourism destinations using mobile tracking data. *Tourism Management*, 57, 202–212.

Muu teadustegevus ja preemiad:

2019–2021	Tartu Ülikooli hea õpetamise grant
2015	3. preemia üliõpilaste teadustööde riiklikul konkursil ühiskonna-teaduste ja kultuuri valdkonnas magistriõppe tasemel
2012–…	Rahvusvahelise konverentsi Mobile Tartu korraldustoimkonna liige

Juhendatud magistritööd:

Kristjan Erik Loik, 2020, (juh) Janika Raun, Välisturistide korduvkülastused Eestis, Tartu Ülikool.

Merli Ilves, 2020, (juh) Janika Raun, Suvekodude külastamine ja seda mõjutavad tegurid Eestis, Tartu Ülikool.

Helen Kruut, 2016, (juh) Heli Tooman; Janika Raun, Raviturismi arendusvõimalused Saaremaa näitel, Tartu Ülikool.

Suulised ettekanded konverentsidel:

Raun, J. (2020, jaanuar). *Measuring tourism destinations with tracking data*. 27th Annual International eTourism Conference (ENTER) PhD Workshop, Surrey, Suurbritannia.

Saluveer, E., Raun, J., Tiru, M., Altin, L., Kroon, J., Snitsarenko, T., Aasa, A., Silm, S. (2019, september). *Generating national tourism statistics from mobile positioning data.* NECTAR Cluster 5 Conference "Sustainable Tourism in the Digital World", Visby, Rootsi.

Raun, J., Shoval, N. (2019, juuni) *The role of major gateways on national tourism flows*. 15th biannual NECTAR conference Towards Human Scale C ities – Open and Happy, Helsingi, Soome.

Raun, J., Shoval, N. (2019, aprill). *The role of major gateways on national tourism flows: A tale of two countries.* Association of American Geographers (AAG) Annual Meeting, Washington, Ameerika Ühendriigid.

Raun, J. (2018, mai). *The role of a major gateway in national tourism.* Workshop for invited participants on: Challenges for European City Tourism, Jeruusalem, Iisrael.

Raun, J., Ahas, R. (2016, november). *Defining usual environment with mobile positioning data.* Global Tourism Forum, Venetsia, Itaalia.

Raun, J. (2016, september). *Measuring tourism destinations using mobile tracking data.* NECTAR Cluster 5 Leisure, Recreation and Tourism Launch workshop "Tourism and Transport: exploration of interdependencies", Lugano, Šveits.

Raun, J. (2016, märts). *Analysing the Influence of Weather Conditions on Everyday Activities in Estonia.* Association of American Geographers (AAG) Annual Meeting, San Francisco, Ameerika Ühendriigid.

Raun, J., Ahas, R., Tiru, M. (2015, detsember). *Defining usual environment with mobile positioning data*. Consumer Behaviour in Tourism Symposium (CBTS), München, Saksamaa.

Raun, J., Ahas, R., Tiru, M. (2015, juuni). *Defining usual environment with mobile positioning data.* Nordic Geographers Meeting (NGM), Tallinn, Eesti.

Raun, J. (2015, aprill). *Defining usual environment with mobile positioning data.* Association of American Geographers (AAG) Annual Meeting, Chicago, Ameerika Ühendriigid.

Raun, J., Ahas, R. (2014, detsember). *Distinguishing tourism destinations with behavioural data.* Consumer Behaviour in Tourism Symposium (CBTS), Brunico; Itaalia.

Raun, J., Ahas, R. (2014, aprill). *Distinguishing Estonian tourism destinations with space-time tracking data.* Association of American Geographers (AAG) Annual Meeting, Tampa, Ameerika Ühendriigid.

Raun, J., Ahas, R. (2013, oktoober). *Distinguishing tourism destinations with behavioural data.* Mobile Ghent, Gent, Belgia.

Raun, J., Ahas, R. (2013, märts). *Distinguishing tourism destinations with behavioural data.* New Techniques and Technologies for Statistics (NTTS), Brüssel, Belgia.

DISSERTATIONES GEOGRAPHICAE UNIVERSITATIS TARTUENSIS

1. **Вийви Руссак.** Солнечная радиация в Тыравере. Тарту, 1991.
2. **Urmas Peterson.** Studies on Reflectance Factor Dynamics of Forest Communities in Estonia. Tartu, 1993.
3. **Ülo Suursaar.** Soome lahe avaosa ja Eesti rannikumere vee kvaliteedi analüüs. Tartu, 1993.
4. **Kiira Aaviksoo.** Application of Markov Models in Investigation of Vegetation and Land Use Dynamics in Estonian Mire Landscapes. Tartu, 1993.
5. **Kjell Weppling.** On the assessment of feasible liming strategies for acid sulphate waters in Finland. Tartu, 1997.
6. **Hannes Palang.** Landscape changes in Estonia: the past and the future. Tartu, 1998.
7. **Eiki Berg.** Estonia's northeastern periphery in politics: socio-economic and ethnic dimensions. Tartu, 1999.
8. **Valdo Kuusemets.** Nitrogen and phosphorus transformation in riparian buffer zones of agricultural landscapes in Estonia. Tartu, 1999.
9. **Kalev Sepp.** The methodology and applications of agricultural landscape monitoring in Estonia. Tartu, 1999.
10. **Rein Ahas.** Spatial and temporal variability of phenological phases in Estonia. Tartu, 1999.
11. **Эркки Таммиксаар.** Географические аспекты творчества Карла Бэра в 1830–1840 гг. Тарту, 2000.
12. **Garri Raagmaa.** Regional identity and public leaders in regional economic development. Tartu, 2000.
13. **Tiit Tammaru.** Linnastumine ja linnade kasv Eestis nõukogude aastatel. Tartu, 2001.
14. **Tõnu Mauring.** Wastewater treatment weltlands in Estonia: efficiency and landscape analysis. Tartu, 2001.
15. **Ain Kull.** Impact of weather and climatic fluctuations on nutrient flows in rural catchments. Tartu, 2001.
16. **Robert Szava-Kovats.** Assessment of stream sediment contamination by median sum of weighted residuals regression. Tartu, 2001.
17. **Heno Sarv.** Indigenous Europeans east of Moscow. Population and Migration Patterns of the Largest Finno-Ugrian Peoples in Russia from the 18th to the 20th Centuries. Tartu, 2002.
18. **Mart Külvik.** Ecological networks in Estonia — concepts and applications. Tartu, 2002.
19. **Arvo Järvet.** Influence of hydrological factors and human impact on the ecological state of shallow Lake Võrtsjärv in Estonia. Tartu, 2004.
20. **Katrin Pajuste.** Deposition and transformation of air pollutants in coniferous forests. Tartu, 2004.

21. **Helen Sooväli.** *Saaremaa waltz*. Landscape imagery of Saaremaa Island in the 20th century. Tartu, 2004.
22. **Antti Roose.** Optimisation of environmental monitoring network by integrated modelling strategy with geographic information system — an Estonian case. Tartu, 2005.
23. **Anto Aasa.** Changes in phenological time series in Estonia and Central and Eastern Europe 1951–1998. Relationships with air temperature and atmospheric circulation. Tartu, 2005.
24. **Anneli Palo.** Relationships between landscape factors and vegetation site types: case study from Saare county, Estonia. Tartu, 2005.
25. **Mait Sepp.** Influence of atmospheric circulation on environmental variables in Estonia. Tartu, 2005.
26. **Helen Alumäe.** Landscape preferences of local people: considerations for landscape planning in rural areas of Estonia. Tartu, 2006.
27. **Aarne Luud.** Evaluation of moose habitats and forest reclamation in Estonian oil shale mining areas. Tartu, 2006.
28. **Taavi Pae.** Formation of cultural traits in Estonia resulting from historical administrative division. Tartu, 2006.
29. **Anneli Kährik.** Socio-spatial residential segregation in post-socialist cities: the case of Tallinn, Estonia. Tartu, 2006.
30. **Dago Antov.** Road user perception towards road safety in Estonia. Tartu, 2006.
31. **Üllas Ehrlich.** Ecological economics as a tool for resource based nature conservation management in Estonia. Tartu, 2007.
32. **Evelyn Uuemaa.** Indicatory value of landscape metrics for river water quality and landscape pattern. Tartu, 2007.
33. **Raivo Aunap.** The applicability of gis data in detecting and representing changes in landscape: three case studies in Estonia. Tartu, 2007.
34. **Kai Treier.** Trends of air pollutants in precipitation at Estonian monitoring stations. Tartu, 2008.
35. **Kadri Leetmaa.** Residential suburbanisation in the Tallinn metropolitan area. Tartu, 2008.
36. **Mare Remm.** Geographic aspects of enterobiasis in Estonia. Tartu, 2009.
37. **Alar Teemusk.** Temperature and water regime, and runoff water quality of planted roofs. Tartu, 2009.
38. **Kai Kimmel.** Ecosystem services of Estonian wetlands. Tartu, 2009.
39. **Merje Lesta.** Evaluation of regulation functions of rural landscapes for the optimal siting of treatment wetlands and mitigation of greenhouse gas emissions. Tartu, 2009.
40. **Siiri Silm.** The seasonality of social phenomena in Estonia: the location of the population, alcohol consumption and births. Tartu, 2009.
41. **Ene Indermitte.** Exposure to fluorides in drinking water and dental fluorosis risk among the population of Estonia. Tartu, 2010.

42. **Kaido Soosaar.** Greenhouse gas fluxes in rural landscapes of Estonia. Tartu, 2010.
43. **Jaan Pärn.** Landscape factors in material transport from rural catchments in Estonia. Tartu, 2010.
44. **Triin Saue.** Simulated potato crop yield as an indicator of climate variability in Estonia. Tartu, 2011.
45. **Katrin Rosenvald.** Factors affecting EcM roots and rhizosphere in silver birch stands. Tartu, 2011.
46. **Ülle Marksoo.** Long-term unemployment and its regional disparities in Estonia. Tartu, 2011, 163 p.
47. **Hando Hain.** The role of voluntary certification in promoting sustainable natural resource use in transitional economies. Tartu, 2012, 180 p.
48. **Jüri-Ott Salm.** Emission of greenhouse gases CO_2, CH_4, and N_2O from Estonian transitional fens and ombrotrophic bogs: the impact of different land-use practices. Tartu, 2012, 125 p.
49. **Valentina Sagris.** Land Parcel Identification System conceptual model: development of geoinfo community conceptual model. Tartu, 2013, 161 p.
50. **Kristina Sohar.** Oak dendrochronology and climatic signal in Finland and the Baltic States. Tartu, 2013, 129 p.
51. **Riho Marja.** The relationships between farmland birds, land use and landscape structure in Northern Europe. Tartu, 2013, 134 p.
52. **Olle Järv.** Mobile phone based data in human travel behaviour studies: New insights from a longitudinal perspective. Tartu, 2013, 168 p.
53. **Sven-Erik Enno.** Thunderstorm and lightning climatology in the Baltic countries and in northern Europe. Tartu, 2014, 142 p.
54. **Kaupo Mändla.** Southern cyclones in northern Europe and their influence on climate variability. Tartu, 2014, 142 p.
55. **Riina Vaht.** The impact of oil shale mine water on hydrological pathways and regime in northeast Estonia. Tartu, 2014, 111 p.
56. **Jaanus Veemaa.** Reconsidering geography and power: policy ensembles, spatial knowledge, and the quest for consistent imagination. Tartu, 2014, 163 p.
57. **Kristi Anniste.** East-West migration in Europe: The case of Estonia after regaining independence. Tartu, 2014, 151 p.
58. **Piret Pungas-Kohv.** Between maintaining and sustaining heritage in landscape: The examples of Estonian mires and village swings. Tartu, 2015, 210 p.
59. **Mart Reimann.** Formation and assessment of landscape recreational values. Tartu, 2015, 127 p.
60. **Järvi Järveoja.** Fluxes of the greenhouse gases CO_2, CH_4 and N_2O from abandoned peat extraction areas: Impact of bioenergy crop cultivation and peatland restoration. Tartu, 2015, 171 p.

61. **Raili Torga.** The effects of elevated humidity, extreme weather conditions and clear-cut on greenhouse gas emissions in fast growing deciduous forests. Tartu, 2016, 128 p.
62. **Mari Nuga.** Soviet-era summerhouses On homes and planning in post-socialist suburbia. Tartu, 2016, 179 p.
63. **Age Poom.** Spatial aspects of the environmental load of consumption and mobility. Tartu, 2017, 141 p.
64. **Merle Muru.** GIS-based palaeogeographical reconstructions of the Baltic Sea shores in Estonia and adjoining areas during the Stone Age. Tartu, 2017, 132 p.
65. **Ülle Napa.** Heavy metals in Estonian coniferous forests. Tartu, 2017, 129 p.
66. **Liisi Jakobson.** Mutual effects of wind speed, air temperature and sea ice concentration in the Arctic and their teleconnections with climate variability in the eastern Baltic Sea region. Tartu, 2018, 118 p.
67. **Tanel Tamm.** Use of local statistics in remote sensing of grasslands and forests. Tartu, 2018, 106 p.
68. **Enel Pungas.** Differences in Migration Intentions by Ethnicity and Education: The Case of Estonia. Tartu, 2018, 142 p.
69. **Kadi Mägi.** Ethnic residential segregation and integration of the Russian-speaking population in Estonia. Tartu, 2018, 173 p.
70. **Kiira Mõisja.** Thematic accuracy and completeness of topographic maps. Tartu, 2018, 112 p.
71. **Kristiina Kukk.** Understanding the vicious circle of segregation: The role of leisure time activities. Tartu, 2019, 143 p.
72. **Kaie Kriiska.** Variation in annual carbon fluxes affecting the soil organic carbon pool and the dynamics of decomposition in hemiboreal coniferous forests. Tartu, 2019, 146 p.
73. **Pille Metspalu.** The changing role of the planner. Implications of creative pragmatism in Estonian spatial planning. Tartu, 2019, 128 p.