HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Master's thesis in Geography

Geoinformatics

Multi-local living: a comparison between mobile phone and electricity consumption data

Iivari Laaksonen

2022

Supervisors:
Janika Raun
Olle Järv
Tuuli Toivonen

Master's Programme in Geography

Faculty of Science

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Faculty	Department
Faculty of Science	Department of Geosciences and Geography

Author
Iivari Laaksonen
Title of thesis
Multi-local living: a comparison between mobile phone and electricity consumption data
Programme and study track
Master's programme in geography, Geoinformatics

Level of the thesis	Date	Number of pages
Master's thesis, 30 credits	May 2022	59

Abstract

Multi-local living is a complex social phenomenon that is tightly connected to human mobility. In previous research, the phenomenon has been mainly researched with official statistics that fail to capture the dynamic nature of people's mobilities and dwelling. This thesis approaches multi-locality in Finland and in the county South Savo from the perspective of second homes with novel data sources like mobile phone data and electricity consumption data. These spatially and temporally accurate big data sources can be used to ensure sufficient coverage of population and geographic area.

I approach multi-local living by analyzing the spatiotemporal changes in people's presence with mobile phone data, and by examining how the changes relate to second homes in different areas separately for workdays and weekends. This is examined both for the whole country and by comparing different counties. In the thesis, mobile phone data is utilized as the ground truth to assess the performance of household occupancy detection methods for electricity consumption, and to examine how electricity consumption data captures the spatiotemporal dynamics of second home users in South Savo.

The results indicate that people are generally more mobile during the summer, and the seasonal growth in people's presence correlates strongly with second homes. This shows a prominent seasonal effect for multi-local living in Finland. Additionally, it is shown that the results vary spatially as there is variation in the results both between counties and within South Savo.

The best performing second home occupancy detection method is revealed by correlation analyses between mobile phone data and electricity consumption data. Moreover, it is shown that electricity data correlates better with mobile phone data during the summer, and that the data captures the monthly dynamics of second home users well. This further highlights the seasonal effect of multi-local living.

The thesis provides valuable insight into how the seasonal variation of population in different areas is connected to multi-local living in Finland. Furthermore, it is shown that novel data sources can capture the changes in people's presence at multiple spatial levels with high temporal accuracy, and that they can be utilized to study multi-local living.

Keywords

Multi-local living, multi-locality second homes, mobile phone data, electricity consumption data, big data, GIS, South Savo

Where deposited

University of Helsinki electronic theses library E-thesis/HELDA

Additional information

HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

Tiedekunta	Osasto
Matemaattis-luonnontieteellinen tiedekunta	Geotieteiden ja maantieteen osasto

Tekijä
Iivari Laaksonen

Tutkielman otsikko
Monipaikkainen asuminen: vertailu matkapuhelin- ja sähkönkulutusaineistojen välillä

Koulutusohjelma ja opintosuunta
Maantieteen maisteriohjelma, geoinformatiikka

Tutkielman taso	Aika	Sivumäärä
Maisterintutkielma, 30 opintopistettä	Toukokuu 2022	59

Tiivistelmä

Monipaikkainen asuminen on monitahoinen yhteiskunnallinen ilmiö, joka liittyy vahvasti ihmisten liikkuvuuteen. Ilmiötä on aiemmin tutkittu pääosin virallisilla tilastotiedoilla, jotka eivät ota huomioon ihmisten asumisen ja liikkuvuuden monimuotoisuutta. Tässä maisterintutkielmassa monipaikkaisuutta tutkitaan Suomessa ja erikseen Etelä-Savon maakunnassa kesämökkien näkökulmasta uudenlaisilla aineistolähteillä: matkapuhelin- ja sähkönkulutusaineistoilla. Sen lisäksi, että nämä massa-aineistot ovat spatiaalisesti ja ajallisesti tarkkoja, niillä voidaan varmistaa tutkimusalueen kattavuus ja väestön riittävä otanta.

Tutkin monipaikkaisuutta analysoimalla ihmisten läsnäolon spatiotemporaalisia vaihteluita matkapuhelinaineistoilla. Vaihtelut kytketään monipaikkaiseen asumiseen ja erityisesti kesämökkien määrään eri alueilla erikseen työpäivien ja viikonloppujen osalta. Matkapuhelinaineistoja pidetään tässä tutkielmassa ”ground truth” aineistona, jonka avulla arvioidaan sähkönkulutusdataan liittyviä käyttöasteen havaitsemisen menetelmiä. Lisäksi vertailemalla näitä aineistoja arvioidaan, miten tarkasti sähkökulutusaineistoilla saadaan selville kausiväestön ja mökkeilijöiden vaihtelu Etelä-Savon alueella.

Tulokset osoittavat, että ihmiset liikkuvat enemmän kesäkuukausina kuin talvella. Tämä vaihtelu korreloi vahvasti kesämökkien määrän kanssa, mikä osoittaa, että monipaikkainen asuminen Suomessa on vuodenajoista riippuva ilmiö. Lisäksi alueelliset erot sekä eri maakuntien välillä että Etelä-Savossa korostuvat tuloksissa.

Korrelaatioanalyysillä matkapuhelin- ja sähkönkulutusaineistojen saadaan selville tehokkain menetelmä vapaa-ajan asuntojen käyttöasteen havaitsemiseen. Tulokset osoittavat myös, että nämä kaksi aineistoa korreloivat vahvemmin kesällä kuin muina kuukausina, ja että kausiväestön kuukausittainen vaihtelu korreloi useimmilla postinumeroalueilla Etelä-Savossa vahvasti. Tämä korostaa monipaikkaisen asumiseen liittyvän kausittaisen vaihtelun merkitystä.

Tämän maisterintutkielman tulokset osoittavat, että kausiväestön vaihtelu eri alueilla on yhteydessä monipaikkaiseen asumiseen Suomessa. Lisäksi tulokset osoittavat, että uudenlaiset aineistolähteet soveltuvat monipaikkaisen asumisen tutkimiseen hyvin eri aluetasoilla ja mahdollistavat ilmiön tarkan temporaalisen tarkastelun.

Avainsanat

Monipaikkainen asuminen, vapaa-ajan asuminen, matkapuhelinaineistot, sähkönkulutus, big data, GIS, Etelä-Savo

Säilytyspaikka
Helsingin yliopiston sähköinen tietokanta: E-thesis/HELDA

Muita tietoja

Table of contents

Table of figures

Table of tables

1. Introduction

Simply defined, multi-local living can be defined by individuals or families having access to more than one residence in their everyday lives. There are many possible reasons for people's multi-local practices, which may be related to working life or education, leisure (e.g., stays at free-time residences), family reasons (e.g., children with parents living in separate households) or a combination of various reasons (Schier et al., 2015).

The lives of multi-local people are structured by movements between multiple residences (Lehtonen, Muilu, & Vihinen, 2019; SYKE, 2021). This is connected to human mobility as people travel between "anchor points" that are meaningful to them to sustain close ties to places, to properties located there or to people living there (Hiltunen & Rehunen, 2014; Schier et al., 2015). Multi-locality causes changes in population at many spatiotemporal levels, and it is dependent of the spatial accessibility between regions. For instance, daily multi-locality may refer to people mostly moving within a certain municipality or subregion, whereas weekly and seasonal multi-locality may often be related to cross-county mobility or cross-border mobility (SYKE, 2021).

The scope of this thesis is examining multi-local living in Finland from the perspective of second homes or free-time residences (kesämökki in Finnish). The terms are used interchangeably in the thesis, since free-time residence is the term used in official Finnish statistics, while the term second home is the widely used in scientific literature. Despite this scope, it should be noted that multi-local living is a highly complex social phenomenon that requires interdisciplinary approaches as it can be studied from different perspectives with a range of methodologies. In other words, the phenomenon does not have a single consistent and comprehensive theory in sight (Weichhart, 2015).

Previously, multi-local living has been commonly researched with static information about people's residence, migration, and overall population. The assumption for official statistics and governance is that people anchor their lives in a single place which may lead to simplistic views on people's mobility and dwelling (Adamiak, Pitkänen, &

Lehtonen, 2017). This viewpoint has proven to be inadequate in capturing how seasonal shifts in people's mobilities alter regional population structures (Lehtonen, Kotavaara, Muilu, Vihinen, & Huovari, 2020). I aim to address these issues in the thesis by approaching the changes in people's presence and multi-local living with novel big data sources like mobile phone data from Telia Crowd Insights and electricity consumption data from Suur-Savon Sähkö Oy. Along with national level analyses of the phenomenon, the thesis will specifically focus on the county of South Savo, which has the highest relative proportion of seasonal population in the country (Ruralia Institute, 2021) .

The study period of the thesis ranges from November 2018 to August 2019 due to the data period for the mobile phone activity data. While this period does not cover the COVID-19 pandemic, its effects on multi-local living cannot be overlooked as the response to the crisis has caused one of the biggest disruptions to individual mobilities in modern times (Poom, Järv, Zook, & Toivonen, 2020). The implications of the pandemic for multi-local living are examined in a separate subchapter of the background section of the thesis.

This thesis is a part of the *MOPA* research project (*The rhythms, places, and customer groups of multi-local dwelling)* (Ruralia Institute, 2021). The project is carried out by the Ruralia Institute in collaboration with the Digital Geography Lab utilizing electricity consumption data from Suur-Savon Sähkö Oy. Overall, the research project aims to provide accurate information about the dynamics of multi-local living for the benefit of local businesses and stakeholders in South Savo.

Main research questions for the thesis are:

1. What can mobile phone activity data reveal about people's presence and multi-local living in Finland and in South Savo?
2. How do different occupancy detection methods for electricity consumption data correlate with mobile phone presence data?
3. How does electricity consumption data capture the spatiotemporal dynamics of second home users in South Savo?

I answer the first research question of the thesis by analyzing people's presence with mobile phone data, and by conducting correlation analyses of the results with free-time residence data from official statistics. The second and third research questions are answered by comparing presence data derived from mobile phone data to electricity consumption data.

2. Background

2.1. Multi-local living

2.1.1. Second homes as form of multi-local living

Having a second home is considered a specific form of multi-local living which should be distinguished from migration, since people do not abandon a primary place of residence, but rather move repetitively and cyclically between multiple residences (Schier et al., 2015). The definition of second homes has evolved depending on time and culture: second homes are considered a major part of (often domestic) travel and tourism, for instance, in the Nordics, Southern Europe, Russia, North America, Australia, and New Zealand. A significant proportion of the population have access to a second home in these areas (Pitkänen et al., 2020; Zoğal, Domènech, & Emekli, 2020). Additionally, globalization and the blurring of national borders has caused the phenomenon of second homes to increasingly have a cross-border dimension (Zoğal et al., 2020).

Second homes vary from region to region, but they are commonly located in rural areas, for instance due to rural roots of families and cottage inheritance (Pitkänen, 2008; Pitkänen et al., 2020). People may want to continue using their inherited second homes as an expression of their identity, which may also increase their willingness to travel long distances and spend considerable amounts of time to access their second homes. This is because place attachment and the sense of belonging to second home areas appear to be persistent and multi-generational (Hall, 2014; Tjørve, Flognfeldt, & Tjørve, 2013). Also, drivers pushing people to buy a second home in rural areas may be the longing for more diverse natural environments and attractions compared to their primary residential areas. Furthermore, especially in warm countries such as Turkey, people seek refuge from second homes in higher altitudes to escape summer heat (Zoğal et al., 2020).

From a public planning perspective, administrators usually plan services for the static, registered population based on the assumption that people have only one home (Willberg, Järv, Väisänen, & Toivonen, 2021) leaving second home users out of the equation. Additional information about second homes and multi-local practices could be utilized as an additional resource for regional development in rural areas if the phenomenon was more visible in governance (Lehtonen et al., 2019). For instance, Back & Marjavaara (2017) refer to second-home users as the invisible population from a public planning perspective. To address this issue, Adamiak, Pitkänen & Lehtonen (2017), have proposed two alternative measures for population: average population that indicates the average number of people present at an area throughout a year, and seasonal population that indicates the number of people expected to be present in an area during the highest tourist season. These measures are examples of tools that may be helpful for public planning and policy perspectives to ensure adequate resources, infrastructure, and services for both the seasonal and permanent population in an area.

People move between their primary and second homes at varying usage-patterns since second homes are places for both spare time and remote working (Czarnecki, Dacko, & Dacko, 2021). Some may spend entire summers at their second homes while others may visit almost every weekend throughout the course of the year. For instance, in Finland, people's overall time spent at their second homes has been increasing during recent years along with the number of winter-habitable summer homes since more and more people demand amenities at their free-time residences comparable to their primary homes (Hiltunen & Rehunen, 2014; Voutilainen, Korhonen, Ovaska, & Vihinen, 2021). The increase in time spent at second homes is more common in buildings suitable for year-round use rather than buildings suited for summer use only (Voutilainen et al., 2021), which indicates that the boundaries between primary and secondary residences have become less distinct (Lehtonen et al., 2020). This highlights the need for more accurate information on people's multi-local living and related mobility practices as second homes are becoming more important to people instead of just being places that are only used in the summer. The information could be used in different municipalities and counties for more accurate governance in terms of the second home population.

2.1.2. The effect of COVID-19 on multi-local living

The COVID-19 disease, which rapidly spread throughout the world in the beginning of 2020, reached the status of pandemic on March 11th of the same year. This caused governments to implement various non-pharmaceutical intervention measures (Pepe et al., 2020), including restricting people's movements cross country borders and domestically, to contain the spread of the virus by social distancing, and to ensure the functioning of health care systems (Santamaria et al., 2020). What's to note is that all aspects of the pandemic, including the initial outbreak of the virus and its spread since, as well as the impacts of the containment measures on society take place somewhere across geography. In other words, the crisis is inherently spatial, and the response has caused one of the biggest disruptions to individual mobilities in modern times (Poom et al., 2020).

The initial outbreak and the mitigation measures of COVID-19 caused people to escape to their rural second homes in many countries in search of safety and a more meaningful isolation environment. This was contrasted by recommendations by governments either for or against visiting second homes influenced by health risk questions and questions about whether second home population has the right to be in rural areas (Pitkänen et al., 2020). Zogal et al. (2020) refer to this as a privileged escape from cities by those who have access to second homes and, for instance, in the UK, the pandemic has sparked conversations about housing inequality since some people own multiple properties while others suffer from poor housing situations with rising rents. Whereas, in the Nordics, the discourse around second homes has mainly dealt with the health care capacity of rural regions as second-home owners from large cities pose a risk of spreading the disease to the rural communities (Pitkänen et al., 2020). This may lead to an overload in the need for health care and emergency services due to service providers not being prepared for the population increase (Willberg et al., 2021).

Due to the pandemic, the role of second home mobilities has been shifting from primarily leisure-driven to work-oriented mobility. That is, people have searched for alternative places to work remotely resulting in the temporary conversion of second homes into primary residences (Czarnecki et al., 2021; Zoğal et al., 2020). For instance, research by Willberg et al. (2021) showed that people’s presence in areas with many workplaces decreased while areas with second home concentrations saw the most increase in people’s presence during the first wave of the pandemic.

Since the initial outbreak, people have continued remote working while living multi-locally. This may provide development possibilities to areas with decreasing population and pessimistic demographic projections. However, COVID may either represent an isolated instance of change to people’s dwelling and working locations, or it may represent a wider course of development regarding how people can choose their residence location more freely due to flexible working conditions (Lehtonen & Kotavaara, 2021).

The changing role of second homes during the pandemic further highlights the need for more information on the phenomenon of multi-local living (Willberg et al., 2021), which could be studied with ever-growing number of big data sources (Poom et al., 2020).

2.2. Novel big data sources for mobility research

There is a need for more detailed information to understand dynamic social phenomena like multi-local living, as traditional static data sources often fail to capture people’s flexible lifestyles and mobilities (Willberg et al., 2021). Traditional data to support human mobility and travel behavior research have been derived from conventional travel surveys and diaries which are considered costly to collect. To supplement or replace traditional data, various novel big data approaches have recently been applied to ensure extensive coverage of population and geographical area (Wang, He, & Leung, 2018), while enabling mid- or long-term research perspectives on human mobility (Järv, Ahas, & Witlox, 2014).

Big data has many varying definitions, but a common way to define it is through three terms: volume, velocity, and variety, which set these data apart from traditional data sources (Kitchin, 2014). Volume refers to there being a multitude of data available, while velocity refers to new data being constantly generated in or near real-time. Variety refers to significant differences in the types of data available that are often spatially and temporally referenced.

These novel data sources are used as proxies for people. For instance, mobile phones and smart watches, as well as their applications, are used to analyze people's behavior. These specific data with spatial and temporal information are referred to as mobile big data (Poom et al., 2020). Mobile big data sources include mobile network operators' call detail records and other communication data, data from smart phone operating systems (accurate data only available to system developers), geolocated posts from social media platforms (e.g., Instagram or Twitter), data from transactional smart cards, and other mobile applications utilizing location features (e.g., sports applications and ride-sharing platforms) (Järv, Tominga, Müürisepp, & Silm, 2021; Poom et al., 2020). According to Järv, Tominga, Müürisepp & Silm (2021), recent academic research has demonstrated the suitability of these data to uncover mobility flows and people's spatial behavior at the societal level.

2.2.1. Mobile phone data

The focus in this thesis is on mobile phone data which, according to Silm, Järv & Masso (2020), can be collected from each mobile phone in a mobile phone operator's network with varying positioning methods. The data typically consist of geographical location and moment in time allowing for similar analyses to more conventional research methods. Compared to conventional methods, such as questionnaires, GPS-tracking, and travel surveys, mobile phone data can help to obtain more objective information on mobility since memory gaps and selectivity in filling in travel diaries can cause inaccurate results. Moreover, data can be collected nearly in real time with higher spatial and temporal accuracy (Silm et al., 2020).

These data are collected through mobile positioning methods ranging from network-based to mobile phone-based approaches. In network-based positioning the location of users is derived through network performance at aggregate level, whereas in mobile phone-based approaches the devices are examined individually through passive or active means of mobile positioning (Silm et al., 2020). Mobile phones provide a unique opportunity for human mobility studies because of their ubiquity as over five billion people own a mobile phone worldwide (GSMA Intelligence, 2021) while, in Finland, 83% of people own a smartphone (Official Statistics of Finland (OSF), 2019).

However, sometimes the methodologies used in collecting mobile phone data are concealed or black boxed. Especially when the data are aggregated and anonymized by private companies such as mobile network operators, the potential of their usage in scientific research may be limited (Poom et al., 2020). This is the case for commercial datasets like Telia Crowd Insights products that are used in this thesis as spatial aggregations to the scale of the entire population in addition to temporal aggregations have already been performed by the data provider. This may restrict the possibilities for further custom spatial and categorical aggregations needed for detailed demographical analyses that could be possible with individual level raw data.

The concerns regarding data aggregation are related to privacy and security issues that are among the most important aspects to consider for both mobile phone users and data providers (Silm et al., 2020). This is because data containing information on individuals' locations and mobilities can be used to reconstruct their unique movements making them one of the most sensitive data being collected (de Montjoye, Hidalgo, Verleysen, & Blondel, 2013). The use of mobile phone data should be deemed necessary based on the relevance of the use purpose and the need for timely data (Jansen et al., 2021).

According to Silm et al. (2020), people's privacy can be protected at different stages: first by sufficient security in data collection and storing, second by ensuring individuals' anonymity when processing and analyzing data, and finally by generalizing the results to a scale where individuals cannot be recognized. Furthermore, privacy concerns can be mitigated with clear communications to all stakeholders, which is especially important when using mobile phone data in research. Poom et al. (2020), also highlight the need to develop collaborative platforms between stakeholders to ensure that privacy issues are being met when using individual level raw data instead of aggregated products. This would help enable scientific developments and to foster relations between companies providing mobile big data.

Applications for mobile phone data include analysis of people's origins and destinations, traffic monitoring, and analyzing spatiotemporal human mobility patterns (Rojas, Sadeghvaziri, & Jin, 2016). Mobile phone data enables analyses of people's activity locations and movements to reveal presence of people and mobilities of the entire population. This can contribute to assessing population dynamics and changes both temporally and spatially (Silm et al., 2020).

Themes related to multi-local living have previously been researched with these novel data sources. Silm & Ahas (2010) developed means to monitor the short-term mobility of population to analyze seasonal population in Estonia using mobile phone data. More recently, multi-locality has been examined in articles regarding the impacts of the COVID-19 pandemic. Namely, Järv et al. (2021) utilized means of longitudinal smartphone tracking in their article regarding how the pandemic affected the lives of transnational Estonians residing in Finland. People's movements were collected with an application, which allows for long term tracking of people's mobility and their social network interactions (information on incoming and outgoing phone calls and text messages). This is an example of smartphone-based tracking, which utilizes device-integrated sensors, such as GPS to position devices.

Network-based mobile phone positioning data has been utilized in recent research to capture people's overall mobility together with traditional statistical data on people's net migration during the pandemic (Lehtonen & Kotavaara, 2021). Willberg et al. (2021) utilized the Telia data to capture people's mobilities to their second homes to assess the impacts of multi-local living and second home mobilities on population dynamics in Finland during the initial COVID-19 outbreak. These studies highlight the applicability of novel mobile phone-based data sources to study people's mobility and multi-local practices.

2.2.2. Electricity consumption data

Electricity consumption data is another novel data source applicable for second home and multi-local living research. The usage of consumption data for research has been made possible in recent years by the increased prevalence of smart meters in buildings, which allow for spatiotemporal monitoring of electricity consumption and its amounts meaning up to near real-time data collection on consumption data (Kavousian, Rajagopal, & Fischer, 2013; Yildiz, Bilbao, Dore, & Sproul, 2017). However, the data is collected and handled by electricity companies, so it is not openly available.

Data analytics in the electricity sector can be applied in many phases ranging from the forecasting of renewable energy generation, trading and distribution, monitoring, and both forecasting and analysis of electricity consumption. This allows for the classification of customer groups based on consumption patterns and habits (Scheidt et al., 2020). Moreover, electricity data can be used to assess appliance ownership and usage, as well as detecting people’s presence and occupancy of a household, which also means that the data should be secured to ensure people’s privacy (Yildiz et al., 2017).

Previous research has shown that temporally detailed consumption data derived from smart meters can be used to detect and predict a household’s occupancy (Razavi, Gharipour, Fleury, & Akpan, 2019), which may be used as an indicator for people’s presence at different points in time. Electricity consumption data has been used to gain understanding about regional differences in household electricity consumption (Sun, Zhou, & Yang, 2018), and to reveal long-term consumption growth trends in second homes (Andersen, Christensen, Jensen, Kofoed, & Morthorst, 2008).

The article by Sun et al. (2018), showed the potential of electricity consumption data for capturing temporary mobilities from cities in Jiangsu province, China during national holidays, while Andersen et al. (2008) used the annual consumption data together with economic parameters to assess reasons for the increase in second home electricity consumption in two regions in Denmark. These studies are related to population dynamics and show the suitability of electricity consumption data for assessing both household-level and larger scale spatiotemporal rhythms of second home usage.

3. Case study: overview of second homes

3.1. Distribution of free-time residences in Finland

Multi-local living is common in Finland and in other Nordic countries where 50% of all households have access to a second home (Müller, 2007, 2021). According to recent research (Willberg et al., 2021), multi-local living is tightly connected to seasonal residences in Finland as there are over half a million summer cottages in the country (Official Statistics of Finland (OSF), 2021a) and over 1.5 million Finns belong to a household with ownership to a second home (SYKE, 2021). Second homes in Finland are usually cottages located by water bodies in the countryside (Hiltunen & Rehunen, 2014). These environmental factors, along with the rural roots of families' and cottage inheritance, affect the locations and distribution of second homes (Pitkänen, 2008).

Additionally, the number of second homes in rural areas has been growing while permanent population has been decreasing during the past decades (Adamiak et al., 2017). Notably, the development of Finnish population structure has been previously mostly researched from the perspective of static statistical information on people's residence, migration, and population in different areas (Lehtonen et al., 2020). The reason for this is the prevailing assumption that people reside in a single location and public services are planned from the starting points of the permanent dwelling structure.

However, while people's mobilities cannot be captured with these static data sources, the spatial distribution, and the number of second homes in Finland can be explored. Figure 1 shows the development of the free-time residence stock in Finland from 1970 to 2020 with data from Statistics Finland (Official Statistics of Finland (OSF), 2021a). The number of free-time residences has been growing quite steadily in recent decades, but notably the number has nearly tripled in 50 years. However, the statistic from 2020 uses a different building classification than in other years: detached houses reported for leisure-time use are no longer included as free-time residential buildings and rental holiday cottages are included instead. Consequently, the latest statistic is not comparable to the previous ones, but nevertheless the overall trend is clear.

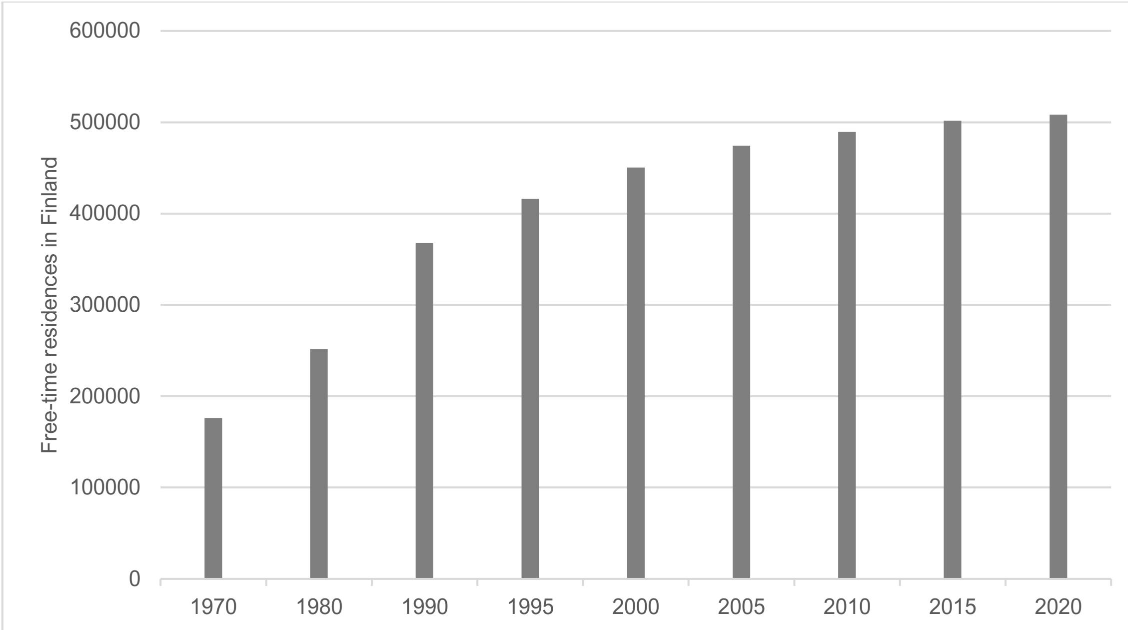


Figure 1. Development of Finnish free-time residence stock 1970-2020 (Official Statistics of Finland (OSF), 2021a)

Overall, the top 5 counties with the most free-time residences in Finland are: Southwest Finland, Pirkanmaa, South Savo, Uusimaa, and Lapland. However, the absolute number needs to be proportional (e.g., compared to permanently occupied residences) to accurately describe how predominant free-time residences are in an area.

To assess the distribution of free-time residences in a more balanced way, Figure 2 shows comparisons between the number free-time residences and permanently occupied dwellings in different Finnish counties (maakunta in Finnish). The ratio is calculated by dividing the number of free-time residences by the number of permanently occupied dwellings. If there are 500 free-time residences, and 100 permanently occupied dwellings in an area, the ratio is 5.0, indicating that the building stock in mainly comprised of free-time residences.

The statistics show that there are quite notable differences between the counties. For instance, South Savo has the highest ratio overall followed by Åland and Lapland. Whereas, in Uusimaa the ratio is the lowest, although overall it has the fourth highest number of free-time residences.

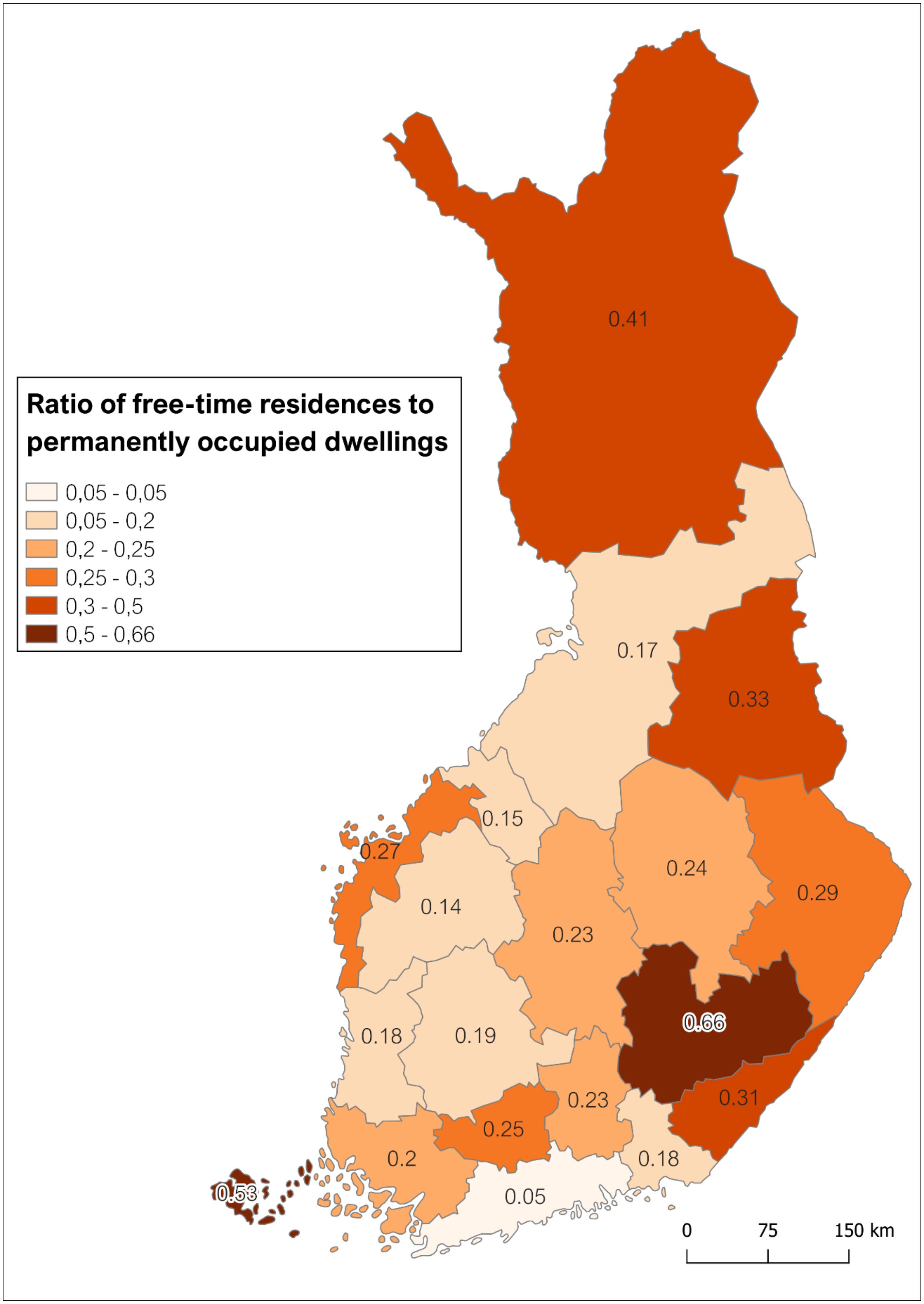


Figure 2. Ratio of free-time residences to permanently occupied dwellings by county in 2020 (Official Statistics of Finland (OSF), 2021b, 2021c)

Table 1. Municipalities with the highest ratio of free-time residences to permanently occupied dwellings in 2020 (Official Statistics of Finland (OSF), 2021b, 2021c)

	Municipality	**Free-time residences**	**Permanently occupied dwellings**	**Free-time residences/Permanently occupied dwellings**
1.	Kustavi	3269	505	6.47
2.	Puumala	3998	1143	3.50
3.	Vårdö	562	203	2.77
4.	Föglö	698	253	2.76
5.	Hirvensalmi	2959	1140	2.60
6.	Kuhmoinen	3039	1212	2.51
7.	Taivassalo	1970	812	2.43
8.	Kökar	273	116	2.35
9.	Luhanka	833	375	2.22
10.	Geta	485	226	2.15
11.	Kumlinge	340	161	2.11
12.	Lumparland	355	170	2.09
13.	Pertunmaa	1808	911	1.98
14.	Sottunga	117	59	1.98
15.	Sysmä	4001	2060	1.94
16.	Padasjoki	2867	1516	1.89
17.	Enontekiö	1557	838	1.86
18.	Eckerö	788	462	1.71
19.	Pelkosenniemi	825	484	1.70
20.	Kolari	3009	1774	1.70

Table 1 shows a similar comparison at the municipality level. It shows the top twenty municipalities with the highest ratio of free-time residences to permanent dwellings. Similarly, there are notable differences between the municipalities with Kustavi having the most free-time residences compared to permanently occupied dwellings.

3.2. Distribution of free-time residences in South Savo

This thesis will, along with national level analyses, especially focus on the county of South Savo located in Eastern Finland. As Figure 2 shows, South Savo is the county with the highest ratio of free-time residences to permanently occupied dwellings in Finland. It also ranks third highest in the number of free-time residences overall. Table 2 zooms into the county and shows population comparisons with statistics about free-time residences and permanent dwellings in the municipalities of South Savo. Puumala is the municipality with the most free-time residences per permanently occupied buildings, while Mikkeli has the most free-time residences overall. Figure 3, on the other hand, shows the ratio of free-time residences to permanently occupied dwellings in South Savo on a map.

Table 2. Overview on the population development in South Savo with comparisons between free-time residences and permanent dwellings in 2020 (Official Statistics of Finland (OSF), 2021b, 2021c, 2021e)

	Municipality	Population	Population change from 2019 (%)	Free-time residences	Permanently occupied dwellings	Free-time residences/Permanently occupied dwellings
1.	Puumala	2137	-0.7	3998	1143	3.50
2.	Hirvensalmi	2156	0.9	2959	1140	2.60
3.	Pertunmaa	1654	-2.1	1808	911	1.99
4.	Sulkava	2482	-0.4	2193	1310	1.67
5.	Mäntyharju	5676	-2	4791	3046	1.57
6.	Rantasalmi	3364	-2	2185	1693	1.29
7.	Kangasniemi	5312	-0.8	3605	2817	1.28
8.	Enonkoski	1369	0.6	755	700	1.08
9.	Juva	5932	-3	2099	3128	0.67
10.	Savonlinna	32662	-0.9	8765	17690	0.50
11.	Mikkeli	52583	-1	10345	27477	0.38
12.	Pieksämäki	17375	-1.7	3069	9458	0.32

Figure 4 shows the absolute number of free-time residences by postal code area. There appear to be differences between postal code areas inside municipalities. For instance, the centers of Kangasniemi, Hirvensalmi, Mäntyharju, and Puumala have the highest number of free-time residences overall, while the distribution is more even in other municipalities.

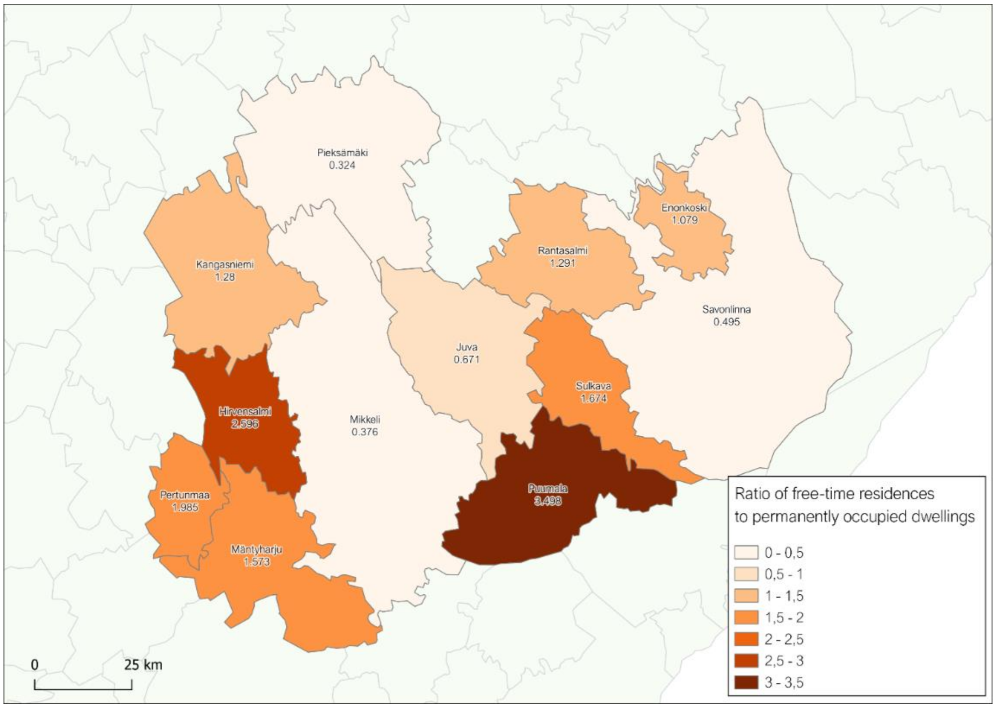


Figure 3. Ratio of free-time residences to permanently occupied dwellings in South Savo municipalities in 2020 (Official Statistics of Finland (OSF), 2021b, 2021c)

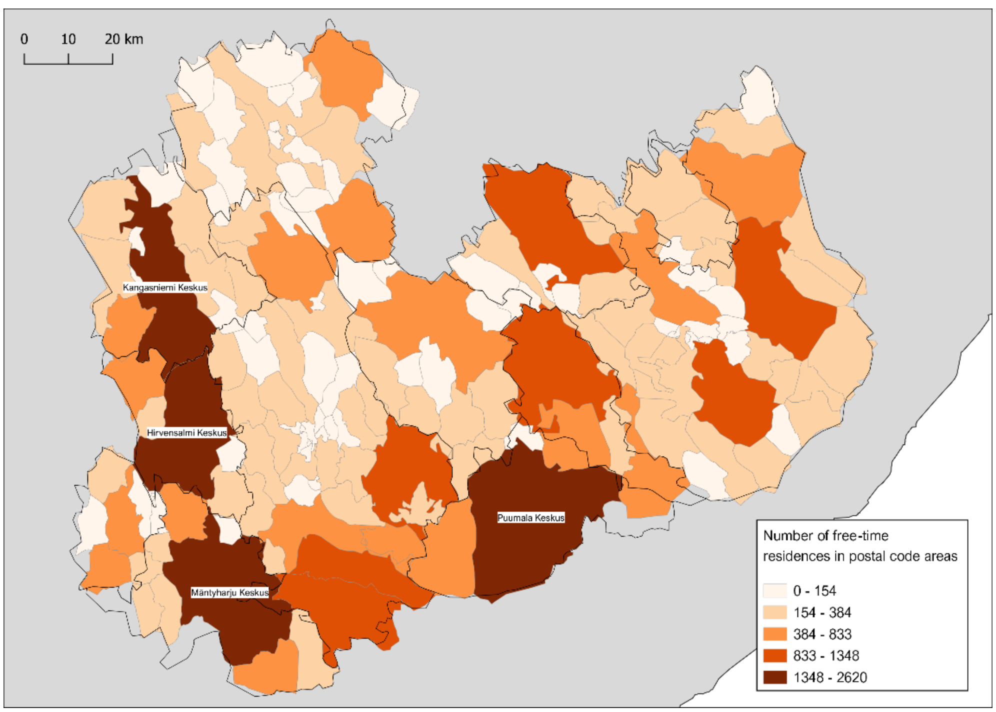


Figure 4. The absolute number of free-time residences in South Savo by postal code area in 2019 (Official Statistics of Finland (OSF), 2021d)

4. Data

Several datasets are used in this thesis to examine multi-local living. Table 3 shows the descriptions, uses, and sources for the datasets.

Table 3. Data overview

Dataset	**Description**	**Use**	**Source**
Mobile phone activity data	Mobile phone activities in a postal code unit for each month (11/2018 - 9/2019) in Finland	People's activities are considered a proxy for people's presence in a postal code area. Aggregations to larger areas possible.	Telia Crowd Insights
Electricity consumption data	Aggregated consumption and household occupancy levels in South Savo postal code areas at the monthly level from (2018 – 2019)	Correlation analyses between mobile phone data to assess occupancy detection methods and how the data captures people's spatiotemporal dynamics	Suur-Savon Sähkö Oy
Paavo postal code area statistics 2021	Inhabitants and the number of free-time residences in each postal code area (statistics from the year 2019)	Connecting the presence data to official statistics, which links it to multi-local living	Statistics Finland

For the purposes of this thesis, mobile phone activity data is considered the ground truth data for people's presence on which the electricity consumption data is compared.

The attributes of the mobile phone activity dataset are:

- Postal code area where the activities were recorded
- Home locations of people (previous night location)
- Weekdays divided into groups: workdays (Monday-Thursday), Fridays, Saturdays, and Sundays
- Hour of day for the activity
- Sum of all activities in a postal code unit for each weekday for each month
    - Activity is recorded in if a person spends 20 minutes or more there
- Month and year of the activities

The electricity consumption data consists of aggregated electricity consumption and property occupancy levels for properties classified as second homes by postal code area in South Savo, which have been previously calculated in the MOPA research project (Ruralia Institute, 2021). The data contains information about properties contracted to Suur-Savon Sähkö Oy. The aggregations have been made from monthly raw electricity consumption values per property (kWh) using five detection methods to calculate the number of properties occupied:

- > 0 kWh consumption
- > 2.5 kWh consumption
- > 5 kWh consumption
- lower 95% confidence interval: the range where the true population mean is located within the 95% confidence interval
- > 2.5 kWh consumption and the lower 95% confidence interval methods combined

For the first three methods, if the monthly consumption was lower than the threshold kWh-number, the property was considered not occupied. Furthermore, for the lower 95% CI method, monthly data from the period 2015-2020 have been used in the calculations to detect the true population mean instead of using a threshold value for detection.

The number of properties occupied (consumers) have been used together with the total number of properties to calculate use rates for the areas. Use rates in postal code areas have been calculated by dividing the consumers by the total number of properties multiplied by 100.

5. Methods

Because of the file size and hourly temporal accuracy of the mobile phone activity dataset, it is necessary to pre-process and further aggregate the data. Main tool for aggregations and analyses of people's presence is Python, specifically the Pandas library in a Jupyter Notebook environment. Correlation analyses for comparing the mobile phone data to registry data and electricity consumption data are also made in Python using the scipy.stats package. Cluster analysis, on the other hand, was conducted with the Skikit learn library in Python. The Jupyter Notebook files used to carry out the analyses can be found on Github.

Furthermore, map visualizations were made in QGIS, and figures were visualized with Microsoft Excel.

5.1. Country-level mobile phone data analysis methods

The workflow for analyzing the mobile phone activity data at the country level consists of pre-processing and analyses. First, the mobile phone data consisting of multiple separate CSV-files are combined into one and aggregated from hourly level to weekday level. Next, the presence values are normalized based on the number of days in each weekday group in each month. Subsequently, workdays and weekends are separated from each other. Following these steps, analyses for both workdays and weekends on people's presence between areas, as well as correlation analyses between people's presence and the number of free-time residences are conducted. The workflow is described in detail in Figure 5.

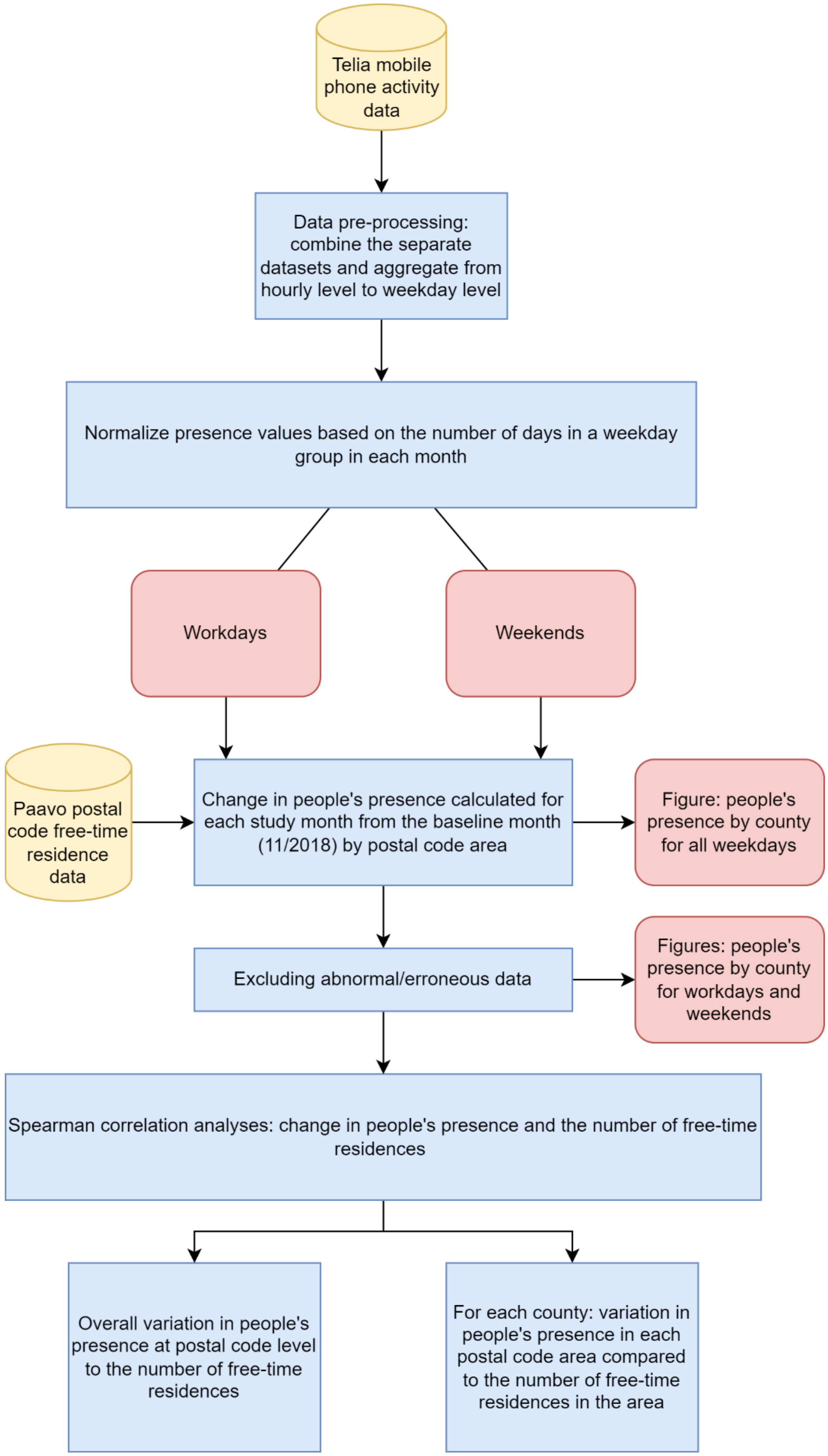


Figure 5. Country level mobile phone data workflow

Figure 6 shows an exploration of the change in people's presence in each county for each study month from the first month in the activity dataset (11/2018), which is considered the baseline month to which the other months are compared. The figure reveals irregularities in the data that must be dealt with. For instance, the presence values overall for each county in September are significantly lower compared to other months. Also, the summer values (June-August) for Kymenlaakso and South Karelia seem abnormal, as they are significantly lower compared to other counties and other months. Based on these observations, the abnormal data are excluded from further analyses.

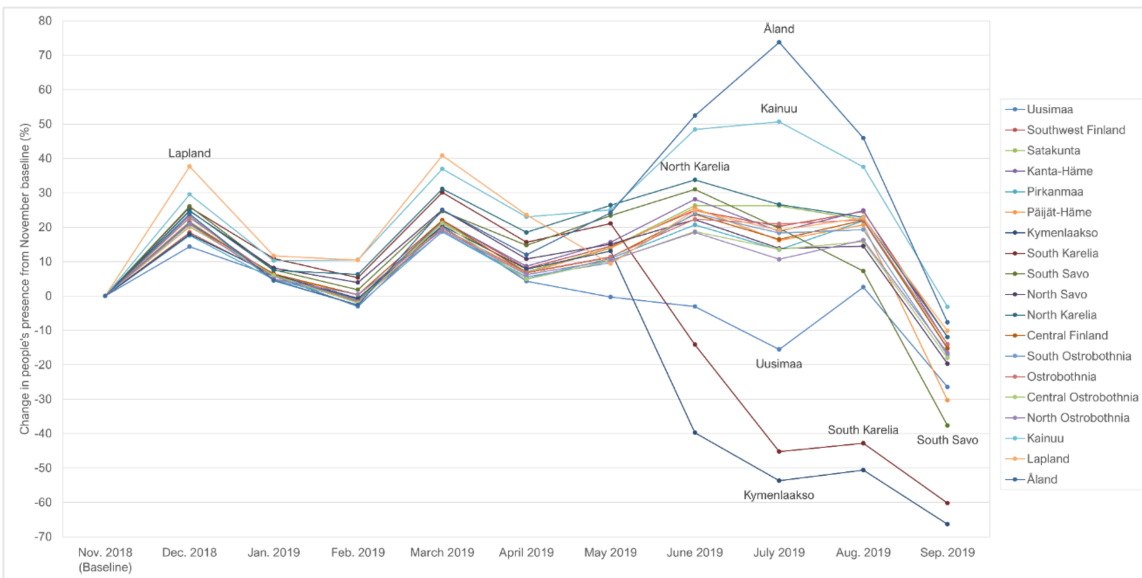


Figure 6. Change in people's presence from the November baseline for all weekdays by county

The abnormalities for Kymenlaakso and South Karelia are confirmed in Figure 7B. People's presence in July in South-Eastern Finland is shown to be extremely low for a summer month. Figure 7A shows people's presence by postal code area in January and no prominent abnormalities can be spotted. The figures show the variation in people's presence from the November baseline for each postal code area in Finland taking all weekdays into account.

a) b)

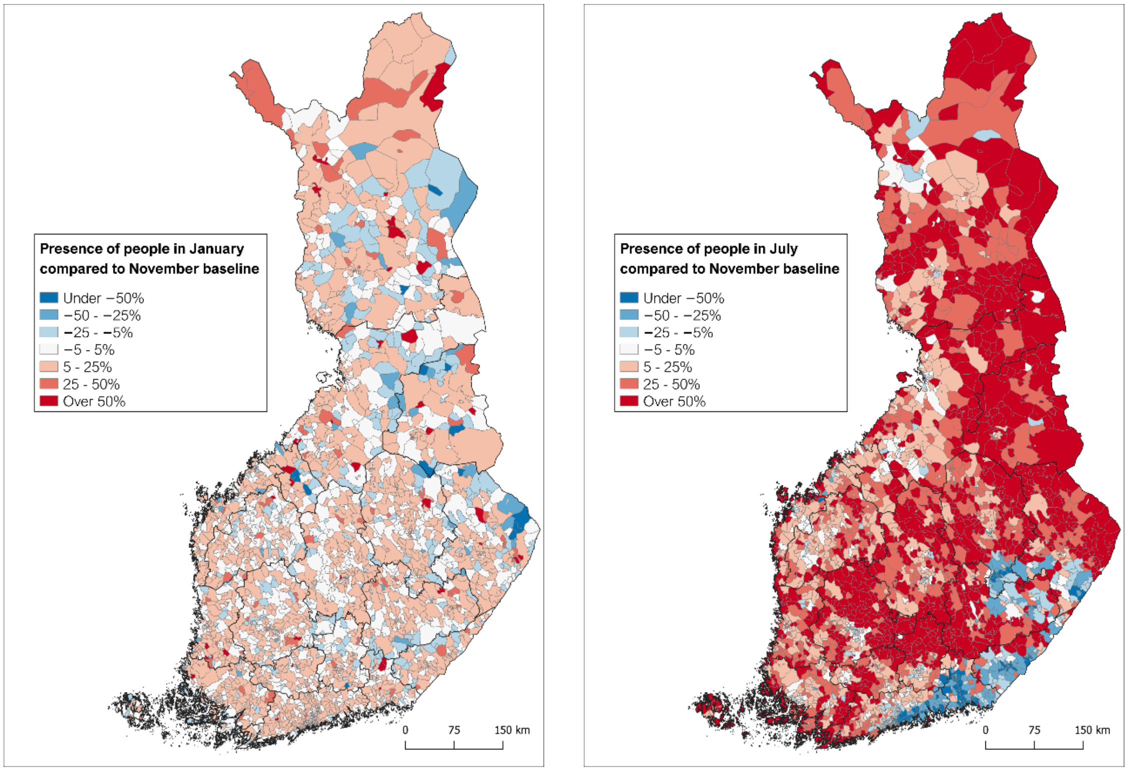


Figure 7. Presence of people a) in a January and b) in July compared to the November baseline by postal code area (all weekdays)

Further analyses are made separately for workdays (Monday to Thursday values), and weekend values (Saturday and Sunday values). The separations are made to capture the differences for workday and weekend populations. This can be helpful in the assessment of people’s use of free-time residences and multi-local practices. Fridays are excluded from the analyses for clarity as people’s movements could represent both workday and weekend mobilities. Furthermore, people’s presence and multi-local living are assessed overall for the entire country taking all postal code areas in Finland into account and separately for postal code areas in each county to assess differences within the country.

Correlation analyses are performed to assess the relationship between the mobile phone presence data and the number of free-time residences in the postal code areas in Finland (Official Statistics of Finland (OSF), 2021d). More specifically, the analysis assesses the change of the Spearman correlation coefficient between change in people's presence and the number of free-time residences. As both datasets rejected normality in the Shapiro-Wilk tests, Spearman correlation was chosen as the appropriate method.

Correlation analyses for workdays and weekends were made between the change in people's presence from the established November baseline in each postal code area for each month and the number of free-time residences in each area for the whole country. First, the analysis was conducted on the raw number of free-time residences, but to make the data proportional to the population, free-time residences per 1000 inhabitants in each postal code area was chosen as the more appropriate statistic. Next, the correlation analysis was repeated separately for each county to compare how correlation differs between different areas of the country within the study period.

5.2. Mobile phone data analysis methods for South Savo

The workflow for South Savo analyses is described in detail in Figure 8. In a similar manner to the country-level analyses, Spearman correlation analyses between the change in people's presence from the baseline and the number of free-time residences per 1000 inhabitants were performed for postal code areas in South Savo. Additionally, to capture variations within the county, three subregions were separated: Mikkeli area (Hirvensalmi, Kangasniemi, Mikkeli, Mäntyharju, Pertunmaa, Puumala), Savonlinna area (Enonkoski, Rantasalmi, Savonlinna, Sulkava) and Pieksämäki area (Pieksamäki and Juva). Presence comparisons to the baseline and correlation analysis were made separately for these areas.

The mobile phone data also contains information about people's origins (previous night location), which allows for spatiotemporal assessment of people's mobilities to South Savo. To accurately analyze mobility rather than immobility, locals were excluded from the analysis. Locals are defined here as data points where the activity postal code location and the origin location are the same since a person's presence in a postal code area is registered regardless of them constantly being in the area rather than moving to the area. In other words, an activity is considered an instance where a person spends 20 minutes or more in a postal code unit.

After excluding the locals, people were assigned to home counties based on the origin postal code area information. Finally, the presence values by each origin county were summed and proportions of the origin counties were calculated based on the activity sums of each home county.

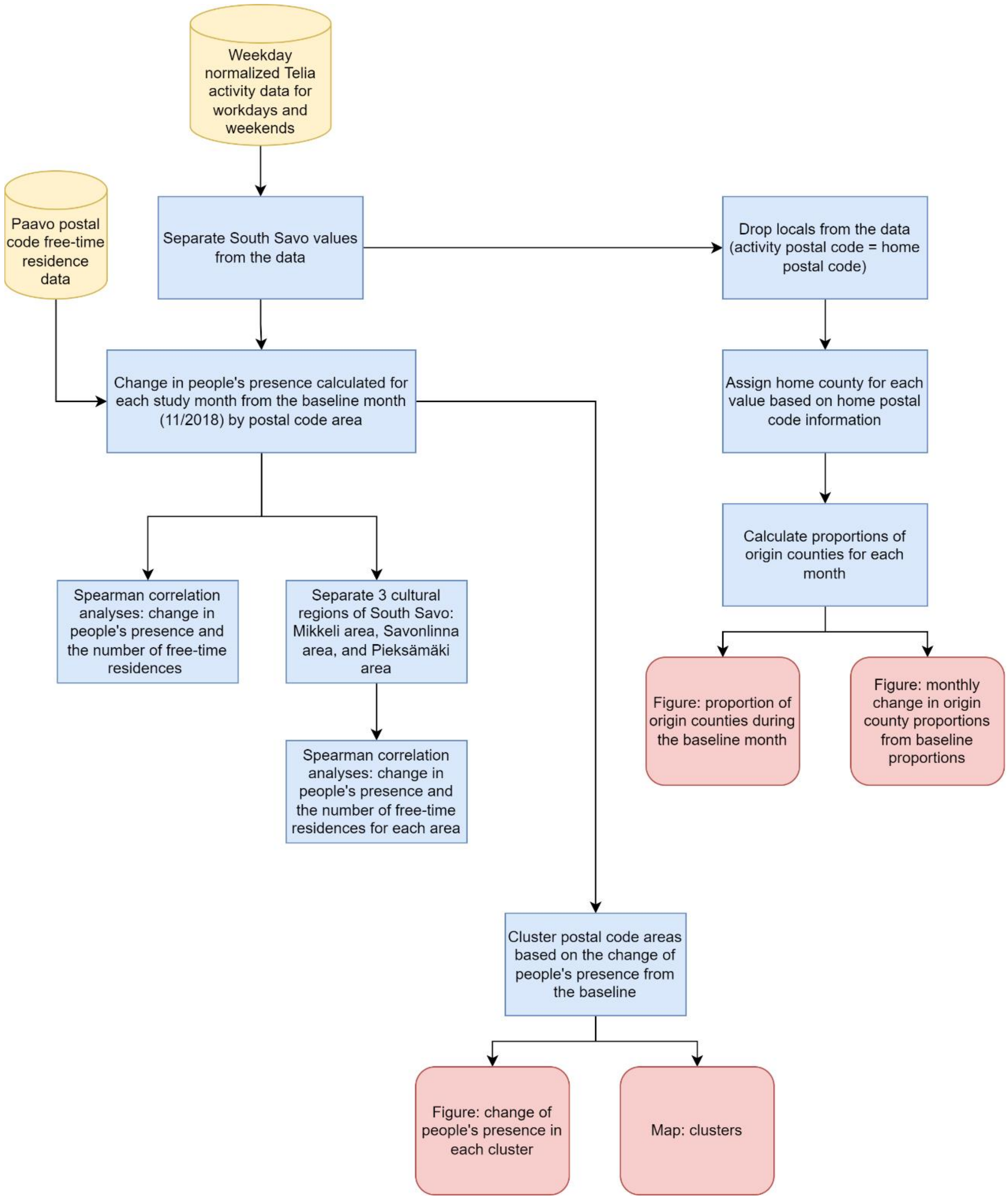


Figure 8. South Savo mobile phone data workflow

Next, postal code areas in South Savo were assigned to clusters based on the change in people's presence from the baseline in each month. K-means clustering was chosen as the method, but before generating the clusters, the elbow method was used to determine the optimal number of clusters to generate.

a) b)

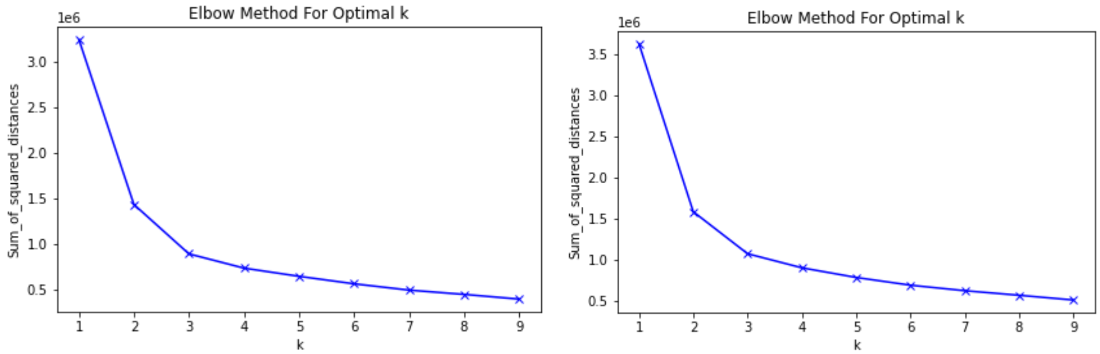


Figure 9. Elbow method results to find out the optimal number of clusters to generate in the K-means analysis a) for workdays, b) for weekends

Based on the results shown in Figure 9A for workdays, and Figure 9B for weekends, three clusters were chosen to be generated in the K-means clustering.

5.3. Electricity consumption analysis methods

The workflow for comparisons between mobile phone activity data, which is considered the "ground truth" data in the analysis, and electricity consumption data is described in Figure 10. The aim is to find out which occupancy detection method for electricity consumption correlates best with the mobile phone presence data. Other aims are to find out how the electricity consumption data captures the spatial distribution of second home users for each month, and how it captures the monthly dynamics of second home users for each postal code area.

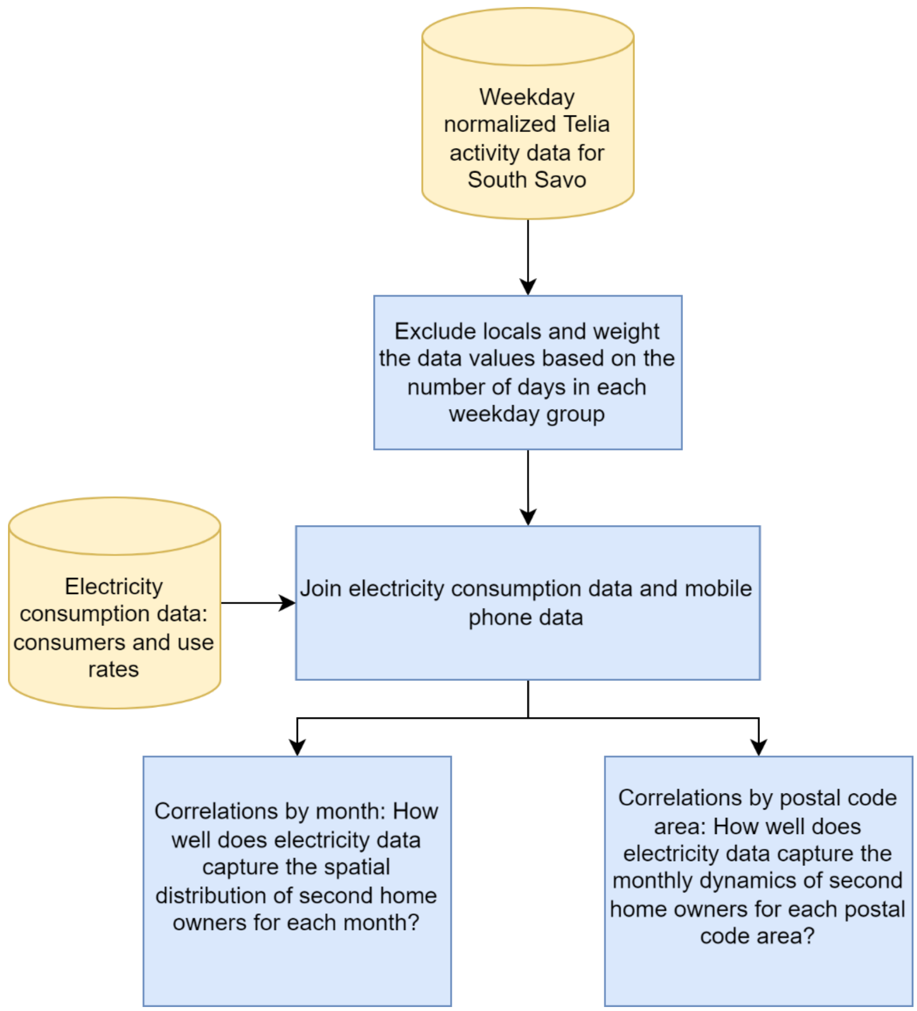


Figure 10. Workflow for comparing mobile phone data to electricity consumption data

First, the mobile phone data was pre-processed by excluding locals (postal code for the activities = origin postal code) from the data to assess second home users. Moreover, the presence values were weighted based on the number of days in each weekday group. Since the workday group consists of four different weekdays, the values in that group were multiplied by four. Other groups contain information from single weekdays (e.g., Saturdays), so for those no weighting was needed. Next, the number of properties occupied (consumers) and the use rates for each occupancy detection method were joined with the modified mobile phone data presence values.

Next, two types of correlation analyses were performed using the Spearman correlation method.

Correlation between both use rates and consumers (electricity data) and the presence values (mobile phone data) in each postal code area for each month were calculated. This part of the analysis answers the question about which occupancy detection method correlates best with mobile phone data, and how the electricity data captures the spatial distribution of second home users for each month.

Based on the results, one occupancy detection method was chosen for the next step to analyze the correlation between the consumers (lower 95% CI occupancy detection method) and the presence values in each month for each postal code area. Through this, it is assessed how the electricity data captures the monthly dynamics of second home users in a postal code area.

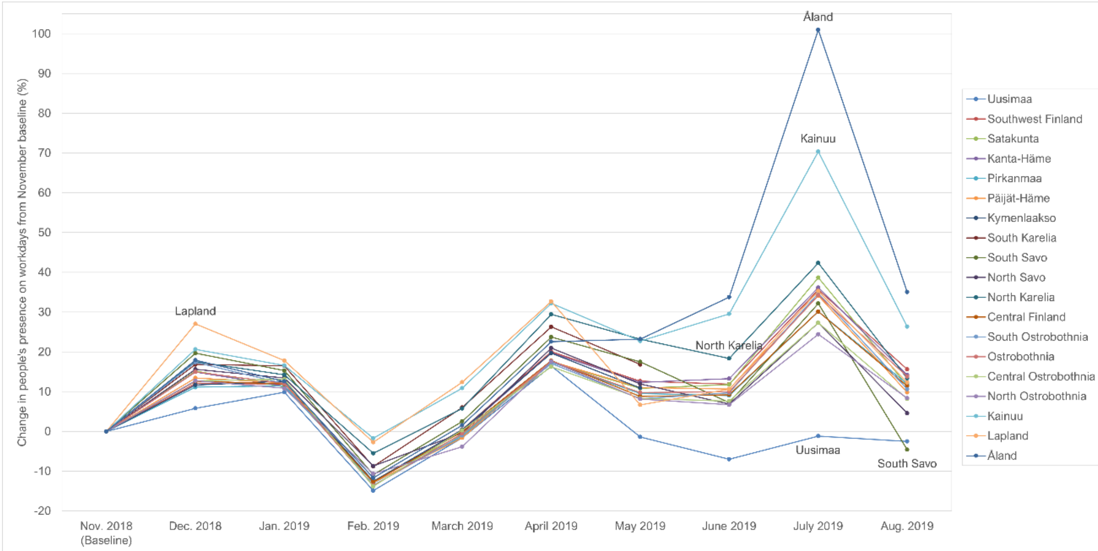
6. Results

6.1. Country level mobile phone data analysis

6.1.1 Change in people's presence in Finland

To illustrate the spatiotemporal differences in people's presence in Finland, Figure 11A and B show the change in people's presence by county during the period 11/2018 – 8/2019 for workdays and weekends. The presence values are shown relative to the established baseline month (November 2018) to make them comparable throughout the study period.

a)



b)

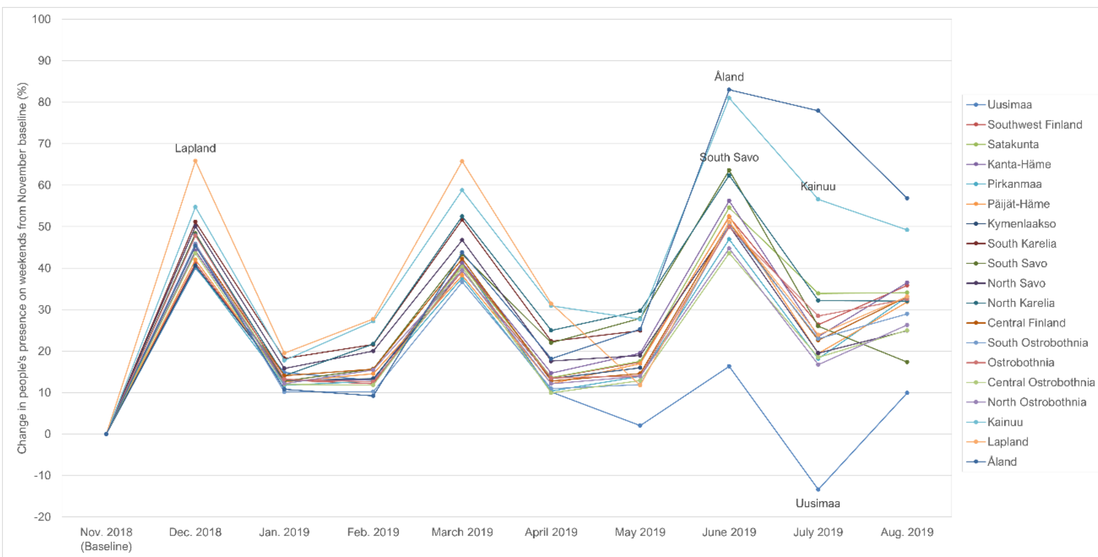


Figure 11. Change in people's presence from the November baseline by county a) for workdays, and b) for weekends

People's presence in most counties peak during the summer, but notably for workdays the peaks appear in July while, for weekends, the peak appears in June. However, for Uusimaa, no distinct summer peak is shown, but in fact, a drop in presence appears for in July for weekends. The most prominent instances of growth in presence appear in the workday summer values for Kainuu and Åland where presence values grow over 100 % from November. Also, for Lapland, distinct peaks for weekends values are shown in December and March, while for workdays, the peak appears in April.

6.1.2. Relation between the change in presence and the number of free-time residences

The change of the Spearman correlation coefficient in all Finnish postal code areas between the change in people’s presence from the baseline and the number of free-time residences (per 1000 inhabitants) is shown in Figure 12A for workdays, and in 12B for weekends. For workdays, the correlation coefficient appears to be quite high (~0.65) in December. In January the coefficient drops significantly and starts increasing during spring and summer months reaching a peak in July (~0.75) with a slight drop in August.

In contrast to the high correlation for workdays in December, for weekends, the correlation is almost non-existent. The coefficient reaches a low point in January (~-0.40) with quite steady growth appearing towards the spring and summer where the coefficient reaches a peak in June (~0.80). The coefficient starts decreasing slightly towards the end of summer months in July and August. There appears to be more drastic seasonal variation in the correlation for weekends compared to workdays.

a) b)

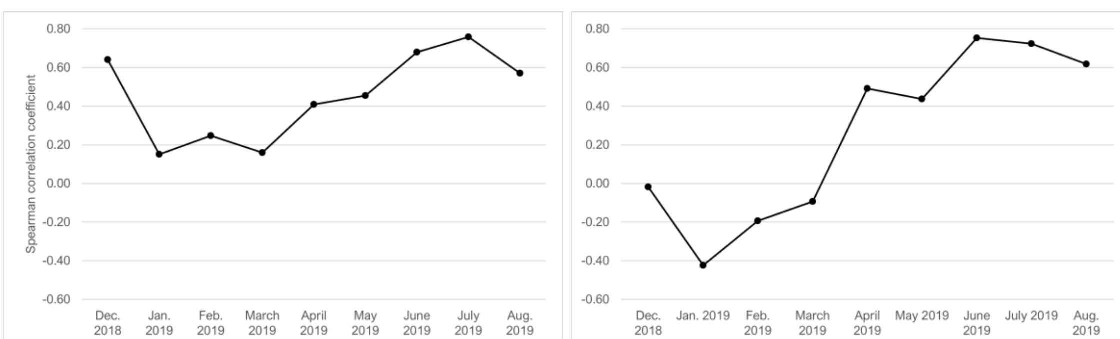
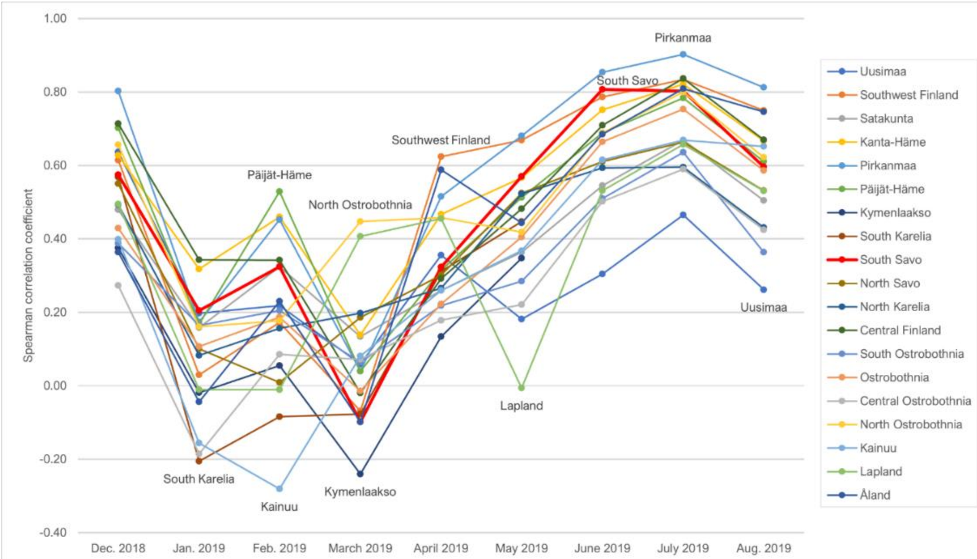


Figure 12. Change of Spearman correlation between change in people's monthly presence from the November baseline and the number of free-time residences per 1000 inhabitants a) for workdays, b) for weekends (all statistically significant, $p < 0.05$)

Similar development in correlation coefficient for each county in Finland is assessed next for workdays (Figure 13A) and weekends (13B). Notably, there are quite drastic differences between different counties. For instance, the correlation coefficients for Lapland in May are significantly lower compared to all other counties. Uusimaa, on the other hand, has the lowest coefficients during summer months in both figures.

a)



b)

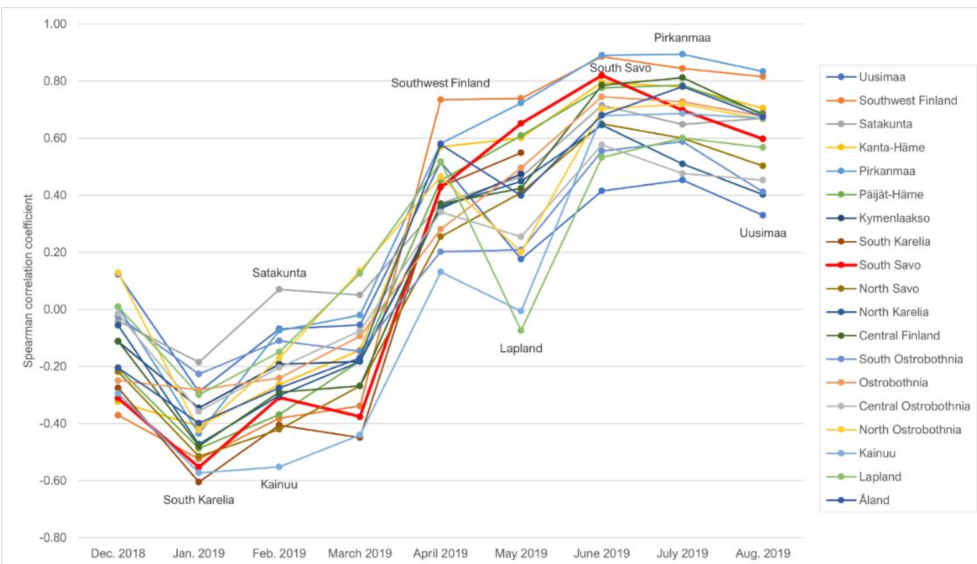


Figure 13. Change of Spearman correlation between change in people's monthly presence from the November baseline and the number of free-time residences per 1000 inhabitants by county a) for workdays, b) for weekends

Figure 14 A-D show the differences in coefficients on a map for workdays and weekends in January and July. What is to note, in addition to the overall differences between the two months, are the differences in the same month between the weekday groups. This shows that there are differences in people’s multi-local practices both between months and between different weekdays.

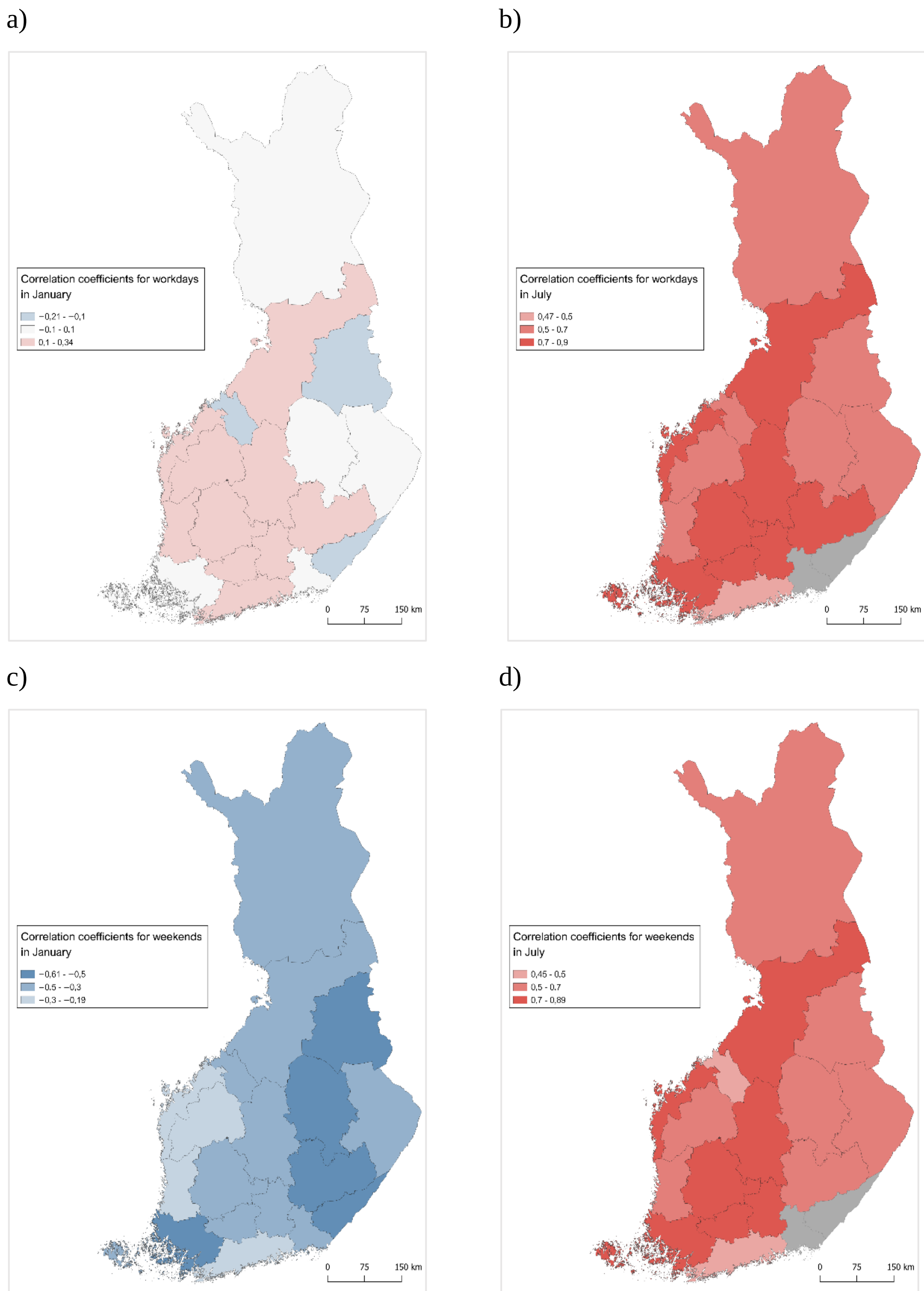


Figure 14. Correlations between the change in people's presence from the November baseline and the number of free-time residences per 1000 inhabitants by county for a) workdays in January, b) workdays in July, c) weekends in January, and d) weekends in July.

6.2. South Savo mobile phone data analysis

6.2.1. Change in people's presence

To gain an understanding for the development in people’s presence in South Savo, three subregions (Mikkeli area, Savonlinna area, and Pieksämäki area) are assessed. The most glaring differences appear during the summer months: postal code areas in the Mikkeli area have the highest peak in presence during the summer. The workday peak occurs in July (Figure 15A), while the weekend peak takes place in June (Figure 15B). For both weekday groups, a high point also appears in December. Overall, higher presence values also appear in February and March for the weekends compared to workdays.

a) b)

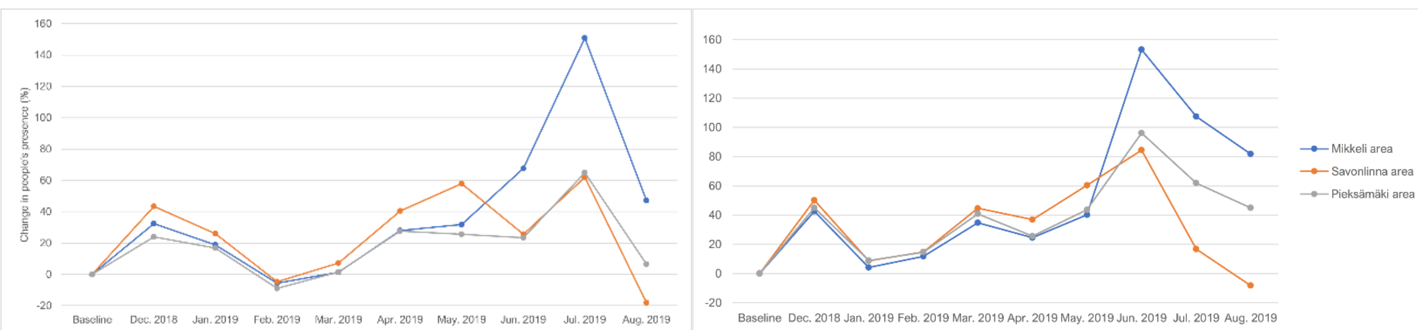


Figure 15. Monthly change in people’s presence compared to the baseline in the subregions of South Savo for a) workdays, and b) weekends

Next, people’s presence is assessed in more detail in South Savo with comparisons of people’s presence in two months (January and July) between workdays and weekends (Figure 16) to gain more perspective on the spatiotemporal dynamics in the county. People’s overall presence is much higher during July than in January, but during both months there are differences between postal code areas during workdays and weekends. The difference between weekday groups appears to be more drastic in January, although some differences can be identified in July as well.

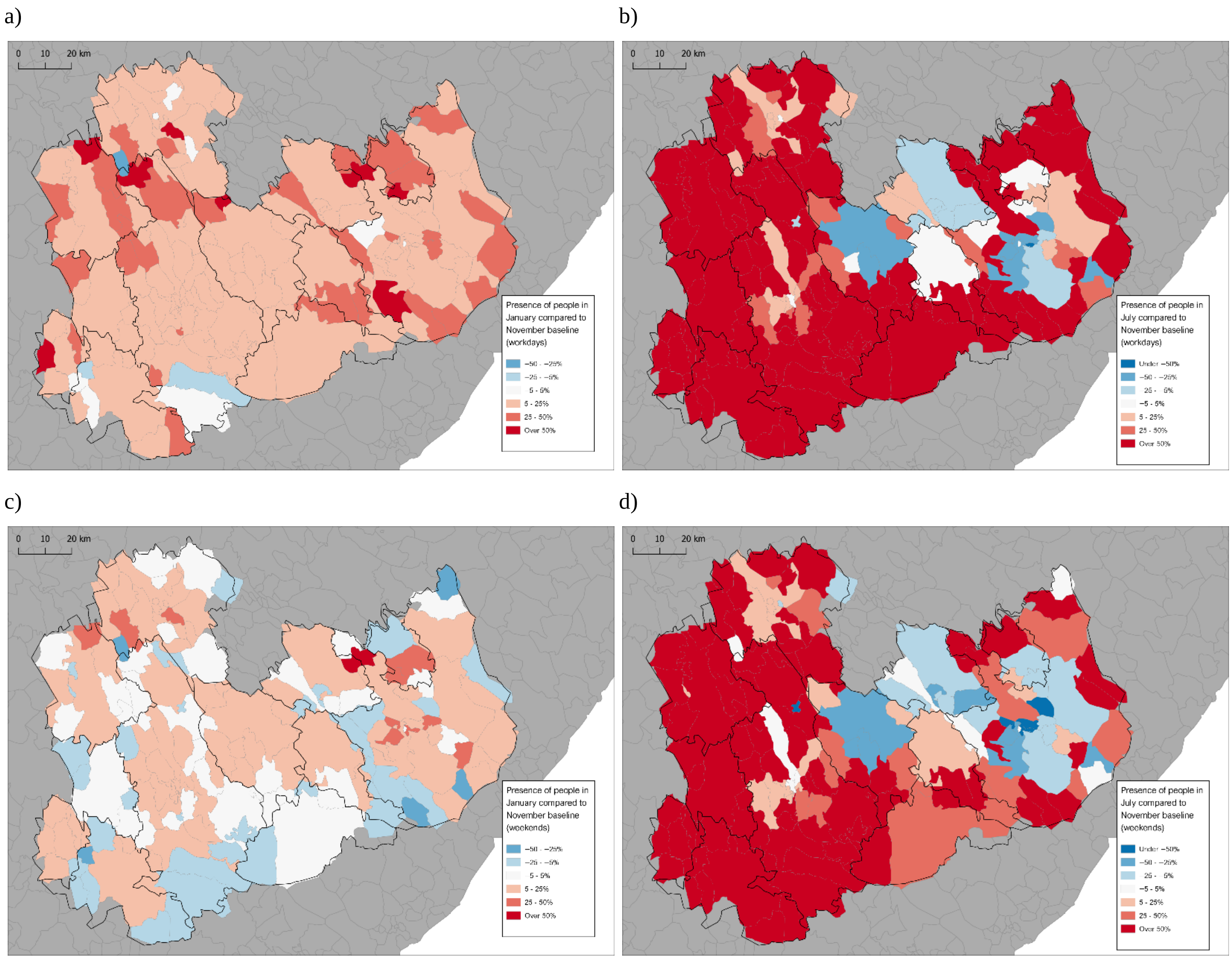


Figure 16. People's presence in South Savo compared to the November baseline on a) workdays in January, b) workdays in July, c) weekends in January, d) weekends in July

6.2.2. Relation between the change in presence and the number of free-time residences

The correlation between the change in people's presence and the number of free-time residences (per 1000 inhabitants) in South Savo is assessed in two ways: For all postal code areas in the county, and by comparing three subregions within the county.

Figure 17 shows the development of the correlation between the change in people's presence from the baseline and the number of free-time residences in South Savo and in the country for comparison. For workdays (17A), the developments of the coefficient for both follow a similar trend, although there is a more prominent low point in March in the South Savo values where the correlation is almost non-existent (~-0.10). Also, the p-value of the correlation in March is not significant ($p > 0.05$), while in all the other months, the correlations were statistically significant ($p < 0.05$). The coefficient reaches a two-month peak during the summer (June and July, ~0.80), while in the case of the whole country, the peak is shown in July.

For weekends (17B), the development for South Savo and the whole country follows a very similar trend. The coefficient reaches a low point in January (~-0.60), next, increasing rapidly as the coefficient changes drastically from February (~-0.40) to March (~0.40) towards the summer peak in June (~0.80). Again, the development of correlation coefficients shows higher seasonal variation for weekends than for workdays.

a) b)

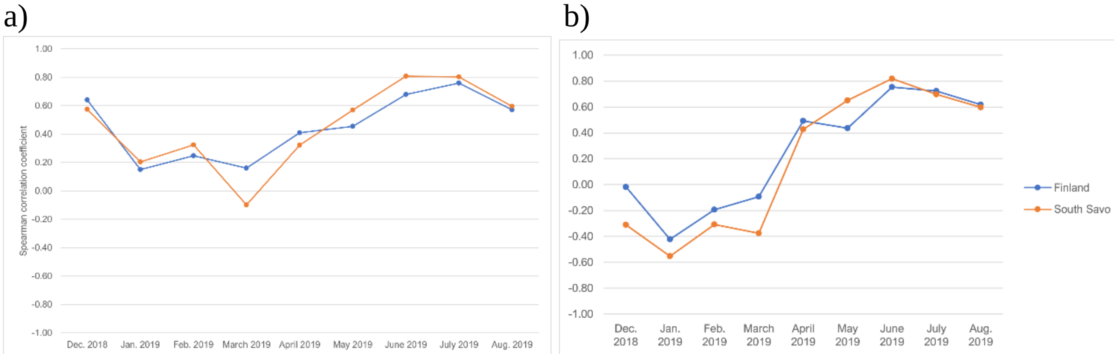


Figure 17. Change of Spearman correlation between change in people's monthly presence from the November baseline and the number of free-time residences per 1000 inhabitants in South Savo compared to Finland overall a) for workdays, b) for weekends (all but March workdays statistically significant, p < 0.05)

The monthly development of the Spearman correlation coefficient between the change in people’s presence and the number of free-time residences in analyzed next for workdays and for weekends in the three subregions of South Savo (Mikkeli area, Savonlinna area, and Pieksämäki area). Overall, the changes between months are the most moderate in the Pieksämäki area, while the changes are more extreme in the other two areas.

For workdays (Figure 18A), the lowest points in the coefficients for all subregions appear in March and the coefficients begin to increase in spring reaching peak values in the summer months. However, there is a distinct increase in the coefficient for Mikkeli area from January to February.

For weekends (Figure 18B), the coefficient is higher for the Pieksämäki area for most of the study period - Mikkeli and Savonlinna coefficients are higher during the summer months. For all areas, there appears drastic growth for the coefficients from March to April and the coefficients remain high throughout the summer.

The statistical significance of the results is also indicated in the figures. If the point has a red outline, the correlation is not statistically significant ($p > 0.05$).

a) b)

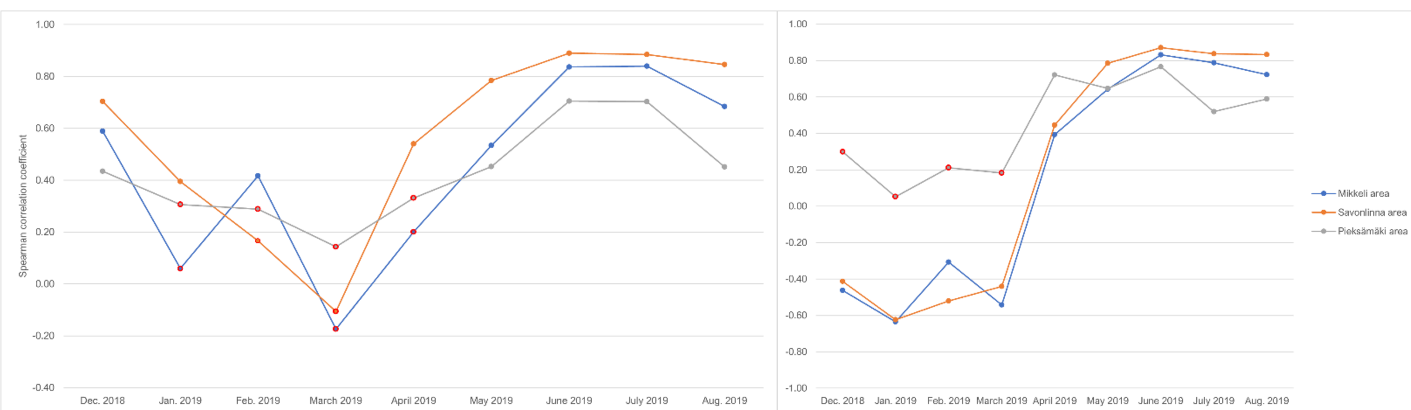


Figure 18. Change of Spearman correlation between change in people's monthly presence from the November baseline and the number of free-time residences per 1000 inhabitants in the subregions of South Savo a) workdays, b) weekends

6.2.3. Origins of people

The proportions of origin counties for people who have been registered as being present in South Savo are examined to understand people's mobility to the county. The proportions of people's origins during the baseline month are assessed along with differences of origin county proportions from the baseline. The differences throughout the study period are examined for the eight counties that had the highest proportion during the baseline month.

During workdays (Figure 19A), over 85% of people in South Savo have their origins in the same county. North Savo has the second highest share of origins at ~4% with all the other counties following. Notably, there are differences in the proportions of origins throughout the year (Figure 19B). For instance, in December, the proportion of people from Uusimaa is over 80% higher compared to the proportion in November. Moreover, the differences in proportions are overall higher during spring and summer. For example, the proportions of Päijät-Häme and North Karelia origins are over 150% higher in July compared to November, and the proportions of Uusimaa origins are ~500% higher.

a) b)

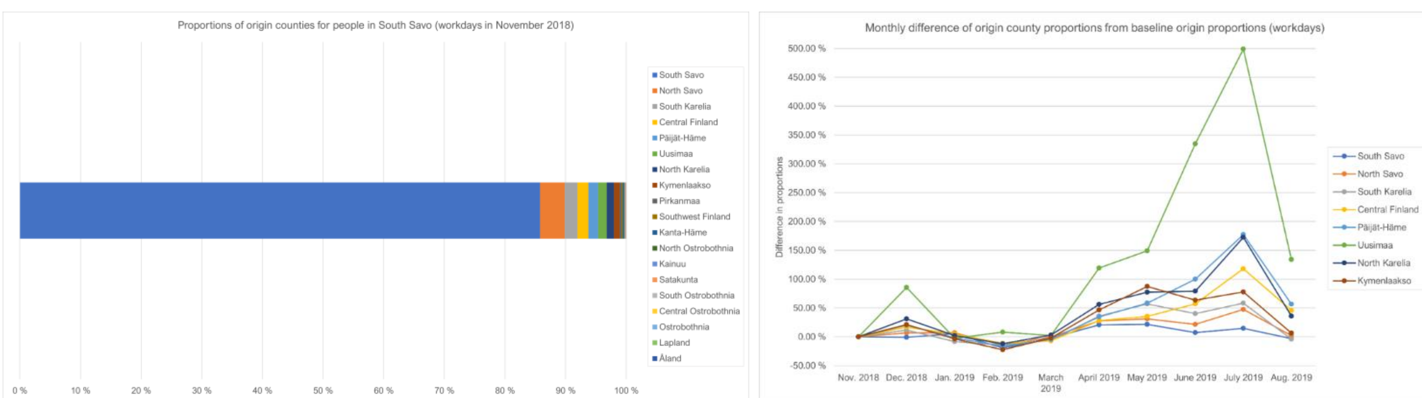
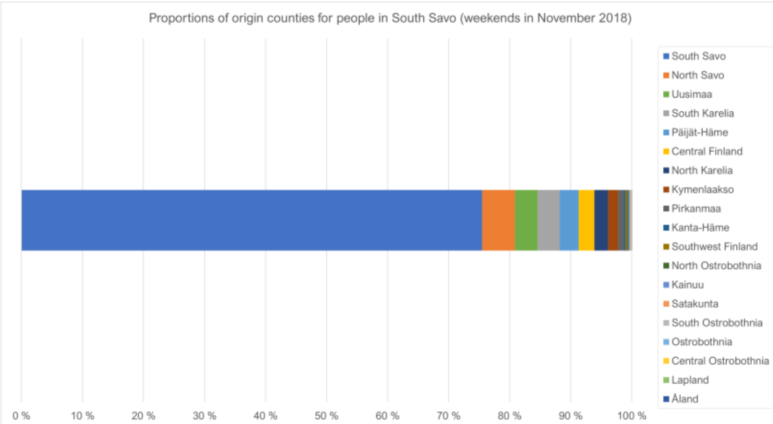


Figure 19. Proportions of origin counties of people in South Savo for workdays a) during the baseline month (November), b) monthly differences from the baseline for 8 counties

a)



b)

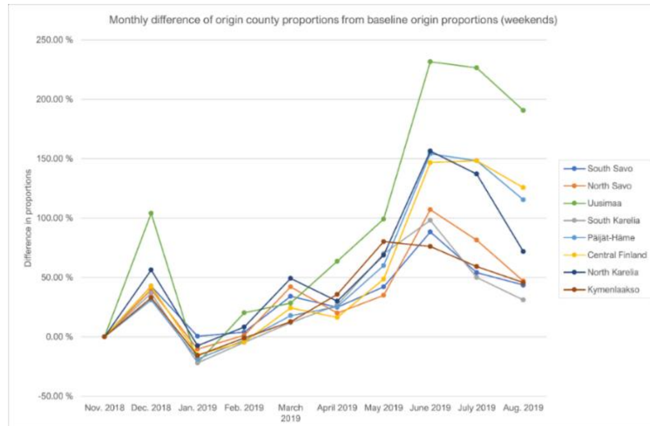


Figure 20. Proportions of origin counties of people in South Savo for weekends a) during the baseline month (November), b) monthly differences from the baseline for 8 counties

The proportion of people whose origin is in South Savo during the baseline month is ~75%, so during the weekends there are more people visiting from other counties compared to the workdays (Figure 20A). North Savo, Uusimaa, and South Karelia follow South Savo with the next highest proportions of origins. The differences in origin proportions from the baseline month (Figure 20B) also show differences throughout the period. For instance, the proportion of Uusimaa is over 100% higher in December compared to the baseline, while the proportion is over 200% higher in June. Overall, the differences in proportions are higher in the summer months with the largest differences from the baseline being in June.

Compared to the workdays, the peaks in the change of origin county proportions are less extreme. For instance, during the summer months, the origin proportions from counties stay consistently high, while the workday proportions show significant high points during July specifically.

6.2.4. Clusters based on the change in presence

The postal code areas in South Savo were clustered based on the change in people's presence from the baseline for both workdays and weekends. This resulted in three types of clusters: areas where people's presence peaks extremely in the summer, areas with a high peak in presence in the summer, and areas where no distinct peak is evident.

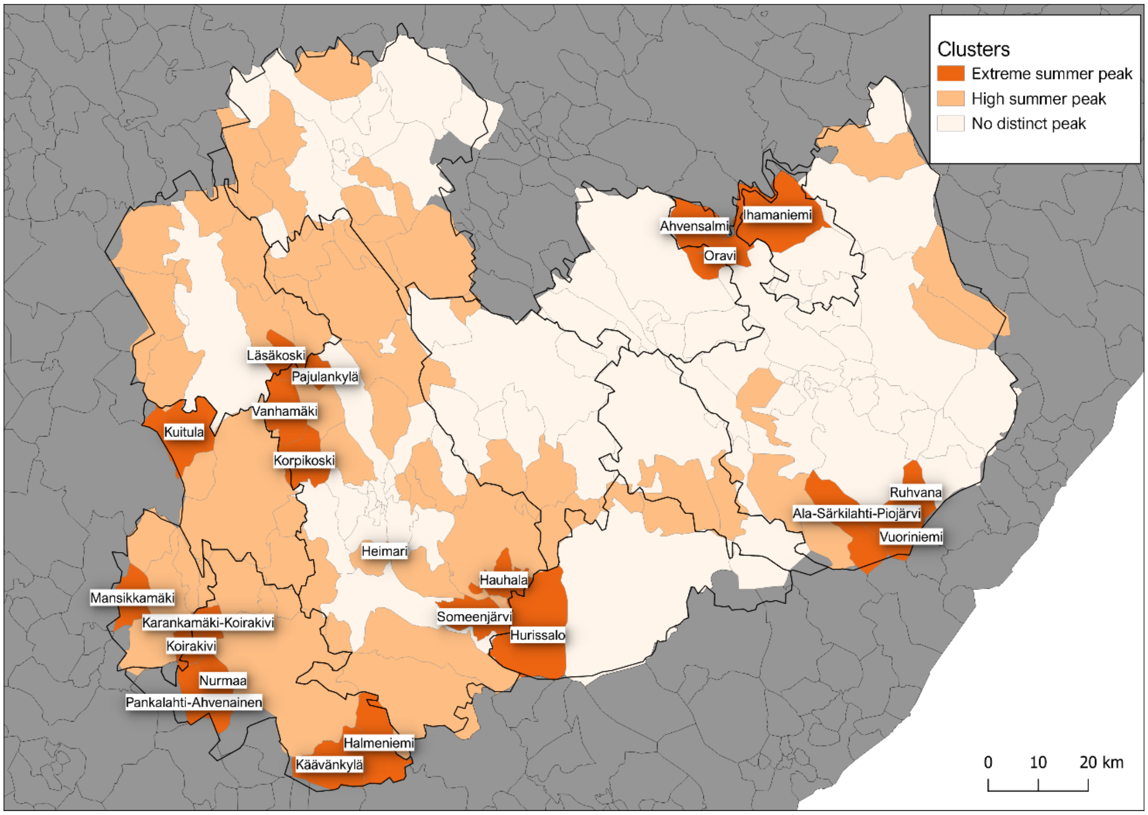
Table 4. Share of postal code areas in the clusters for workdays and weekends

	Share of postal code areas in the cluster (workdays)	**Share of postal code areas in the cluster (weekends)**
Extreme summer peak	13.8% (22 areas)	15.1% (24 areas)
High summer peak	36.5% (58 areas)	40.3% (64 areas)
No distinct peak	49.7% (79 areas)	44.7% (71 areas)

Most of the postal code areas in both weekday groups were assigned to the "No distinct peak" cluster, while the least postal code areas were assigned to the "Extreme summer peak" cluster (Table 4). However, for weekends, more postal code areas were assigned to the summer peak clusters compared to workdays.

For workdays, the areas in the extreme summer peak clusters are located mostly in the western, north-eastern, and southern parts of South Savo. Whereas most of the areas in the "No distinct peak" cluster are in the east (Figure 21A). The peak in people's presence for workdays occurs in July for the summer peak clusters, whereas for the third cluster has only moderate change in presence throughout the study period (Figure 21B)

a)



b)

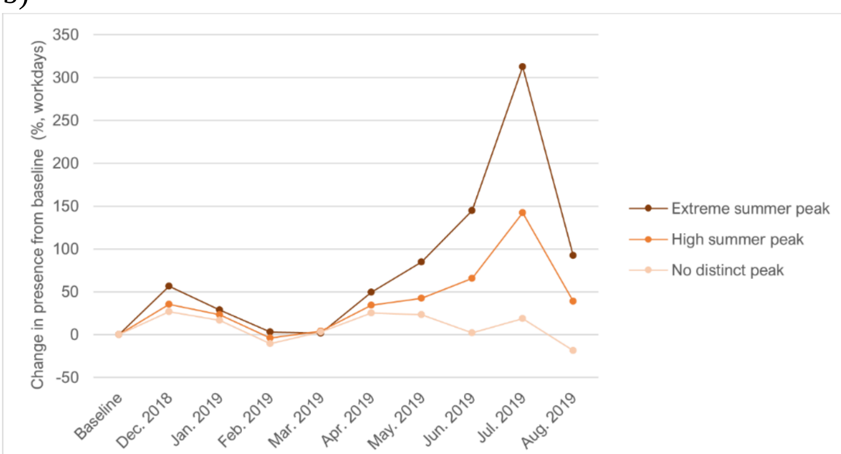
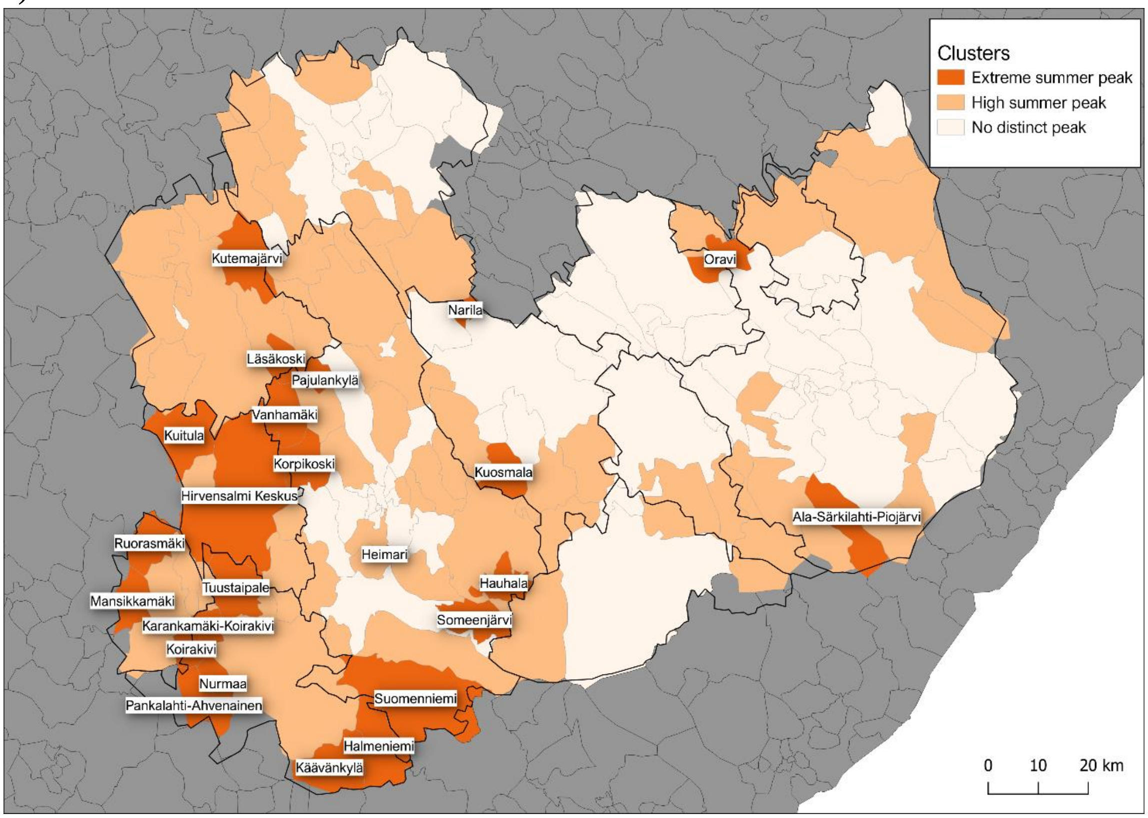


Figure 21. a) Postal code areas by cluster type based on the change of people's presence from the baseline, b) Change in presence from the baseline by cluster (workdays)

For weekends, the spatial distribution of areas by cluster follows a similar pattern to the workdays (Figure 22A), but there are more areas classified to the summer peak areas in the west of South Savo. However, some of the areas in the north-eastern and south-eastern parts of South classified as the "Extreme summer peak cluster" differ in classifications here.

Notably, the peaks in people’s presence occur in a different month compared to the workdays as June is the peak month for the summer peak clusters (Figure 22B).

a)



b)

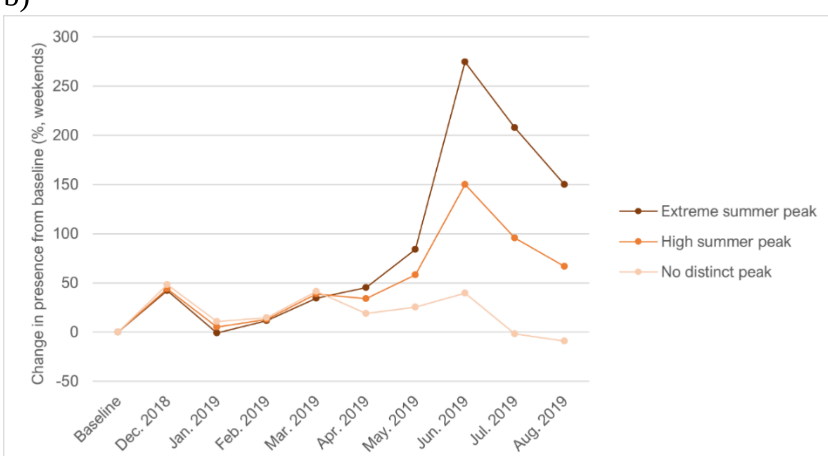


Figure 22. a) Postal code areas by cluster type based on the change of people’s presence from the baseline, b) Change in presence from the baseline by cluster (weekends)

6.3. Comparison of electricity consumption and mobile phone data

The representativeness of the electricity consumption data to examine second homes has been previously explored in the MOPA project (Ruralia Institute, 2021) as it was determined that the data correlates strongly with Statistics Finland data in terms of the number free-time residences. To assess the different occupancy detection methods used in the aggregation of the electricity consumption data, consumers and use rates were compared to mobile phone presence data for postal code areas in each month. The consumers-variable (total number of properties occupied) was chosen as the most appropriate variable in initial testing. Mobile phone data were compared using Spearman correlation analysis with the use rates and consumers (aggregated with the lower 95% confidence interval method) for January and July. The consumers-variable correlated better with the presence values both in January (consumers: 0.314, use rate: -0.005), and in July (consumers: 0.456, use rate: 0.244). Furthermore, the p-values for use rates were only significant ($p < 0.05$) in July, but not in January, whereas for consumers, the results were statistically significant in both months.

The results of the Spearman correlation analysis between the consumers (electricity data) and the presence values (mobile phone data) in each postal code area for each month are shown in Table 5. Based on the results, the lower 95% confidence interval method correlates slightly better with the mobile phone data compared to the other methods. Consequently, the best correlating method was chosen for the next step of the analysis.

Table 5. Spearman correlation coefficients between people's presence (mobile phone data) and consumers (electricity consumption data) for different months (all statistically significant, p < 0.05)

	Nov. 2018	Dec. 2018	Jan. 2019	Feb. 2019	March 2019	April 2019	May 2019	June 2019	July 2019	Aug. 2019	Mean
>0 kWh	0.312	0.313	0.285	0.314	0.307	0.334	0.367	0.444	0.454	0.412	0.354
>2.5 kWh	0.306	0.308	0.277	0.311	0.303	0.332	0.364	0.445	0.454	0.413	0.351
>lower 95% CI	0.313	0.315	0.314	0.309	0.336	0.369	0.447	0.456	0.456	0.415	0.373
>5 kWh	0.306	0.306	0.276	0.308	0.301	0.331	0.361	0.444	0.454	0.413	0.350
>2.5 kWh & >lower 95% CI	0.310	0.311	0.277	0.311	0.307	0.334	0.367	0.447	0.455	0.415	0.353

Next, the monthly dynamics of second home users were assessed by correlating the presence values in each month for each postal code area with the consumers-variable for the lower 95% confidence interval method. Based on this, the postal code areas were ranked by the Spearman correlation coefficients. Notably, 58% of the postal code areas had strong correlation (> 0.75), while 9% of areas correlated negatively (Figure 23).

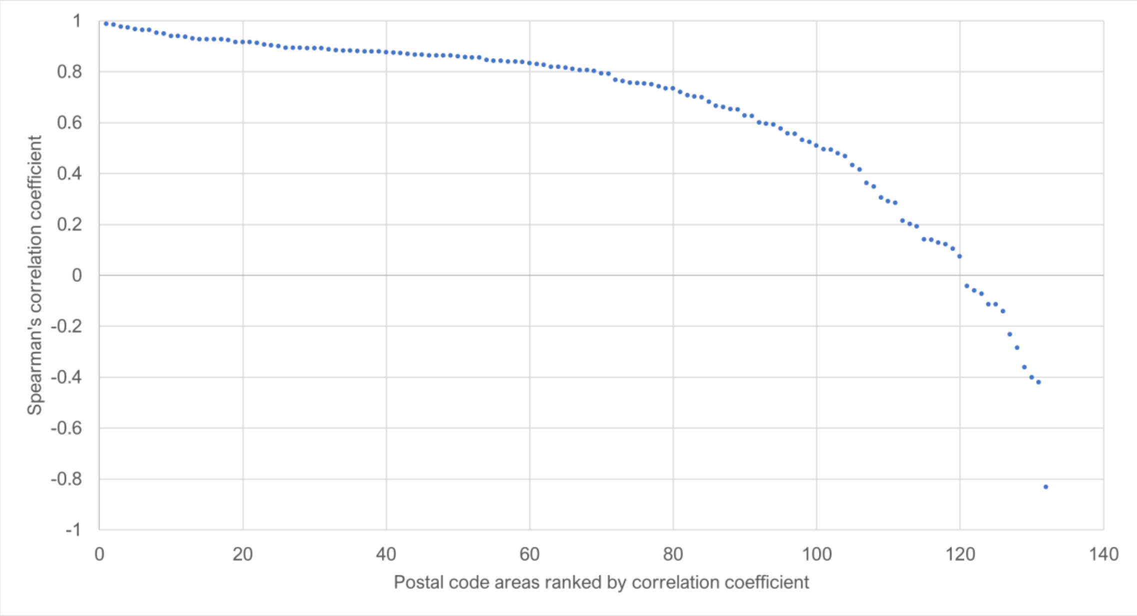


Figure 23. Postal code areas in South Savo ranked by the correlation between people's presence (mobile phone data) and consumers (electricity consumption data)

Most of the postal code areas classified in the higher classes of correlation coefficients are in the western and southern parts of South Savo while, for most of the eastern areas, either the correlation is classified in the lower positive classes or negative ones (Figure 24).

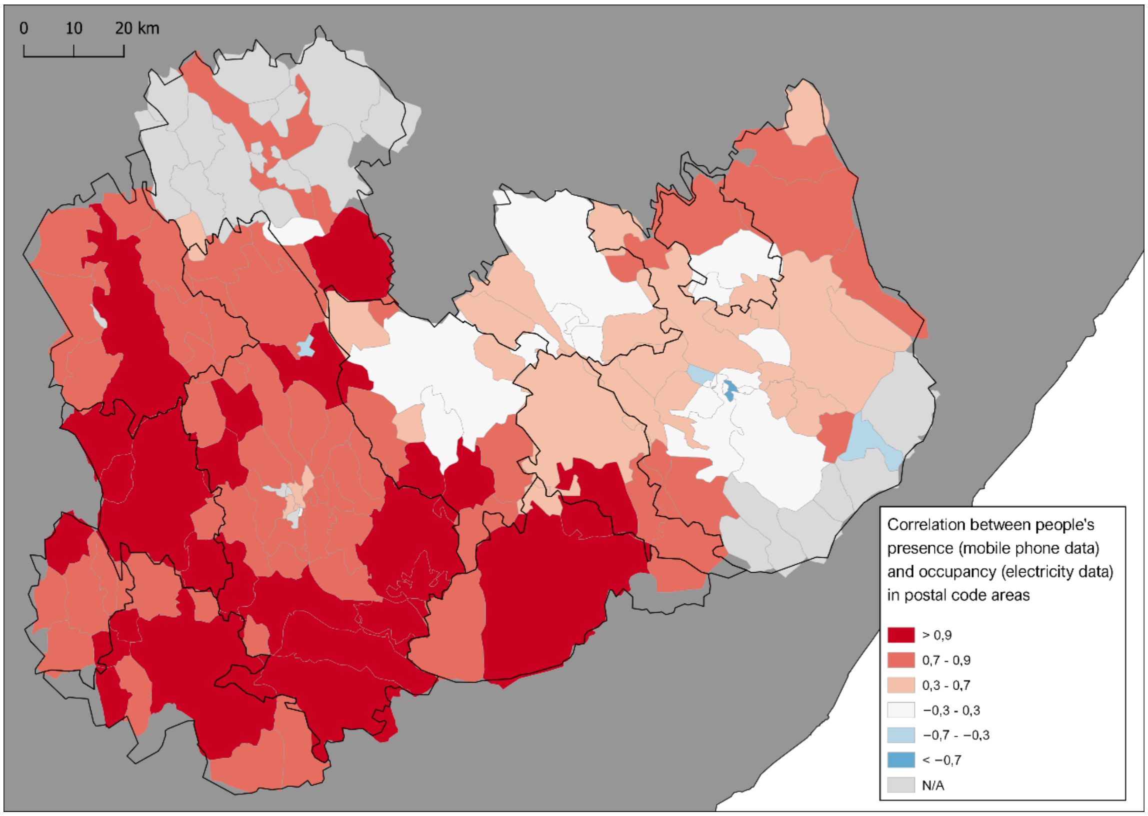


Figure 24. Postal code areas ranked by Spearman correlation

7. Discussion and conclusions

7.1. Mobile phone -based analysis

7.1.1. Presence of people and correlation with free-time residences

Mobile phone activity data was analyzed to answer the first research question of the thesis: *What can mobile phone activity data reveal about people's presence and multi-local living in Finland and in South Savo?* People's presence was assessed solely with the mobile phone data. To examine how the results relate to multi-local living, correlation analyses with official statistics for the number of free-time residences were performed.

People's overall presence derived from mobile phone data varies spatially and temporally across counties and months. Comparisons between counties reveal that people move within the country more during the summer most likely due to warmer weather and summer vacations. Furthermore, people's mobility varies between workdays and weekends due to commuting patterns and weekend mobilities differing from each other. As July is the most popular summer vacation month in Finland, people's presence during workdays peaks in various counties during the month. High presence across counties is shown throughout warm summer months for weekends. This indicates that there are more consistent mobilities between counties during weekends. Mobile phone data also captures other seasons. People's presence in Lapland is shown to be high during the Christmas season and winter holidays in March.

Similar findings were also detected for South Savo across three subregions in the county as presence peaks are evident in the summer and Christmas time. More dramatic growth in presence is shown in the Mikkeli region during the summer compared to other subregions indicating that the area is a more popular summer destination. Furthermore, spatially more accurate presence comparison between winter and summer revealed the postal code areas in which presence changes are most evident. Distinctly, higher presence is detected in the western parts of South Savo compared to the east.

The examinations of people’s presence both in the country and in South Savo raise the question about how the changes in presence are tied to multi-local living. This was explored by comparing the presence values derived from mobile phone data to the number of free-time residences from official statistics.

Overall, high correlations between the change in people’s presence and the number of free-time residences were detected for the summer months both at the national scale and in South Savo. For workdays the correlations are consistently low during winter months, while negative correlation showed extreme seasonal variation for weekends. These variations show that the growth in presence in the summer is more connected to free-time residences than in winter months, which indicates a prominent seasonal effect for multi-locality from a free-time residence perspective.

Furthermore, correlation and presence values for Uusimaa were consistently lower compared to other counties in the summer, whereas counties considered more rural showed higher correlation and peaks in presence values. This is in line with prior research indicating that second homes are commonly located in the countryside (Hiltunen & Rehunen, 2014; Pitkänen, 2008; Pitkänen et al., 2020). It appears that people from Uusimaa especially tend to “escape” their urban homes to rural areas during the summer. Taking these mobilities and the seasonal population (Adamiak et al., 2017) into account in public planning could prove to be useful for both the local community and the seasonal population. However, recent research by Greinke & Lange (2022) showed that due to multi-locals being present in areas for a limited time, it is unlikely for them to become locally involved.

While the data period for the mobile phone data is dated before the COVID-19 pandemic, it is important to note that recently the boundaries between primary and secondary residences have become less distinct (Lehtonen et al., 2020), and that the seasonal rhythms in people’s presence may change due to remote working (Lehtonen & Kotavaara, 2021). Notably, Willberg et al. (2021) showed that people’s presence in workplace concentrations decreased while second home concentrations saw the most increase in people’s presence during the first wave of the pandemic. However, it is yet to be seen if this pandemic-induced change in people's dwelling represents a general course of development or a more isolated instance (Lehtonen & Kotavaara, 2021).

7.1.2. Origins of people in South Savo

One of the key advantages of mobile phone data is that, in addition to detecting people's activities, it enables analyses of people's origins. For South Savo, the analysis highlighted differences in people's origin counties, and additionally, mobility differences between workdays and weekends. For weekends during the baseline month, there was a higher proportion of people originating from different counties than for workdays indicating that more people not local to the county tend to visit then. Analysis for other study months highlights that people are more mobile during the summer as the proportions of origins from various counties are manyfold compared to the baseline.

While previous research has shown that people may spend considerable amounts of time to access their second homes (Hall, 2014), the results demonstrate that the highest proportions of origins were nevertheless from counties close to South Savo. This indicates that more people tend to come to South Savo from counties that are close. There are, however, differences in the origins from the neighboring counties as well. For instance, counties close to South Savo to the south and west had a higher proportion of origins than North Karelia which borders the county to the east. This may explain the higher presence values detected in the western postal code areas of South Savo as people from counties to the south and west of South Savo seem to visit the county more.

However, the origins in this case indicate the location of people from the previous night, so it does not capture the length of the stay for people visiting. In other words, a person will become a local to South Savo according to the data once they spend more than a day there. As people tend to use their second homes with varying usage patterns (Czarnecki et al., 2021), the length of stays at second homes would be important to capture when assessing people's origins. However, due to the format of the mobile phone data, this was not possible.

7.1.4. Clusters based on people's presence in South Savo

Clustering postal code areas in South Savo based on the change in people's presence resulted in classifications that vary between their presence peaks in the summer. While most of the high summer presence areas in the western and southern parts of the county in the Mikkeli subregion, some exceptions are evident. For instance, some postal code areas in the Savonlinna subregion (e.g., Figure 21A), are classified in the summer peak clusters. The peaks in north-eastern postal code areas can be explained as there is a popular spa resort called Järvisydän in the area. Whereas summer peak areas in the south-east could be explained by there being popular holiday home resorts there (e.g., Tynkkylän Lomaniemi). This highlights that specific tourist destinations have a substantial effect people's mobilities, and that the locations of these "hot spots" can be captured with mobile phone data.

Additionally, the proximity to large cities such as Helsinki, Tampere, and Jyväskylä may affect people's presence in postal code areas as city dwellers "escape" their second homes in the summer. Another explaining factor could be the distribution water bodies since usually free-time residences in Finland are located by lake shores (Hiltunen & Rehunen, 2014). Some of the postal code areas classified in the no distinct peak -cluster appear to be areas with few lakes. However, further research is needed to examine these factors in more detail.

7.2. Comparison between mobile phone data and electricity consumption

Mobile phone data and electricity consumption data were compared in the thesis to answer two research questions: "*How do different occupancy detection methods for electricity consumption data correlate with mobile phone presence data?*" and "*How does electricity consumption data capture the spatiotemporal dynamics of second home users in South Savo?*"

In this thesis, presence derived from mobile phone data is considered the ground truth that includes all activities from people in each postal code area, whereas electricity consumption data only includes information about the second homes. This means that the two datasets have unique information about people's whereabouts and patterns. The correlation analysis revealed the best performing occupancy detection method (95% confidence interval), in which consumption data from multiple years has been used to detect the true population mean instead of using a simple threshold value for detection. This information can be utilized in the MOPA research project going forward. Furthermore, results assessing the temporal dynamics of second home users showed that the electricity consumption data correlated best with mobile phone data during summer months. This further highlights the seasonal dynamics in people's use of free-time residences.

Assessing the monthly dynamics of second home users in the second correlation analysis between mobile phone data and electricity consumption data demonstrated that most of the postal code areas correlated positively while only a small share of the areas correlated negatively. The areas with high correlation are in the western and southern parts of South Savo in the Mikkeli subregion.

These results are in line with the results of the presence comparison and clustering analysis as most of the high summer presence areas were in the west and south of the county. Moreover, this ties to the results gained from the origin analysis where it was shown that more people visit South Savo from counties to the west and south of South Savo than from the east. As the consumption data only includes electrified properties classified as second homes, houses with no electricity are not included in the analysis. Consequently, there may be regional differences in the distribution of electrified second homes, which may influence the results.

7.3. Data and analysis methods

For the thesis, mobile phone data proved to be a suitable data source for studying people's presence and mobility as it enabled a relatively long study period with extensive coverage of population. The affirms the reasonings given by Wang et al. (2018) and Järv et al. (2014) on the merits of novel data sources for supplementing or replacing more traditional methods in mobility research. Furthermore, as activity data enabled studying people's presence with multiple spatial and temporal aggregations. For instance, due to the presence values being recorded at the postal code level, it allowed for accurate spatial analyses while the values could also easily be generalized to municipality or county-level.

Moreover, two temporal levels of classification were used to assess people's presence: monthly level data and information about different weekdays within these months. It would have been possible to also use the data in its original form at the hourly level. The weekday/monthly level was deemed appropriate for the purposes of the thesis.

Mobile phone data was used in the thesis to assess multi-local living through correlation analyses with official statistics data on free-time residences. The analyses showed the suitability of mobile phone data in capturing the seasonal dynamics of people's presence and its connections to multi-local living as growth in presence during holiday seasons at various spatial scales were detected. Particularly for the summer, the changes in presence showed strong correlation to the number of free-time residences. This supports findings from previous research that has shown the suitability for the use of mobile phone data to assess the seasonal variability population (e.g., Silm & Ahas, 2010).

Instead of mobility, activity data here represents "stillness" as people's presence is registered in a postal code area once they spend 20 minutes there. This allows for analysis of presence dynamics in Finland. But to fully assess people's mobility, locals had to be eliminated from the data. This elimination was conducted in the analysis of people's origins. Locals, or people whose origin and destination is in the same postal code area, were overrepresented in the data since people's presence in a postal code area is registered regardless of them constantly being in the area rather than moving to the area. This, however, did not influence the overall results in presence, but had to be dealt with in the analysis of people's origins.

The mobile phone activity dataset is a commercial product by Telia Crowd Insights in which spatiotemporal aggregations and anonymization have already been performed by the data provider. These aggregations made it easier to understand the structure of the data, but also restricted the possibilities for customized spatial and categorical aggregations that would have been possible with raw data. As the methodology behind the data generation is concealed by the data provider, it may restrict the full understanding needed to fully grasp its potential. This, in turn, may limit the use of the data in scientific research (Poom et al., 2020). For instance, because of the concealed methods behind the data, abnormalities and erroneous data spotted in the beginning of the analysis needed to be removed as the reasons behind the abnormalities were unclear.

To make the presence values comparable throughout the study period, it was necessary to normalize the presence values for each study month based on the number of days in each workday group and to compare the normalized values to an established baseline period. These modifications allowed for time series analysis that would not have been as scientifically valid with the "raw" presence values. This further highlights the effect of aggregations for the validity of the use of this data.

Electricity consumption data used in this thesis was previously generated in the MOPA research project. Notably, the validity of the data for the purposes of second home research has been explored previously in the research project as the number of free-time residences in the electricity data correlated strongly with official statistics. Also, based on the customer groupings by the data provider, second homes had previously been selected from the data and electricity consumption patterns were used to further analyze the use patterns of second home users. The use of electricity consumption patterns is one of the examples given by Scheidt et al (2020) on the potential of implementing data analytics in the electricity sector.

The purpose for the correlation analyses between mobile phone data and electricity consumption was to assess the performance of different occupancy detection methods. Furthermore, as previous research (Andersen et al., 2008; Sun et al., 2018) has shown the potential for the use of electricity consumption data to examine population dynamics at various spatial scales, it was assessed how the data can capture the spatiotemporal dynamics of second home users in South Savo. The correlation analyses with mobile phone presence data revealed the best performing occupancy detection method and showed that the data captures the temporal dynamics of second home users well. While the spatial dynamics of second home users for each month had lower correlation, the performance was overall better for summer months indicating good suitability for the research of seasonal population and multi-local living.

7.4. Possibilities for future research

This thesis and other studies using mobile phone data (e.g., Willberg et al., 2021) have demonstrated the suitability of the data for studying multi-local living. The study period for this thesis was limited because it was carried out within the confines of the mobile phone data. To assess the dynamics of multi-local living and the spatiotemporal changes in people's presence more comprehensively, analyzing mobile phone data from a longer time would provide a potential future research avenue. A study period assessing multiple years pre-pandemic together with data from pandemic times and post-pandemic periods could yield valuable insight into the long-time implications of change in people's behavior. Furthermore, analyzing data on a daily level instead of monthly level could provide valuable information for local stakeholders and planners through revealing highly accurate patterns on the behavior of tourists and multi-local residents. This temporally accurate perspective could be taken with both mobile phone data and electricity consumption data.

The results of this thesis and the MOPA research project in general provided valuable insight on second homes and multi-local living for local stakeholders, such as decision-makers and local businesses in South Savo. Consequently, the results can be utilized to provide better services for both the locals and the multi-local population. Moreover, country level results presented in the thesis provide valuable insight into the seasonal variability of population and its connection to multi-local living in Finland.

8. Acknowledgements

This thesis is a part of the research project conducted by the Ruralia Institute in collaboration with the Digital Geography Lab called *MOPA (The rhythms, places, and customer groups of multi-local dwelling)*. Thank you to everyone involved in the MOPA project who have shared their knowledge with me. Moreover, thank you to Telia Crowd Insights and Suur-Savon Sähkö Oy for providing the data for this thesis.

I would like to thank my supervisors for their valuable guidance throughout the thesis process, and especially for helping me push through the final analyses.

Thank you to everyone who have supported me during this process, particularly to Matias and Eero who were there every step of the way. A special shoutout to everyone at the geography sofas for making sure that there was never a shortage of quizzes to be taken. Most importantly, thank you to Varpu for everything. Especially for putting up with me nonsensically ranting about data aggregations and the like.

9. Literature

Adamiak, C., Pitkänen, K., & Lehtonen, O. (2017). Seasonal residence and counterurbanization: the role of second homes in population redistribution in Finland. *GeoJournal*, *82*(5), 1035–1050. https://doi.org/10.1007/s10708-016-9727-x

Andersen, F. M., Christensen, M. S., Jensen, O. M., Kofoed, N. U., & Morthorst, P. E. (2008). Second-home electricity consumption. *Energy Policy*, *36*(1), 280–289. https://doi.org/10.1016/j.enpol.2007.09.013

Back, A., & Marjavaara, R. (2017). Mapping an invisible population: the uneven geography of second-home tourism. *Tourism Geographies*, *19*(4), 595–611. https://doi.org/10.1080/14616688.2017.1331260

Czarnecki, A., Dacko, A., & Dacko, M. (2021). Changes in mobility patterns and the switching roles of second homes as a result of the first wave of COVID-19. *Journal of Sustainable Tourism*. https://doi.org/10.1080/09669582.2021.2006201

de Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the Crowd: The privacy bounds of human mobility. *Scientific Reports*, *3*. https://doi.org/10.1038/srep01376

Greinke, L., & Lange, L. (2022). Multi-locality in rural areas–an underestimated phenomenon. *Regional Studies, Regional Science*, *9*(1), 67–81. https://doi.org/10.1080/21681376.2021.2025417

GSMA Intelligence. (2021). *The Mobile Economy 2021*.

Hall, C. M. (2014). Second Home Tourism: An International Review. *Tourism Review International*, *18*(3), 115–135. https://doi.org/10.3727/154427214x14101901317039

Hiltunen, M. J., & Rehunen, A. (2014). Second home mobility in Finland: Patterns, practices and relations of leisure oriented mobile lifestyle. *Fennia*, *192*(1), 1–22. https://doi.org/10.11143/8384

Jansen, R., Kovacs, K., Esko, S., Saluveer, E., Sõstra, K., Bengtsson, L., … Arai, A. (2021). Guiding principles to maintain public trust in the use of mobile operator data for policy purposes. *Data & Policy*, *3*.

Järv, O., Ahas, R., & Witlox, F. (2014). Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies*, *38*, 122–135. https://doi.org/10.1016/j.trc.2013.11.003

Järv, O., Tominga, A., Müürisepp, K., & Silm, S. (2021). The impact of COVID-19 on daily lives of transnational people based on smartphone data: Estonians in Finland. *Journal of Location Based Services*, *15*(3), 169–197. https://doi.org/10.1080/17489725.2021.1887526

Kavousian, A., Rajagopal, R., & Fischer, M. (2013). Determinants of residential electricity consumption: Using smart meter data to examine the effect of climate, building characteristics, appliance stock, and occupants' behavior. *Energy*, *55*, 184–194. https://doi.org/10.1016/j.energy.2013.03.086

Kitchin, R. (2014). *The data revolution: Big data, open data, data infrastructures and their consequences*. Sage.

Lehtonen, O., & Kotavaara, O. (2021). Havaintoja ihmisten liikkuvuudesta ja nettomuuttoliikkeestä koronapandemian aikana. *Maaseutututkimus*, *29*(1), 32–59.

Lehtonen, O., Kotavaara, O., Muilu, T., Vihinen, H., & Huovari, J. (2020). The seasonal population diversifies the picture of development of the regional structure. *Terra*, *132*(2), 69–84. https://doi.org/10.30677/terra.85022

Lehtonen, O., Muilu, T., & Vihinen, H. (2019). Multi-local living-An opportunity for rural health services in Finland? *European Countryside*, *11*(2), 257–280. https://doi.org/10.2478/euco-2019-0013

Müller, D. K. (2007). Second Homes in the Nordic Countries: Between Common Heritage and Exclusive Commodity. *Scandinavian Journal of Hospitality and Tourism*, *7*(3), 193–201. https://doi.org/10.1080/15022250701300272

Müller, D. K. (2021). 20 years of Nordic second-home tourism research: a review and future research agenda. *Scandinavian Journal of Hospitality and Tourism*, *21*(1), 91–101. https://doi.org/10.1080/15022250.2020.1823244

Official Statistics of Finland (OSF). (2019). Use of information and communications technology by individuals, Suomalaisten internetin käyttö 2019. *Use of Information and Communications Technology by Individuals, Suomalaisten Internetin Käyttö 2019*. Retrieved from https://www.stat.fi/til/sutivi/2019/sutivi_2019_2019-11-07_kat_001_fi.html

Official Statistics of Finland (OSF). (2021a). Buildings and free-time residences, Number of free-time residences by region 1970-2020. *Buildings and Free-Time Residences, Number of Free-Time Residences by Region 1970-2020*. Retrieved from http://www.stat.fi/til/rakke/index_en.html

Official Statistics of Finland (OSF). (2021b). Dwellings and housing conditions. *Dwellings and Housing Conditions*. Retrieved from http://www.stat.fi/til/asas/index_en.html

Official Statistics of Finland (OSF). (2021c). Free-time residences 2020. *Buildings and Free-Time Residences, Free-Time Residences 2020*. Retrieved from https://www.stat.fi/til/rakke/2020/rakke_2020_2021-05-27_kat_001_en.html

Official Statistics of Finland (OSF). (2021d). Paavo postal code area statistics - Buildings and free-time residences. *Buildings and Free-Time Residences (2019)*. Retrieved from https://www.stat.fi/tup/paavo/index.html

Official Statistics of Finland (OSF). (2021e). Population structure. *Population Structure 2020*. Retrieved from http://www.stat.fi/til/vaerak/index_en.html

Pepe, E., Bajardi, P., Gauvin, L., Privitera, F., Lake, B., Cattuto, C., & Tizzoni, M. (2020). COVID-19 outbreak response, a dataset to assess mobility changes in Italy following national lockdown. *Scientific Data*, *7*(1). https://doi.org/10.1038/s41597-020-00575-2

Pitkänen, K. (2008). Second-home landscape: The meaning(s) of landscape for second-home tourism in Finnish Lakeland. *Tourism Geographies*, *10*(2), 169–192. https://doi.org/10.1080/14616680802000014

Pitkänen, K., Hannonen, O., Toso, S., Gallent, N., Hamiduddin, I., Halseth, G., … Nefedova, T. (2020). Second homes during corona - safe or unsafe haven and for whom? Reflections from researchers around the world. *Matkailututkimus*, *16*(2), 20–39. https://doi.org/10.33351/mt.97559

Poom, A., Järv, O., Zook, M., & Toivonen, T. (2020, July 1). COVID-19 is spatial: Ensuring that mobile Big Data is used for social good. *Big Data and Society*, Vol. 7. SAGE Publications Ltd. https://doi.org/10.1177/2053951720952088

Razavi, R., Gharipour, A., Fleury, M., & Akpan, I. J. (2019). Occupancy detection of residential buildings using smart meter data: A large-scale study. *Energy and Buildings*, *183*, 195–208. https://doi.org/10.1016/j.enbuild.2018.11.025

Rojas, M. B., Sadeghvaziri, E., & Jin, X. (2016). Comprehensive review of travel behavior and mobility pattern studies that used mobile phone data. *Transportation Research Record*, Vol. 2563, pp. 71–79. National Research Council. https://doi.org/10.3141/2563-11

Ruralia Institute. (2021). *Monipaikkaisen asumisen rytmit, paikat ja asiakasryhmät (MOPA)*. Retrieved from https://www2.helsinki.fi/fi/ruralia-instituutti/tutkimus/vapaa-ajan-asuminen-tutkimuskohteena#section-108290

Santamaria, C., Sermi, F., Spyratos, S., Iacus, S. M., Annunziato, A., Tarchi, D., & Vespe, M. (2020). Measuring the impact of COVID-19 confinement measures on human mobility using mobile positioning data. A European regional analysis. *Safety Science*, *132*. https://doi.org/10.1016/j.ssci.2020.104925

Scheidt, F. vom, Medinová, H., Ludwig, N., Richter, B., Staudt, P., & Weinhardt, C. (2020, August 1). Data analytics in the electricity sector – A quantitative and qualitative literature review. *Energy and AI*, Vol. 1. Elsevier B.V. https://doi.org/10.1016/j.egyai.2020.100009

Schier, M., Hilti, N., Schad, H., Tippel, C., Dittrich-Wesbuer, A., & Monz, A. (2015). Residential Multi-Locality Studies - The Added Value for Research on Families and Second Homes. *Tijdschrift Voor Economische En Sociale Geografie*, *106*(4), 439–452. https://doi.org/10.1111/tesg.12155

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

Silm, S., Järv, O., & Masso, A. (2020). Tracing human mobilities through mobile phones. In *Handbook of research methods and applications for mobilities* (pp. 182–192). Edward Elgar Publishing.

Sun, L., Zhou, K., & Yang, S. (2018). Regional difference of household electricity consumption: An empirical study of Jiangsu, China. *Journal of Cleaner Production*, *171*, 1415–1428. https://doi.org/10.1016/j.jclepro.2017.10.123

SYKE. (2021). Elinvoimainen ja kestävä monipaikkainen Suomi. Retrieved November 18, 2021, from https://www.syke.fi/fi-FI/Tutkimus__kehittaminen/Tutkimus_ja_kehittamishankkeet/Hankkeet/Elinvoimainen_ja_kestava_monipaikkainen_Suomi_VN_TEAS_Monipaikkaisuus

Tjørve, E., Flognfeldt, T., & Tjørve, K. M. C. (2013). The Effects of Distance and Belonging on Second-Home Markets. *Tourism Geographies*, *15*(2), 268–291. https://doi.org/10.1080/14616688.2012.726264

Voutilainen, O., Korhonen, K., Ovaska, U., & Vihinen, H. (2021). Mökkibarometri 2021. *Luonnonvara- Ja Biotalouden Tutkimus*, *47*.

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

Weichhart, P. (2015). Residential Multi-Locality: In Search of Theoretical Frameworks. *Tijdschrift Voor Economische En Sociale Geografie*, *106*(4), 378–391. https://doi.org/10.1111/tesg.12156

Willberg, E., Järv, O., Väisänen, T., & Toivonen, T. (2021). Escaping from cities during the COVID-19 crisis: Using mobile phone data to trace mobility in Finland. *ISPRS International Journal of Geo-Information*, *10*(2). https://doi.org/10.3390/ijgi10020103

Yildiz, B., Bilbao, J. I., Dore, J., & Sproul, A. B. (2017, December 15). Recent advances in the analysis of residential electricity consumption and applications of smart meter data. *Applied Energy*, Vol. 208, pp. 402–427. Elsevier Ltd. https://doi.org/10.1016/j.apenergy.2017.10.014

Zoğal, V., Domènech, A., & Emekli, G. (2020). Stay at (which) home: second homes during and after the COVID-19 pandemic. *Journal of Tourism Futures*. https://doi.org/10.1108/JTF-06-2020-0090