Comparing location data from smartphone and dedicated GPS devices: Implications for population health surveys

Thierry B[1], Stanley K[2], Kestens Y[3], Winters M[4], Fuller D[5]

[1]Département de médecine sociale et préventive, Université de Montréal, benoit.thierry@umontreal.ca
[2]Department of Computer Science, University of Saskatchewan, kstanley@cs.usask.ca
[3]Département de médecine sociale et préventive, Université de Montréal, yan.kestens@umontreal.ca
[4]Faculty of Health Sciences, Simon Fraser University, mwinters@sfu.ca
[5]Department of Community Health and Epidemiology, University of Saskatchewan, daniel.fuller@usask.ca

ABSTRACT

The purpose of this paper is to compare location data from a dedicated Global Positioning Systems (GPS) device with location data from smartphones. Data from the INTerventions, Equity, and Action in Cities Team (INTERACT), a Canada-wide study examining the impacts of urban form changes on health in Victoria, Vancouver, Saskatoon, and Montreal, were used. A total of 337 participants contributed data from the Ethica Data smartphone app and the SenseDoc dedicated GPS. Participants recorded an average total of 14,781 Ethica locations (SD=19,353) and 197,167 SenseDoc locations (SD=111,868). Dynamic time warping was used to examine spatial and temporal similarity of GPS point time sequences. Three activity space measures, a convex hull, a 2-standard deviation ellipse, and 500meter radius circular buffers around points, derived the smartphone and dedicated GPS device were compared. The analysis shows that if the outcome data are available at the day or survey level, there may only be small differences in observed associations between exposure and outcome whether using dedicated GPS or smartphone location data. However, if the exposure and outcome data are summarized and analyzed at the hour or minute frequency, dedicated GPS and smartphone devices are likely to provide different substantially different estimates of exposure.

1. Introduction

The advent of cheap, portable location sensing devices, typically based on Global Positioning System, has revolutionized the study of human mobility (Zhao et al., 2016). The ability to continuously locate an individual has made detailed trajectory analysis possible, improving our capacity to explore how our daily mobility relates to health. The advent of smartphones with embedded GPS capabilities, and later combined GPS-Wifi-Cellular localization provides an even greater potential to collect mobility data in large populations (Zhao et al., 2016).

Mobility and health research that uses GPS location has mainly relied on dedicated GPS receivers (Duncan et al., 2009). Dedicated location loggers have a longer battery life than smartphones and typically provide one-second epoch GPS data (Stopher et al., 2018a). Various factors can affect GPS signal accuracy (Düking et al., 2018; Lee et al., 2020; Ueberham & Schlink, 2018), including

 Volume DOI: 10.17605/OSF.IO/N6DU5

device type, chipset, and environmental conditions. Researchers are increasingly using smartphones to collect location data to study behavior in everyday life (Harari et al., 2017), but their use presents a series of drawbacks including battery life (Hashemian et al., 2012) and participant compliance (Keusch et al., 2022). While research using GPS data has surged, few studies have compared location data obtained from dedicated GPS and smartphone devices, worn by human participants, over longer periods (Adamakis, 2017; Stopher et al., 2018b).
In this paper we compare location data from a dedicated GPS device with location data from smartphones. We further examine similarities between these location measures for common indicators of environmental exposure.

2. Methods & Data

The INTerventions, Equity, Research, and Action in Cities Team (INTERACT) is a Canada-wide study examining the impacts of urban changes on health and equity. Since 2017, it has been following a cohort of participants in four Canadian cities with data collection every 2 years. The study protocol (Kestens et al., 2019) and baseline results (Fuller et al., 2021) have been published. Ethical approval from the ethics boards of Simon Fraser University (2017s0158, 2017s0531, and 2018s0127), the University of Saskatchewan (17-347), and the Centre de Recherche du Centre Hospitalier de l'Université de Montréal (CÉR CHUM 16.397) was obtained.

Participants completed an online survey about their health and well-being, and demographic information. They had the choice to run the Ethica App (*Ethica Data*, 2020), a smartphone application that collects location and accelerometry data, on their personal phone for 30 days. To save battery and reduce the impact on participants, data acquisition occurred for 1 minute every 5 minutes. Some participants who opted into Ethica data collection were also invited to wear a SenseDoc (*The SenseDoc*, 2020), a dedicated device that records GPS and accelerometry data for 10 days. To compare data from the dedicated GPS device and data collected using the smartphone App, we took a two-step approach by (i) measuring the location data similarity and (ii) measuring the characteristics of various activity space measures regularly encountered in the health geography literature (namely convex hulls, deviational ellipses, and buffers).

2.1 Track Similarity

Tracks were compared within subjects and aligned in time. Ethica tracks were filtered to remove all fixes with an accuracy above 100m, as computed by the smartphone locational subsystem, and duplicated timestamps were resolved by keeping the most accurate fix. SenseDoc tracks did not undergo any specific filtering. Finally, all GPS tracks were clipped to the corresponding Census Metropolitan Area (CMA – a Canadian Census spatial unit delineating large, densely populated centers made up of adjacent municipalities that are economically and socially integrated(Statistics Canada, n.d.)). Once cleaned, GPS track similarity of the two data sources was assessed using two metrics: spatial cross-correlation and dynamic time warping.

Spatial cross-correlation is commonly used in image processing to evaluate how much one image resembles another.(Gonzalez et al., 2009) We rasterized the projected GPS locations by binning them into 2D histograms, with varying temporal and spatial scales to evaluate point density patterns similarity. We used temporal epochs ranging from the whole survey (i.e., the longest common period of data from both sensors), day, 6 hours, 1 hour, 15 minutes and down to 5 minutes (which is the smallest temporal resolution for Ethica data) and spatial resolutions of 1km, 500m, 250m, 125m, 50m, 25m and 10m. We computed the pair-wise cross-correlation coefficient of corresponding histograms using function `signal.correlate2d` from the Scipy Python module.

2.2 Activity Space Similarity

To assess aggregate measures of exposure using location data, we computed a number of different activity space measures common in the health geography literature (Smith et al., 2019). For each participant and each data source, we built:

1. a convex hull (the smallest convex polygon encompassing all the GPS points);
2. a 2-standard deviation ellipse (defined by the x and y-axis distributions centered on the mean location of all GPS points);
3. a 500m radius circular buffer around points;
4. a 500m radius buffer along linear-shaped tracks.

We compared the resulting polygons in terms of areal overlap, measured as a percentage (Shareck et al., 2013). The Ethica and SenseDoc comparison metrics inform us about which of the Ethica or SenseDoc activity space includes, possibly only partly, the other.

3. Results

Among the INTERACT participants who contributed locational data using both Ethica App and SenseDocs, 337 of them had at least partially overlapping data collection periods. This cohort subsample comprises a majority of women (61.1%), with participants in the mid-forties (mean age=43.6 years, SD=14.8), highly educated (78.1% have a university degree) and predominantly white (87.5%). The average duration of the common period of data collection totals 8 days, 21 hours and 17 minutes, which is somewhat shorter than the target length of the SenseDoc data collection (10 days).

3.1 Track Similarity – Spatial Cross-Correlation

Spatial cross-correlation (CC) of GPS point density for SenseDoc and Ethica tracks is shown in Figure 1, with subpanels corresponding to aggregation epoch (ranging from 5 minutes to the entire survey period) and the x-axis corresponding to grid cell size (ranging from 10m to 1000m).

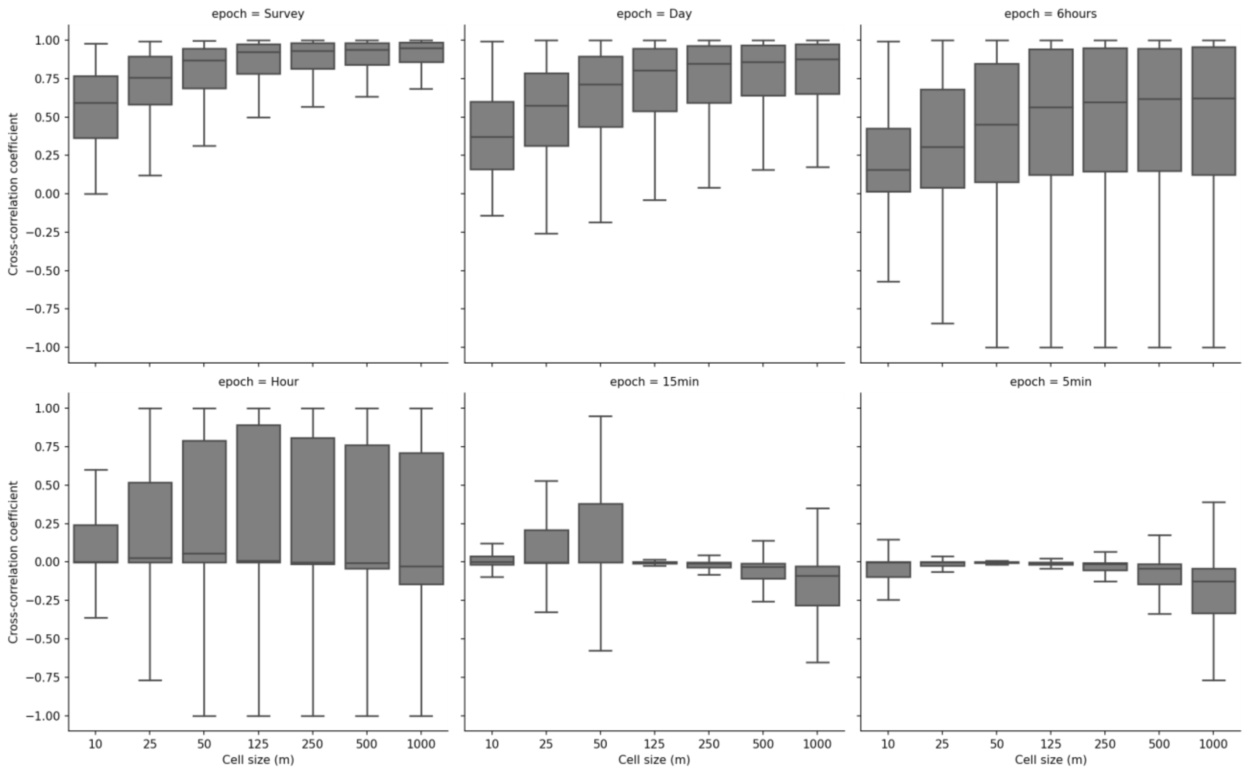

Figure 1. Cross-correlation analysis comparing 2D histograms of each track at various temporal and spatial scales, ranging from 5-minute epoch to the whole survey for the time dimension and from 10m to 1km for the spatial dimension.

Figure 1 shows that Ethica and SenseDoc capture the same kind of information at coarser scales. For the survey and 24 hour aggregation period using a cell size of 125m or greater, cross correlations between the SenseDoc and Ethica app are above 0.75. Cross correlation median values at 125m resolution for survey and day levels are respectively 0.92 and 0.81 and improve as cell size increases. Cross correlation begins to degrade below 125m or on timescales shorter than a day. At the 1 hour, 15 minute, and 5 minute aggregation the cross-correlation between the SenseDoc and Ethica is effectively zero.

3.1 Activity Space Similarity – Bland-Altman Plots

The Bland-Altman plots (Figure 2) show absolute differences in km^2 between the SenseDoc and Ethica activity space metrics. Overall, the SenseDoc based activity spaces are larger than Ethica derived spaces. Across all cities the overall agreement between the activity space measures was 80% for the convex hull, 54% for the 2SD ellipse, 61% for the 500m buffer around fixes, and 75% for the 500 buffer around tracks. The absolute differences between the SenseDoc and Ethica app was -43.7km for the convex hull, -21.7km for the 2SD ellipse, -22.3km for the 500m buffer around fixes, and -12.8km for the 500 buffer around tracks.

 Volume DOI: 10.17605/OSF.IO/N6DU5

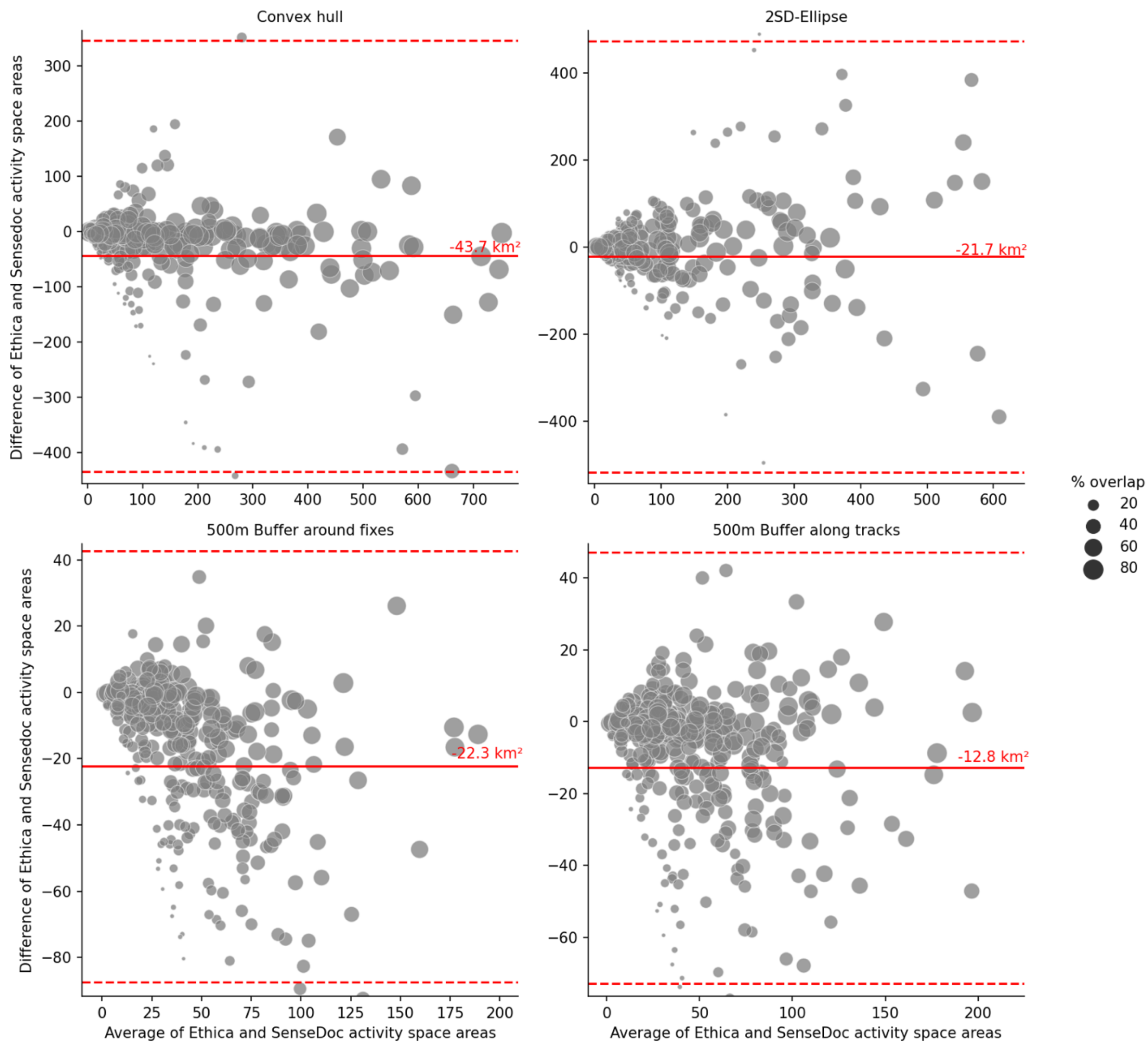

Figure 2. Bland-Altman plots of agreement between activity space areas calculated using Ethica and Sensedoc data using four different activity space metrics, a) convex hull, b) 2 Standard Deviation (2SD)-Ellipse, c) 500m buffer around fixes, d) 500 buffer along tracks.

4. Discussion & Conclusion

The purpose of this paper was to compare location data and activity space metrics obtained from a dedicated GPS device (Sensedoc) and a smartphone (using the Ethica Data app) to understand differences and similarities between the data collection methods and inform future research. We specifically focus our discussion on implications for health research as INTERACT is a health research study. Our results showed that for time periods of 6 hours or more, the point level data captured by dedicated GPS devices and smartphones are similar. Based on our findings, if the health outcome data for a study are available at the day or survey level, there may only be small differences in observed associations between exposure calculated using location data and outcome whether using dedicated GPS or smartphone location data. However, if the exposure and outcome data are available and analyzed frequencies less than the day level, dedicated GPS and smartphone devices are likely to provide different substantially different estimates of exposure.

Our findings are similar to past research (Adamakis, 2017; Stopher et al., 2018b). Participants with dedicated GPS devices had better quality data overall with fewer periods when the device was turned off compared to smartphones (Stopher et al., 2018b). It is not clear from this study if the difference in data quality are due to differences in the smartphone app operation, or based on true differences in participant behaviours. In a study comparing walking and running, total trip distance errors were estimated to be 0.30% vs 3.28% for walking, and 0.74% and 4.43% for running, when comparing the dedicated GPS and smartphone data respectively (Adamakis, 2017). In a study comparing dedicated GPS versus smartphone location data collection for football, results showed no statistically significant differences in total distance travelled during a match between devices (Tierney & Clarke, 2019). In static field tests, smartphones are shown to be slightly more accurate than dedicated GPS devices, however, these tests do not reflect participant use cases of differences in device fix time, battery charging, and participant use (Klimaszewski-Patterson, 2010). As well, in real-world comparisons such ours, it is impossible to have a gold-standard ground truth, as a result, we are limited to assessing how the results may differ between the devices rather than some comparison to a gold-standard.

This paper examined similarities and differences between location data collected using a dedicated GPS device with those collected using a smartphone app. At a day, week, or longer time scale, the location data collected between the devices is highly correlated. However, at time scales of less than a day, the correlation between the data location approaches zero with decreasing time scale. Activity spaces calculated at 10-day time scale show that the dedicated GPS produces larger activity spaces.

Acknowledgments

Funding for INTERACT is provided two Canadian Institutes of Health Research Grants.

- The Team Grant: Environments and Health: Intersectoral Prevention Research / Subvention d'équipe: Santé et environnements: Recherche intersectorielle en prevention. Competition Code: 201607IP5. Funding Reference Number (FRN): IP2-1507071C. $2 million.
- Project Grant Spring 2021. Funding Reference Number (FRN): PJT 178309. $918 000.

Yan Kestens (NPI) and **Benoît Thierry** (Coll) hold shares in Mobysens Technologies Inc., a spin-off company that markets the SenseDoc 2.0. The SenseDoc is a multisensor device used for mobility (GPS) and physical activity (accelerometer) tracking in the INTERACT study. The SenseDoc was filed as an invention in 2013 at Univalor (www.univalor.ca), the valorisation company affiliated with Université de Montréal and Centre de Recherche du CHUM.

Kevin Stanley (Co-I) hold shares in EthicaData Inc., a company marketing the Ethica Health smartphone app that will be used to collect data on mobility and physical activity through participants' smartphones.

References

Adamakis, M. (2017). Comparing the Validity of a GPS Monitor and a Smartphone Application to Measure Physical Activity. *Journal of Mobile Technology in Medicine*, *6*(2), 28–38. https://doi.org/10.7309/jmtm.6.2.4

Düking, P., Fuss, F. K., Holmberg, H. C., & Sperlich, B. (2018). Recommendations for assessment of the reliability, sensitivity, and validity of data provided by wearable sensors designed for monitoring physical activity. *JMIR MHealth and UHealth*, *6*(4), e9341. https://doi.org/10.2196/mhealth.9341

Duncan, M. J., Badland, H. M., & Mummery, W. K. (2009). Applying GPS to enhance understanding of transport-related physical activity. In *Journal of Science and Medicine in Sport* (Vol. 12, Issue 5, pp. 549–556). https://doi.org/10.1016/j.jsams.2008.10.010

Ethica Data. (2020). https://ethicadata.com/

Fuller, D., Bell, S., Firth, C. L., Muhajarine, N., Nelson, T., Stanley, K., Sones, M., Smith, J., Thierry, B., Laberee, K., Stephens, Z. P., Phillips, K., Kestens, Y., & Winters, M. (2021). Wave 1 results of the INTerventions, Research, and Action in Cities Team (INTERACT) cohort study: Examining spatio-temporal measures for urban environments and health. *Health and Place*, 102646. https://doi.org/10.1016/j.healthplace.2021.102646

Gonzalez, R. C., Woods, R. E., & Masters, B. R. (2009). *Digital Image Processing* (4th ed.). Pearson.

Harari, G. M., Müller, S. R., Aung, M. S., & Rentfrow, P. J. (2017). Smartphone sensing methods for studying behavior in everyday life. In *Current Opinion in Behavioral Sciences* (Vol. 18, pp. 83–90). Elsevier Ltd. https://doi.org/10.1016/j.cobeha.2017.07.018

Hashemian, M., Knowles, D., Calver, J., Qian, W., Bullock, M. C., Bell, S., Mandryk, R. L., Osgood, N. D., & Stanley, K. G. (2012). iEpi: An end to end solution for collecting, conditioning and utilizing epidemiologically relevant data. *Proceedings of the International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*. https://doi.org/10.1145/2248341.2248345

Kestens, Y., Winters, M., Fuller, D., Bell, S., Berscheid, J., Brondeel, R., Cantinotti, M., Datta, G., Gauvin, L., Gough, M., Laberee, K., Lewis, P., Lord, S., Luan, H., McKay, H., Morency, C., Muhajarine, N., Nelson, T., Ottoni, C., … Wasfi, R. (2019). INTERACT: A comprehensive approach to assess urban form interventions through natural experiments. *BMC Public Health*, *19*(1), 51. https://doi.org/10.1186/s12889-018-6339-z

Keusch, F., Wenz, A., & Conrad, F. (2022). Do you have your smartphone with you? Behavioral barriers for measuring everyday activities with smartphone sensors. *Computers in Human Behavior*, *127*, 107054. https://doi.org/10.1016/j.chb.2021.107054

Klimaszewski-Patterson, A. (2010). Smartphones in the field: Preliminary study comparing GPS capabilities between a smartphone and dedicated GPS device. *Applied Geography Conferences*, 270–279.

Lee, T., Bettinger, P., Cieszewski, C. J., & Garzon, A. R. G. (2020). The applicability of recreation-grade GNSS receiver (GPS watch, Suunto Ambit Peak 3) in a forested and an open area compared to a mapping-grade receiver (Trimble Juno T41). *PLoS ONE*, *15*(4). https://doi.org/10.1371/journal.pone.0231532

Shareck, M., Kestens, Y., & Gauvin, L. (2013). Examining the spatial congruence between data obtained with a novel activity location questionnaire, continuous GPS tracking, and prompted recall surveys. *International Journal of Health Geographics*, *12*(1), 40. https://doi.org/10.1186/1476-072X-12-40

Smith, L., Foley, L., & Panter, J. (2019). Activity spaces in studies of the environment and physical activity: A review and synthesis of implications for causality. *Health & Place*, *58*, 102113.

Statistics Canada. (n.d.). *Illustrated Glossary—Census metropolitan area (CMA) and census agglomeration (CA)*. Retrieved July 5, 2022, from https://www150.statcan.gc.ca/n1/pub/92-195-x/2016001/geo/cma-rmr/cma-rmr-eng.htm

Stopher, P. R., Daigler, V., & Griffith, S. (2018a). Smartphone app versus GPS Logger: A comparative study. *Transportation Research Procedia*, *32*, 135–145. https://doi.org/10.1016/j.trpro.2018.10.026

Stopher, P. R., Daigler, V., & Griffith, S. (2018b). Smartphone app versus GPS Logger: A comparative study. *Transportation Research Procedia*, *32*, 135–145. https://doi.org/10.1016/j.trpro.2018.10.026

The SenseDoc. (2020). MOBYSENS.

Tierney, P., & Clarke, N. (2019). A Comparison of a Smartphone App with Other GPS Tracking Type Devices Employed in Football. *Exercise Medicine*, *3*, 4. https://doi.org/10.26644/em.2019.004

Ueberham, M., & Schlink, U. (2018). Wearable sensors for multifactorial personal exposure measurements – A ranking study. *Environment International*, *121*, 130–138. https://doi.org/10.1016/j.envint.2018.08.057

Zhao, K., Tarkoma, S., Liu, S., & Vo, H. (2016). Urban human mobility data mining: An overview. *Proceedings - 2016 IEEE International Conference on Big Data, Big Data 2016*, 1911–1920. https://doi.org/10.1109/BigData.2016.7840811