MyDynamicForest: citizen data on spatial patterns and motives of recreational use in Helsinki's Central Park

Silviya Korpilo	Tarmo Virtanen	Susanna Lehvävirta
Department of Environmental Sciences	Department of Environmental Sciences	Department of Environmental Sciences
University of Helsinki, Finland	University of Helsinki, Finland	University of Helsinki, Finland
silviya.korpilo@helsinki.fi	tarmo.virtanen@helsinki.fi	susanna.lehvavirta@helsinki.fi

Abstract

In this paper we present a method for gathering up-to-date participatory data on soft mobility for various purposes, using a case study from Central Park in Helsinki, Finland. We describe the use of an innovative Public Participation GIS tool (MyDynamicForest) that combines smartphone GPS tracking, drawing of routes, and a questionnaire for collecting citizen data on recreational use for adaptive planning and management purposes. Our main finding is that by applying this method, together with information campaigns, informative data can be generated with relatively low effort, and postulate that specific groups could be targeted when needed. While this study focuses on movement in urban green areas, we suggest that a variety of planning and maintenance challenges could be addressed using this kind of data. Patterns and drivers of soft mobility (e.g. commuting) in cities or visitor movement in national parks and protected areas could be equally targeted. We encourage the use and testing of this methodology in various user-centred research and planning approaches.

Keywords: PPGIS, VGI, human movement, urban planning, smartphone GPS tracking.

1 Introduction

Understanding how citizens use urban space is crucial to user-centred planning of our cities as liveable places. One key element in liveability is movement, especially soft (i.e. non-motorized) mobility through the urban environment, the promotion of which has become a target in many modern cities. The increasing popularity of Volunteered Geographic Information (VGI) and Public Participation GIS (PPGIS) tools presents new opportunities for research on spatial behaviour as citizens become important actors in the co-production and use of geographic information (Feick & Roche, 2013; Goodchild, 2007). In this paper we describe a methodology combining VGI and PPGIS for gathering up-to-date data on human movement, using recreational use in urban forests as a showcase.

Recreational use consists of complex behavioural and spatial patterns that can vary on an individual and group level, and change over time (Arnberger, 2006; Wolf, Hagenloh, & Croft, 2012). Spatial technologies such as Geographic Information Systems (GIS) and Global Positioning Systems (GPS) have provided useful ways to gather accurate, detailed and timely data on visitor behaviour for a variety of natural resource applications (e.g. Beeco, Hallo, & Brownlee, 2014; D'Antonio et al., 2010; Meijles, de Bakker, Groote, & Barske, 2014). Moreover, understanding the spatial and social aspects of visitor use is essential to planners and managers in order to balance between high demand for quality nature experiences and ecological preservation (Cole & Daniel, 2003; Orellana, Bregt, Ligtenberg, & Wachowicz, 2012).

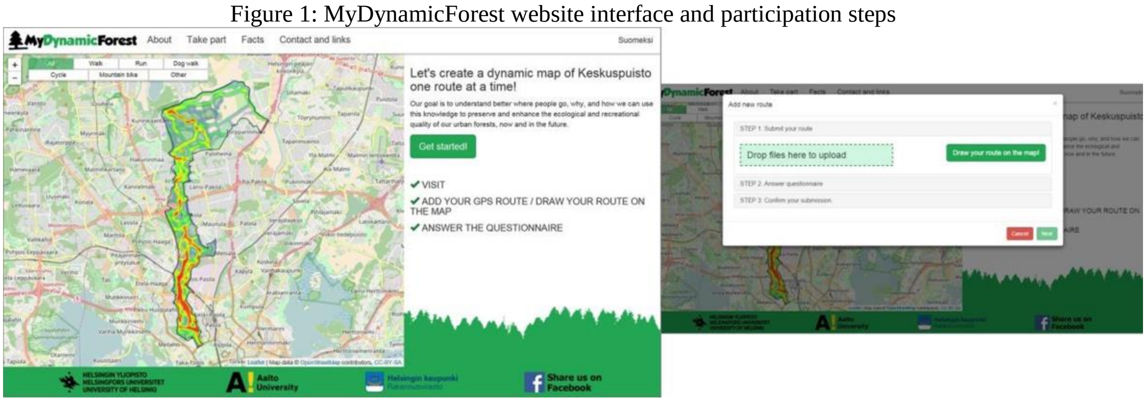
Here we present experiences in data generation using a web-based PPGIS tool called "MyDynamicForest" (MDF) that contains georeferenced movement information and a questionnaire. We introduced the tool in Central Park (Keskuspuisto), Helsinki, Finland, with the aim to collect and analyse data on visitor spatial behaviour i.e. spatial patterns and motives of movement, for planning and management purposes.

2 Description of MyDynamicForest tool

The MDF tool was pilot launched in the summer of 2015 as a collaborative effort between researchers (University of Helsinki, Aalto University) and city officials (Public Works Department, City of Helsinki). MDF provides a virtual space for citizens to participate in an easy and interactive manner by submitting a GPS-tracked or self-drawn route, and by answering an online questionnaire (Fig.1). The MDF study involved three consecutive steps of participation:

1) Movement data from GPS and drawn tracks. GPS tracks could be uploaded to the MDF online database from any smartphone GPS tracking source. Study participants were asked to share routes that they had already tracked voluntarily for personal reasons using any kind of sports tracking application on their smartphones. To decrease bias towards sports tracking enthusiasts, participants could alternatively submit a route by drawing it in the MDF website over an Open Street Map base. Multiple track submissions were allowed and each track was classified by participants according to their primary activity e.g. running, dog walking, cycling etc. The MDF tool created dynamic heat maps that portrayed movement patterns of different activities. The heat maps were automatically updated every time a route was added, which provided instant and evolving visualizations of the spatial data.

Figure 1: MyDynamicForest website interface and participation steps



Source: www.mydynamicforest.fi

2) Visitor and activity profile: Each route submission was followed by a short questionnaire consisting of 12 questions related to the socio-cultural background of participants (e.g. age, gender, education), their activities (type and frequency), and route choice motivations. At the end of the questionnaire, respondents could provide free-form comments e.g. recommendations for improving current and future forest management practices. The questionnaire was designed to offer complementary social information to the digital tracks.

3) Principles of informed consent: To complete the process, all participants were asked to sign a letter of "Consent to Participate in Research" providing clear terms and conditions of voluntary participation. Since there may be various ethical issues related to using smartphone tracking data (Meijles et al., 2014; Taczanowska, Muhar, & Brandenburg, 2008), privacy protection of human subjects in this study was carefully addressed according to the National Advisory Board on Research Ethics in Finland. All personal identifying information was processed so as to guarantee confidentiality and anonymity. In addition, to avoid possible tracing of citizens to their home or work location, the route data was cut so that only intra-site tracks were used in the analysis (Korpilo, Virtanen, & Lehvävirta, 2017).

3 Data collection and visualization

The site for piloting MDF, Central Park, is an urban forest and the largest single green area in Helsinki, the capital of Finland. It includes approximately 100 km of formal trails and stretches over 1100 ha of land, 64% of which is covered by mature forest (City of Helsinki Urban Facts, 2005). Central Park is intensively used for a variety of recreational activities, as well as for commuting.

In total, 366 tracks (139 GPS and 227 draw tracks) and 340 questionnaire responses from 233 participants were gathered during the six-month data collection period from June to December 2015. The study was advertised widely through radio, local newspapers and social media (Facebook, the website and Twitter page of City of Helsinki) from the beginning of July. The highest level of participation was recorded in July (58% of all data submitted), while 76% of submissions during that month was collected immediately (1-3 days) after the newspapers and radio coverage.

In general, 38% of participants submitted GPS tracks and the rest used the draw-on-the map tool. The GPS routes were tracked by 12 different sports tracking applications (e.g. Sports Tracker, Strava, HeiaHeia, Garmin, Endomondo, Runkeeper, Cycle Tracks) that are generally free to use and have different functionality and target users.

Some immediate differences in background characteristics became visible among GPS and draw users (Table 1). The GPS dataset included mostly mountain biking, cycling and running tracks, and was biased towards middle-aged men as 72% of all the GPS contributors were male and 46% were in the 35-44 age group. On the other hand, the draw data portrayed relatively more equal gender, age and activity distribution compared to the overall dataset, but was under-represented for mountain biking (Table 1).

Table 1: Participant background in relation to the two types of movement data (GPS and drawn tracks) received from Helsinki's Central Park users.

	GPS data (n=68)	**Draw data (n=165)**	**Overall questionnaire data (n=233)**
Gender			
Male	72.1%	46.7%	54.1%
Female	27.9	53.3	45.9
Age group			
18-24	0%	5.6%	4.0%
25-34	32.3	35.0	34.1
35-44	46.2	25.6	31.4
45-54	16.9	20.6	19.5
55-74	4.6	13.1	11.1
Education			
High school or less	10.3%	10.9%	10.8%
Professional/technical degree	29.3	29.7	29.6
Bachelor's degree	13.8	16.4	15.6
Master's degree or higher	46.6	43.0	44.1
Activity tracks	**(n=139)**	**(n=227)**	**(n=340)**
Mountain biking	21.6%	2.6%	6.5%
Running	40.3	29.1	32.4
Cycling	30.9	31.7	33.8
Walking	4.3	19.8	15.0
Dog walking	2.9	16.7	12.0

Initial cleaning of the data was conducted by deleting identical GPS points at the same location fix (due to a pause in the movement) and removing drawn tracks that were visually considered too coarse to be informative (7/227 were removed). To have an overall estimation of the spatial accuracy of the data, first we calculated the average deviation of tracks from the formal trail network. Proximity analyses in ArcGIS were conducted with both datasets in order to calculate the average distance of all on-trail tracks (from participants who stated to have followed only the formal trails) to the formal trail line features within a search radius of 20 m (Korpilo et al., 2017). The trail network was acquired from the topographic database of National Land Survey of Finland (scale 1:10 000). On-trail GPS data was sufficient only for runners and cyclists, while calculations for the draw data were performed for the entire on-trail dataset (drawn routes portray only perceived use that could not be affected by GPS signal errors or activity speed). The location accuracy of the data was satisfactory, ranging from 4 to 6 m (Table 2).

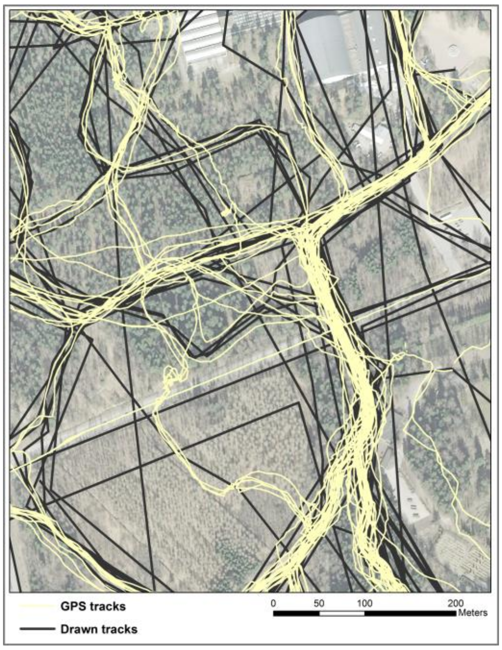
Table 2: Estimation of location accuracy of on-trail tracks based on point data for runners and cyclists, and line data for drawn tracks. On-trail refers to tracks reported by respondents as containing only on-trail movement. Distances refer to distance from the formal trail network.

On-trail tracks	Mean distance (m)	Standard deviation of distance (m)	Observations within mean distance
Running GPS tracks (n=43)	5.60	4.46	59.38%
Cycling GPS tracks (n=43)	5.28	4.60	60.96%
Drawn tracks (n=69)	3.95	5.43	65.36%

Then, the spatial variation between and within the two datasets was explored. Visually, the GPS and drawn tracks differed substantially in veracity and level of detail (Fig.2), however, some similarities were also observed. Figure 3 shows a runner's drawn route that follows similar spatial trajectory as a mountain biking GPS track. These tracks illustrate similarities between visitor-reported and actual spatial behaviour since both were reported and observed as containing off-trail movement. Yet, Figure 4. also demonstrates inconsistencies between reported behaviour and spatial representation with drawing of a walking route.

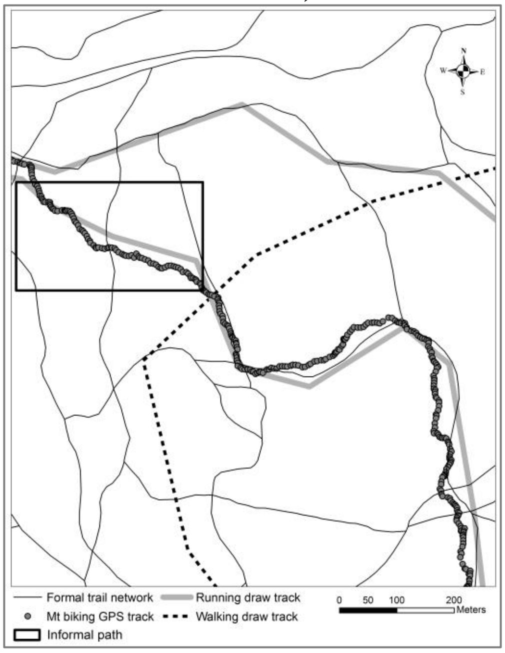
An important step towards better understanding of spatial behaviour is to examine the distribution and intensity of movement at the landscape scale. Kernel Density Analysis (using 10 x 10 raster cell size, 20 m search radius) of the GPS and draw line data displayed similar spatial patterns and provided a clear visual representation of the highest intensity of use (Fig.3). The spatial overlap between the GPS and draw tracks outside the formal trail network indicates the existence of popular off-trail routes.

Figure 2: Example of spatial variation between the GPS and draw data.



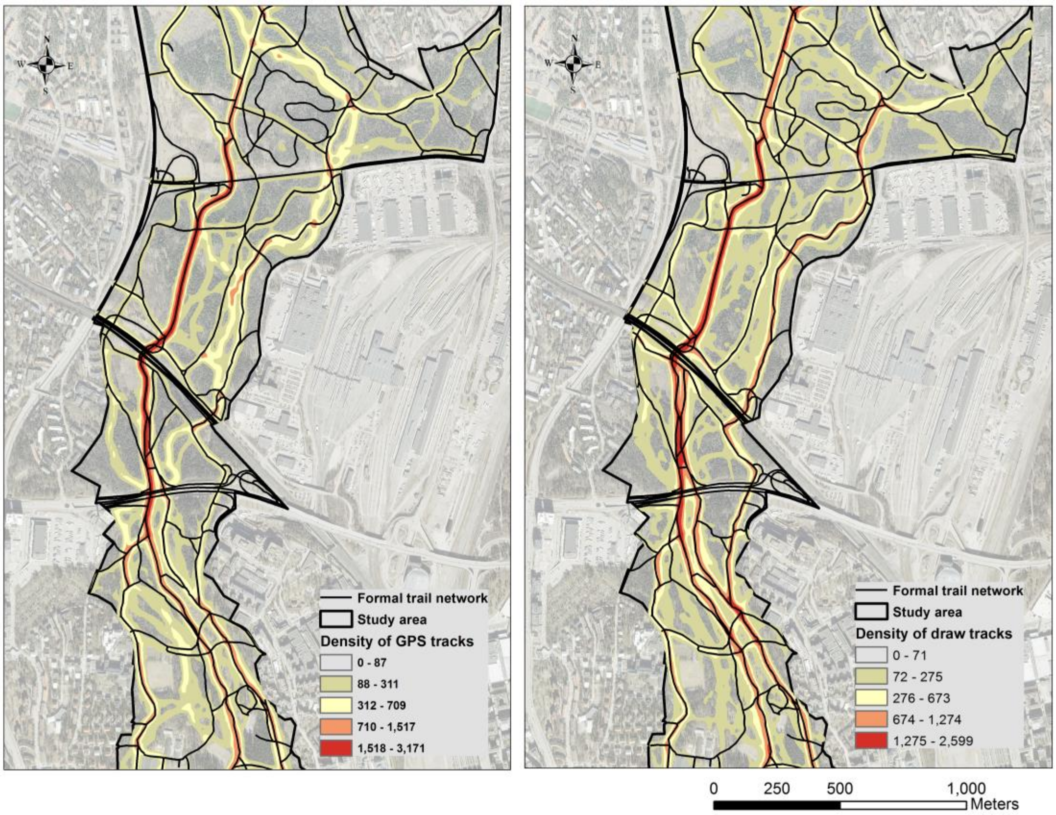
Source: Ortho background aerial image from the National Land Survey of Finland, 2014.

Figure 3: Example of typical GPS and drawn tracks in relation to the formal trail network (National Land Survey of Finland, 2014).



Note: The walking track was drawn as an off-trail route although it was reported by the respondent as an on-trail track.

Figure 4: Kernel density analysis (line density m/km²) of recreational movement based on collected GPS (n=139) and drawn (n=227) tracks.



Source: Ortho background aerial image from the National Land Survey of Finland, 2014.

4 Discussion and future work

In this paper we presented a participatory data collection method combining smartphone GPS tracking, drawing of routes and a questionnaire for gathering accurate and diverse movement information. Adopting a mixed-methods approach allows cross-validation of the data and results. For example, the questionnaire data provided self-reported on-trail movement as a reference for estimating location accuracy of the VGI data. Similar to previous studies that showed 5-10 m GPS positioning accuracy of smartphones (Hess, Farahani, Tschirschnitz, & von Reischach, 2012; Zandbergen, 2009), here the average spatial deviation of the on-trail GPS tracks was found to be up to 6 m. However, as illustrated by the example with a drawn walking track (Fig.3), on-trail claims may not reflect actual spatial behaviour as they are largely affected by visitor recall and the ability to distinguish formal from informal trails while inside the forest (Korpilo et al., 2017). A commonly discussed challenge of VGI, which was also showed here, is that user-generated data exhibits heterogeneity and variation in quality within and between datasets, as well as within individual records (Feick & Roche, 2013; Flanagin & Metzger, 2008). The drawing of routes is generally less accurate and detailed, and it is highly scale-dependent, nevertheless, the density maps demonstrated that it can be used to complement and strengthen the GPS mapping.

Providing multiple modes of participation can also help increase response rates and reduce bias associated with each data collection method (Brown & Reed, 2009). Even though our GPS data was generated by different applications, it portrayed bias towards high-activity level uses (mountain biking, cycling and running) and male users, indicating similar gender and activity-dependent trends noted for sports tracking applications in general (Hirsch et al., 2014; Oksanen, Bergman, Sainio, & Westerholm, 2015). The additional option to draw on the map allowed for collecting data on the spatial behaviour of other user groups. The draw data proved to be valuable as it increased socio-demographic representativeness of the volunteer public - it portrayed more equal gender and age distribution and better captured some of the most popular outdoor activities in Central Park, i.e. walking and dog-walking (Ilvesniemi & Saukkonen, 2015).

In addition, the utilization of a web-based tool helped to reach younger age groups that were under-represented in previous research. In a Central Park visitor survey conducted by city officials in 2007-2009, which combined postal-based (random sample of local residents) and on-site survey techniques (Ilvesniemi & Saukkonen, 2015), the oldest age group was over-represented (53% of respondents were >50 years-old), while the youngest age group was under-represented (17% were <30 years-old). Similar age distribution bias was noted in a large PPGIS study by Kyttä, Broberg, Tzoulas, et al. (2013) that used postal invitations in the city of Helsinki and neighbouring Espoo to encourage participation in a web-questionnaire. By contrast, experiences from this study differed as the younger age group was slightly over-represented (38% of respondents were <34 years-old) and the oldest age group was under-represented (11% were >54 years-old). This agrees with findings by Brown & Kyttä (2014) indicating that under and over-representation in PPGIS research may not be systematic, here even in the same geographical and socio-cultural context. Our results show that the choice of advertising techniques in PPGIS may play an important role in socio-demographic representation.

From a practical perspective, the methodology presented here involves little time and investment costs (e.g. compared to on-site labour-intensive surveys, handling GPS devices or developing a new app), and reduces the burden on participants, researchers and planners on data collection. Movement data is easily understandable and it allows for visualizing and studying patterns of use at various spatial and temporal scales. For example, the density maps portrayed location and intensity of use on formal trails, which can assist managers to determine mismatch between trail infrastructure supply and demand. In addition, the smartphone GPS tracking data captured actual off-trail behaviour and located informal paths, which is crucial information to natural resource management. Density mapping of intensively used off-trail routes can point out areas where heavy use and impacts occur (Korpilo et al., 2017).

We envision the use of similar web platforms as MyDynamicForest for a wide variety of planning purposes involving movement, ranging from commuter traffic planning (e.g. cycling) to leisure uses of urban green spaces. Future-oriented perspectives could be also performed by collecting both actual routes (where people move at present; GPS and drawing) and desired routes (drawing).

While co-production of knowledge can enhance citizen engagement into planning and decision-making, and help better understand the use of space and the factors that affect it, it needs to be highly feasible for all involved. Cost-effective methods are needed, while online services provide smart and easily-adaptable technologies that are not tied to a certain time. Participation websites could be a common routine of e.g. city planning and management departments or wilderness and parks administration. This might involve advertising campaigns to recruit various stakeholders with different levels of participation linked to the problem type, governance context and required knowledge (Hurlbert & Gupta, 2015), and re-campaigns based on analysis of the representativeness of the sample and obtained results.

Acknowledgements

This study has been funded by Maj and Tor Nessling Foundation and University of Helsinki. We are very thankful to Tiina Saukkonen from Public Works Department, City of Helsinki for the continuous collaboration on this project. We would also like to thank Markettä Kyttä and Kamyar Hasanzadeh for their valuable advice on the development and design of the MyDynamicForest tool.

References

Arnberger, A. (2006). Recreation use of urban forests: An inter-area comparison. *Urban Forestry and Urban Greening*, *4*(3-4), 135–144.

Beeco, J. A., Hallo, J. C., & Brownlee, M. T. J. (2014). GPS Visitor Tracking and Recreation Suitability Mapping: Tools for understanding and managing visitor use. *Landscape and Urban Planning*, *127*, 136–145.

Brown, G., & Reed, P. (2009). Public participation gis: A new method for use In national forest planning. *Forest Science*, *55*(2), 166–182.

Brown, G., & Kyttä, M. (2014). Key issues and research priorities for public participation GIS (PPGIS): A synthesis based on empirical research. *Applied Geography*, *46*, 126–136.

City of Helsinki Urban Facts. (2005). *Nature, Recreation and Sport in Helsinki* (Web publications 2005:23). Retreived from http://www.hel.fi/hel2/Tietokeskus/julkaisut/pdf/05_06_27_ranto_vj23.pdf

Cole, D. N., & Daniel, T. C. (2003). The science of visitor management in parks and protected areas: from verbal reports to simulation models. Journal for Nature Conservation, 11(4), 269–277.

D'Antonio, A., Monz, C., Lawson, S., Newman, P., Pettebone, D., & Courtemanch, A. (2010). GPS-Based Measurements of Backcountry Visitors in Parks and Protected Areas : Examples of Methods and Applications from Three Case Studies. *Journal of Park and Recreation Administration*, *28*(3), 42–60.

Feick, R., & Roche, S. (2013). Understanding the Value of VGI. In Sui, D., Elwood, S., & Goodchild, M. (Eds.), *Crowdsourcing Geographic Knowledge: Volunteered Geographic Information (VGI) in Theory and Practice*, Dordrecht: Springer Netherlands, 15-21.

Flanagin, A. J., & Metzger, M. J. (2008). The credibility of volunteered geographic information. *GeoJournal*, *72*(3-4), 137–148.

Goodchild, M. F. (2007). Citizens as sensors: The world of volunteered geography. *GeoJournal*, *69*(4), 211–221.

Hess, B., Farahani, A. Z., Tschirschnitz, F., & von Reischach, F. (2012). Evaluation of fine-granular GPS tracking on smartphones. In Proceedings of the First ACM SIGSPATIAL International Workshop on Mobile Geographic Information Systems, California, USA

Hirsch, J. a, James, P., Robinson, J. R. M., Eastman, K. M., Conley, K. D., Evenson, K. R., & Laden, F. (2014). Using MapMyFitness to place physical activity into neighborhood context. *Frontiers in Public Health*, *2*(March), 1–9.

Hurlbert, M., & Gupta, J. (2015). The split ladder of participation: A diagnostic, strategic, and evaluation

Ilvesniemi, S., & Saukkonen, T. (2015). Keskuspuiston käyttäjätutkimus 2007–2009 (Central Park visitor survey 2007-2009). City of Helsinki: Public Works Department. Retrieved from http://www.hel.fi/www/hkr/fi/palvelut/suunnitelmat/keskuspuisto

Korpilo, S., Virtanen, T., & Lehvävirta, S. (2017). Smartphone GPS tracking: Inexpensive and efficient data collection on recreational movement. *Landscape and Urban Planning*, *157*, 608–617.

Kyttä, M., Broberg, A., Tzoulas, T., & Snabb, K. (2013). Towards contextually sensitive urban densification: Location-based softGIS knowledge revealing perceived residential environmental quality. *Landscape and Urban Planning*, *113*, 30–46.

Meijles, E. W., de Bakker, M., Groote, P. D., & Barske, R. (2014). Analysing hiker movement patterns using GPS data: Implications for park management. *Computers, Environment and Urban Systems*, *47*, 44–57.

Oksanen, J., Bergman, C., Sainio, J., & Westerholm, J. (2015). Methods for deriving and calibrating privacy-preserving heat maps from mobile sports tracking application data. *Jtrg*, *48*, 135–144.

Orellana, D., Bregt, A. K., Ligtenberg, A., & Wachowicz, M. (2012). Exploring visitor movement patterns in natural recreational areas. Tourism Management, 33(3), 672–682.

Taczanowska, K., Muhar, A., & Brandenburg, C. (2008). Potential and limitations of GPS tracking for monitoring spatial and temporal aspects of visitor behaviour in recreational areas. In Management for protection and sustainable development (pp. 451–455).

Wolf, I. D., Hagenloh, G., & Croft, D. B. (2012). Visitor monitoring along roads and hiking trails: How to determine usage levels in tourist sites. *Tourism Management*, *33*(1), 16–28.

Zandbergen, P. A. (2009). Accuracy of iPhone locations: A comparison of assisted GPS, WiFi and cellular positioning. *Transactions in GIS*, *13*(sup. 1), 5–25.