Validation of a new Meta database usage to understand population's mobility: the February 6th, 2023, Türkiye event case study

Auteur : Gosselin, Constance
Promoteur(s) : Hubert, Aurelia; Devillet, Guénaël
Faculté : Faculté des Sciences
Diplôme : Master en sciences géographiques, orientation global change, à finalité approfondie
Année académique : 2024-2025
URI/URL : http://hdl.handle.net/2268.2/22269

Faculté des sciences
Département de géographie

Validation of a new Meta database usage to understand population's mobility: the February 6th, 2023, Türkiye event case study

Master thesis presented by: Constance Gosselin

Master in Geography, Global Change,
Human aspects of global environmental changes

Academic Year : 2024-2025

Presentation and defense : January 2025

Jury President : Pr Xavier Fettweis

Promoter : Pr. Aurelia Hubert

Co-Promoter : Pr. Guénaël Devillet

Lecture Jury : Pr. Caroline Zickgraf
Pr. Pierre Ozer

Acknowledgments

First of all, I would like to express my deepest gratitude to my two co-promoters, Prof. Hubert and Prof. Devillet. They took the time to guide me through new technical analyses and pushed me to think beyond my initial understanding. Their professionalism and kindness have sincerely helped me evolve and complete this master's thesis. They also provided valuable direction, helping me find the path to present this work today.

Secondly, I would like to thank Sébastien Dujardin, who introduced me to the idea of using the Meta dataset during the internship I had under his guidance.

I am also grateful to Tao Beaufils, a dear friend of mine, who taught me Python and provided invaluable help during long hours of coding challenges, which enabled me to achieve the results presented in this thesis.

My sincere thanks also go to Elise Faulx, who introduced me to LaTeX and suggested I use the same first page template as hers.

I would also like to express my gratitude to my friends and family, who have been by my side during these six months of writing. Their support has been essential in helping me stay focused and motivated.

I want to thank my colleagues from both my bachelor's university (University of Namur) and my master's university. They have been a great source of motivation and energy, and I cherish the moments we shared together, which were both inspiring and thought-provoking.

Lastly, I would like to thank all the professors I had during my studies. Their passion for geographical sciences has been contagious. As I write the final sentences of my master's thesis, I can confidently say that my enthusiasm for geography has only grown. I sincerely hope that my future professional journey will be as fulfilling and exciting as my academic studies.

Abstract

This study explores the utility of the MDM dataset for analyzing mobility patterns during natural disasters. Unlike traditional Big Data sources, MDM provides unique advantages, such as finer spatial and temporal scales, enabling the examination of daily mobility variations.

The dataset was validated against known mobility patterns, revealing sensitivity to noise and fluctuations in user activity on the Facebook platform. Significant differences were observed in mobility responses between regular and test days, highlighting the role of socio-demographic factors.

The study also investigates population behavior during disasters, showing that individuals either stayed within affected zones or relocated quickly to safer areas.

Although the study has limitations, such as the exclusion of motivations behind mobility and noise resilience, it lays the groundwork for using MDM in future disaster mobility research. Future studies could explore its application across diverse disaster types, urban settings, and community mobility patterns.

Résumé

Cette étude explore l'utilité du jeu de données MDM pour analyser les schémas de mobilité pendant les catastrophes naturelles. Contrairement aux sources traditionnelles de Big Data, MDM présente des avantages uniques, tels que des échelles spatiales et temporelles plus fines, permettant l'examen des variations quotidiennes de la mobilité.

Le jeu de données a été validé par rapport aux schémas de mobilité connus, mettant en évidence sa sensibilité au bruit et aux fluctuations de l'activité des utilisateurs sur la plateforme Facebook. Des différences significatives ont été observées dans les réponses de mobilité entre les jours réguliers et les jours de test, soulignant le rôle des facteurs socio-démographiques.

L'étude examine également le comportement de la population pendant les catastrophes, montrant que les individus restaient soit dans les zones affectées, soit se déplaçaient rapidement vers des zones plus sûres.

Bien que l'étude présente des limites, telles que l'exclusion des motivations derrière la mobilité et la résilience au bruit, elle jette les bases de l'utilisation de MDM dans les futures recherches sur la mobilité en cas de catastrophe. Les études futures pourraient explorer son application à travers différents types de catastrophes, d'environnements urbains et de schémas de mobilité communautaire.

Contents

List of abbreviations

EQ	**Earthquake**
FP	**Fraction of the population**
MDM	**Movement Distribution Map**
PGA	**Peak Ground Acceleration**
GPS	**Global Positioning System**

List of Figures

List of Tables

I Introduction

Big Data is playing an increasingly significant role in academic research, offering insights that surpass those obtainable through traditional data sources. Its greatest strengths lie in its capacity to integrate vast amounts of data, encompassing diverse formats and spanning broad spatial and temporal scales. Examples include remote sensing data, geo-tagged datasets, and social media information.

Despite its widespread use, the definition of Big Data remains a topic of debate within the scientific community. Among the various approaches to defining Big Data, the "Vs" framework (Variety, Velocity, and Volume) stands out as the most widely accepted. For the purposes of this study, we adopt the definition provided by the European Commission, which aligns with this framework.

In geographical research, Big Data has emerged as a powerful tool due to its affordability and ability to provide spatio-temporal insights across interdisciplinary subjects. It has enabled significant advancements in fields such as climate modeling, medical geography, and human geography. Specifically, in the context of human geography and natural disaster research, Big Data facilitates the study of human mobility patterns, behavioral responses, and temporal trends. These insights are invaluable for optimizing land-use planning and disaster preparedness strategies.

Human mobility studies utilizing Big Data often rely on text mining and geo-tagged attributes derived from social media or mobile phone data. These datasets capture sentiment and location-based movements, offering unique perspectives on mobility dynamics. However, such studies face limitations at smaller spatial scales, such as districts or towns. Researchers must often choose between a detailed spatial analysis with limited temporal scope or broader temporal studies constrained to larger geographic scales.

In natural disaster research, social media data is particularly valuable for identifying local clusters of activity and analyzing time series of human mobility. Yet, these datasets are insufficient for understanding mobility impacts at finer spatial scales. Similarly, mobile phone datasets, while offering detailed insights, are often inaccessible due to privacy concerns and commercial restrictions.

The Movement Distribution Map (MDM), a dataset provided by Meta for Good (Meta, 2024), offers a promising alternative. This open dataset captures daily mobility patterns at the district level, reflecting the distances traveled by fractions of the population. Unlike traditional datasets, MDM is publicly available and has yet to be extensively utilized in scientific research. Preliminary studies, such as (Heydari et al., 2023), have used earlier versions of the dataset for validation purposes, but its full potential remains unexplored.

The February 6, 2023 EQ (EQ) in Turkey and Syria, which affected over 16 million people, provides a critical context for evaluating the relevance of the MDM dataset. Traditional datasets, whether derived from social media or mobile phone records, are limited in their ability to capture the nuanced mobility changes resulting from such large-scale events. By leveraging the MDM dataset, this study aims to fill this gap and provide new insights into population mobility during extreme events.

Therefore the reasearch question of this paper is :

> **How relevant and effective is the use of the Meta database in understanding population mobility during extreme events, such as the February 6, 2023 earthquake in Turkey?**

The main hypothesis is that MDM is accurate in natural disasters with the sensibility of national-wide events like national holidays and summer holidays. In the context of natural disasters, MDM provides a net change in the daily mobility of the population. This mobility is mostly changing mainly in the EQ-affected area. This change depends on the vulnerability of the people, the intensity, and the community resilience.

This study is divided into two main parts: Validation of the Meta Database and the analysis of the February 6, 2023 EQ event.

Validation of the Meta Database implies the methodology and results of these part:

- Overview of population mobility patterns.
- Examination of the impact of Laplace noise on the dataset.
- Tests to assess how different events influence daily mobility.

Analysis of the February , 2023 EQ is divided in three main part with their methodologies and their results. Here are the part:

- National Scale: Assessing population reactions to the EQ across Turkey.
- District Scale: Exploring localized mobility responses in affected areas.
- Urban vs. Rural Comparison: Evaluating variations in mobility responses based on geographic and demographic factors.

The study begins with a review of the relevant scientific literature, followed by an overview of the event and study area. After presenting and processing the data, the two main analyses are conducted. The findings are then discussed in relation to existing literature, highlighting the study's limitations and implications. Finally, the conclusion summarizes the key insights and contributions of this research.

II State of the Art

The state-of-the-art part elaborates on research in the scientific literature on the two main concepts of this study: Natural disasters and Big Data. It integrates definitions surrounding the two concepts and their link between them.

II.1 Big Data in Geographical Sciences

Big data arrived with the evolution of new technologies in our lives. They increased the amount of information thanks to the rising storage space. Such changes opened a new branch of the use of data. New technologies pair new perceptions of information with new methods, and technical capacities. These discern new under layers (Venkatram and Geetha, 2017). In this section, the concepts of big data have their place in geographical sciences.

II.1.a Concept of Big Data

By its unanimously accepted terminology, the Big Data concept is described in diverse terms. Researchers who tried to propose a universal definition of Big Data had issues with what is alludes to. It can be referred to as a technology, a platform, or even a dataset. However, most of the descriptions relate to 'Vs' rules. There are at least 3 'Vs' cited in the literature reviews: Variety, Velocity and Volume (Li Cai, 2015; Maddalena Favaretto, 2020; Rob Kitchin, 2016; Xiaoyao Han, 2024; Chuan Liao, 2018; Borko Furht, 2016; Hrehova, 2018; Shukla et al., 2015; Alvarez-Dionisi, 2016).

Borko Furht (2016); Chen and Yu (2018) develops a more concise definition of Volume, Velocity, and Variety:

- Volume needs to be at least Terabytes and includes the structure of the data found.
- Velocity integrates the resolutions, the storage, and the computation behind it. It also needs to be at a high speed.
- Variety concerns the type of data. It can be text, images, tables, geospatial data, etc.

Meanwhile, the scientific community diverges on the number of 'Vs' accepted in such definition. Over the literature, ten 'Vs' can be considered in that definition:

- Value: interprets the will of having wide datasets, proposes near real-time data, and is broad,
- Veracity: includes the conformity of data and the value of its density with a fine temporal resolution and has a large spatial scale,
- Vocabulary: the new language developed to speak about Big Data and the coding language linked with this concept,
- Vagueness: is the challenges it faces,
- Venue: comprehend the diversity of the platforms,
- Variability: is about the changes in the data source,

- Validity: globalize the reliability of the data,

As remarks, Velocity, and Value are the two sides 'Vs' the most cited in the literature. The type and the number of 'Vs' differs depending on the discipline. The integration of open sources availability is often mentioned as a branch of the concept (Li Cai, 2015; Maddalena Favaretto, 2020; Rob Kitchin, 2016; Xiaoyao Han, 2024; Chuan Liao, 2018; Borko Furht, 2016; Hrehova, 2018; Tiwari, 2019; Venkatram and Geetha, 2017; Watson, 2014; Riahi and Riahi, 2018; Shukla et al., 2015; Alvarez-Dionisi, 2016).

The aforementioned perception only integrates datasets with their technology and analytics methods. This suggests a lack of its use and its purposes. In addition, some peers express this definition as not enough specific. They add that such absences cannot be enough relevant. Indeed for them, Big Data should be defined by the type of data itself which is a link to how it is used (Mark Graham, 2013; Rob Kitchin, 2016).

As the description of the concept is eclectic it is interesting to include a formal European Commission (nd) statement:

> Big data refers to large amounts of data produced very quickly by a high number of diverse sources. Data can either be created by people or generated by machines, such as sensors gathering climate information, satellite imagery, digital pictures and videos, purchase transaction records, Global Positioning System (GPS) signals, and more. It covers many sectors, from healthcare to transport to energy.

This quote proposes a similar concept of the 3 'Vs' and proposes the playground field of Big Data uses and dataset types. It might be interesting to and in this summary that the United Nations are using the 3 'Vs' terminology as their main definition (United Nations, 2022). As explained earlier, Big Data is not only about databases but also about analysis, data mining, and analysis technologies. Therefore Big Data is about the whole process from the acquirement of the data to its visualization.

The difference between data called *traditional data* and Big Data is in terms of:

- volume
- time resolution
- storage
- structure
- integration
- access

Traditional data is in gig-bites instead of terabytes. The rate is at a maximum per hour versus near-real-time in Big Data. Big Data is not structured or semi-structured. The data source is distributed in Big Data, while it is centralised in traditional. In terms of data sources cited from Big Data acquisition: public data, social media, sensors data, and company data (Borko Furht, 2016; Bao, 2023; Shukla et al., 2015). Data access is called interactive in traditional data but in Big Data, there is time resolution and a batch.

In order to understand the confusion of consensus over a definition, the same author explains that the structure of social media data can be seen as Big Data. But Big Data should be unstructured / semistructured (Borko Furht, 2016).

However, Big Data has challenges mostly related to privacy, such as video camera use, outliers, noise, data processing, and efficiency that should be considered while using such data (Bao, 2023; Borko Furht, 2016; Poullet, 2020; Watson, 2014; Shukla et al., 2015; Macnish and Galliott, 2020; Alvarez-Dionisi, 2016).

Because Big Data is massive, processes and models of traditional data are limited. Therefore, new technologies are developed. These models fill the need to manage high-capacity data, near-real-time information, and unstructured data (Bao, 2023). Big Data has analytics elements for the data sources and type. Its use is interesting because it is a costless methodology, and the data set in the research is bigger than a data sample that traditional data could have.

Big Data analytics can be divided into four categories (Riahi and Riahi, 2018):

- descriptive: study of the present situation,
- diagnostic: understanding of the event's causes,
- predictive: about the future situations,
- prescriptive: relate to the past situation and the changes to do.

Here are some examples of analytics methods: text mining, audio analytics, video content analysis social media analytics, and predictive analytics (Borko Furht, 2016).

The methodologies of Big Data are wide. The use of Machine Learning and Artificial Intelligence is common in the collection of data but also in analyses. Statistical methods, cluster, and visualization methods are also used in the context of Big Data. However, in the statistical analysis, correlations, regressions, and trends are used to discover unseen patterns (Bao, 2023).

Since Big Data is dedicated to a variety of data, its use and interdisciplinary impact almost every human activity. Here are a few examples of disciplinaries using Big Data: gaming, economics, public administration, social studies, healthcare (and public care), risk management, physics, meteorology, and urban planning (Borko Furht, 2016; Bao, 2023; Shukla et al., 2015).

These datasets are creating new technologies that are faster to obtain results linked with decision-making. It can also integrate sentimental and cultural perceptions over decision making(Venkatram and Geetha, 2017; Watson, 2014; Borko Furht, 2016). As this study focuses on social media, here are examples of its use: geographies, economics, and marketing, in Health Care and social sciences (Chuan Liao, 2018; Schwaller et al., 2024; Muniz-Rodriguez et al., 2020; Wu and Ma, 2022; Yan et al., 2024)

This study focuses on the capacity of geospatial open source data from social media use in order to understand the impact of a disaster by its effect on daily mobilities. Therefore the following concept is the use of social media as Big Data in geography.

II.1.b Big data use in Geographical sciences

As explained earlier Big Data is used in several sciences. Big Data is integrated into physical and human geographies. This study focuses on the utility of social media in a geographical research context and social aspects of the discipline. It limits the state-of-the-art coverage of Big Data in physical geographies. Therefore, it is relevant to integrate the use of social media in geographical sciences.

Big Data has interdisciplinarity which leads to cross overs of patterns and processes like social behaviours and human environments (Mark Graham, 2013; Chun et al., 2019). The sources are broad and concerns social media, phone operator, vehicle movements, crowed sources platforms, remote sensing and sensors (Chuan Liao, 2018; Chun et al., 2019; Ivan et al., 2017).

In human geography, Big Data integrates social analysis throughout numerous type of analytics. Thanks to its open access to massive information and its spatio-temporality, Big Data resolves complexities and long lasting data collection from social sciences methodologies. It reveals side perspectives of already existing theories and new branch of researches (Chen and Yu, 2018; Chun et al., 2019).

Geo-tagged information are broadly used in geographical context. Data from the internet helps regarding technical problems. Such heterogenous in that science leads to an expansion of possibilities like interest in socio-economics impact over a certain spatial context or evolution. In addition, studies are more statistically relevant (Feizizadeh et al., 2024).

In the context of disaster risk assessment and post-disaster studies, the variety and small spatial resolution help to understand the causes and prevent hazards with Remote Sensing technologies in terms of physical geography aspects and social media use in human geography (Ivan et al., 2017; Yan et al., 2024; Huo et al., 2024).

This study's research is based on social media use in migration linked to a disaster. In such research, the main social media use is location-based social media data. It implies social media as sensors. It completes traditional data by is volume and its spatio-temporality complexity. Its design innovates in the analysis of patterns and the comparison of indicators to outline features. As a remark, Located based social media opens a new type of methodologies mostly in the human mobility market 2023 collection (Ivan et al., 2017; Yan et al., 2024; Huo et al., 2024).

Analytics techniques used in such research are visualizations of the data and statistical analysis like correlation, linear regression, or geographically weighted regression, and clustering. Here are some clustering techniques k-means and Density-based Spatial Clustering of Applications with Noise (McKitrick et al., 2023; Huo et al., 2024). In human migration research, this type of analysis provides results on the spatio-temporal changes, specific spatial movements in some cases, hot-spots, and social networks/behaviours (McKitrick et al., 2023; Huo et al., 2024).

Here are some articles using social media in a spatial and spatio-temporal human geographies context:

- Wu and Ma (2022) uses this to understand clusters on site-parking occupancy

- Ginzarly et al. (2018) maps the tourism values of historic urban landscapes by photo analysis
- Longley et al. (2015) uses X[1] in a demographical spatial analysis
- Huang and Wong (2015); Arjona and Palomares (2020) use X in order to model human mobility patterns
- Carvalho et al. (2021); Heydari et al. (2023) are using the social media to understanding human mobilities.

As remarks, Heydari et al. (2023) is the only researcher that have for now used the type of data mobility used in this master thesis. This shows the wideness of social media data use.

In Geographical Information Systems, Big Data interfere in the spatio-temporal resolutions, data quality, and uncertainty. As an example, the use of crowed sources informations, GPS, and volunteering informations require the consideration of missing values, errors on data and real positions.

Furthermore, quality cannot be verified in data collection or production. In case of spatio-temporal analysis, non use of daily mobility can lead to errors. Big Data deals with Geographical issues: Modifiable Areal Unit Problem, Spatial Autocorrelation and Uncertain Geographical Context Problems (Chun et al., 2019; Chuan Liao, 2018; Feizizadeh et al., 2024).

In geography, social media data have limitations on their use. The users are over represented which excludes the non-user. Social media data are secondary data that links to noises and not precise values for the specific research. The spatial accuracy of geo-tagged information is entirely relatable since the people might post or indicate wrong GPS (McKitrick et al., 2023).

II.1.c Social Media's population trend behaviours

Social media assumed passive crowdsourcing generates information on users' lives (Lam et al., 2023). In the context of social media use, it is relevant to highlight the demographic impacts of different social media use. As an example Facebook have 79% of their users being from 45 to 55 years old (Singh et al., 2019) which is higher than X. Singh et al. (2019) covers the use behaviours of social media.

Smartphone development helps social media to enter the daily life of the population. At dinner time, people are more likely to use it (70%) compared to breakfast and work time. In Singh et al. (2019) study, 57% of the population used social media right after waking up. In terms of periodicity of social media use, 35% of the population opens them at least five times a day, only 3% use them once a week, and 20% do not use them in a day but more than once a week. Applications are the main interfaces to use social media for more than 68% of the population.

[1] previously Twitter

II.2 Natural Disaster

This section elaborates on natural disasters through the definitions of related concepts. The migration literature review is specific on that subject. And the link between social media use and natural disasters.

II.2.a Concepts linked with natural disasters

Natural disasters affect most of the population on earth. The impact of natural disasters develops emergency responses. This is followed by addressing the hidden needs of the population. People must have access to basic needs such as shelter, security, and health (W Sharp and W Beadling, 2013). To better understand the link between displacement and migration, it is essential to define the concepts related to natural disasters.

In the context of natural disasters as physical events, the concept of a hazard can be defined by its intensity, likelihood of occurrence, and type. The occurrence implies a frequency with which the hazard may happen over time. When hazards reach a certain level of impact, they are classified as natural disasters (Ozer, 2019).

That concept is linked with vulnerability which depends on exposure, sensitivity, and adaptation. Exposure is closely related to the concept of hazard, as it encompasses the spatiotemporal dimension, frequency, and intensity of the hazard's impact on the affected population. The sensitivity is the potential condition of the individual to be more or less affected (Yu et al., 2021). As examples:

- the socioeconomic status of the person affects the person in terms of economic capacity to adapt,
- its social network is also impacting the ability of the person to seek shelter outside of the affected area,
- its risk perception,
- knowledge of the hazard.

The adaptation capacity is the ability to cope and recover after and during a hazard. It also identifies the capability of the individual to reduce the next disaster's impact (Ozer, 2019).

High exposure and high sensitivity increase the vulnerability but the adaptative decreases it. Vulnerability can be measured by the ability to resorb the hazard and its limitation of loss and damages (Ozer, 2019).

The concept of risk is fundamentally linked to both hazard and vulnerability. It provides the potentiality for a person to be more likely highly impacted or less impacted. If the vulnerability or the hazard increases, the risk increases (Ozer, 2019).

The vulnerability concept resides in another theme broadly used in natural disaster studies and climate change: Resilience. This notion is expressed as the ability to respond to the hazard. It is discerned to reveal responses of the individuals, communities, and populations. It includes societal, community, and organization intervention on the risk (Ozer, 2019). The

wideness of the resilience concept is linked to the number of sub-units like the seismic, communitarian, economic, and urban resilience (Aguilera and Villagra, 2023; Wang et al., 2024).

For the context of the master's thesis, it is necessary to define community, urban, and seismic resilience. Communitarian resilience is linked with the dynamics of preparedness and the capacity to resorb/adapt to a hazard, not as an individual but as a community (Aguilera and Villagra, 2023). Urban resilience mostly considers the infrastructure systems to adapt/resorb the hazard (Wang et al., 2024). Seismic resilience is the adaption and recovery of a person/organization by improving its urban resilience, economic loss absorption, preparedness in health care, emergency seek, and research rescue. The seismic resilience is mostly linked with the *earthquake culture* a type of risk perception specialized in EQ (Karasözen et al., 2023).

II.2.b Displacement of people in Natural Disaster studies

In high-impact events, emergencies situation creates mobility, displacement, and migration (W Sharp and W Beadling, 2013). Displacement is a human mobility engendered by a shock. Migration is on the contrary provoked by slow changes in their environments (for Migration , IOM).

Displacement is related to vulnerability and exposure of the individual. People not living in their homes but staying in their country after a shock (like a natural disaster) are referred to as internally displaced (W Sharp and W Beadling, 2013).

There is a duality in human mobility in preparedness and recovery. In the case of prevention of a natural disaster, organizations and institutions could lead to evacuating or warning the population of their exposure to the hazard as a disaster risk reduction (Wisner et al., 2012; Karampotsis et al., 2024). Human mobility is highly dispersed, as some population displacement serves as an adaptive strategy that reduces vulnerability (Schwaller et al., 2024). As a remark, evacuation differs with vulnerabilities of individuals (Castillo Betancourt and Zickgraf, 2024).

Displacement is part of the recovery and resilience of individuals by temporary settlements or long-term resettlement. While recovering, the displacement of individuals can increase their vulnerability to other hazards and dangers. Such trends are visible mostly in complex emergencies (W Sharp and W Beadling, 2013; Castillo Betancourt and Zickgraf, 2024).

Some researchers question the immobility in the context of environmental changes. Indeed, immobility relates to strong disparities in a community. Some people are willing to stay in the hazard area because of their cultural habits, risk perception level, and the importance of the places in their lives. Some people need to leave but do not have the capacity. They are called trapped population (Courtoy, 9 01; Castillo Betancourt and Zickgraf, 2024).

More technically, the preparedness of the evacuation and temporary housing is affecting the mobility and so, displacement of the population (Komatsuzaki et al., 2022).In high-density populations, such as urban areas, narrow streets often make it difficult to integrate shelters within them.

The population resettlement in the affected area is linked with socio-cultural and sentimental perceptions of the hazard. As an example, Komatsuzaki et al. (2022) explained that the population seemed to be less likely to go back into a flooded area and were more willing to stay in the temporary resettlement location.

II.2.c Social Media in Natural Disasters Research

In the context of this research, it is necessary to understand the use of social media in a (post-)disaster context.

Social media are proactively used in natural disaster migration because they provide geo-tagged information and possible movement effects before a disaster. Such tools could help prevent a large influx of individuals in one area or designate specific locations to create accommodations and shelters for others (Schwaller et al., 2024).

As the migration is disaster response, social media can be used to understand insight into vulnerabilities like (Schwaller et al., 2024; Lam et al., 2023; Muniz-Rodriguez et al., 2020; Han et al., 2024; Li et al., 2024):

- social networks lacks,
- people being able to leave because they have social connections outside of the place impacted,
- risk perception culture and resilience.

In emergencies and resilience, social media are primarily used to understand the information exchanged within a user's community(Muniz-Rodriguez et al., 2020; Lam et al., 2023). It integrates resources :

- basic needs post-disasters,
- volunteering propositions,
- evolution of weather conditions (mostly for natural disasters related to weather) and,
- evolutions of the emergencies (like places the most devastated, the infrastructures damages).

The time and spatial component of these data release information on the sentimental impact of the disaster, effective disaster responses, and recoveries. It is intergrade in the adaptation of this type of event. Post-disaster studies look at such adaptation by behavioral research, risks, and risk perceptions analysis(Lam et al., 2023; Muniz-Rodriguez et al., 2020; Han et al., 2024; Li et al., 2024).

The main social media used in natural disasters is to understand the population's needs, their sentimental impact, and the post-migration by geolocated texts. This engendered uncertainty about the information, missed location issues, and rumors dissemination. The type of users is also a limitation in such studies The disparities of the social use reduce the validation of the social network data. These needs to be considerate throughout social media uses (Muniz-Rodriguez et al., 2020; Lam et al., 2023).

III Study Area

III.1 General situation

Located at the crossroads of Europe and the Middle East, Türkiye is home to 85,372,377 people and boasts a Growth Rate Annual of 2.1% (Turkish statistical institue, ndb; Dilek Yildiz, 2023). The country spans six tectonic plates: Anatolian, Iranian, Arabian, African, Aegean, and Eurasian, positioning it in one of the world's most tectonically active regions. This study examines the devastating earthquakes that impacted Türkiye in February 2023.

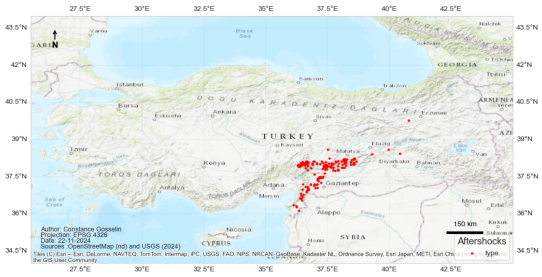


Figure 1: Eathquakes aftershocks

On February 6, 2023, the East Anatolian Fault system ruptured, triggering an earthquake of magnitude 7.8. This catastrophic event was immediately followed by a second earthquake of magnitude 7.5 within 24 hours (Sarı et al., 2023). Both earthquakes resulted from strike-slip mechanisms, with the first rupture occurring along the East Anatolian Fault and the second along a subsidiary fault connected to it. This cascade of seismic events caused widespread destruction across nine Turkish provinces and parts of northeastern Syria. The epicenter of the first EQ was near Gaziantep city, a densely populated urban area.

In total, approximately 17 million people were affected by these earthquakes, resulting in 55,000 deaths, 100,000 injuries, and 6.5 million individuals requiring immediate necessities. Over 264,389 rescue workers were mobilized, but the scale of the disaster overwhelmed initial relief efforts. Temporary shelters housed approximately 845,000 people in the first weeks, with 181,264 individuals still relying on these shelters by August 2023 (United Nations Office for the Coordination of Humanitarian Affairs, 2023; British Red Cross, 2024). Psychological support was also provided to 498,225 people to address the trauma caused by the disaster. The World Health Organization declared the event a Level 3 emergency (its highest level) underscoring the severe challenges faced in coordinating relief efforts (Sarı et al., 2023).

The crisis prompted mass displacement, including 2 million Syrians relocating near the Turkish border to receive aid. Urban hubs like Adana, which were less severely impacted, served as logistical centers for relief operations, distributing supplies to more affected areas. These urban centers provided crucial support in the aftermath of the disaster.

III.2 Challenges During Rescue Operations

Despite rapid mobilization, the first 24 hours of response were hindered by blocked roads and severe winter weather, including heavy snowfall that obstructed access to affected regions. Many medical staff were unable to assist due to personal injuries, loss of family members, or lack of shelter and basic necessities. Airports and other critical infrastructure were damaged, further delaying aid delivery. High numbers of injured individuals strained limited medical resources, leaving many without timely care (Yılmaz, 2023; Şenol Balaban et al., 2024; Tayfur et al., 2024).

While rescue teams arrived within 24 hours in most cities, the evacuation of severely injured individuals was delayed until 72 hours post-disaster. It was not until the fourth day that evacuees could find shelter outside the disaster zones, and by the seventh day, most logistical issues were resolved (Yılmaz, 2023; Şenol Balaban et al., 2024; Tayfur et al., 2024).

III.3 Economic and Social Impact

The earthquakes caused the destruction of approximately 202,000 buildings, many of which were non-compliant with building codes revised after 2018. The total economic cost of the destruction is estimated at $39 billion, with an additional $5 billion required for immediate humanitarian aid, including shelter for both Turkish and Syrian populations. The region's production experienced an economic loss equivalent to $7.5 billion per capita (DEMIRALP, 2023).

The agricultural sector faced significant setbacks, with crop production dropping by 15% at the national level. Key industries such as steel and textiles, concentrated in Hatay and Kahramanmaraş provinces, were heavily affected. Overall, the earthquakes reduced Türkiye's GDP by approximately 1%, amounting to $10 billion (Turkish statistical institue, ndb; DEMIRALP, 2023). These multifaceted impacts underscore the catastrophic consequences of the February 2023 earthquakes, highlighting the urgent need for improved disaster preparedness and resilient infrastructure.

IV Data

This section is dedicated to the understanding of the various datasets used in this thesis.

IV.1 Meta dataset

Meta, formerly known as Facebook, has established a dedicated branch to address social challenges and support societal advancements through data-driven initiatives (Meta, 2024). As part of its commitment to contributing to the humanitarian community, Meta has made several datasets publicly accessible. Notable examples include the High-Resolution Population Density Map and the Social Connectedness Index. Additionally, certain datasets, such as the Business Activity Trends, require special permissions for access. Meta's open-source datasets are hosted on the Humanitarian Data Exchange platform, offering greater accessibility to the global humanitarian community. Among these resources is the MDM dataset, which provides insights into daily distances traveled by Meta users. This dataset is a continuum of the previous datasets so-called Movement Range Maps, based on two metrics: *stay in put* and *Change in mobility* (Meta, 2022).

The dataset studied is generated by analysing the movement of users who have consented to location-sharing via the Facebook mobile app. As outlined in the accompanying README file (Meta, 2022), this dataset represents aggregated and anonymized distributions of user movement patterns. The Movement Distribution Maps categorize distances travelled into four bins (Figure 2):

- No movement
- 0 km to 10 km
- 10 km to 100 km
- 100 km and beyond

This data is derived from GPS-enabled devices, and the movement is calculated based on the distance between a user's *home location*. It is defined as their most frequent location during nighttime hours (typically 8 PM to 6 AM). A randomly selected location during the day when the app is accessed (Figure 2). The data is organized into monthly folders, with daily records specified in Pacific Time. To ensure privacy, the data is obfuscated by applying Laplace noise, with an epsilon of 1 as a differential privacy mechanism. Notably, individual polygons are limited to a maximum of 10 users to further safeguard anonymity (Meta, 2022).

Geospatially, this dataset relies on GADM Level 2 administrative boundaries, providing district-level granularity for most countries, including Türkiye (Figure 2. However, the spatial granularity varies, with some areas exhibiting lower resolution. Furthermore, updates between 2018 and 2022 may have introduced inconsistencies in regions where administrative boundaries have changed. It is also important to note that GADM employs a toponymic naming system, which may differ from official names, complicating dataset integration. The limitations of Meta's coverage also stem from the platform's unequal global availability, restricting the dataset's completeness in some regions (Meta, 2022).

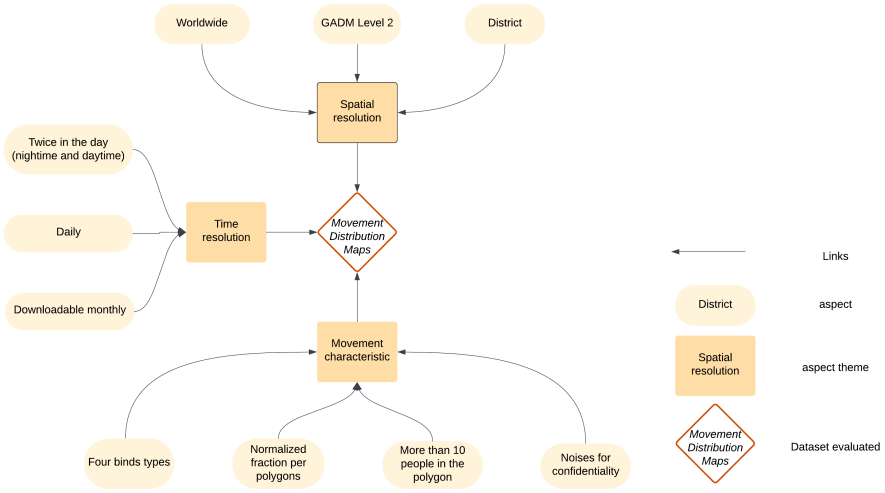


Figure 2: Characteristics of *Movement Distribution Map*

This dataset represents a valuable tool for understanding mobility trends, offering critical insights for humanitarian efforts, urban planning, and disaster response. By sharing these resources, Meta's Data for Good team continues to foster collaboration and innovation within the global humanitarian community(Meta, 2022).

IV.2 Dataset used in the validation of the Meta datase : socio-demographic indicators

The socio-demographic indicator datasets utilized in this study are sourced from the (Turkish statistical institue, ndb) a leading provider of comprehensive and thematically diverse datasets. These datasets cover a wide array of domains, including migration, demography, economy, and agriculture. However, variations exist in both spatial granularity and temporal resolution. Most datasets are updated annually, with spatial coverage ranging from national to regional scales.

For this master's thesis, the following key indicators have been selected for analysis (Table 1):

- Literacy rate by district
- Total population
- Sex ratio
- Median age
- Urban population
- Rural population
- Agricultural land area

These indicators offer spatial resolutions that extend from the district level to the commune level (Table 1). The datasets are updated on an annual basis. They can be accessed via TUIK's central dissemination platform [2]. Table 1 summarize the type of datasets used from TUIK and what socio-demographical indicators they provide. It also defines their characteristics as well as their spatial and temporal resolution.

[2]The database can be found at (Turkish statistical institue, nda): emoire/t1cmtt.fd" `https://biruni.tuik.gov.tr/medas/?kn=130&locale=en`.

Name of the Dataset	Indicators	Characteristics	Spatial Resolution	Temporal Resolution
Pop Age	Median Age	Population per age range per district in 2023	Provincial and districts	Yearly
Population Of SRE-1 SRE-2 Province and District	Total population, ratio of population living in villages compared to cities, population density	Population per village and city per district in 2023	Provincial and districts	Yearly
Illiterate Literate	Percentage of illiterate population per district	Numbers of illiterate, literate, and unknown individuals per age range and per district in 2023	Provincial and districts	Yearly
Pop Sex Ratio	Sex Ratio	Sex ratio per district in 2023	Provincial and districts	Yearly
Agricultural Area Decare	Agricultural area density	Agricultural area (in decares) per district	Provincial and districts	Yearly

Table 1: Dataset Characteristics

IV.3 Dataset used in the validation of the Meta database: change of daily mobility tests

To integrate the change of daily mobility in the validation, this thesis tests the MDM dataset with four types of events that could impact more or less the change of movement of population :

- workdays/weekends
- Summer holidays
- Hazard registered
- National holidays

The summer dataset and school holidays are from grey literature (Feiertagskalender, nd; American Association of Collegiate Registrars and Admissions Officers, 2023)

The validation part includes the change of movement due to National Holidays. Therefore the use of data from the Tourism and Culture Minister have been used (Ministre de la culture et du tourisme de Turquie, 2024).

The EM-DAT dataset, maintained by the Centre for Research on the Epidemiology of Disasters (Centre for Research on the Epidemiology of Disasters, 2024), serves as a comprehensive repository for global disaster events, cataloging incidents at the country level.

The dataset includes detailed records encompassing disaster type, geographic location, start and end dates, associated hazards, intensity, impacts, and estimates of losses and damages. Despite its extensive scope, the dataset is characterized by inherent limitations, including incomplete data, inconsistencies in geographical resolution, and significant gaps in key variables such as losses, damages, intensities, and impacts. As the same idea of EM-DAT, *Disasters Global Internal Displacement Dataset* (Internal Displacement Monitoring Center, nd) enumerates hazards per country and provides information about the internal displacement. Because of the sources used, this database is less relevant than EM-DAT.

IV.4 Dataset used in the case study

The dataset used to understand the intensity and impact of the earthquakes in this study is sourced from the United States Geological Survey[3]. Specifically, this study focuses on the Peak Ground Acceleration (PGA) to measure the severity of ground shaking. These data were cross-referenced with damage assessments from the literature, including Demir Aydin (2024). This integration allowed for precise identification of high-impact areas and correlation with observed damages. The datasets also included aftershock sequences and hazard-specific metrics for further analysis (United States Geological Survey, nd; United Nations Office for the Coordination of Humanitarian Affairs, nd). The integration of these diverse datasets enabled a holistic, multi-dimensional evaluation of the socio-economic and environmental impacts of the 2023 Eastern Anatolia earthquakes, spanning immediate physical damages, socio-economic disruptions, and the effectiveness of response mechanisms.

The datasets about the roads and settlements are downloaded from (United Nations Office for the Coordination of Humanitarian Affairs, nd) and are from Humanitarian OpenstreetMap Team (nd). They are crowdsourced datasets used in a humanitarian stake. The road dataset contains the main road in that order :

- highways,
- primary road,
- secondary road,
- tertiary road,
- tracks

These only take into account car mobilities.

The settlement dataset provides information in the order :

- city
- town
- villages
- hamlet

[3]Please have a look at their manual (United States Geological Survey, nd): `https://cbworden.github.io/shakemap/manual4_0/tg_philosophy.html`

V Preprocessing

Big Data are secondary data, and their structures vary depending on their sources. Consequently, preprocessing is required to adapt them to the study's needs. Different types of data require distinct processing approaches.

V.1 *Movement Distribution Map (MDM)*

The data used in this study were collected from Meta (2024) through the Humanitarian Data Exchange[4] and are available as monthly CSV files with a global spatial distribution. These data lack spatial geometry but include an ID column that aligns with GADM[5] geometries.

The preprocessing begins by compiling all raw data files into a single file. This is achieved by creating a function to list all filenames, reading each file iteratively, and concatenating them using the `pandas.concat()` function. The dataset is then filtered to include only values relevant to the study area. The column `GID_0` is used to identify the Turkish national-level ID, ensuring only relevant data are retained for further analysis.

To geolocate the MDM data, the `pandas.merge()` function in Python is employed to join the `GID_2` values from the MDM dataset with the `gadm_id` values from the GADM dataset. The resulting output is a vector file containing all MDM values.

Some indicators from Turkish statistical institue (ndb) are also used throughout the study. However, these data lack unique IDs, and several districts share identical names (e.g., *Merkez*, which refers to city centers). To resolve this, a new column is created in the MDM dataset using the format `NAME_1(NAME_2)`, where `NAME_1` represents the province name and `NAME_2` the district name.

The MDM dataset includes temporal data in the `ds` column. When necessary, these values are converted into Python's date type using the `datetime` package. Additional features, such as the day of the week and its name, are also derived using this package.

V.2 Indicator Datasets

Socio-demographic indicators are sourced from Turkish statistical institue (ndb). These datasets often contain unused columns, which are removed during preprocessing to create a streamlined file that can be easily joined with other datasets.

Due to differences in the usage of Latin and Turkish alphabets across socio-demographic and MDM datasets, inconsistencies in district and province names are common (Global Administrative Area organisation, 2022; Turkish statistical institue, ndb). To address this, a new mapping file is created, containing columns for both MDM and socio-demographic indicator names, formatted as `Province(District)`.

The area of each region is calculated using Global Administrative Area organisation (2022). Since the original coordinate reference system (CRS) is in degrees(ESPG, nd), it is converted to a metric CRS (`EPSG:5636`). The area is computed using a custom function that reprojects the CRS, creates new columns for the calculated area, and scales the values to square kilometers by dividing by 1,000,000. Additional indicators, such as sex ratio, urban-to-rural ratio, illiteracy rate, and agricultural activity, are computed. Illiteracy and agricultural activity percentages are multiplied by 100 to standardize their representation.[6]

[4]Website: `https://data.humdata.org/dataset/movement-distribution`
[5]Downloadable from: `https://gadm.org/download_country.html`
[6]Normalization during later analysis made this step unnecessary for subsequent indicators.

Finally, all indicators are merged into a unified dataset using consistent district and province names, facilitating seamless integration in the methodology.

V.3 PGA Dataset

PGA data, sourced from USGS (United States Geological Survey, nd), are provided as shapefiles. These are processed using the `sjoin()` function from the `geopandas` package in Python. A spatial join is performed to compute the mean PGA value for each district.

VI Validation of Meta Database

To assess the usability of the dataset, it is essential to evaluate its sensitivity, and overall reliability. This section focuses on the statistical examination of the MDM, aiming to determine its applicability in analyzing displacement patterns associated with natural disasters. The validation process is systematically structured, encompassing a detailed methodology section to outline the statistical techniques employed, a results section presenting the findings, and a concise summary providing a critical overview of the outcomes

VI.1 Methodology

This section delves into the diverse statistical analyses employed to validate the dataset. The methodology is systematically structured into five distinct tests: Sum of Fractions, Workday and Weekend, Summer, National Holidays, and Natural Disaster. Each test presents unique characteristics while also exhibiting areas of overlap, offering nuanced insights into the robustness of the validation framework. Before these tests, MDM is analyzed. The majority of this analytical work has been executed using Python, leveraging Jupyter Notebook through the Anaconda Navigator environment for seamless coding and data manipulation. Additionally, QGIS has been utilized for spatial analyses, and its specific applications will be explicitly noted where relevant.

VI.1.a Overall of the Movement Distribution Maps

Movement Distribution Maps is the main aspect of this thesis. Therefore, it has a special thematic and is in need to be tested (Figure 3). The visualization is separated throughout two aspects: temporally and spatially. In order to define what causes the changes throughout the years (2022, 2023, and 2024), time series of all districts and the mean of all districts per day are plotted. These time series integrate dates of National holidays, summer, and disasters (from Centre for Research on the Epidemiology of Disasters (2024)) to assess variations. However, before the next steps of this methodology, MDM is visualized by several plots and graphs but it is also tested with basic statistics:

- Video plots : These represent animate visualizations of the time series, showing the evolution of movement data over time. Each frame corresponds to a specific day, providing a dynamic view of how movement patterns evolve across districts. This approach is particularly helpful for identifying trends or anomalies that static graphs may miss. Due to technical limitations, the video plots include data only for the first 100 days of the database.

- Boxplots: used to summarize the distribution of data points within specific categories. They display key statistical measures such as the median, quartiles, and potential outliers. In this study, boxplots are particularly valuable for visualizing variability in movement distances and comparing data across different temporal or spatial subsets.

- Spatial Map outlier : it shows where the outliers are in Turkÿe and per bind.

- sort mean distribution graph : It shows the districts with the most different fraction of mobility.

- Cumulative distribution plot : shows the increase of the FP in the bind and compared to the others.

- Cluster map of local moran index : present spatial cluster and outliers with the use of autocorrelation coefficient.

A table of statistical descriptions complements these visualizations, offering numerical summaries such as mean, variance, and skewness for different categories. To analyze weekly patterns, boxplots and kernel density plots are computed for each day of the week. This is managed using the `datetime` package in Python, which assigns each day a number from 0 (Monday) to 6 (Sunday).

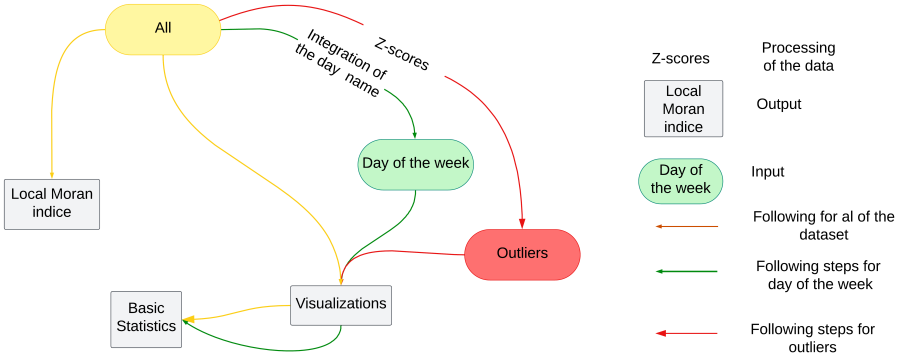


Figure 3: Scheme of Overall

VI.1.b Sum of Fraction Test

This section focuses on understanding the impact of Laplace noise on the aggregated fractions of mobility data (Figure 4). The MDM divides population movements into four categories:

- no movement,
- 0 –10 km,
- 10–100 km,
- 100 km and over.

As Facebook anonymizes data using Laplace noise, it is essential to assess how this noise affects the sums of these fractions. The initial step involves calculating the total sum of all four fractions for each district and day. The resulting values are analyzed through various statistical and visualization methods, including :

- boxplots,
- kernel density plots,
- and cumulative distribution plots

These tools help identify trends, variability, and anomalies in the dataset (Figure 4). Key statistical metrics, such as: mean, variance, skewness, kurtosis, mean absolute error, are summarized in a comprehensive table. An anova is conducted to understand the relation of Laplace noises by the socio-demographic indicators (Turkish statistical institue, ndb).

Outliers are identified using Z-scores, and maps are generated to visualize their spatial distribution across districts. To further explore spatial patterns, local Moran's I and cluster maps are utilized, highlighting areas with significant spatial autocorrelation or deviations caused by the Laplace noise (Figure 4).

These analyses provide a detailed understanding of the extent to which the Laplace noise influences the reliability of the aggregated mobility data, ensuring its suitability for subsequent tests.

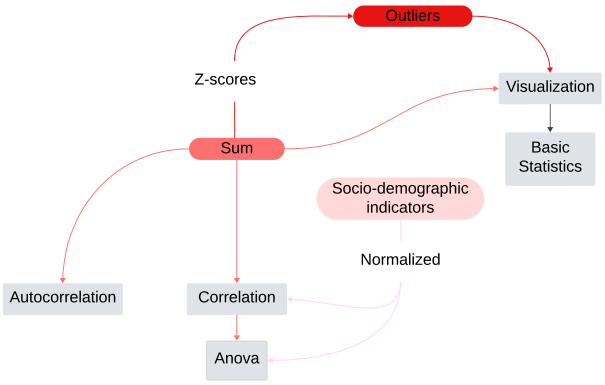


Figure 4: Scheme for Laplace Noise senstibility

VI.1.c Sensitivity Tests

To understand the sensitivity of the MDM dataset to various temporal and contextual factors, several tests are conducted (Figure 5):

- Working days and weekends: This test evaluates differences in mobility patterns between weekdays and weekends, integrating Friday into the weekend category to account for its unique distribution.

- Summer and regular days: This test compares population movements during summer vacations to regular days, using official school holiday dates American Association of Collegiate Registrars and Admissions Officers (2023) .

- National holidays: The impact of national holidays is examined by comparing mobility patterns on holidays to those on equivalent weekdays from preceding weeks. The dates are selected by the use of *datetime* package Feiertagskalender (nd).

- Disasters: The effect of natural disasters on daily movements is studied by analyzing data from three weeks before, during, and three days after disaster events.

For all tests, the visualizations include kernel density plots, and boxplots, cumulative distribution plots emphasizing differences in movement patterns. These are complemented by linear regressions, correlations, and ANOVA tests, which quantify the impact of each factor on mobility (Turkish statistical institue, ndb). Outliers and anomalies are specifically investigated to assess the dataset's robustness under different conditions.

The final test evaluates the impact of disasters on daily movement. The analysis considers three distinct periods:

- same day of the disaster over the three weeks before the disaster,
- the disaster days themselves,
- three days following the disaster.

This method helps isolate the immediate and short-term effects of disasters on mobility patterns. For this test, two natural disasters with significant internal displacement (other than earthquakes) are selected (Figure 5). Internal displacement data is sourced from EM-DAT and IOM (Centre for Research on the Epidemiology of Disasters, 2024; Internal Displacement Monitoring Center, nd) and supplementary datasets. Visualizations, correlations, and ANOVA tests are conducted using Python and QGIS to ensure a comprehensive assessment of disaster impacts.

The disaster part will concerned to two natural disasters (except earthquakes) that provoked the highest internal displacement. The internal displacement data is from the dataset of Internal Displacement Monitoring Center (nd). The two disasters chosen are a storm and a wildfire.

The wildfire does not have any type of day difference, therefore, the district touch by the wildfire is not separated (Figure 5). However, the storm is starting on a week end day which permits to check on the effect of the type of the day even over an extreme event. The storm event has three spatial scales : provincial, districts and villages. As the MDM has the highest scale as districts, the village data are converted into the district scale. In this test, the disasters FP are presented in tables in the provincial and district scale (Figure 5).

To assess the usability of the dataset, it is essential to evaluate its sensitivity and overall reliability. This section focuses on the statistical examination of the MDM, aiming to determine its applicability in analysing displacement patterns associated with natural disasters. The validation process is systematically structured, surround a detailed methodology section to outline the statistical techniques employed, a results section presenting the findings, and a concise summary providing a critical overview of the outcomes (Figure 5).

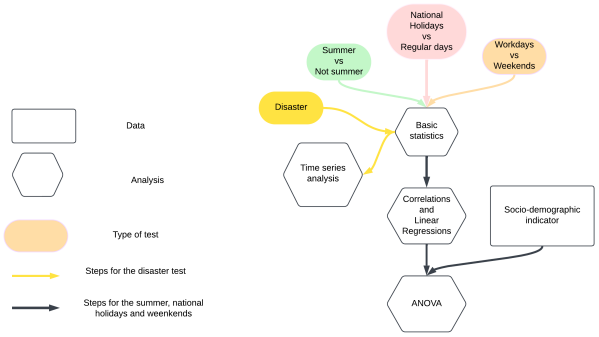


Figure 5: Scheme of the sensitivity test analysis

VI.2 Results

VI.2.a Overall of the *Movement Distribution Maps*

The visualization of the MDM is presented as a time series across all districts and days (Figure 7). This figure offers a short-term perspective on changes, illustrating the nationwide impact and the variations across specific districts.

It is also insightful to observe that the majority of the population travels a distance between 0 and 10 km per day, while only a small proportion covers 10 to 100km and 100 km and over daily. The video[7] representation of the data reflects similar results, but with notable changes in specific districts and a higher concentration of the population travelling longer distances in the northern regions.

Figure 7 displays the daily changes on a national scale. In this figure, the MDM reveals the daily mobility patterns of the population. An oscillation is evident across all movement fraction. The largest portion of the population travelling is between 0 and 10 km daily. The second largest group is comprised of individuals with no movement at all.

Facebook users, in particular, appear to cover fewer distances over 10 km per day. This oscillation is visible in the population's movement trends (Figure 7), and the periodicity is disrupted by factors such as the studied earthquake (which occurred on February 6th), the summer season, national holidays, and other natural disasters. The smallest daily distances observed are those over 100 km, and the time series (Figure 7) indicates a spike in this fraction of the population just before national holidays.

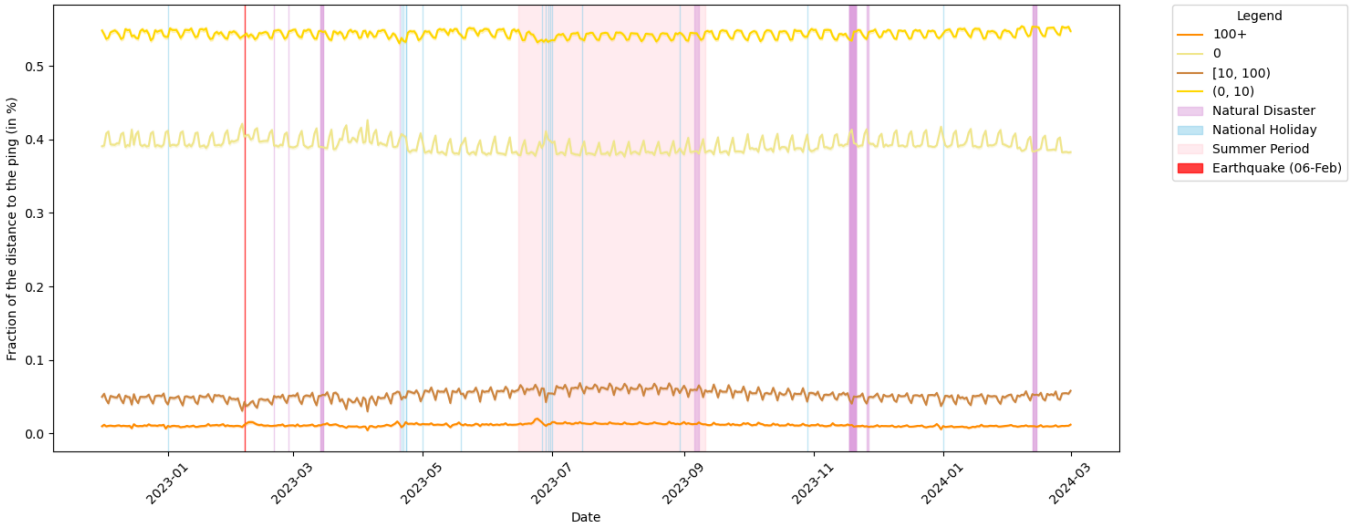


Figure 7: Time series of daily movement of Facebook users per bind of MDM distance.

[7] (*Please the video by clicking on this text*)

The boxplots (Figure 8) and their corresponding statistical descriptions (Table 4) offer a statistical perspective on the data. Due to the presence of minimum values below 0 and maximum values above 1, there is some noise in the fractions. This will be addressed in the following section, which focuses on the outliers.

The figures and table reveal that the largest fraction of the population is moving between 0 and 10 km daily (54%), while the second largest fraction is not moving at all (0.39%). On average, the FP travelling more than 10 km per day is 6%.

Except for the FP in the 0 to 10 km range, the distributions tend to be slightly skewed above the mean. The distribution tails are high, which, when compared to a Gaussian distribution, indicates a broader spread of observations. The FP for distances over 100 km shows the greatest variability and spread.Figure 9 illustrates the differences between these fractions.

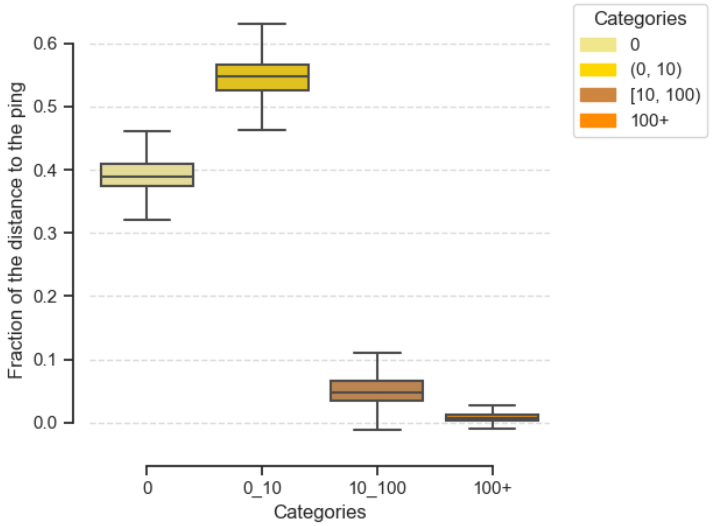
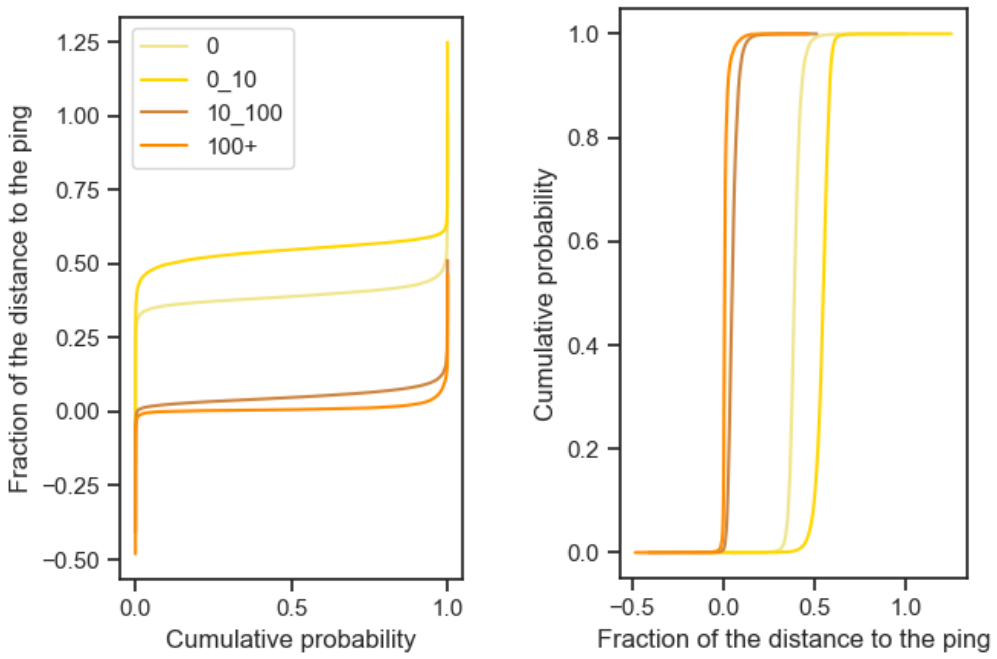


Figure 8: Overview of the distribution dataset studied.

	Minimum	**Maximum**	**Mean**	**Variance**	**Skewness**	**Kurtosis**
No movements	-0.17	1.00	0.39	0.00	1.08	10.17
0 and 10km	-0.07	1.25	0.54	0.00	-1.03	6.19
Between 10 and 100km	-0.41	0.51	0.05	0.00	1.52	9.38
100 km and over	-0.48	0.46	0.01	0.00	2.95	30.20

Table 4: Statistical descriptions of the four intervals.



(a) Cumulative probability over FP (b) FP over cumulative probability

Figure 9: Fraction to the ping (in percent) compared with the cumulation probability of having a district that such such porportion.

As observed in the time series (Figure 7), there is a noticeable oscillation that appears to follow a weekly pattern. Therefore, it is insightful to examine the basic statistics for each FP across the days of the week (Figure 10, Table 5).

For all FPs, Sundays and Saturdays stand out as being different from the workdays. It is also notable that Fridays exhibit distinct behaviour compared to Mondays, Tuesdays, Wednesdays, and Thursdays. The means for the 0 km and 0 to 10 km categories differ by 0.02 and 0.01, respectively, between weekdays and weekends. In the no movement FP, Fridays and Saturdays show a different asymmetry, with more observations falling below the mean.

The kurtosis test indicates that the distributions for these FPs have sharper peaks compared to those in Table 4, which can be linked to the changes in the means. However, there is one exception: the FP for distances over 100 km does not appear to be impacted by the changes across the days of the week.

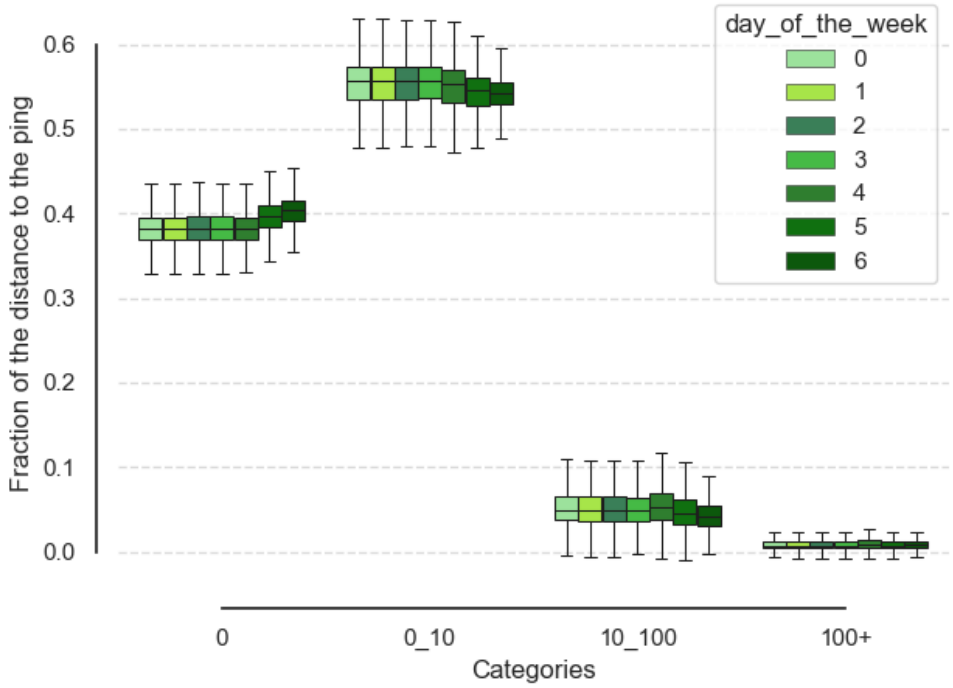


Figure 10: General overview of MDM distribution per type of week day.

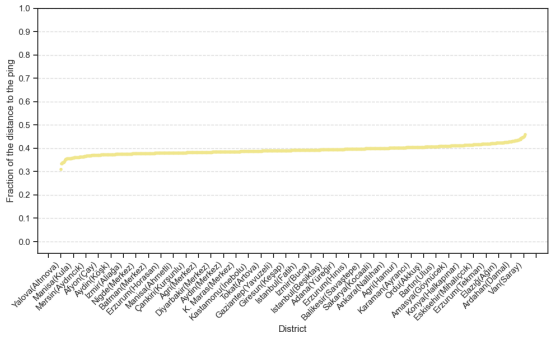
As mentioned earlier, it is important to analyze the outliers and identify which FP is most impacted by them. In Figure 11 and Figure 12, it is evident that the outliers are distributed across different regions. The FPs corresponding to distances of 10 km and over are particularly affected by these outliers. Specifically, the FP for distances of 100 km and over shows outliers primarily in districts located along highways, as shown in Figure 13.

Figure 11 reveals that the increase in the FP performing these long-distance movements is relatively consistent across all districts. However, it is worth noting that the outliers, especially those related to longer distances, are more concentrated in areas with better transportation infrastructure, such as highways. This suggests that the presence of highways might influence the mobility patterns, allowing for higher travel distances in certain districts.
The next step is to analyze spatially whether there are any spatial clusters or outliers, as shown in Figure 14.

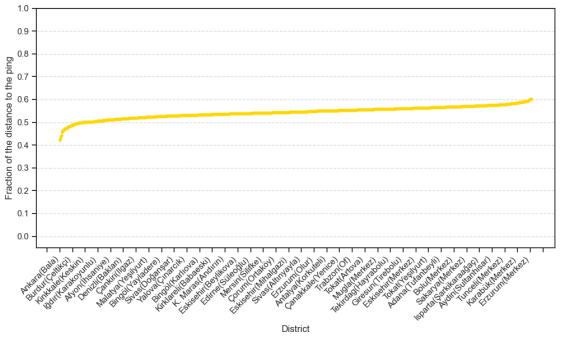
The no movement fraction of MDM reveals low outliers in areas with low population density and settlements smaller than villages, particularly in the western part of Türkiye. These outliers seem to follow the flatlands in that region, as depicted in Figure 13. In contrast, the eastern part of Türkiye exhibits high auto-correlation clusters, which are likely linked to people living in villages or smaller settlements that are still relatively close to towns. This region is also characterized by a more mountainous terrain.

	Monday	**Tuesday**	**Wesdneday**	**Thursday**	**Friday**	**Saturday**	**Sunday**
No movement distance bind							
Minimum	0.22	0.23	0.23	0.24	0.20	0.24	0.16
Maximum	0.60	0.64	0.57	0.63	0.58	0.70	0.62
Mean	0.38	0.38	0.38	0.38	0.38	0.40	0.40
Variance	$5.56\text{x}10^{-4}$	$5.59\text{x}10^{-4}$	$5.58\text{x}10^{-4}$	$5.57\text{x}10^{-4}$	$5.47\text{x}10^{-4}$	$5.31\text{x}10^{-4}$	$4.82\text{x}10^{-4}$
Skewness	0.60	0.71	0.59	0.71	-0.51	-0.42	0.27
Kurtosis	3.01	3.56	2.19	3.49	1.91	3.33	4.07
Displacements from 0 to 10 km bind							
Minimum	0.32	0.31	0.32	0.24	0.28	0.09	0.31
Maximum	0.74	0.71	0.66	0.67	0.67	0.86	0.83
Mean	0.55	0.55	0.55	0.55	0.54	0.54	0.54
Variance	$9.60\text{x}10^{-4}$	$9.52\text{x}10^{-4}$	$9.34\text{x}10^{-4}$	$9.31\text{x}10^{-4}$	$9.53\text{x}10^{-4}$	$7.88\text{x}10^{-4}$	$5.60\text{x}10^{-4}$
Skewness	-0.86	-0.87	-0.86	-0.89	-0.88	-0.94	-0.81
Kurtosis	1.89	1.92	1.83	2.15	1.91	3.85	3.81
Displacements from 10 to 100 km bind							
Minimum	-0.11	$-3.49\text{x}10^{-2}$	$-4.46\text{x}10^{-2}$	$-2.03\text{x}10^{-2}$	-0.10	$-4.76\text{x}10^{-2}$	$-6.58\text{x}10^{-2}$
Maximum	0.28	0.32	0.31	0.32	0.32	0.39	0.32
Mean	$5.34\text{x}10^{-2}$	$5.31\text{x}10^{-2}$	$5.27\text{x}10^{-2}$	$5.27\text{x}10^{-2}$	$5.59\text{x}10^{-2}$	$4.98\text{x}10^{-2}$	$4.38\text{x}10^{-2}$
Variance	$6.10\text{x}10^{-4}$	$5.94\text{x}10^{-4}$	$5.79\text{x}10^{-4}$	$5.79\text{x}10^{-4}$	$6.30\text{x}10^{-4}$	$7.88\text{x}10^{-4}$	$3.54\text{x}10^{-4}$
Skewness	1.31	1.32	1.31	1.29	1.10	1.17	1.32
Kurtosis	3.62	3.80	1.83	3.58	2.69	3.65	6.91
Displacements from 100 and over bind							
Minimum	-0.11	$-4.74\text{x}10^{-2}$	$-3.61\text{x}10^{-2}$	$-6.59\text{x}10^{-2}$	$-9.36\text{x}10^{-2}$	$-8.91\text{x}10^{-2}$	$-8.80\text{x}10^{-2}$
Maximum	0.22	0.22	0.30	0.30	0.27	0.24	0.23
Mean	$1.11\text{x}10^{-2}$	$1.11\text{x}10^{-2}$	$1.11\text{x}10^{-2}$	$1.14\text{x}10^{-2}$	$1.27\text{x}10^{-2}$	$1.11\text{x}10^{-2}$	$1.17\text{x}10^{-2}$
Variance	$2.72\text{x}10^{-4}$	$2.69\text{x}10^{-4}$	$2.53\text{x}10^{-4}$	$2.73\text{x}10^{-4}$	$3.06\text{x}10^{-4}$	$2.37\text{x}10^{-4}$	$2.27\text{x}10^{-4}$
Skewness	4.44	4.37	4.54	4.46	4.38	4.58	4.68
Kurtosis	26.57	25.46	28.76	27.58	26.19	30.29	31.32

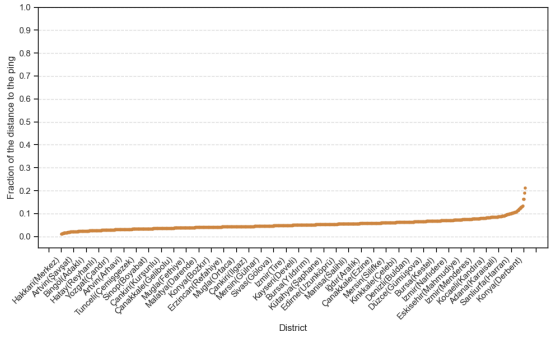
Table 5: Statistical Description of MDM per type of week day.



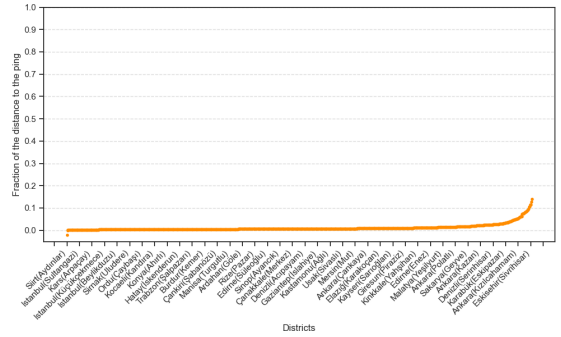
(a) No movement



(b) Displacements from 0 to 10 km



(c) Displacements from 10 to 100 km



(d) Displacements over 100 km

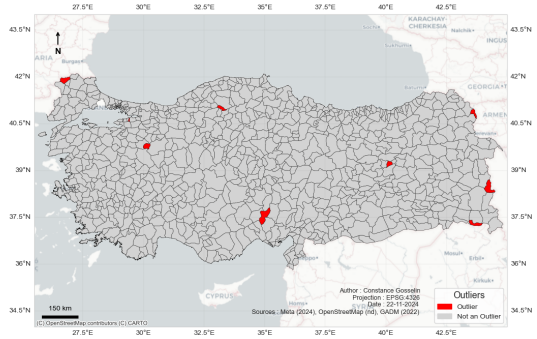
Figure 11: Daily displacement sorted per districts and per ascending order. (*Please find a zoomed and interactive figure by clicking on this text*)

For the fraction of the population travelling 0 to 10 km per day, high-high clusters are observed in areas close to towns but are more dispersed across Türkiye. In the central part of Türkiye, low clusters and outliers are visible, corresponding to areas with fewer people travelling these distances. This can be associated with regions that have larger towns and the proximity of the capital. Additionally, the highway and primary roads in this area may influence the movement patterns.

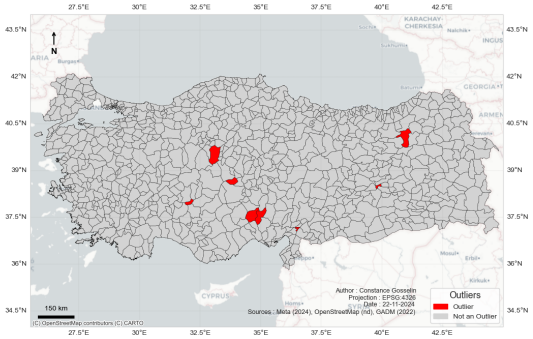
The 10 to 100 km per day FP shows high clusters primarily in the western part of the country. Looking at Figure 13, the population that seems to travel the most in this range are those in settlements smaller than villages but close to towns. On the eastern side of Türkiye, low outliers follow settlement patterns, particularly in areas with towns. The low clusters in these districts are also linked to such towns.

For the 100 km and over FP fraction, high clusters are observed along highways and primary roads, with high outlier districts typically located close to these high clusters, suggesting a strong relationship between them. Low outliers are mainly found on the east coast and in the southwestern part of the country. In the western region, these low outliers appear to be linked to the flat lands between mountainous areas, while in the east, the low outliers seem to follow settlements that consist mostly of villages and are located near primary and secondary roads.

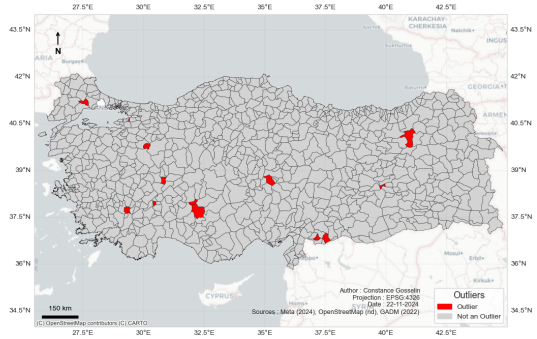
In summary, the spatial analysis highlights that transportation infrastructure, such as highways and primary roads, plays a significant role in shaping mobility patterns, with clusters and outliers largely corresponding to areas with specific settlement types and geographic features.



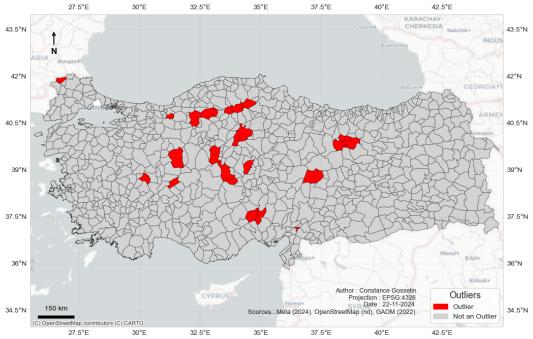
(a) FP doing no movement a day



(b) FP doing 0 to 10 km a day

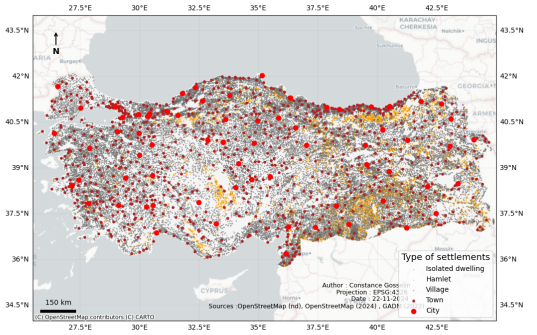


(c) FP doing 10 to 100 km a day



(d) FP doing 100 km and over a day

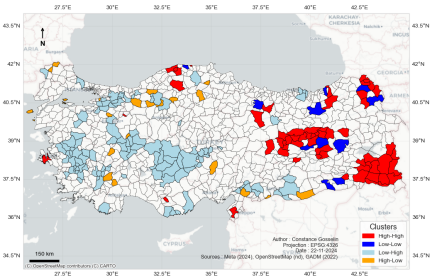
Figure 12: Turkish districts identified as outliers based on the mean of the FP of selected distance bind.



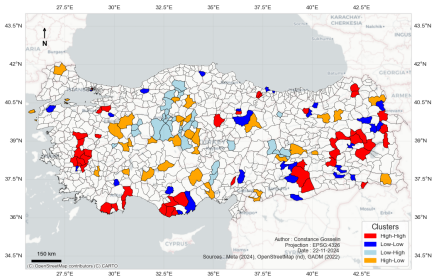
(a) Type of settlement in Türkiye

(b) Main roads in Türkiye

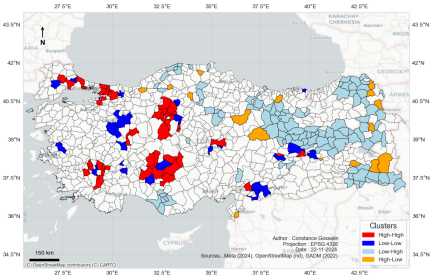
Figure 13: Settlements and Road infrastructure in Türkiye.



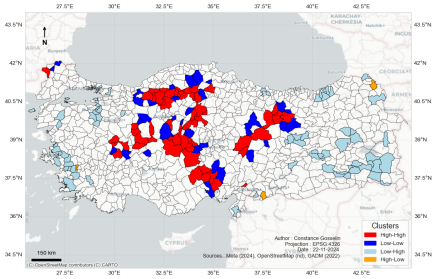
(a) No movements category



(b) 0 to 10 km over distance category



(c) 10 to 100 km over distance category



(d) 100 km over distance category

Figure 14: Cluster of local Moran per distance category.

Summary

The overall results provides informations on the mobility habit and the sensibility of the movements.

The first main information is its weekly oscillation and the disturbance of some events in the oscillation. Figure 10 verifies the changes over a type of a week. It shows that Mondays, Tuesdays, Wednesdays and Thursdays are not much disturbances but Fridays, Saturdays and Sundays do. Table 5 contribute to add proof in that ways with the description of the spreader queue and a change of the mean.

THe population are mostly doing no movement or 0 to 10 km. 39% of the population are no moving in at a daily mobility and 54% of the population are doing 0 to 10 km in a daily basis.

In terms of the distribution of the FP binds, all of them have more values of FP than the mean FP. The exception in the bind of 0 to 10km which have more FP values doing less than the mean of the FP distance bind.

In the outliers of the FP distance binds, the distance interval the lost impacted is the 100 km and over. It is related with the highway going to the capital and Iran.

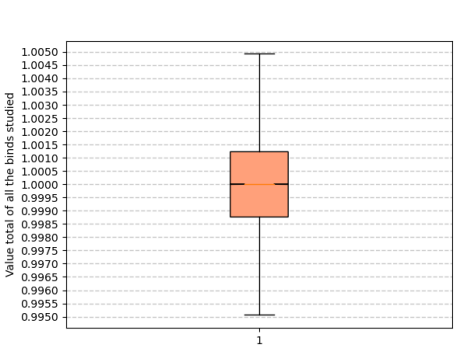
VI.2.b Laplace Noise test *Movement Distribution Maps*

This section examines the influence of Laplace noise on the dataset, focusing on its statistical implications and accuracy.

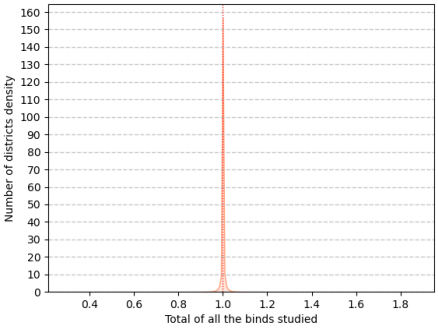
From Figure 15 and Table 6, it is evident that the Laplace noise minimally affects the dataset's integrity. The mean value of 1.00 aligns with the theoretical expectation, and the skewness indicates that the MDM sum of all fractions is symmetric. However, the kurtosis is exceptionally high (247.00), suggesting the presence of a substantial number of high-value observations, which could be classified as outliers.

Figure 15 illustrates that most values are tightly clustered around the mean, reinforcing the symmetry and stability of the dataset despite the noise. The high kurtosis, as noted in Table 6, implies a significant concentration of extreme observations, which may result in a higher proportion of values being labeled as outliers.

The relative error, calculated at 1.35×10^{-8}, demonstrates that MDM maintains a high degree of accuracy. This minimal error indicates that the Laplace noise does not significantly distort the dataset or impair the accuracy of the MDM results, as confirmed in Table 6.



(a) Overall of the sums of all fraction



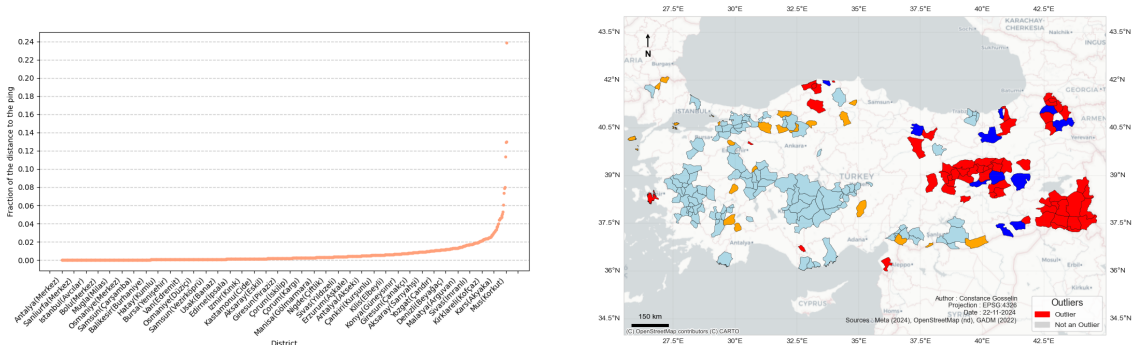
(b) Density of the disctricts for the total of all fractions

Figure 15: Overall of the sums of all binds.

Statitical Description	Total of all binds
Mean	1.00
Variance	0.00
Skewness	0.85
Kurtosis	247.20
Mean absolute error	5.74×10^{-2}
Mean relative Error	1.35×10^{-8}

Table 6: Basic statistics.

The outliers show under Figure 16 defined the districts with the highest error. They are dispersed in the country.



(a) Asbolute Error displacement sorted per districts*Please find a zoomed and interactive figure by clicking on this text*)

(b) Districts in Türkiye identified as outliers based on the absolute Error of the total of fraction

Figure 16: Outliers.

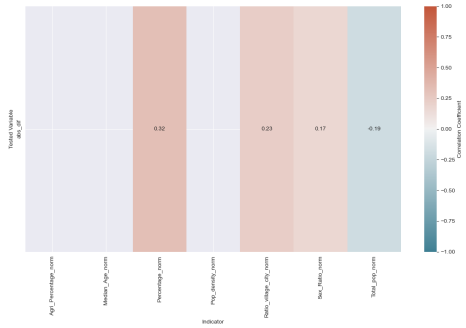
The correlations with socio-demographic indicators (Figure 17) illustrate the influence of these variables on the aggregate fractions. While the data reveals a range of dispersed linear correlations, a notable exception arises with the percentage of illiteracy[8] within the district. A distinct and robust relationship is observed with an alternative form of correlation.

Specifically, the total number of FP exhibits a strong negative correlation with both the total population and population density, alongside a weak positive correlation with the sex ratio, the percentage of illiteracy, and the proportion of the population residing in urban or rural areas. In contrast, no discernible correlation is evident with the percentage of individuals living in agricultural area.

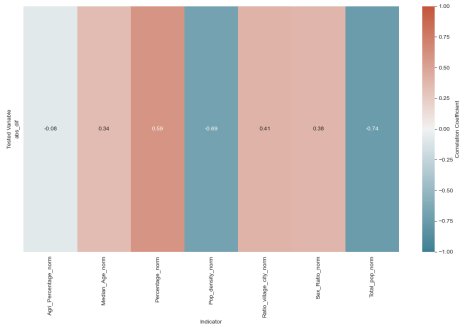
Building on the correlation analysis, an ANOVA was conducted to assess the relationship between the dataset sums and socio-demographic indicators. The analysis identified significant associations with the sex ratio, total population, percentage of illiteracy, and rural-urban ratio. These indicators exhibit a notable influence on the aggregate values of the dataset (Table 7).

The Local Moran index reveals the presence of low spatial outliers predominantly in the western regions, while high spatial clusters, along with adjacent high outliers, are concentrated in the east. As illustrated in Figure 13, it is notable that the majority of low-outlier anomalies are located in the west, often in areas characterized by settlements smaller than villages. In contrast, the high autocorrelation clusters in the east align closely with a mountainous range and a secondary road network.

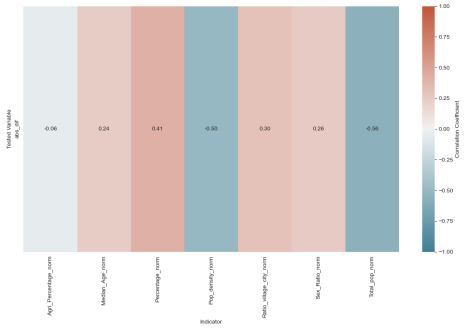
[8]Also referred to as *Percentage_norm*



(a) Correlation with Pearson



(b) Correlation with Spearman



(c) Correlation with Kendall

Figure 17: Correlation tests of absolute differences between the observed sum and the theoretical sum.

Indicators	**F-statistic**	**p-value**	**Significative**
Median Age Normalized	2.67	0.10	Non-significant
Sex Ratio Normalized	44.64	$3.62\text{x}10^{-10}$	Significant
Total Population Normalized	32.78	$1.30\text{x}10^{-8}$	Significant
Percentage of Agriculture Area Normalized	0.01	0.90	Non-significant
Percentage of Illiterate	87.73	$3.61\text{x}10^{-20}$	Significant
Population Density Normalized	0.12	$0.73\text{x}10^{-1}$	Non-significant
Ratio Village and City Normalized	17.01	$3.98\text{x}10^{-5}$	Significant

Table 7: Test Anova of the sum with socio-demographic indicators.

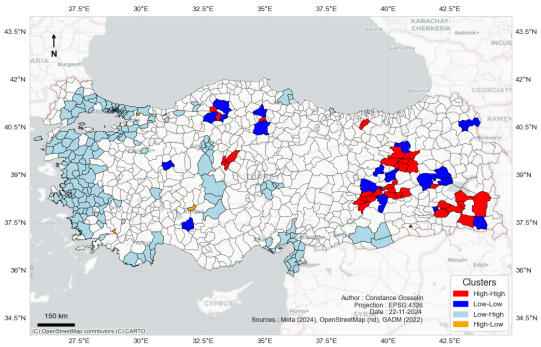


Figure 18: Cluster Noise of the sum

.

Summary

This section contributes to understanding the impact of Laplace noise on the sum of the four intervals.

The values presented in Table 6 confirm that the dataset maintains its accuracy despite the presence of noise. However, the sums are influenced by outliers, as evidenced by the kurtosis and z-score outlier analysis (Figure 16 and Table 6).

Further analysis using ANOVA and spatial autocorrelation (Table 7 and Figure 18) demonstrates that MDM noise is significantly affected by socio-demographic factors, including the sex ratio, illiteracy rate, total population, and the rural-urban ratio.

These findings suggest that MDM process is influenced by the number of individuals capable of using the Facebook app on mobile devices.

VI.2.c Sensitivity tests

As observed in the overall analysis, daily mobility exhibits sensitivity not only across different days but also in response to seasonal factors, national holidays, and disasters. To explore these dynamics further, sensitivity tests are applied to the MDM dataset. This section presents the results of these tests.

VI.2.d Workday vs. Weekend Test: *Movement Distribution Maps*

The first test involves examining the differences between workdays and weekends. The statistical analysis follows a series of steps: basic statistical analysis, correlation between workdays and weekends, linear regression, correlation with socio-demographic indicators, comparison of the differences between workdays and weekends, and ANOVA with the indicators.

The basic statistical results are presented in Figure 19, Figure 20, and Table 8. The analysis suggests a higher number of individuals remaining stationary during weekends, which appears to influence the three other FP categories. The distribution of movement is broader on weekdays. The kurtosis indicator is higher across all FP categories, particularly in the no movement and 0 to 10 km groups, where distribution differences are more pronounced. The largest maximum difference is observed in the FP category 0 to 10 km, with a difference of 0.12. Minimum values also show notable differences, especially in the 0 to 10 km category, where the difference between days is 0.15. The FP category most affected by asymmetry differences is no movement, with a difference of 0.37.

As it has been seen in the Overall observations, the daily mobility seems to be sensitive over the days but also over summers, national holidays and disasters. Therefore, sensitivity tests on these subjects are applied on MDM. This sections follows the results of each tests.

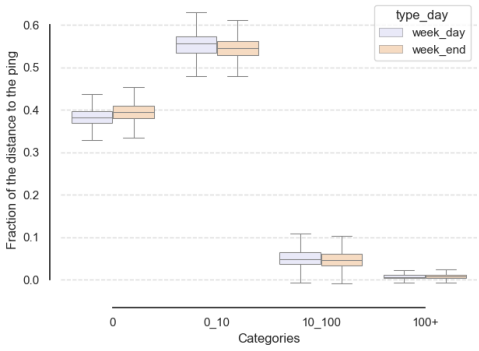
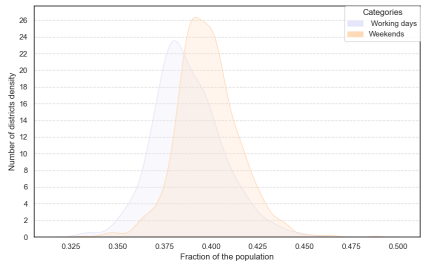
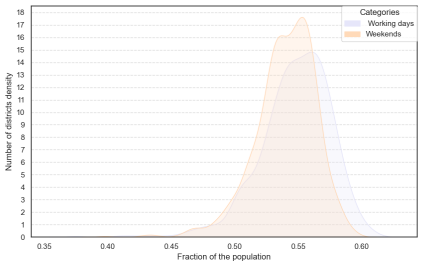


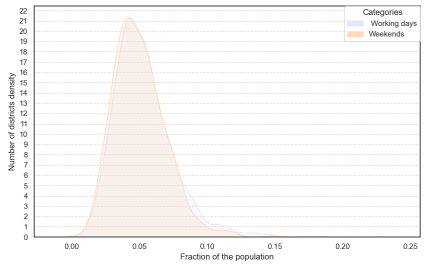
Figure 19: General overview of the distribution of the database comparing workdays and weekends.



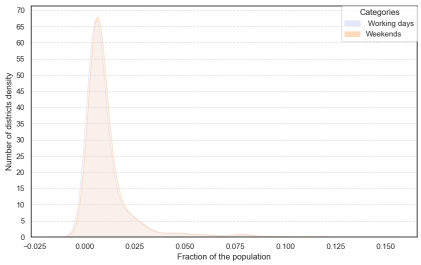
(a) No movements category



(b) 0 to 10 km distance category



(c) 10 to 100 km distance category



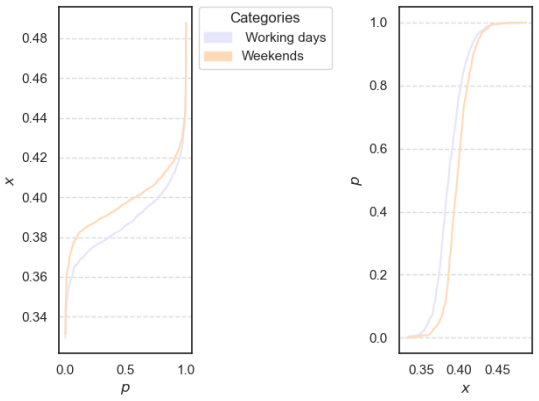
(d) 100 km and over distance category

Figure 20: Kernel density plot of the distribution of MDM per category and comprising workays and weekends.

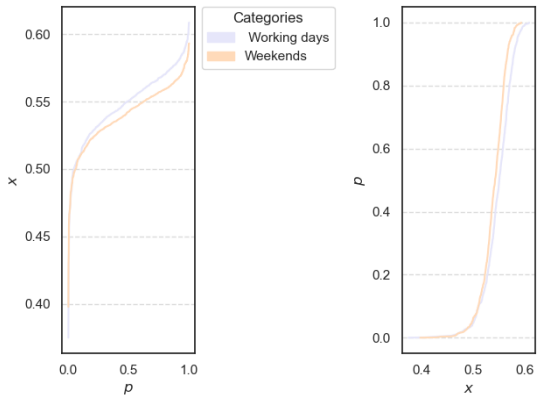
Phase	**Minimum**	**Maximum**	**Mean**	**Variance**	**Skewness**	**Kurtosis**
Distance: 0						
workday	0.22	0.64	0.38	5.57×10^{-4}	0.65	1.95
Weekend	0.16	0.70	0.39	5.91×10^{-4}	0.28	2.49
Distance: 0-10						
workday	0.24	0.74	0.55	9.44×10^{-4}	-0.87	1.95
Weekend		0.86	0.54	7.90×10^{-4}	-0.83	2.95
Distance: 10-100 and week end						
workday	-0.11	0.32	5.29×10^{-2}	5.93×10^{-4}	1.31	3.66
Weekend	-0.10	0.40	4.99×10^{-2}	5.36×10^{-4}	1.24	3.88
Distance: 100						
workday	-9.75×10^{-2}	0.30	1.12×10^{-2}	2.72×10^{-4}	4.46	27.12
Weekend	-9.36×10^{-2}	0.27	1.17×10^{-2}	2.57×10^{-4}	4.56	29.26

Table 8: Statistical summary of distribution of MDM comparing workdays and weekends.

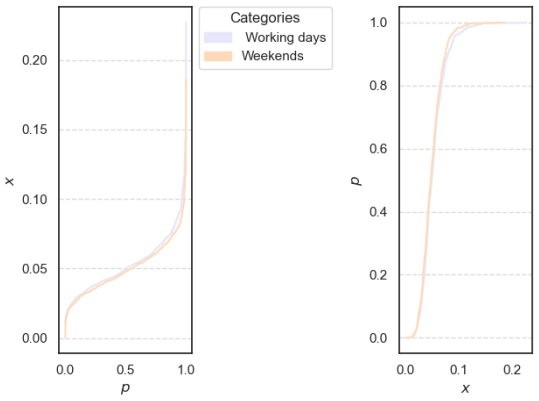
In order to continue with the differences between the workdays and week ends, Figure 21 presents the differences over a cumulative probability. FP no movements expresses a net difference within the first probabilities (0.01). FP betwwen 0 and 10 km starts the change by the 0.3 and the gap increases slightly. The two lasts does not show such differences.



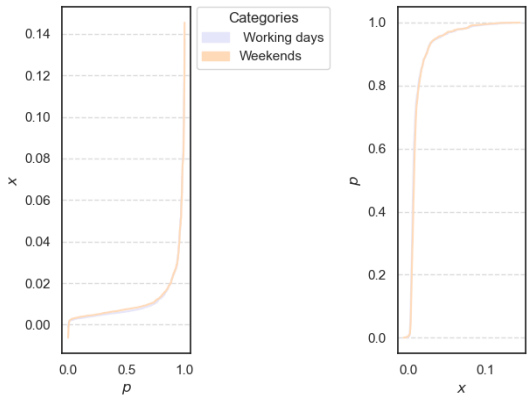
(a) MDM's FP not moving



(b) MDM's FP doing 0 to 10 km a day



(c) MDM's FP doing 10 to 100 km a day



(d) MDM's FP doing 100 km and over a day

Figure 21: The fraction of pings compared with the cumulative probability of a district exhibiting such a proportion and per workdays and weekends categories.

Linear regressions (Table 9) are conducted to examine the relationships between workdays and weekends. Since the data originates from the same dataset, all results are statistically significant, with narrow confidence intervals for the estimates. The 0 to 10 km FP category shows the strongest effects, followed by the 10 to 100 km, no movement, and 100 km or more categories.

	No movement	*From 0 to 10 km*	*From 10 to 100 km*	*From 100km and over*
Coefficient	1.0760	1.1279	1.0862	1.0159
Standard Error	0.013	0.010	0.011	0.008
t	82.507	114.163	97.326	122.489
p-value	0.00	0.00	0.00	0.00
Minimum intervalle of confidence (0.025)	1.050	1.109	1.064	1.000
Maximal intervalle of confidence (0.975)	1.0159	1.147	1.108	1.032

Table 9: Linear Regression of workday over weekends per distance categories.

Given that the sum is relatively influenced by socio-demographic indicators, it is im-

(a) Correlation with Pearson

(b) Correlation with Spearman

(c) Correlation with Kendall

Figure 22: Correlation tests of absolute difference (abs_dif) and difference (dif) of workday (WD) and weekends (WE) FP by socio-demographic indicators.

portant to investigate whether these indicators are related to and correlated with both the absolute differences and the differences in the dataset (Figure 22).

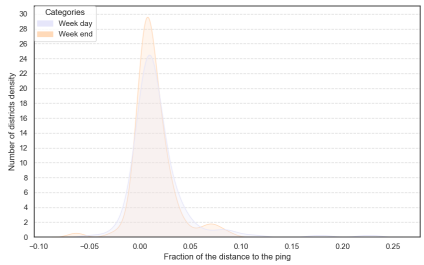
Spearman and Kendall correlations reveal a moderate positive correlation between population density and the absolute difference in the no movement FP category. The median age exhibits a moderate negative correlation with the no movement FP and a moderate positive correlation with the FP category of 100 km and over. Other significant correlations show coefficients close to zero.

ANOVAs (Table 10) provide insight into the influence of socio-demographic indicators on daily movement patterns within districts. The four fractions of the population's daily movement are significantly impacted in different ways by these indicators. Notably, median age and total population exhibit similar significance patterns. Median age consistently shows the highest F-statistic across most FP categories. For the FP category between 10 to 100 km, the percentage of illiteracy produces the highest F-statistic, while for the no movement FP category, illiteracy also stands out as the most significant factor. An interesting finding is the impact of the urban-rural ratio on both the no movement FP and the 100 km and over FP categories. Additionally, sex appears to play a notable role in the differences observed in the no movement and 0 to 10 km FP categories.

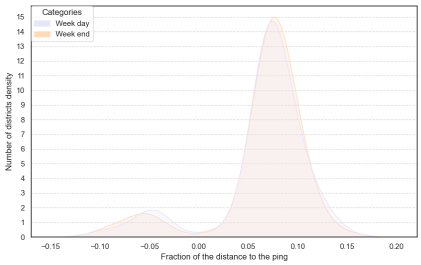
Indicators	**F-statistic**	**p-value**	**Significative**
Distance: 0			
Median Age Normalized	60.61	1.03×10^{-14}	Significant
Sex Ratio Normalized	186.52	6.01×10^{-41}	Significant
Total Population Normalized	7.72	5.51×10^{-3}	Significant
Percentage of Agriculture Area Normalized	1.27	2.59×10^{-1}	Non-significant
Percentage of Illiterate	221.57	5.4×10^{-48}	Significant
Population Density Normalized	5.0×10^{-5}	0.998	Non-significant
Ratio Village and City Normalized	19.71	9.40×10^{-6}	Significant
Distance: 0 - 10			
Median Age Normalized	290.379818	1.52×10^{-61}	Significant
Sex Ratio Normalized	116.337929	1.62×10^{-26}	Significant
Total Population Normalized	47.495373	7.02×10^{-12}	Significant
Percentage of Agriculture Area Normalized	0.642593	0.4228530	Non-significant
Percentage of Illiterate	8.660061	3.28×10^{-3}	Significant
Population Density Normalized	0.378908	0.538	Non-significant
Ratio Village and City Normalized	36.312006	1.94×10^{-9}	Non-significant
Distance: 10 - 100			
Median Age Normalized	41.88	num1.17e-10	Significant
Sex Ratio Normalized	0.29	0.589	Non-significant
Total Population Normalized	6.76	9.39×10^{-3}	Significant
Percentage of Agriculture Area Normalized	2.68	0.101	Non-significant
Percentage of Illiterate	91.04	3.34×10^{-21}	Significant
Population Density Normalized	1.9610^{-2}	0.888	Non-significant
Ratio Village and City Normalized	0.741	0.389	Non-significant
Distance: 100 km and over			
Median Age Normalized	96.42	2.44×10^{-22}	Significant
Sex Ratio Normalized	7.22	7.25×10^{-3}	Significant
Total Population Normalized	19.51	1.04×10^{-5}	Significant
Percentage of Agriculture Area Normalized	5.40	2.02×10^{-2}	Significant
Percentage of Illiterate	1.37	0.241	Non-significant
Population Density Normalized	0.658	0.417	Non-significant
Ratio Village and City Normalized	12.27	4.69×10^{-4}	Significant

Table 10: Anova test of workdays and week end by socio-demographic indicators per daily movement fractions.

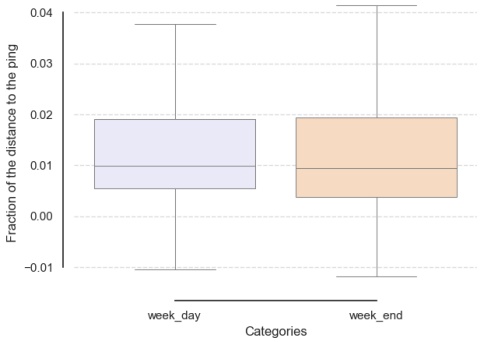
As highlighted in the overall results, outliers play a significant role in the complexity of daily movement patterns within the Facebook user population, particularly in the 10 to 100 km and 100 km and over categories. Observations from Figure 23 suggest distribution asymmetries and a broader spread of the data for weekdays in the 10 to 100 km FP category. The "100 km and over" fraction appears to be more heavily impacted by Laplace noise compared to the 10 to 100 km category. The cumulative probability indicates a notable increase in values beyond 0.13 for the 10 to 100 km category and 0.6 for the 100 km and over category.



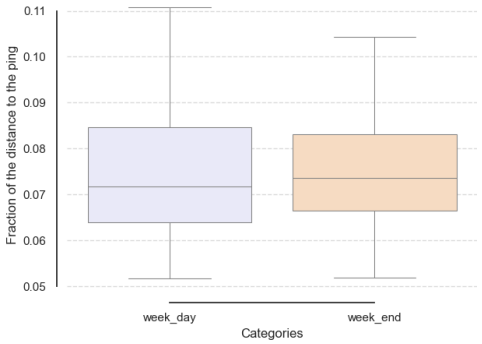
(a) Kernel density plot of the 10 and 100 km category outlier



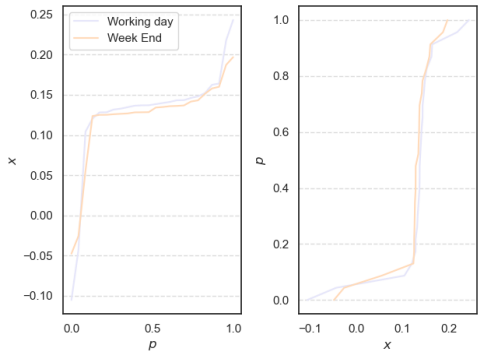
(b) Kernel density plot of the 100 km and over category outlier



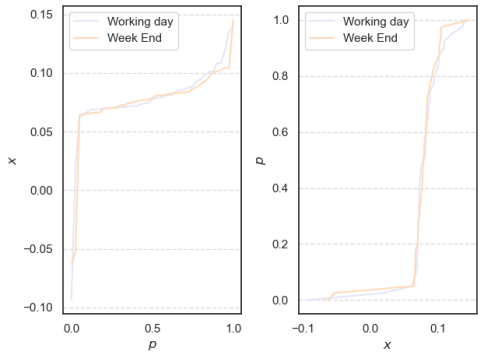
(c) Bowplot plot of the 10 to 100 km category outlier



(d) Boxplot plot of the 100 km and over category outlier



(e) Cumulative probability plot of the 10 to 100 km category outlier



(f) Cumulative probability plot of the 100 km and over category outlier

Figure 23: General visualisation of the outliers.

The correlation is strong and significance which is linked with the fact that they come from the same dataset . The correlation with the daily movement outliers and the socio-demographic indicators shows spearman correlation of 0.51 of the workdays and weekends in the bind of 10 to 100 km outliers. The 100 km and over binds have a spearman correlation of 0.58. The outliers difference from the bin 10 to 100 km have a correlation with spearman of 0.43 with the indicator ratio urban vs rural. These oultiers have a strong correlation with pearson of 0.53 with the percentage of illeterate. The bind of 100km and over has a median age of 0.36 with pearson.

Summary

This section contributes to understanding how the FPs engaging in different distances vary depending on whether the day is a workday or a weekend day (including Fridays).

The no movement FP category is the most affected, with a noticeable increase in the population remaining stationary. However, the distribution in this category is more spread out on weekends compared to workdays. The 100 km and over FP category is less asymmetric on weekends.

Apart from the no mobility FP category, the other intervals show less variation on weekends, which may be linked to the daily pendular mobility during workdays. This is further evidenced by the greater changes observed in the two primary categories: no mobility and 0 to 10 km.

The 100 km and over FP category shows the greatest sensitivity to socio-demographic indicators on a daily basis.

The total population and illiteracy rate are the socio-demographic indicators most strongly associated with changes during the week.

Regarding outliers, illiteracy appears to be significantly affected. Outliers in the 10 to 100 km distance category show a notable relationship with the rural-urban ratio. Additionally, the median age is related to both workday and weekend outliers.

VI.2.d Summer and non summer test *Movement Distribution Maps*

This test examines the variations in daily movement between regular days and those during the summer school period. As part of the analysis, the initial statistical results are presented in Figure 24, Figure 25, and Table 11. Figure 24 reveals a mean difference between the two periods for the no movement, 0 to 10 km, and 100 km and over FP categories.

Except for the 100 km and over category, the distribution is more spread out during the summer, as evidenced by the changes in kurtosis shown in Table 11. However, the asymmetry of the distributions remains relatively stable. The 0 to 10 km FP category is most affected by asymmetry, with more values falling below the mean. The largest change in mean values is observed in the 100 km and over FP category.

Figure 26 relate the impact of summer days in the daily mobility. The graphs shows a difference of 0.02 for FP no movement and 10 to 100km. The two other fraction does not show relevant differences. In Summer, there is a higher FP not moving and a lower FP doing between 10 to 100 km a day.

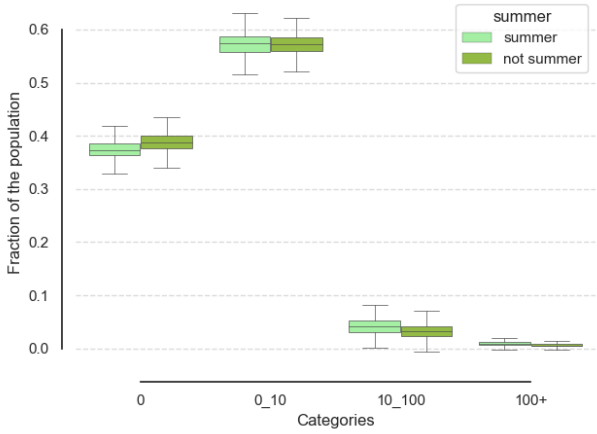
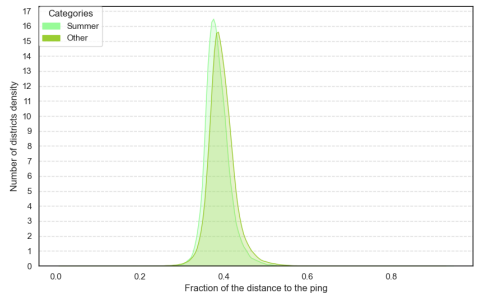
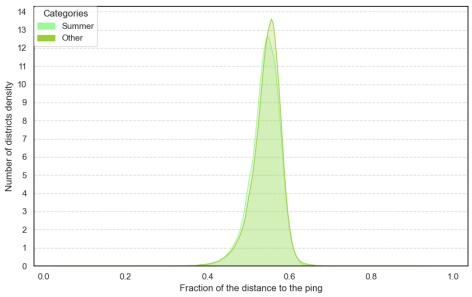


Figure 24: Boxplots of the MDM distributions compared to Summer and not summer dates.

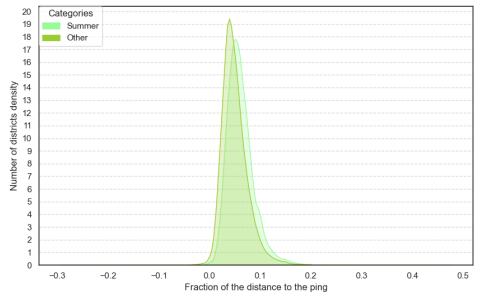
As previously discussed, Figure 27 shows a high positive correlation between regular days and the summer period for the four FP categories. However, there is a slight difference of 0.93 for the no movement FP category, compared to 0.97 for the others. An interesting observation is the strong negative correlation between the 0 to 10 km and 10 to 100 km FP categories. Table 12 presents a lower coefficient for the no movement FP category and a higher coefficient for the 0 to 10 km FP category.



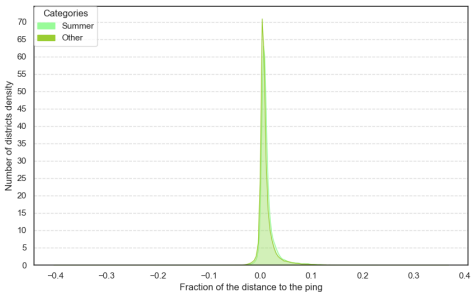
(a) No movements



(b) Displacements from 0 to 10 km



(c) Displacements from 10 to 100 km

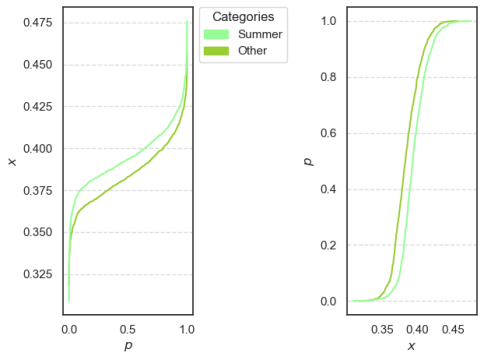


(d) Displacements over 100 km

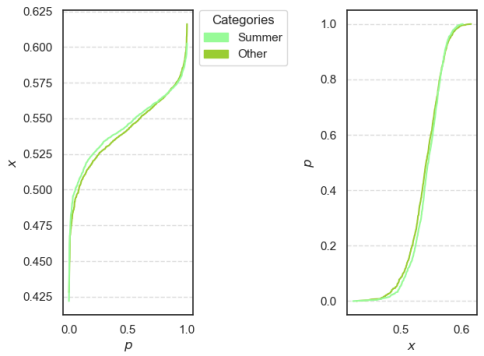
Figure 25: Kernel Density plots of distance categories compared with Summer and not summer.

Phase	**Minimum**	**Maximum**	**Mean**	**Variance**	**Skewness**	**Kurtosis**
Distance: 0						
Other periods	0.33	0.44	0.38	3.48×10^{-4}	0.34	0.12
Summer	0.16	0.70	0.39	3.12×10^{-4}	0.31	0.71
Distance: 0-10						
Other periods	0.43	0.62	0.54	7.51×10^{-4}	-0.56	0.71
Summer	0.42	0.60	0.54	5.940×10^{-4}	-0.78	1.39
Distance: 10-100						
Other periods	1.54×10^{-2}	0.20	4.89×10^{-2}	5.09×10^{-4}	1.13	2.96
Summer	9.88×10^{-3}	0.21	4.90×10^{-2}	4.36×10^{-4}	1.59	6.22
Distance: 100						
Other periods	-7.76×10^{-3}	0.17	1.34×10^{-2}	2.68×10^{-4}	4.19	23.11
Summer	-5.50×10^{-3}	0.13	1.05×10^{-2}	2.24×10^{-4}	4.26	21.83

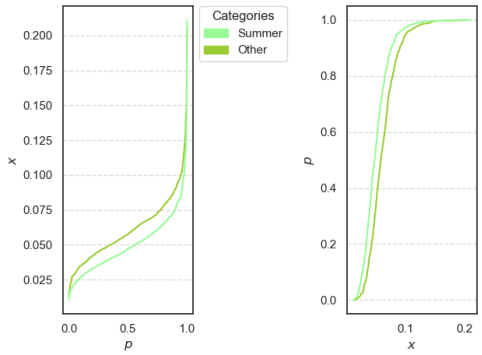
Table 11: Statistical Description of population daily movement comparing Summer and not summer dates.



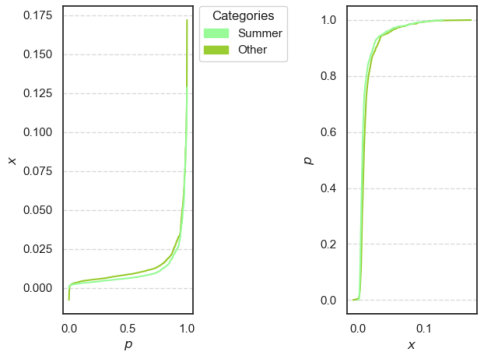
(a) No movement



(b) Daily movement from 0 and 10 km

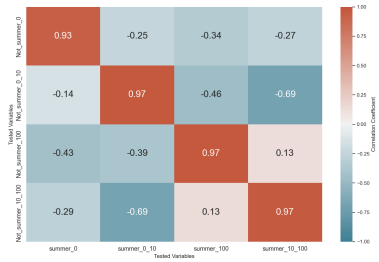


(c) Daily movement from 10 to 100 km

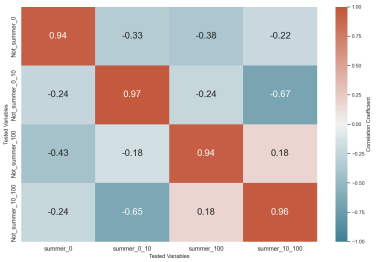


(d) Daily movement from 100 km and over

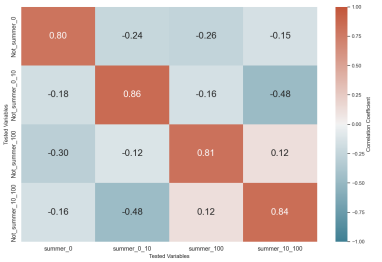
Figure 26: Cumulative probability of the FP comparing summer and not summer values.



(a) Correlation with Pearson



(b) Correlation with Spearman



(c) Correlation with Kendall

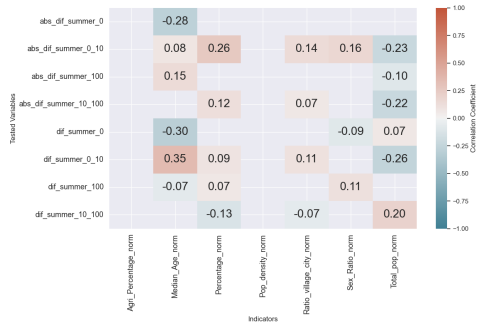
Figure 27: Correlation tests of summer and not summer value categories.

	No movement	*From 0 to 10 km*	*From 10 to 100 km*	*From 100km and over*
Coefficient	0.983	1.0881	1.0401	1.0611
Standard Error	0.013	0.009	0.009	0.009
t	78.643	119.190	109.867	121.758
p-value	0.00	0.00	0.00	0.00
Minimum intervalle of confidence (0.025)	0.959	1.070	1.021	1.044
Maximal intervalle of confidence (0.975)	1.008	1.106	1.059	1.078

Table 12: Linear Regression of Summer and Not summer per distance categories.

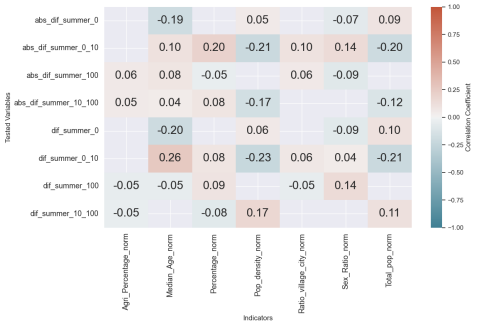
Correlations and ANOVAs are employed to examine the relationship between socio-demographic indicators. Except for the 0 to 10 km FP category and the median age indicator, Figure 28 demonstrates that there is no significant correlation between the absolute differences and the differences.

Moreover, the observed distinction in correlations between the absolute differences and the differences suggests that the relationship between the indicators is not solely determined by their values, but also by their directional trends.



(a) Correlation with Pearson

(b) Correlation with Spearman



(c) Correlation with Kendall

Figure 28: Correlation tests of absolute difference (abs_dif) and difference (dif) of summer and not summer with socio-demographic indicators.

To assess the potential impact of socio-demographic indicators on changes in daily movement distances, Table 13 presents the results of an ANOVA. The analysis reveals a significant influence of literacy and the sex ratio on the no movement FP. Additionally, median age and the urban-rural ratio are also associated with changes in the no movement FP.

For the population engaged in daily movement of 0 to 10 km, the sex ratio and median age show a strong statistical impact, while literacy and the urban-rural ratio affect the displacements to a lesser degree, as indicated by lower F-statistics.

Displacements within the 10 to 100 km category are influenced by age, total population, and literacy, with literacy showing the highest F-statistic. For the 100 km and over category, the median age, total population, and the urban-rural ratio are the most significant factors, with median age yielding the highest F-statistic.

It is worth noting that the highest F-statistics for the 10 to 100 km and 100 km and over categories are lower than those for the "no movement" and "0 to 10 km" categories.

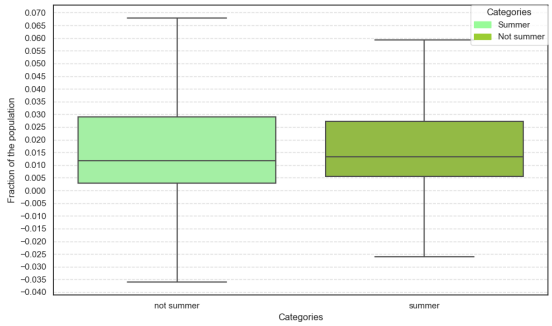
Indicators	**F-statistic**	**p-value**	**Significative**
Distance : No movement			
Median Age Normalized	50.54	1.65×10^{-12}	Significant
Sex Ratio Normalized	193.70	5.30×10^{-42}	Significant
Total Population Normalized	9.31	2.31×10^{-3}	Non-significant
Percentage of Agriculture Area Normalized	0.64	0.422	Non-significant
Percentage of Illiterate	212.34	1.12×10^{-45}	Significant
Population Density Normalized	1.59×10^{-3}	0.968	Non-significant
Distance : 0 - 10 km			
Median Age Normalized	191.15	1.70×10^{-41}	Significant
Sex Ratio Normalized	103.02	1.31×10^{-23}	Significant
Total Population Normalized	44.92	2.69×10^{-11}	Non-significant
Percentage of Agriculture Area Normalized	0.434	0.510	Non-significant
Percentage of Illiterate	9.90	1.67×10^{-3}	Significant
Population Density Normalized	0.189	0.664	Non-significant
Distance : 10 - 100 km			
Median Age Normalized	13.43	2.55×10^{-4}	Significant
Sex Ratio Normalized	0.479	0.489	Non-significant
Total Population Normalized	6.74	x9.52×10^{-3}	Significant
Percentage of Agriculture Area Normalized	1.43	0.232	Non-significant
Percentage of Illiterate	60.82	1.02×10^{-14}	Significant
Population Density Normalized	2.24×10^{-3}	0.962	Non-significant
Ratio Village and City Normalized	0.514	0.473	Non-significant
Distance : 0 - 10 km			
Median Age Normalized	64.35	1.81×10^{-15}	Significant
Sex Ratio Normalized	1.93	0.164	Non-significant
Total Population Normalized	11.11	8.76×10^{-4}	Significant
Percentage of Agriculture Area Normalized	2.46	0.117	Non-significant
Percentage of Illiterate	9.42×10^{-2}	0.759	Non-significant
Population Density Normalized	0.311	0.577	Non-significant
Ratio Village and City Normalized	8.30	4.00×10^{-3}	Significant

Table 13: Anova test of not summer and summer by indicators and per distance categories.

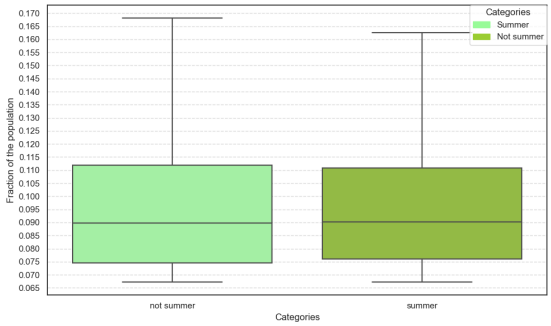
As observed in Figure 29, outliers are not significantly impacted by the summer period compared to regular days. On regular days, the distribution shows a broader spread. Linear regression also indicates a strong relationship between the two periods (Table 14).

However, the correlation between regular and summer days provides some interesting insights. Outliers in the "10 to 100 km" category show a strong positive correlation with Pearson's coefficient (0.64), but a lower positive correlation with Kendall and Spearman. In contrast, the "100 km and over" category exhibits a high positive correlation, except with Kendall. For outliers in the "100 km and over" category, there is no correlation with the socio-demographic indicators. The "10 to 100 km" category shows a weak correlation with the percentage of the population involved in agriculture.

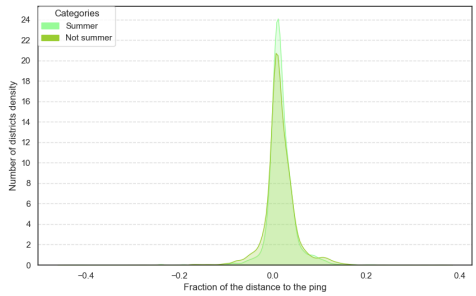
The ANOVA results needs to be taken with great care. They show that the median age indicator is significant in both outliers.



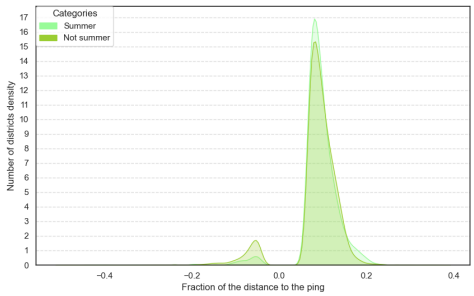
(a) Boxplot of daily movement from 10 to 100km



(b) Boxplot of daily displacements from 100 km and over



(c) Kernel density displacement of displacements from 10 to 100km



(d) Kernel density of displacements over 100 km

Figure 29: Overall of outlier throughout Türkiye districts and per day.

	From 10 to 100 km	*From 100km and over*
Coefficient	0.9416	0.8216
Standard Error	0.021	0.034
t	45.789	23.847
p-value	0.00	0.00
Minimum intervalle of confidence (0.025)	0.901	0.754
Maximal intervalle of confidence (0.975)	0.982	0.889

Table 14: Linear regression of summer and not summer category outliers.

Indicators	**F**	**p-value**	**Significant**
Distance : 10 to 100 km outliers			
Median Age normalized	5.55	1.87×10^{-2}	Significant
Sex Ratio normalized	0.813	0.367583	Non-significant
Total population normalized	2.07	0.150566	Non-significant
Agriculture Percentage normalized	0.540	1.5374×10^{-2}	Significant
Percentage of Illiterate	1.58	0.208678	Non-significant
Population density normalized	0.600	0.438599	Non-significant
Ratio village city normalized	6.77	9.464×10^{-3}	Significant
Distance : 100 km and over outliers			
Median Age normalized	35.96	3.95×10^{-9}	Significant
Sex Ratio normalized	4.85	2.80×10^{-2}	Non-significant
Total population normalized	10.92	1.02×10^{-3}	Significant
Agriculture Percentage normalized	0.78	0.377	Non-significant
Percentage of Illiterate	14.22	1.83×10^{-4}	Significant
Population density normalized	11.97	5.88×10^{-4}	Significant
Ratio village city normalized	0.91	0.340	Non-significant

Table 15: Results of ANOVA for the summer and not summer outliers categories.

Summary

The comparison between summer and regular days highlights changes in habitual mobility from a seasonal perspective.

This test reveals a maximum change of 2% in the Facebook mobility bind. The changes are most prominent in the no movement and 10 to 100 km intervals. The no movement category shows an increase in the number of people staying at home in Turkey, while the 10 to 100 km category experiences a decrease in its share of the population.

Mobility changes associated with the summer versus regular days are strongly linked to literacy rates, particularly for the no movement and 10 to 100 km intervals. Conversely, the 0 to 10 km and 100 km and over categories show a stronger relationship with median age.

However, outliers show limited changes, suggesting that the population continues with their usual routines, but districts with higher illiteracy rates appear to have a higher proportion of people staying at home.

The 0 to 10 km category is influenced by weekends, which could be tied to work-related mobility. The mean in this category does not change significantly, indicating that part of the population remains employed.

VI.2.e National holidays and regular days test *Movement Distribution Maps*

National holidays appear to significantly impact the mobility of the population, as seen in Figure 7. This test examines how national holidays affect population daily movement. The first step involves analysing the basic statistics of the same weekdays before and during the national holidays, as shown in Figure 30 and Table 16.

The mean changes for the 0 to 10 km and 100 km and over categories, with differences of 0.01 and 0.05, respectively. National holidays also impact the maximum and minimum values, with the 10 to 100 km category experiencing the most significant change in these extremes.

The spread of the data also increases, particularly for the no movement category (from 9.08 to 30.05), the 0 to 10 km category (from 7.96 to 9.79), and the 100 km and over category (from 105.34 to 147.35). The 10 to 100 km category shows the most noticeable change in symmetry, with more values exceeding the mean compared to those falling below it.

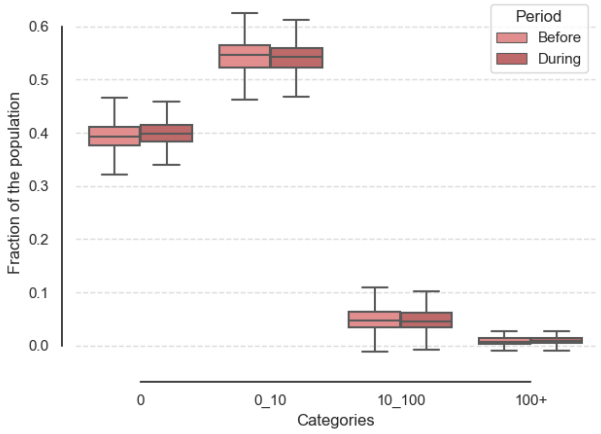
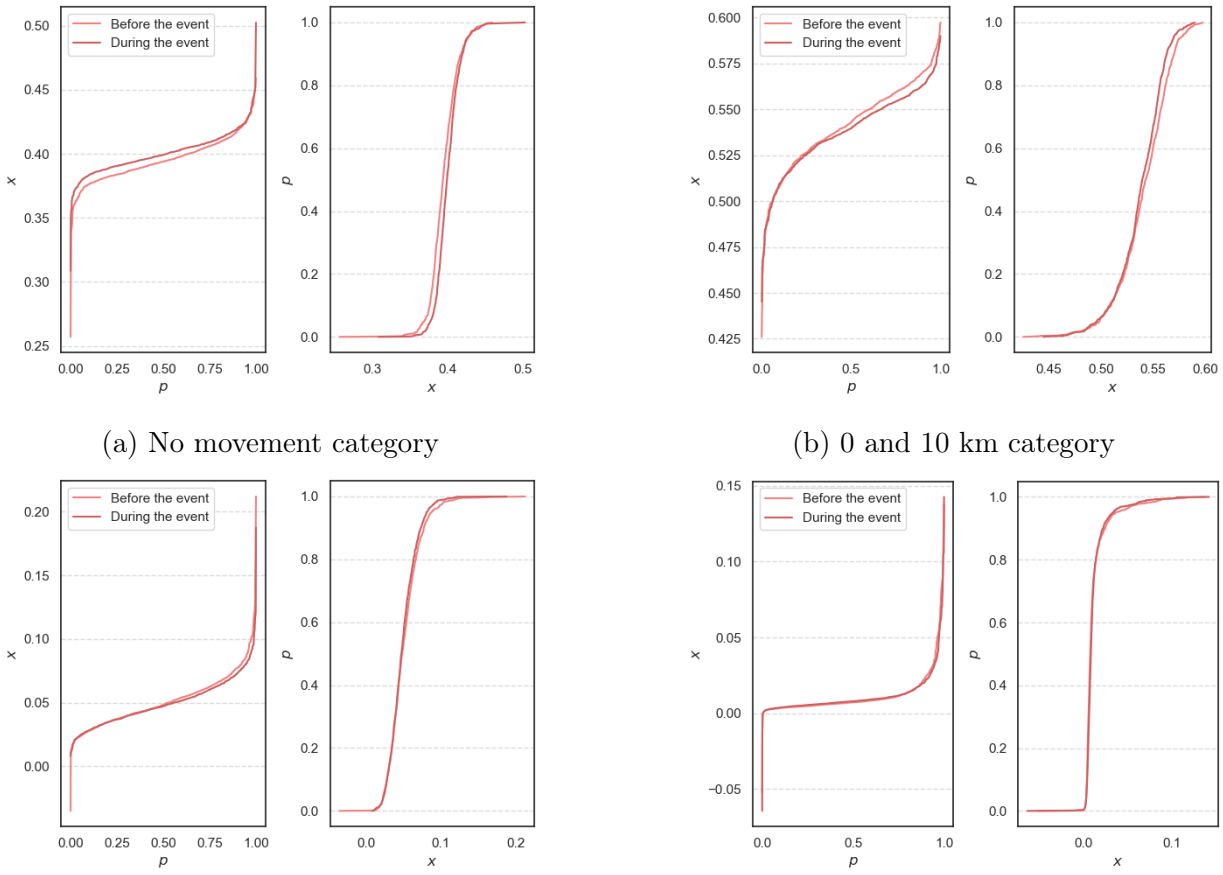


Figure 30: Boxplots of MDM distribution comparing the national holidays and regular days.

Phase	Minimum	Maximum	Mean	Variance	Skewness	Kurtosis
Distance: 0						
Other periods	8.53×10^{-2}	1.00	0.39	4.75×10^{-4}	0.74	9.08
Summer	8.59×10^{-2}	0.99	0.39	3.04×10^{-4}	1.44	30.05
Distance: 0-10						
Other periods	0.15	1.05	0.57		-1.46	7.95
Summer	0.12	0.83	0.56	4.97×10^{-4}	-1.48	9.79
Distance: 10-100						
Other periods	-0.19	1.06	3.60×10^{-2}	3.18×10^{-4}	2.18	14.44
Summer	-0.11	0.36	3.60×10^{-2}	4.97×10^{-4}	-1.48	10.91
Distance: 100						
Other periods	-0.40×10^{-2}	0.24	8.76×10^{-2}	7.20×10^{-4}	4.95	105.34
Summer	-0.34×10^{-3}	0.46	3.60×10^{-2}	7.29×10^{-4}	4.91	147.35

Table 16: Statistical Description of population daily movement per national holidays and regular days category.

To understand the impact of national holidays in values proportion, the Figure 31 interpret the highest values are the most impacted by the differences for the bind 0 to 10 km. The lowest of no movement bind is the one the most affected. The two last binds have little differences. The national holidays increase the fraction of people not moving and decrease the number of people doing 0 to 10 km.



(a) No movement category

(b) 0 and 10 km category

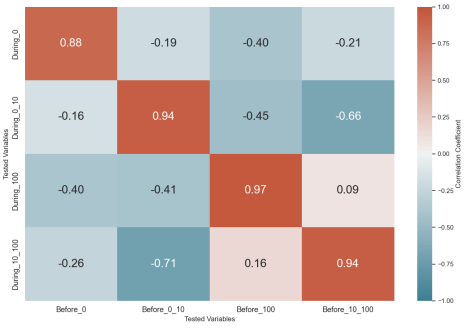
(c) 10 to 100 km category

(d) from 100 km and over category

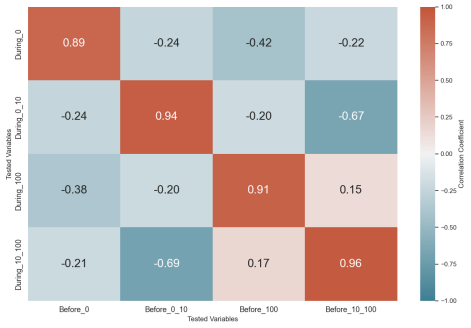
Figure 31: FP compared with the cumulative probability : comparison of normal days and national holidays.

The correlation between the days before and during a national holiday shows a high positive correlation for each of the movement categories (Figure 32). Notably, the 0 to 10 km and 10 to 100 km categories exhibit a strong negative correlation. This suggests that during national holidays, people who are not working are more likely to travel distances between 10 and 100 km.

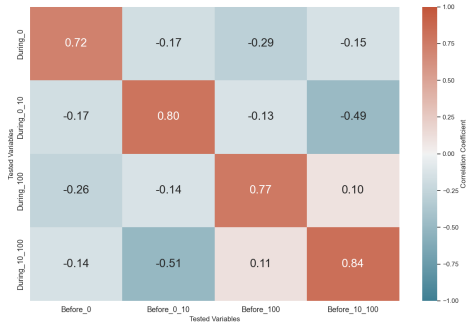
The linear regressions comparing national holidays and regular days show strong coefficients across all categories, with the no movement bind having the lowest coefficient.



(a) Correlation with Pearson



(b) Correlation with Spearman



(c) Correlation with Kendall

Figure 32: Correlation tests of the national holidays and regular days.

	No movement	*From 0 to 10 km*	*From 10 to 100 km*	*From 100km and over*
Coefficient	0.982	1.0437	1.1010	1.09299
Standard Error	0.019	0.012	0.013	0.009
t	52.940	86.881	85.376	117.925
p-value	0.000	0.00	0.00	0.00
Minimum intervalle of confidence (0.025)	0.946	1.076	1.076	1.075
Maximal intervalle of confidence (0.975)	1.0	1.126	1.126	1.111

Table 17: Linear regression of National Holiday by normal day.

Figure 33 shows the correlation between the (absolute) difference in the FP and various socio-demographic indicators. The analysis primarily reveals low positive or negative correlations with total population, median age, population density, and the urban-rural ratio. The absolute difference shows less significant correlations overall. Among the correlation methods, Spearman's correlation provides the highest values.

To evaluate the potential influence of socio-demographic indicators on changes in daily travel distances, Table 18 presents the results of an ANOVA analysis. The findings highlight a strong impact of literacy rates and the sex ratio on the proportion of the population that does not travel. Additionally, median age and the urban-to-rural ratio are significantly associated with variations in this non-moving fraction.

(a) Correlation with Pearson

(b) Correlation with Spearman

(c) Correlation with Kendall

Figure 33: Correlation tests of absolute difference (abs_dif) and difference of national holiday and normal day by socio-demographic indicators.

For individuals travelling daily distances between 0 and 10 km, the sex ratio and median age exhibit the most significant associations, while literacy and the urban-to-rural ratio also play a role, though with lower F-statistics. For travel distances ranging from 10 to 100 km, the key factors include median age, total population size, and literacy, with literacy showing the highest F-statistic among these indicators.

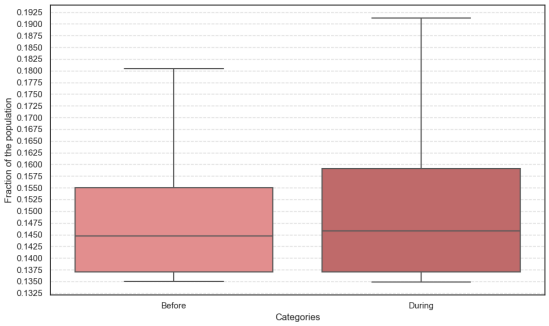
Finally, for distances of 100 km or more, the primary influencing factors are median age, total population size, and the urban-to-rural ratio, with median age yielding the highest F-statistic. Notably, the highest F-statistics for the last two distance categories (10–100 km and 100 km or more) are lower than those observed for the first two categories.

Indicators	**F-statistic**	**p-value**	**Significative**
Distance : No movement			
Median Age Normalized	40.88	1.93 x10^{-10}	Significant
Sex Ratio Normalized	184.63	1.45 x10^{-40}	Significant
Total Population Normalized	0.60	0.44	Non-significant
Percentage of Agriculture Area Normalized	1.59	0.21	Non-significant
Percentage of Illiterate	232.43	3.70x−50	Significant
Population Density Normalized	1.12 x10^{-4}	0.99	Non-significant
Ratio Village and City Normalized	24.05	1.00 x10^{-6}	Significant
Distance : 0 - 10 km			
Median Age Normalized	362.42	2.36 x10^{-75}	Significant
Sex Ratio Normalized	118.32	6.24 x10^{-27}	Significant
Total Population Normalized	36.05	2.21 x10^{-9}	Non-significant
Percentage of Agriculture Area Normalized	3.92 x10^{-2}	0.84	Non-significant
Percentage of Illiterate	6.53	1.07 x10^{-2}	Significant
Population Density Normalized	0.170946	6.793097 x10^{-1}	Non-significant
Ratio Village and City Normalized	39.62	3.65 x10^{-10}	Significant
Distance : 10 - 100 km			
Median Age Normalized	73.86	1.48x10^{-17}	Significant
Sex Ratio Normalized	1.34x10^{-2}	0.907	Non-significant
Total Population Normalized	26.20	3.32x10^{-7}	Significant
Percentage of Agriculture Area Normalized	1.91	0.167	Non-significant
Percentage of Illiterate	96.24	2.66x10^{-22}	Significant
Population Density Normalized	1.56x10^{-2}	0.900	Non-significant
Ratio Village and City Normalized	2.40	0.12	Non-significant
Distance : 100 km and over			
Median Age Normalized	105.52	2.96x10^{-24}	Significant
Sex Ratio Normalized	3.38	6.62x10^{-2}	Non-significant
Total Population Normalized	12.83	3.47x10^{-4}	Significant
Percentage of Agriculture Area Normalized	8.62	3.36x10^{-3}	Significant
Percentage of Illiterate	0.436	0.509	Non-significant
Population Density Normalized	0.684	0.40	Non-significant
Ratio Village and City Normalized	7.76	5.37x10^{-3}	Significant

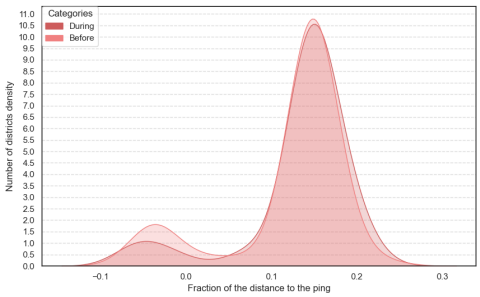
Table 18: Anova test of national holiday and regular days compared to socio-demographic indicator.

The outliers for the 10 to 100 km and 100 km and over categories show interesting results. They indicate that regular days have more limited values compared to summer days, leading to a wider range of values during national holidays. The 100 km and over category is asymmetric, with values lower than the mean (Figure 34).

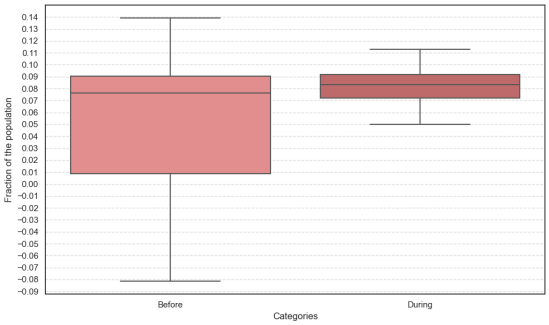
In terms of correlation, all methods reveal strong positive correlations: 0.76 (Spearman), 0.78 (Pearson), and 0.64 (Kendall) for the 100 km and over category. The 10 to 100 km category shows moderate positive correlations: 0.45 (Kendall), 0.52 (Pearson), and a high positive correlation of 0.81 with Spearman. The linear regression results, as shown in Table 19, present relatively low coefficients.



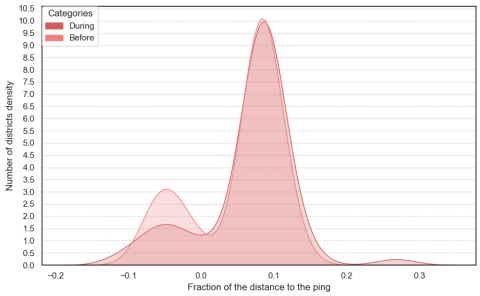
(a) Boxplot of outliers from 10 to 100 km bind



(b) Kendall density of outlier from 10 to 100 km bind



(c) Boxplot of outliers from 100 km and over bind



(d) Kendall Density of outliers from 100 km and over bind

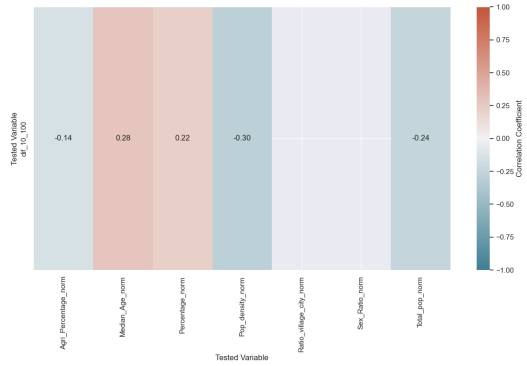
Figure 34: Overall plot of outliers.

	From 10 to 100 km	*From 100km and over*
Coefficient	0.7738	0.8682
Standard Error	0.045	0.047
t	17.343	18.304
p-value	0.00	0.00
Minimum intervalle of confidence (0.025)	0.686	0.775
Maximal intervalle of confidence (0.975)	0.862	0.962

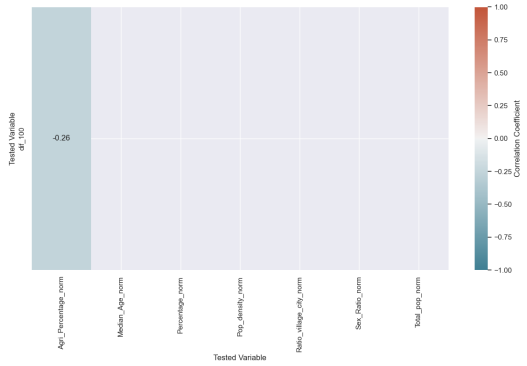
Table 19: Linear regression of national holidays by normal day category outliers.

For the 10 to 100 km bin and the spearman, Agriculture percentage, population density and the total of the population indicators have low negative correlations with the national holidays difference. There are low positive correlation between the percentage of illiterate and the median age (Figure 35). The two other types of correlations do not show any significance.

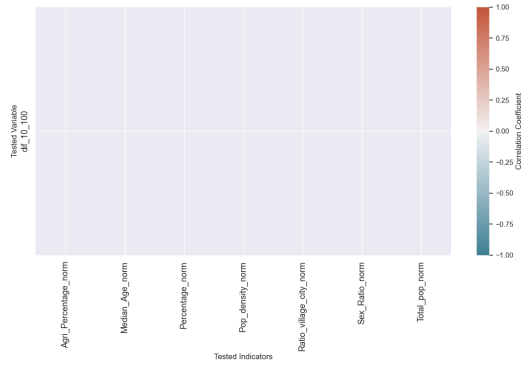
For the 100 km and over outlier, indicators are significant depending on the type of correlation. Spearman shows a low negative correlation with the agriculture ratio. Pearson shows a middle negative correlation with median age and a low positive correlation. Kendal shows correlation with the median age (Figure 35) .



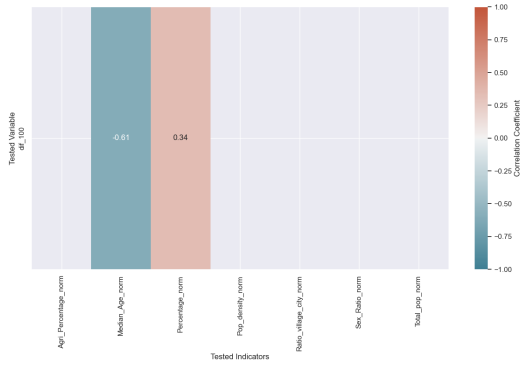
(a) No movements



(b) Displacements from 0 to 10 km

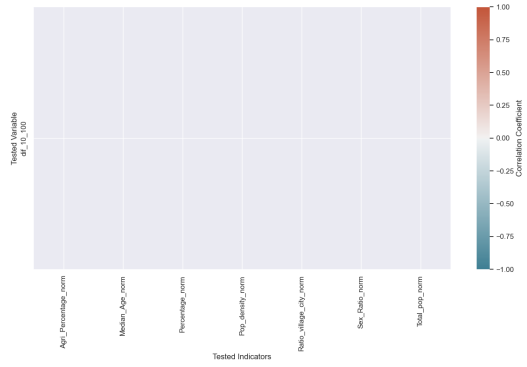


(c) Displacements from 10 to 100 km

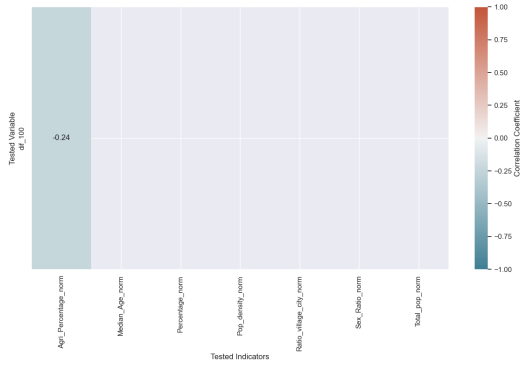


(d) Displacements over 100 km

vspace1em



(e) Displacements from 10 to 100 km



(f) Displacements over 100 km

Figure 35: Correlation of the difference between national holiday and regular days outliers by socio-demographic indicators.

To assess the potential influence of socio-demographic indicators on changes in daily travel distances, Table 20 presents an ANOVA analysis. The results highlight significant impacts of the percentage of illiteracy and sex ratio on the proportion of the population that remains stationary.

Additionally, median age and urban-to-rural ratio exhibit significant associations with changes in this group. For populations travelling short daily distances (0–10 km), the sex ratio and percentage of illiteracy stand out with strong statistical significance, while the urban-to-rural ratio and median age play smaller roles.

In the medium-distance category (10–100 km), the most influential factor is the percentage of illiteracy (F = 10.04), followed by population density and total population. For distances exceeding 100 km, the median age is the most significant indicator, along with sex ratio. However, it is notable that the F-statistics for long-distance categories (10–100 km and 100 km and over) are generally lower than those for stationary or short-distance groups.

Indicators	**F**	**p-value**	**Significant**
Distance : 10 to 100 km			
Median Age normalized	13.34	$2.92\text{x}10^{-4}$	Significant
Sex Ratio normalized	5.41	$2.05\text{x}10^{-2}$	Significant
Total population normalized	10.92	$1.027\text{x}10^{-3}$	Significant
Agriculture Percentage normalized	0.3754	0.540374	Non-significant
Percentage of Illiterate	10.04	1.64210^{-3}	Significant
Population density normalized	7.125	$7.882\text{x}10^{-3}$	Significant
Ratio village city normalized	$1.748\text{x } 10^{-2}$	0.894864	Non-significant
Distance : 100 km and over			
Median Age normalized	0.855	0.355487	Non-significant
Sex Ratio normalized	18.57	2.0e-5	Significant
Total population normalized	0.143	0.705479	Non-significant
Agriculture Percentage normalized	5.14	2.382e-2	Significant
Percentage of Illiterate	11.34	8.23e-4	Significant
Population density normalized	1.84	0.176089	Non-significant
Ratio village city normalized	1.49	0.223372	Non-significant

(a) ANOVA over 100 km outliers (Dataset 2)

Table 20: Anova of national holidays and regular days outliers by socio-demographic indicators.

Summary

The national holidays test refers to a nationwide day off that impacts the entire country, similar to weekends.

Overall, there are noticeable changes in the means for the 0 to 10 km and 100 km and over categories. The 100 km and over category is also affected by both lower and higher maximum FP values. However, the FP for individuals staying at home is more spread out across the distribution. In terms of changes, Figure 31 shows only minimal changes in the movement categories.

The ANOVA provides insights into the effects of national holidays. In general, the sex ratio, literacy, and age indicators are significant factors in these changes.

Regarding outliers, the coefficients are lower compared to normal values, indicating that the impact of national holidays on mobility is less pronounced than the effects observed for other values.

VI.2.f Disasters test on *Movement Distribution Maps*

This section focuses on analyzing the sensitivity of the dataset to natural disasters, using official numbers of internally displaced populations. The two most significant disasters in terms of displaced populations are a wildfire and a storm.

Wildfire test

The first disaster, a wildfire, affects only one district, making it statistically insufficient for broad conclusions. Therefore, the wildfire observation is limited to interpreting the timeline presented in Table 21. From this table, it is notable that only a small percentage of the Facebook user population was impacted. The maximum observed difference is 0.4% of the population that did not move, compared to the pre-event mean for non-movement. This difference appears insignificant in terms of understanding the wildfire's impact.

A similar observation can be made for the 0 to 10 km category, where no significant changes are observed. The only category showing some effect is the 100 km and over displacement. Two values are notably lower than the minimum of the pre-event values. In the 100 km and over category, the minimum pre-event value is 0.71%, while the first day of the event shows a value of 0.10%, and the first day after the disaster records 0.11% of the population engaging in such movement.

Time Line	**Category**	***No Movement***	***Displacement between 0 to 10 km***	***Displacement between 10 to 100 km***	***Displacement for 100 km and over***
Same day Over 3 weeks	Before Mean	0.370	0.582	0.0394	7.87×10^{-3}
2023-08-22	During	0.366	0.586	0.0379	1.03×10^{-3}
2023-08-23	During	0.367	0.588	0.0378	7.24×10^{-3}
2023-08-24	During	0.373	0.579	0.0405	7.25×10^{-3}
2023-08-25	After	0.366	0.583	0.0396	1.11×10^{-3}

Table 21: Time line of the Wildfire disaster.

Storm test

The second disaster, a storm, affected six provinces. Internal displacement data (Internal Displacement Monitoring Center, nd) provides information at two spatial scales: province and district. However, district-level data is incomplete. The storm's impact spans both weekends and weekdays, and observations are divided into three parts: overall province-level visualization, weekend and weekday differences, and district-level analysis (limited to available district data).

The time series (Figure 36) indicates significant changes during and after the storm, with noticeable peaks (Figure 37, Table 22). The no movement fraction's mean is largely unaffected by the storm, showing only a slight increase of 0.1%. The difference in the percentage of the population not moving before and after the disaster is 0.09%. Additionally, the spread in this fraction decreases after the storm.

For the 0 to 10 km fraction, the mean difference between before and during the storm is 0.5%, and between before and after is 0.02%. After the storm, this category exhibits the lowest maximum and highest minimum values.

The 10 to 100 km fraction reveals an increase of 0.07% after the storm and a decrease of 0.3% during the storm. Notably, the maximum population percentage for this fraction occurs before the storm, while the highest minimum is observed after (from 0.5% before to 0.01% after). In this fraction, as well as in the 100 km and over fraction, the lowest maximums and minimums occur during the storm.

For the 100 km and over fraction, the mean percentage before the storm is higher (0.19%) compared to during (0.17%) and after (0.18%). The maximum value before the storm (24.5%) is significantly higher than during (14.1%) and after (14.3%).

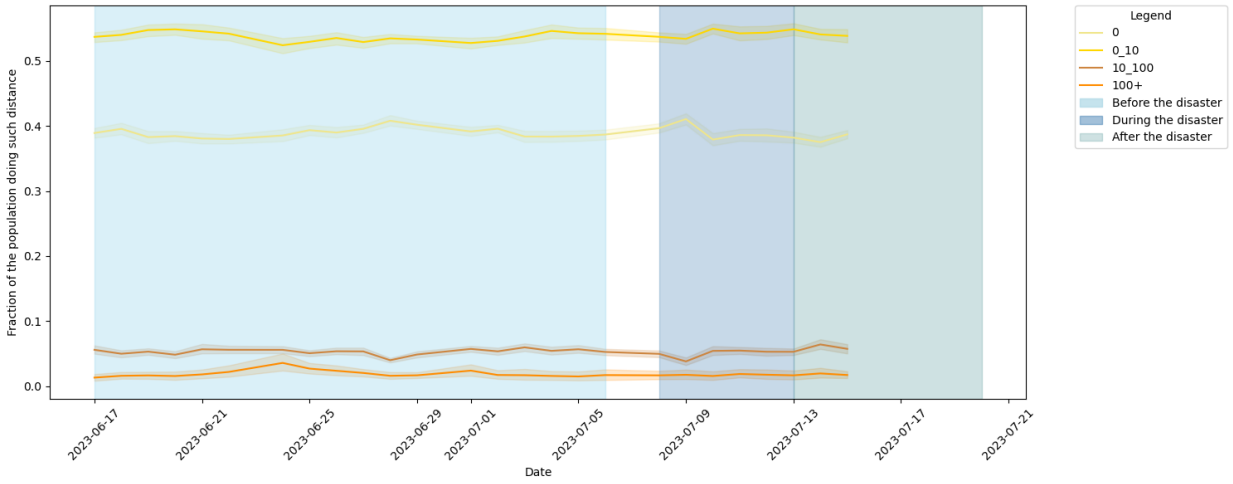


Figure 36: Time series of the storm.

Phase	Minimum	Maximum	Mean	Variance	Skewness	Kurtosis
Distance: 0						
Before	0.2771	0.5220	0.3897	0.0009	0.2597	1.7604
During	0.2969	0.5261	0.3901	0.0012	0.5216	1.0722
After	0.3138	0.4621	0.3812	0.0007	0.3760	0.9196
Distance: 0-10						
Before	0.3547	0.6200	0.5375	0.0012	-0.9020	1.7630
During	0.3786	0.6251	0.5426	0.0011	-0.9733	2.0794
After	0.4015	0.5963	0.5398	0.0011	-0.8910	1.6470
Distance: 10-100						
Before	0.0052	0.1864	0.0531	0.0004	0.8653	2.3249
During	-0.0008	0.1510	0.0504	0.0005	0.7716	1.6014
After	0.0155	0.1807	0.0608	0.0007	1.2234	2.9349
Distance: >100						
Before	-0.0398	0.2453	0.0192	0.0008	3.3784	14.6467
During	-0.0444	0.1410	0.0171	0.0007	2.7397	8.0153
After	-0.0115	0.1428	0.0184	0.0006	2.9205	9.1880

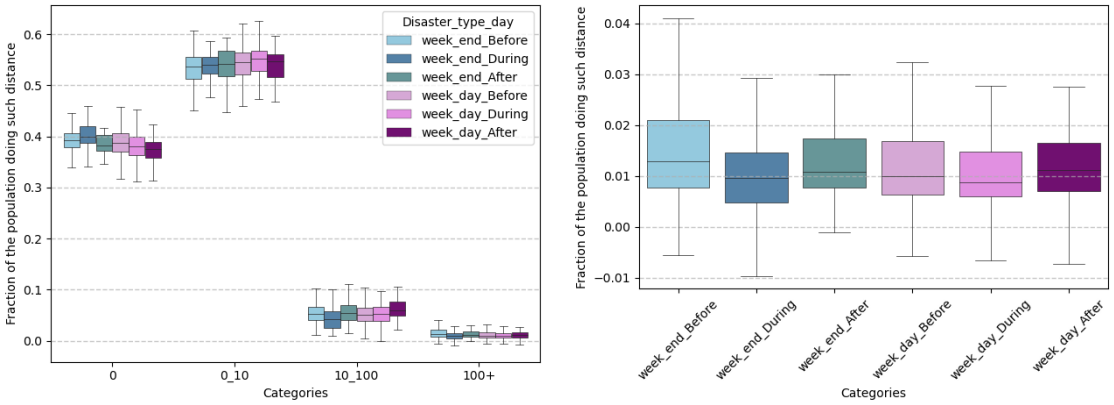
Table 22: Statistical Summary of the storm disaster event.

Figure 37 and Table 23 present the diversity for the storm impact compared with the type of day (weekend or workday). The storm impacted firstly a weekend and then workdays. The distance binds FP of movement proposes that the workdays have been more impacted than the weekends which a decrease through the period selected (difference between before and during the period of 0.8%; the difference between during and after of 0.8%) instead of an increase and a decrease for the weekends (1.5% between before and during and 1.6% between before and after). Mostly in the after phase, the workdays show a high impact on a shorter distribution queue. Such differences can come from the timeline of the event and the habits of the population.

The bind of 0 to 10 km and weekend type does not show real impact 0.5% of the mean of the population changes between during and before and 0.3% between During and after. However, the kurtosis presents a lower spread in the distribution during the event. Meanwhile, the workday shows an increase of the mean by 1.5% of Facebook users between before and during and a slight decrease of 0.6% between during and after. The kurtosis suggests that the first day (weekend) of the storm reduced the range of the fraction of the population doing such movement. On the contrary, workdays show a larger spread during the event and an even smaller distribution than a Gaussian.

The 10 to 100 bind presents a reduction of 0.9% of the weekends (between before and during) and an increase of 0.5% after the event. For the workdays, only after the event is relevant to be highlighted with a rise of 1%. The kurtosis for the weekend presents the same as the two binds studied before. On workdays, the spread of the distribution increased significantly, with kurtosis values rising from 0.38 before the storm to 2.4 during the event and reaching 4.21 after the storm, indicating a broader spread compared to a Gaussian distribution

The 100 km and over bind present a reduction during the storm and a slight increase after the event in the mean for the workday. Weekends have a change of 0.08% before and during the event. The FP range is shorter during the event and after from both types of days with the next change for the workday.



(a) Boxplot of all distances category and per type of day

(b) Zoom of boxplot of 100 km and over category

Figure 37: Overview of the impact of the storm disaster per type of the week.

Phase	Minimum	Maximum	Mean	Variance	Skewness	Kurtosis
Distance: 0 and weekend						
Before	0.2969	0.5192	0.3886	0.0009	0.3245	1.4006
During	0.3404	0.5008	0.4037	0.0009	0.4724	0.8673
After	0.3469	0.4621	0.3872	0.0006	0.9604	1.4452
Distance: 0 and workday						
Before	0.2771	0.5220	0.3919	0.0009	0.1343	2.6325
During	0.2969	0.5261	0.3833	0.0012	0.7576	1.6972
After	0.3138	0.4497	0.3752	0.0008	0.1968	0.2936
Distance: 0-10 and weekend						
Before	0.3547	0.6200	0.5404	0.0012	-0.9016	1.9323
During	0.4609	0.5871	0.5356	0.0007	-0.7652	0.3082
After	0.4015	0.5928	0.5387	0.0014	-1.1511	2.1494
Distance: 0-10 and workday						
Before	0.3935	0.6076	0.5316	0.0011	-0.9825	1.5467
During	0.3786	0.6251	0.5461	0.0013	-1.1565	2.5382
After	0.4673	0.5963	0.5409	0.0009	-0.2862	-0.5443
Distance: 10-100 and weekend						
Before	0.0052	0.1864	0.0528	0.0004	1.0299	3.1131
During	0.0101	0.1093	0.0439	0.0004	0.6266	-0.0135
After	0.0155	0.1279	0.0574	0.0007	0.8401	0.5622
Distance: 10-100 and workday						
Before	0.0105	0.1260	0.0538	0.0004	0.4842	0.3824
During	-0.0008	0.1510	0.0537	0.0005	0.9228	2.4088
After	0.0219	0.1807	0.0641	0.0008	1.4944	4.2176
Distance: 100 and weekend						
Before	-0.0398	0.1660	0.0178	0.0006	2.9294	9.6098
During	-0.0191	0.1358	0.0170	0.0007	2.8229	7.8435
After	-0.0115	0.1058	0.0171	0.0004	2.8750	8.7328
Distance: 100 and workday						
Before	-0.0200	0.2453	0.0221	0.0011	3.6083	16.0691
During	-0.0191	0.1358	0.0170	0.0007	2.8229	7.8435
After	-0.0072	0.1428	0.0196	0.0008	2.7419	7.6550

Table 23: Statistical summary by disaster, type of day, and distance.

In terms of provinces, separating the data into six provinces results in very small sample sizes. Therefore, Table 24 focuses only on the means for the provinces impacted. Based on official data for internally displaced people, the provinces expected to show the greatest changes across the three phases are:

1. Düzce
2. Kastamonu
3. Karabük
4. Bolu
5. Zonguldak
6. Bartın

This highlights the importance of analyzing the spatial scale when assessing disaster impacts. To delve deeper, the next step involves studying districts where the number of internally displaced persons is known. Table 25 presents the mean values for the four distance categories, separated by workdays and weekends.Below is a descending order list of districts with the highest ratio of internally displaced persons to total population:

1. Devrekani
2. Şenpazar
3. Gökeçebey
4. İnebolu
5. Cide
6. Yenice
7. Kurucaşile
8. Amasra
9. Merkez

All types of mobility were affected by a small percentage due to the storm. At the district scale, there is less noise in the population data compared to the province scale, suggesting that the storm's intensity spatially influenced population mobility (Table 24, Table 25 and Figure 38). In both analyses, the districts with the highest percentage of population impacted by the storm exhibited the strongest responses to the disaster.

During the event, populations generally stayed at home, while mobility increased after the storm passed. Responses varied depending on whether the storm occurred on a workday or a weekend, though this variation is partly attributable to the timing of the storm.

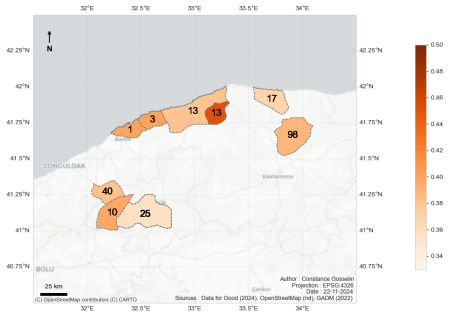
These findings provide valuable insights into the application of MDM in the context of natural disasters and population mobility. The type of disaster significantly influences how populations respond and adjust their mobility.

Phase	**Düzce**	**Bolu**	**Zinguldak**	**Kastamonu**	**Karabük**	**Bartın**
Internal displaced	474	37	40	131	35	4
Total population	409865	324789	591492	383755	255242	207238
Distance: 0 and weekend						
Before	0.386052	0.378493	0.386160	0.399116	0.398973	0.396149
During	0.393348	0.383617	0.400235	0.411576	0.408833	0.427747
After	0.383261	0.386295	0.374201	0.388928	0.397992	0.392256
Distance: 0 and workday						
Before	0.374731	0.381574	0.380068	0.394470	0.400860	0.397435
During	0.365995	0.372252	0.377193	0.389582	0.400790	0.394474
After	0.346570	0.362822	0.372931	0.389214	0.385124	0.379107
Distance: 0-10 and weekend						
Before	0.525072	0.518128	0.553211	0.531296	0.537605	0.535212
During	0.532018	0.521499	0.556873	0.534769	0.541614	0.537998
After	0.525721	0.524597	0.568635	0.546737	0.528683	0.525720
Distance: 0-10 and workday						
Before	0.541078	0.529777	0.564363	0.539970	0.537671	0.533071
During	0.542501	0.540995	0.567163	0.546676	0.540649	0.538997
After	0.536103	0.525810	0.569319	0.537703	0.550074	0.543420
Distance: 10-100 and weekend						
Before	0.057622	0.063185	0.050538	0.051494	0.043575	0.057426
During	0.043274	0.060823	0.034923	0.042768	0.040699	0.030720
After	0.069897	0.065171	0.049038	0.048455	0.057248	0.071949
Distance: 10-100 and workday						
Before	0.056578	0.056504	0.047991	0.051140	0.045526	0.062872
During	0.063113	0.056673	0.048859	0.051452	0.044085	0.060845
After	0.079605	0.075592	0.050659	0.061065	0.048468	0.066494
Distance: 100 and weekend						
Before	0.032461	0.038577	0.009843	0.017257	0.020139	0.010432
During	0.025239	0.034510	0.007468	0.012492	0.013614	0.0.003427
After	0.025487	0.023806	0.008364	0.014783	0.016692	0.011049
Distance: 100 and over and workday						
Before	0.024670	0.032291	0.007587	0.013643	0.017124	0.008085
During	0.028219	0.031603	0.007084	0.012088	0.015259	0.005641
After	0.038055	0.031703	0.007361	0.013503	0.013550	0.012915

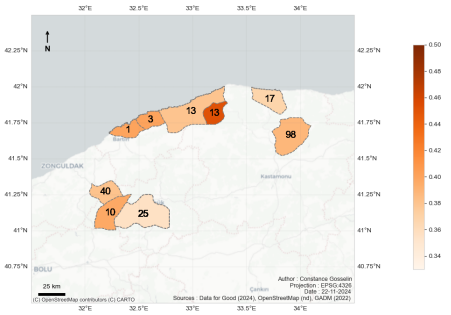
Table 24: Time line of the storm disaster per provinces impacted.

Phase	**Gökçebey**	**Inebolu**	**Şenpazar**	**Cide**	**Devrekani**	**Yenice**	**Merkez**	**Amasra**	**Kurucaşile**
Intern Displaced people	40	17	13	13	98	10	25	1	3
Population	21344	20480	4844	22965	13886	20592	134486	14552	6691
Before	0.386502	0.385779	0.424889	0.409269	0.386992	0.387395	0.386588	0.392621	0.402566
During	0.410253	0.391182	0.409457	0.418188	0.382461	0.413137	0.397957	0.413976	0.469199
After	0.362189	0.359020	0.411832	0.406174	0.402222	0.375432	0.382640	0.373219	0.409979
Distance: 0 and workday									
Before	0.386003	0.386913	0.421991	0.400780	0.390121	0.400882	0.376140	0.384252	0.427716
During	0.395627	0.374173	0.444129	0.396294	0.372556	0.389851	0.367399	0.384252	0.427725
After	0.390193	0.373627	0.423196	0.407304	0.383903	0.382342	0.363314	0.381126	0.388929
Distance: 0-10 and weekend									
Before	0.549510	0.559764	0.521414	0.556672	0.527974	0.560010	0.556174	0.524457	0.539806
During	0.544343	0.566529	0.552513	0.562437	0.533588	0.550120	0.555700	0.552574	0.515637
After	0.580558	0.591545	0.557108	0.559943	0.509679	0.567672	0.562902	0.517954	0.516249
Distance: 0-10 and workday									
Before	0.548854	0.571072	0.531791	0.564936	0.525866	0.545492	0.571065	0.536191	0.507463
During	0.543876	0.582179	0.517772	0.574903	0.541380	0.552957	0.580146	0.539897	0.516357
After	0.548819	0.564862	0.547597	0.556011	0.514435	0.547976	0.578284	0.521890	0.545070
Distance: 10-100 and weekend									
Before	0.053306	0.045568	0.042551	0.022682	0.074973	0.044925	0.041009	0.072431	0.042437
During	0.039709	0.034082	0.030332	0.015973	0.069233	0.032472	0.030864	0.032009	0.015816
After	0.051108	0.035735	0.027759	0.023991	0.075364	0.047716	0.042136	0.093942	0.061899
Distance: 10-100 and workday									
Before	0.057021	0.035402	0.037219	0.027490	0.073564	0.047639	0.039890	0.072243	0.058147
During	0.052983	0.036946	0.039837	0.023510	0.077054	0.048419	0.040281	0.068617	0.049713
After	0.058684	0.056696	0.032488	0.030784	0.094105	0.065308	0.044300	0.081790	0.049203
Distance: 100 and weekend									
Before	0.010521	0.008337	0.008365	0.011043	0.012348	0.008288	0.016197	0.010230	0.011854
During	0.004381	0.006002	0.006700	0.004257	0.009761	0.005592	0.015336	0.003962	0.000749
After	0.008067	0.013735	-0.001022	0.008989	0.009635	0.007201	0.012284	0.016533	0.011649
Distance: 100 and workday									
Before	0.008660	0.007337	0.007103	0.006703	0.009102	0.004824	0.012806	0.008288	0.010166
During	0.008282	0.007235	0.001270	0.004702	0.008759	0.006323	0.012179	0.007009	0.005346
After	0.003104	0.008024	0.001922	0.005222	0.006832	0.006808	0.013947	0.016083	0.020102

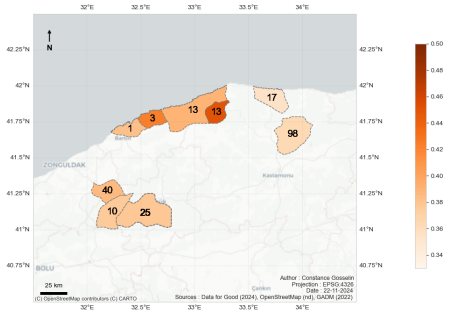
Table 25: FP mean per type of day during the storm event for each district impacted.



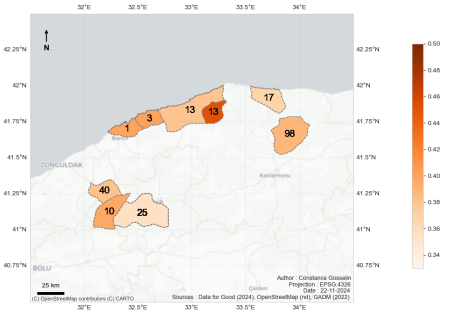
(a) Before the disaster for the workdays



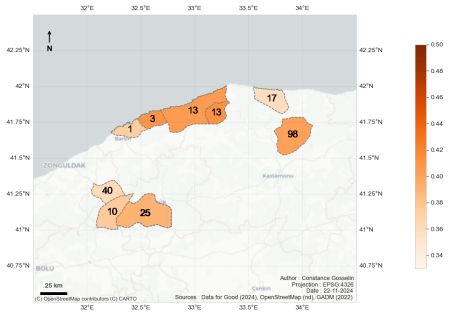
(b) Before the disaster for the weekends



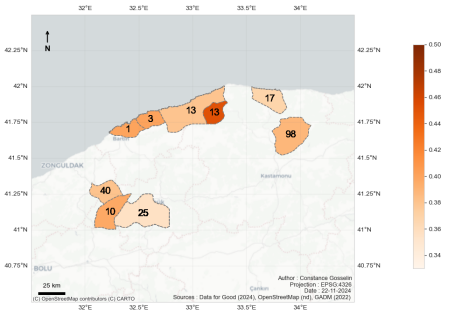
(c) During the disaster for the workdays



(d) During the disaster for the weekends

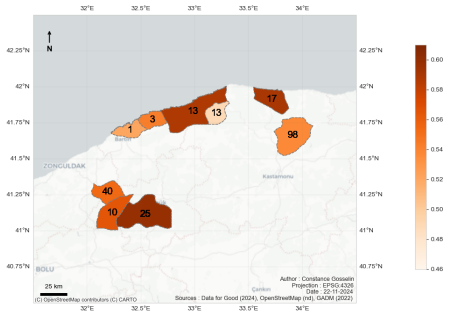


(e) After the disaster for the workdays

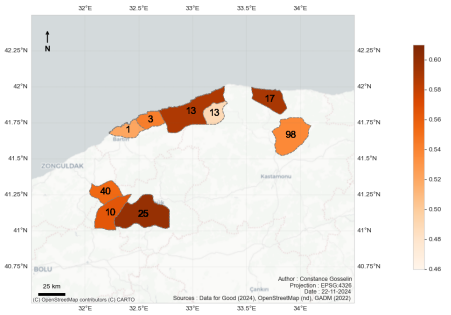


(f) After the disaster for the weekends

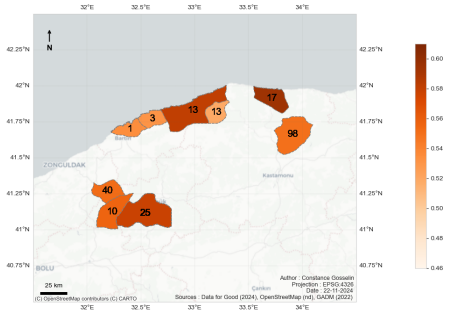
Figure 38: Selected days over the districts touched by the storm of the no movement category.



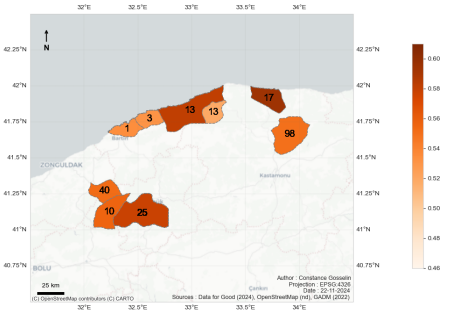
(a) Before the disaster for the working days



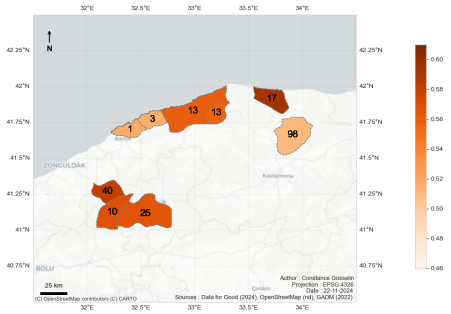
(b) Before the disaster for the week ends



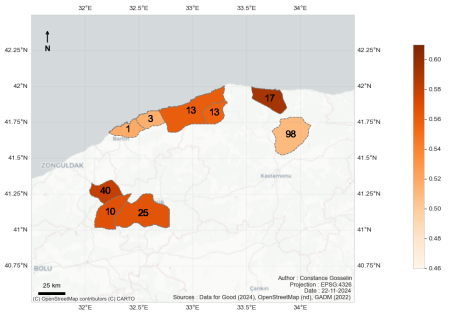
(c) During the disaster for the working days



(d) During the disaster for the week ends

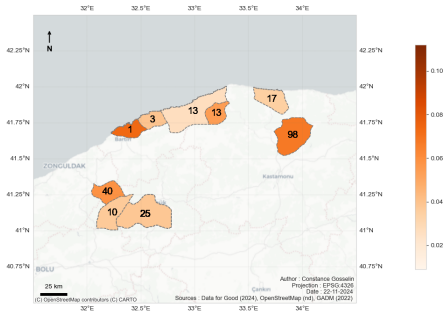


(e) After the disaster for the working days

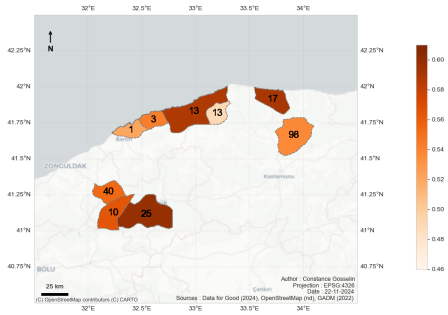


(f) After the disaster for the week ends

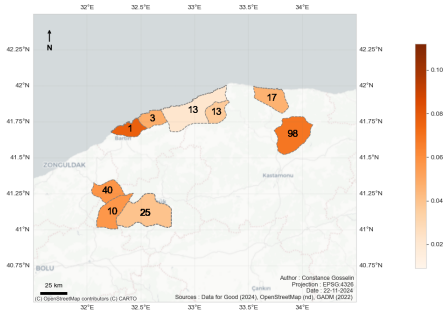
Figure 39: Selected days over the districts touched by the storm doing 0 to 10 km.



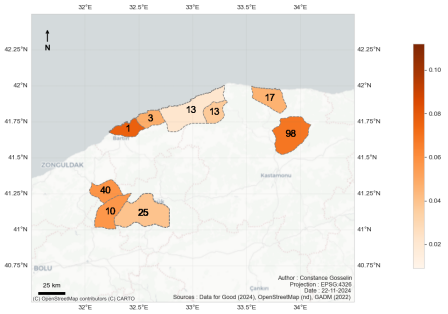
(a) Before the disaster for the working days



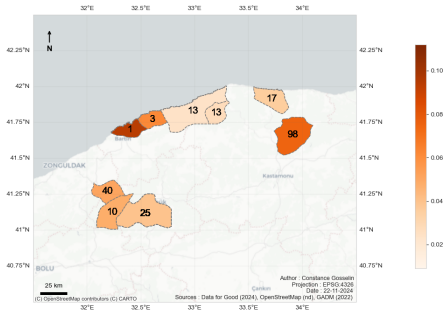
(b) Before the disaster for the week ends



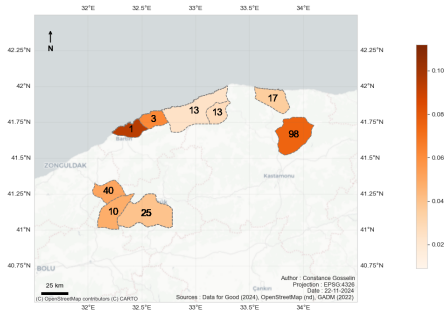
(c) During the disaster for the working days



(d) During the disaster for the week ends

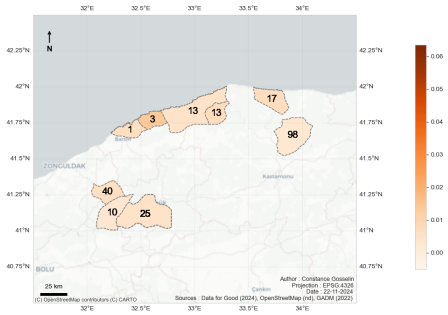


(e) After the disaster for the working days

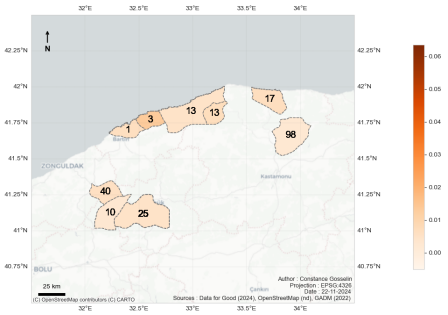


(f) After the disaster for the week ends

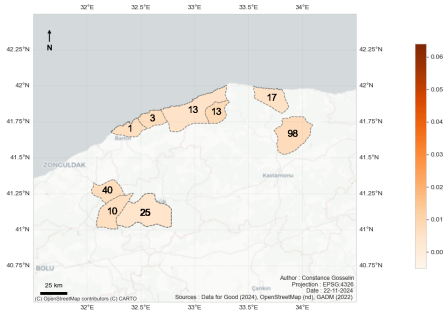
Figure 40: Selected days over the districts touched by the storm doing 10 to 100 km.



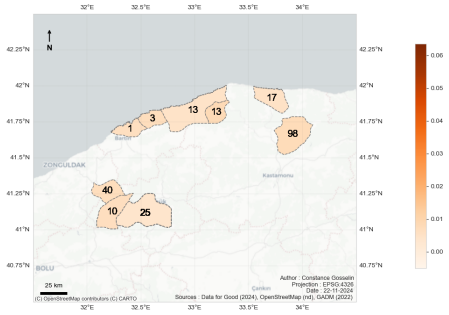
(a) Before the disaster for the working days



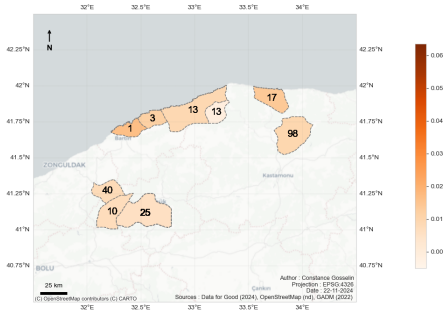
(b) Before the disaster for the week ends



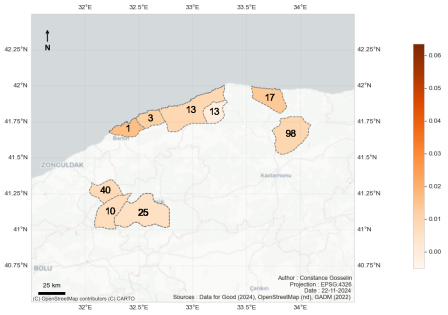
(c) During the disaster for the working days



(d) During the disaster for the week ends



(e) After the disaster for the working days



(f) After the disaster for the week ends

Figure 41: Selected days over the districts touched by the storm doing 100 km and over.

Summary

Disaster tests highlight both the advantages and limitations of using Mobile Data Metrics (MDM) to assess displacement related to natural disasters.

On the positive side, MDM offers valuable insights into changes in population mobility, capturing trends that might not be evident from official internal displacement records. This is particularly evident in the case of the storm, where MDM reveals general mobility changes beyond what is reported through official channels. Furthermore, it demonstrates that population mobility varies depending on the type of hazard, providing a nuanced understanding of disaster impacts.

However, the presence of noise, caused by populations that are either unaffected or only minimally impacted, limits the ability to clearly discern mobility changes. This reduces the visibility of significant shifts and may obscure important patterns in the data.

VII Case Study

VII.1 Methodology

This section outlines the methodology employed to examine the profound impact of recent earthquake events on Turkey. The analysis is conducted at two distinct spatial scales to ensure a comprehensive understanding: provincial and district levels.

VII.1.a Provincial Scale

At the provincial level, the primary objective is to assess the extent to which earthquakes disrupted daily mobility patterns. To this end, temporal segmentation was implemented to analyse mobility data across the following periods:

- Pre-earthquake: One month and two weeks prior to the event
- First week: The day of the earthquake and the subsequent six days
- Second week
- Third week
- First month: Beginning from the date of the earthquake
- Second month
- Third month

The spatial scope of the study encompasses three categories of provinces:

- Impacted provinces: Those directly affected by the earthquake (*Touched*)
- Neighboring provinces: Provinces sharing administrative borders with the impacted provinces (*Neighbors*)
- Unaffected provinces: Provinces neither impacted nor neighboring the affected areas (*Not touched).*

The classification of impacted provinces is informed by a literature review (Demir Aydin, 2024), while neighbouring provinces were identified using spatial analysis tools such as *QGIS*. In total, eleven provinces were classified as impacted, and ten as neighbours Table 26.

The temporal analysis visualizes daily mobility patterns for one month prior and three months post-earthquake. This timeline facilitates an understanding of disruptions to daily routines and subsequent recovery phases.

The analysis of the month over a typical week permits to understand the resorption of the provinces over the three months after the disaster. It also provides information on the impact of other events that could be seen over time.

Additionally, mobility data is examined across days of the week, segmented by province type and compared against pre-earthquake averages. To minimize noise, the analysis excludes school holiday periods (January 22 to February 3) (Feiertagskalender, nd). Notably,

Impacted provinces	**Neighbors provinces**
Adana	Bingöl
Kahramanmaraş (*K. Maras*)	Erzincan
Malatya	Kayseri
Gaziantep	Niğde
Adıyaman	Sivas
Elazığ	Tunceli
Şanlıurfa	Mardin
Kilis	Batman
Hatay	Muş
Osmaniye	Mersin
Diyarbakır	

Table 26: Impacted and neighbour provinces

the earthquake occurred on a Monday, which aids in isolating weekly patterns. Correlation analyses are performed to compare these findings against pre-earthquake baselines.

VII.1.b District Scale

The district-level analysis delves deeper into population mobility within the affected provinces, revealing potential disparities or uniformities in the response. The primary metric evaluates relative changes in mobility using the following formula:

$$\frac{\text{Week before the event date - n Week after the event date}}{\text{mean(Week before the event date - n Week before the event date)}}$$

where n represents the number of weeks after the event.

This approach quantifies deviations in daily mobility for the first three weeks post-earthquake, the period with the most significant impacts, uncontaminated by other external events. Linear regression and correlation analyses are conducted to compare these deviations with earthquake intensity and peak ground acceleration (PGA). Spatial clusters and outliers are further identified using the Local Moran's I statistic, providing geo statistical insights into mobility shifts.

The analysis also contrasts urban and rural districts to understand differential impacts. This includes a time-series analysis of the five largest cities in the impacted provinces and corresponding rural districts selected based on the lowest proportion of urban population (Table 27). The urban and rural districts analyzed are as in Table 27

City	**Urban district**	**Rural district**
Antakya	Hatay(Merkez)	Hatay(Kumlu)
Kahramanmaras	K. Maras(Merkez)	K. Maras(Caglayancerit)
Gazientep	Gaziantep(Sahinbey) and Gaziantep(Sehitkamil)	Gaziantep(Nizip)
Adyaman	Adyaman(Merkez)	Adyaman(Tut)
Malatya	Matalya(Merkez)	Matalya(Yesilyurt)

Table 27: Rural and urban Districts choosen

This comparative analysis elucidates variations in mobility patterns influenced by urbanization and PGA intensity, providing a nuanced understanding of post-earthquake population behaviour.

VII.1.c Interpretation and Contextualization

The findings are interpreted within the context of relevant grey literature and localized studies to establish connections between observed mobility trends and broader socio-economic and environmental impacts of the earthquake.

VII.2 Results

The primary objective of this master's thesis is to evaluate how the daily movement of the population was affected by the February 6th earthquake event. This section presents the results derived from the methodology, focusing on three scales of analysis: the province scale, district scale, and the impact of rural versus urban habitat on mobility.

VII.2.a Province scale analysis

The province-scale analysis is divided into three main components, as outlined in the methodology. The results indicate that all population categories were impacted by the earthquake (Figure 21).

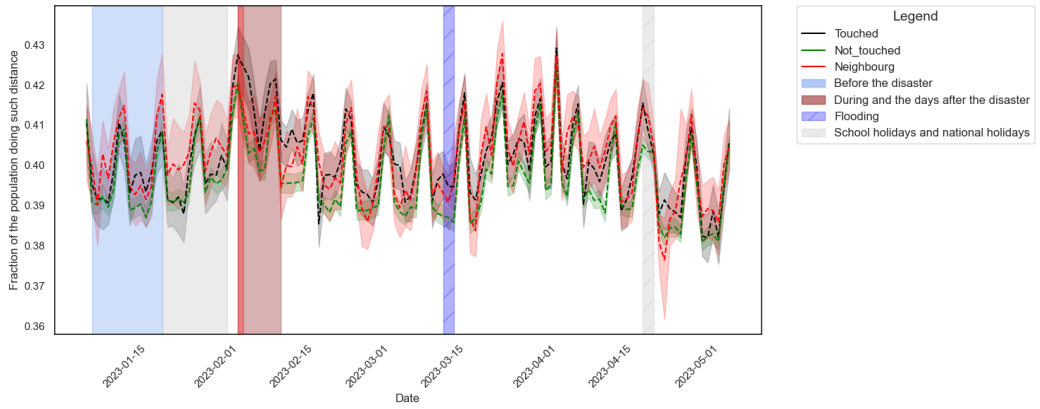
Figure 21 presents a timeline of the event, divided into three phases: the period before the earthquake, the event itself, and the three months following it. The figure clearly illustrates that the earthquake disrupted mobility across the entire country. While neighbouring provinces appeared to recover normal mobility patterns within three weeks, provinces directly affected by the EQ displayed distinct oscillations that differed significantly from the unaffected regions.

The provinces most impacted were those directly affected by the earthquake. However, neighboring provinces also experienced disruptions with less importance. Some EQ affected provinces were further impacted by flooding in March (specific date to be added).

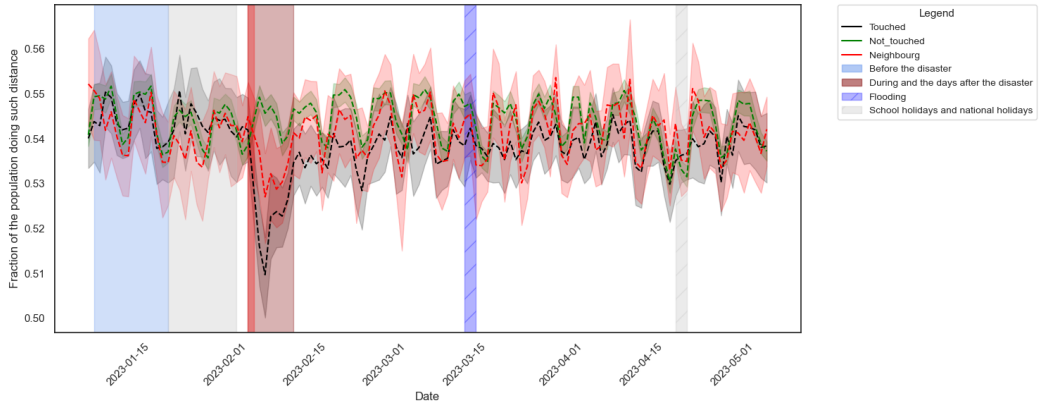
The timeline shows an increase in the fraction of the population not moving during the earthquake, accompanied by a significant peak in long-distance mobility (more than 100 km), even in neighbouring provinces. By the third week, another peak emerged in the fraction of the population moving distances of 10 to 100 km. Conversely, the non-moving and 0 to 10 km distance categories exhibited lower peaks on the same days.

The timeline presented in Figure 21, Figure 22 and Figure 23 illustrate the earthquake's impact on mobility across a regular week. These time series are valuable for understanding the disaster's effect on weekly mobility patterns. Since the earthquake occurred on a Monday, the graphs begin on Monday and end on Sunday.

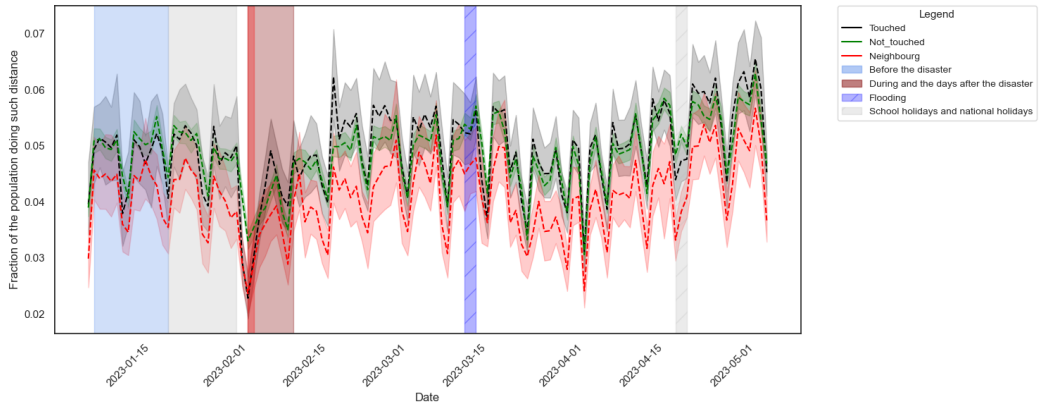
The before data for all provinces (neighbouring, not impacted, and directly affected) generally follow the typical weekly oscillations described earlier in the analysis. However, the mobility patterns in neighbouring and affected provinces show greater variability in the fraction of people across different distance categories compared to the unaffected provinces. The first week following the earthquake displays significant deviations across all distance categories and province types, highlighting the immediate impact of the disaster.



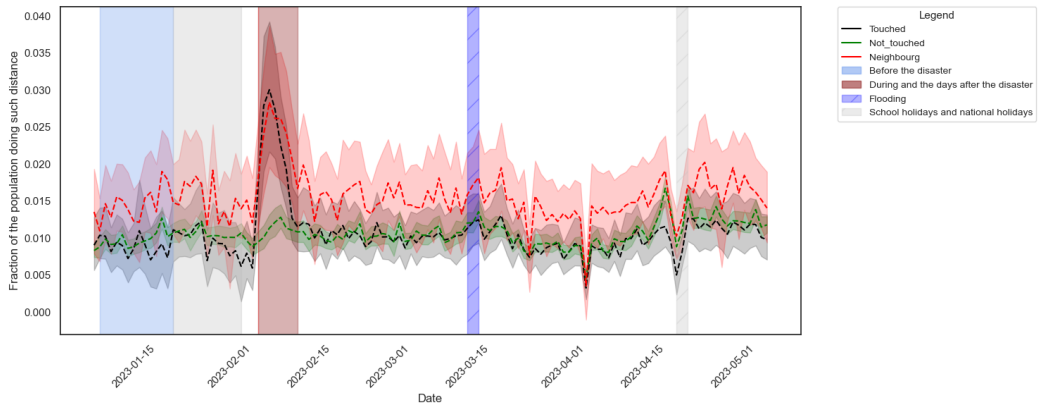
(a) Fraction of the population no moving



(b) Fraction of the population between 0 to 10 km



(c) Fraction of the population 10 km to 100km

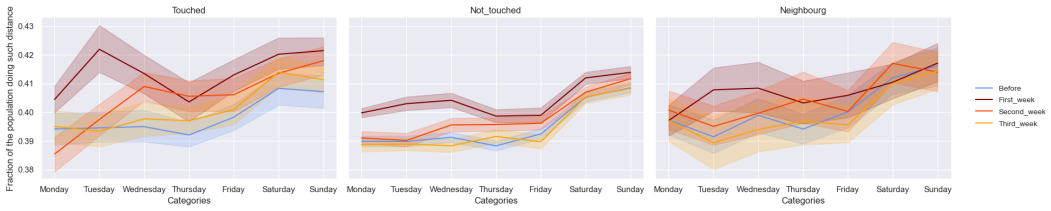


(d) Fraction of the population from 100km and over

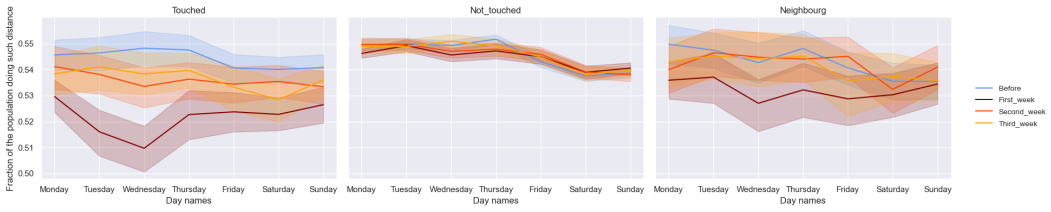
Figure 21: Time series over the three month after the disaster

To complement the timeline presented in Figure 21, Figure 22 and Figure 23 illustrate the earthquake's impact on mobility within a typical week. These time series are particularly useful for understanding how the disaster disrupted the usual weekly oscillations. Since the earthquake occurred on a Monday, the graphs start on Monday and end on Sunday.

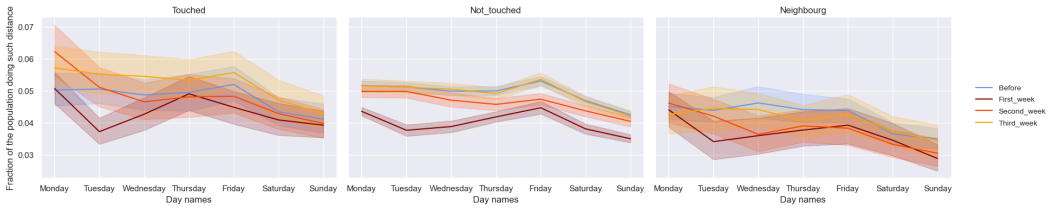
In the before phase, the mobility patterns for neighboring, not impacted, and directly affected provinces follow the expected weekly oscillations described earlier. However, both neighbouring and affected provinces display a greater variability in the fraction of people across different distance categories compared to the unaffected provinces. The first week following the earthquake is markedly different, showing significant disruptions across all distance categories and province types.



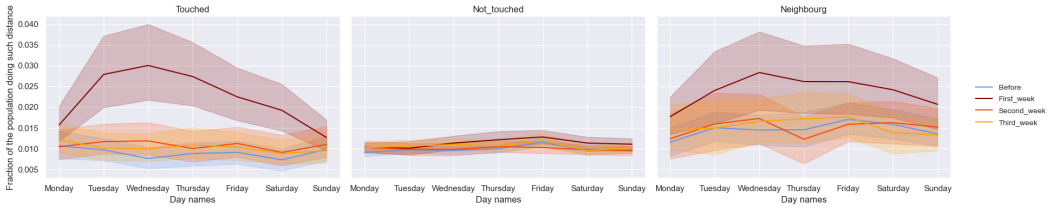
(a) Fraction of the population no moving



(b) Fraction of the population between 0 and 10km



(c) Fraction of the population between 10 and 100km

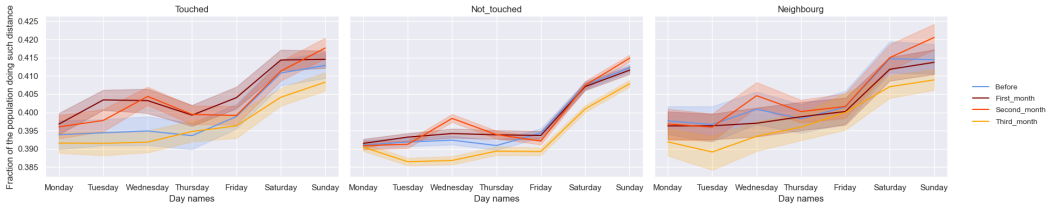


(d) Fraction of the population from 100km and over

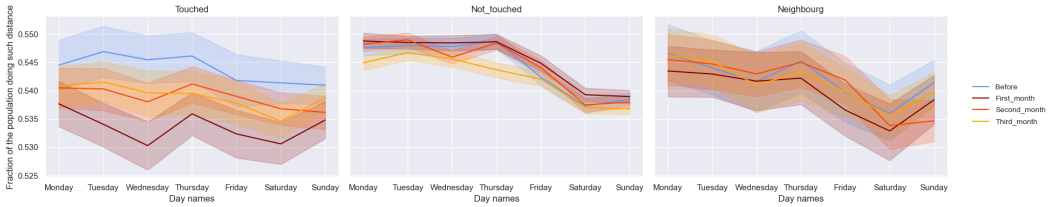
Figure 22: Time Series by days of the week
and the three week (and during) the disaster.

In the touched provinces type of provinces and during the first week, the movement of the population is two-sided with 2.7% on Tuesday (the second day after the event) and 3.0% of the population on Wednesday doing more than 100 km instead of 1% on Tuesday and 0.75% on Wednesday.

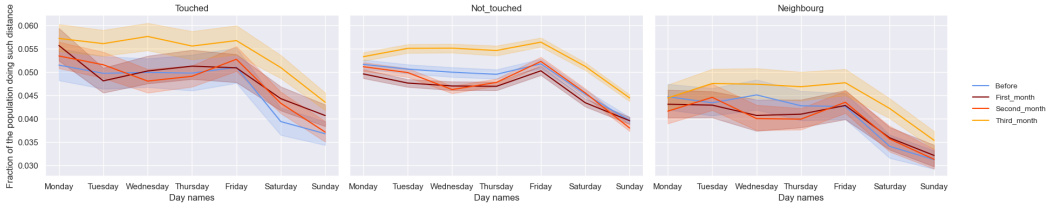
At the same time, on Tuesday, 42.1% of the population studied (compared to 39.5% before the event) was not moving. If the earthquakes had impacted the provinces similarly, the data would show a decrease in both categories, with one experiencing a reduction followed by a corresponding decrease in the other. The two last fractions are impacted by a decrease in their fraction of the population. The population doing 0 to 10 km is decreasing at the lowest with a peak on Wednesday at 51.0% (vs 54.9% before). The bind of the distance between 10 and 100 km decreased from 5% before to 3.9% on Tuesday. In these binds, it is remarkable that a peak on Thursday (Figure 22).



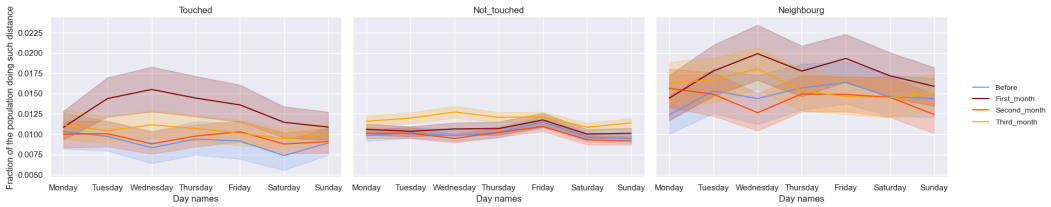
(a) Fraction of the population no moving



(b) Fraction of the population between 0 to 10km



(c) Fraction of the population 10 to 100km



(d) Fraction of the population from 100km and over

Figure 23: Time Series by days of the week
and the three months (and during) the disaster.

In the first week following the earthquake, the population movement in the affected provinces exhibited notable two-sided shifts. On Tuesday (the second day after the event), 2.7% of the population travelled more than 100 km, and on Wednesday, this figure increased to 3.0%, compared to only 1% on Tuesday and 0.75% on Wednesday under normal conditions.

Additionally, on Tuesday, 42.1% of the population remained stationary, a notable increase from the 39.5% observed before the earthquake. Had the earthquake impacted the provinces

in a uniform manner, we would have expected a decrease in population movement in some sectors and an increase in others. The data clearly indicates a reduction in movement across certain distance categories. For instance, the proportion of individuals travelling between 0 and 10 km decreased the least, peaking at 51.0% on Wednesday (compared to 54.9% before the event). The proportion of individuals travelling between 10 and 100 km dropped from 5% before the event to 3.9% on Tuesday. Furthermore, the distance category of 100 km and more showed a peak on Wednesday, signaling a temporary surge in long-distance movement (as seen in Figure 22).

In neighbouring provinces, the response mirrored that of the directly affected regions, though at a lower magnitude. The proportion of the population that did not move reached two peaks during the first week, with 40.9% remaining stationary on Tuesday and Wednesday, compared to a pre-event figure of 39%. The categories that exhibited the most significant shifts were those involving longer distances: the fraction of the population travelling more than 100 km peaked at 2.6% on Wednesday, up from 1.5% in normal times. Similarly, the categories for distances ranging from 0 to 10 km and 10 to 100 km experienced reductions, with the 0 to 10 km category seeing a drop of 1.2% on Wednesday, and the 10 to 100 km category reaching a peak of 3.5% on Tuesday, down from 4.5% before the event.

Provinces not directly impacted by the earthquake, but still sensitive to its effects, showed a similar trend, albeit with only two major shifts: an increase in the proportion of the population not moving, which peaked at 40.4% during the first week (up from 39.1% before the event), and a decrease in the population traveling between 10 and 100 km. On Tuesday, there was a slight overall decline of 0.8% in population movement compared to normal levels.

The second week of analysis revealed further divergences from the baseline. All provinces, whether affected, neighboring, or not impacted, showed a decrease in mobility compared to the first week. In the affected provinces, the percentage of individuals remaining stationary peaked on Friday at 41.9%, reflecting a notable increase from the 38.5% observed on Monday and 39.4% before the event. For long-distance travel, the 10 to 100 km category exhibited a significant decrease, dropping from 6.2% on Monday to 4.0% by Friday. Although the fraction of individuals traveling more than 100 km remained higher than pre-event levels (by about 0.5%), the category of 0 to 10 km did not fully return to its baseline, with a general decline of 0.4% observed over the week.

Neighboring and unaffected provinces displayed similar patterns to the affected regions, though the magnitude of the changes was smaller. Notably, the provinces closest to the earthquake site exhibited larger differences than those farther away. By the third week, the mobility patterns appeared to return to normal, with only minor deviations from pre-event behaviour.

Looking at the monthly trends in Figure 23, we see that the third month post-earthquake showed the greatest impact in the neighboring and unaffected provinces, particularly due to the national holiday period known as the Ramadan Feast Holiday (April 21st to 23rd). The most notable exception to this general trend was the category of long-distance travel (more than 100 km) in neighboring provinces, where the first month following the earthquake saw a noticeable increase. This category peaked twice during the week, with 2% of the population traveling more than 100 km on Wednesday (up from 1.54% pre-event) and 1.73% on Friday.

In the affected provinces, the first month showed the most significant disruptions, es-

pecially in the category of 10 to 100 km, where the third month saw an increase of 0.8% compared to the regular weekly pattern (Figure 23). This shift coincided with the flooding event in Sanliurfa and Adıyaman on March 13th, which caused an uptick in the proportion of individuals not moving on Wednesdays. The fraction of people traveling between 0 and 10 km also decreased slightly (54.1% instead of 54.5%).

In conclusion, Figure 23 (and the corresponding data in Table 28) provides strong evidence that the earthquake had a far more significant impact on population movement than the national holidays or flooding events, particularly in the first week following the disaster. The data also highlight how, over time, the affected provinces showed a return to normalcy, with the exception of certain categories like long-distance travel. Furthermore, the analysis underscores the differing responses between affected, neighboring, and unaffected provinces, with the former showing the most pronounced shifts in mobility.

	First Week	**Second Week**	**Third Week**
Neighbors provinces			
Distance no movement	0.38	0.44	0.52
Distance between 0 to 10	0.59	0.60	0.63
Distance between 10 to 100	0.67	0.74	0.76
Distance of 100 km and over	0.65	0.62	0.65
Impacted provinces			
Distance no movement	0.33	0.40	0.49
Distance between 0 to 10	0.48	0.56	0.61
Distance between 10 to 100	0.56	0.66	0.70
Distance of 100 km and over	0.44	0.49	0.56
Not impacted provinces			
Distance no movement	0.50	0.54	0.57
Distance between 0 to 10	0.61	0.63	0.62
Distance between 10 to 100	0.77	0.80	0.81
Distance of 100 km and over	0.59	0.59	0.59

Table 28: Spearman correlation of the provinces. All of the correlation are statistically significant.

To assess the impact of the earthquake event, Figure 25 and Figure 26 offer valuable insights into the population's response, especially when compared with the mean PGA (Peak Ground Acceleration) values of the earthquake (Figure 24). While Osmaniye exhibits the highest mean PGA, the weekly time series data reveals minimal impact on population movement in this province when compared to others, such as Kilis and Elazığ. Notably, Gaziantep, being located at the epicenter, was expected to show significant changes in mobility. However, the data reveals a surprisingly low difference in mobility patterns for this province, except for the category of individuals traveling more than 100 km.

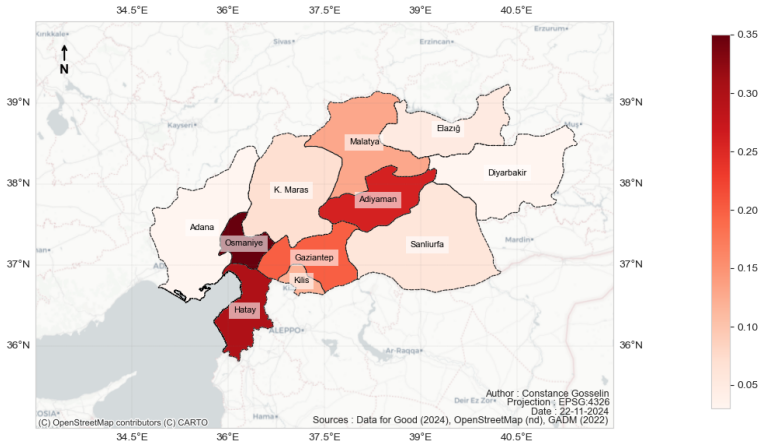


Figure 24: PGA mean per province.

In the fraction of no movement, the first week is mainly impacted with a higher percentage of the population stay at home. The highest difference is Malatya with 45.5% of the population not moving compared to 41.2% as a regular day. Elazığ and Kilis are also showing quite an important peak during the first week. Elazığ non mobility peak is on Thurdays an is up to 44.3%. Kilis's peaks the non mobility on Wednseday with 44.6% instead of 42%. Depending of the non mobility seems to be affected in two different ways the first week. Monday and Tuesday are the most affected then there is a decrease either on Wednesday or Thursday. There are provinces that have and increase of no movement during the week end (which also compared with a regular week) : Elazığ, Osmaniye, Gaziantep, Adana, Dyarbakir and Hatay.

During the second week, a pattern emerges on the no movement category with a peak on Wednesday for Adana, Gaziantep, Hatay, Osmaniye, Kahramanmaraş, and Kilis. In contrast, Adıyaman and Malatya experience peaks on Thursday. Diyarbakır, Elazığ, and Şanlıurfa display distinct timelines compared to other provinces. Şanlıurfa exhibits a steady increase, nearing pre-event levels with only a 0.5% change on Sunday. Notably, Kahramanmaraş and Hatay also show signs of normalization in their population mobility by the end of the second week. Elazığ demonstrates fluctuating mobility patterns, beginning with a lower fraction of the population remaining stationary compared to pre-event levels and experiencing two notable low points: 39% (down from 41%) and 38% (down from 40.5%). Diyarbakır's second week reveals a gradual rise in stationary individuals, increasing from 39.5% on Monday to 43.6% on Friday, with peaks of 42.0% on Tuesday. Although deviations from pre-event levels persist, these differences are diminishing.

More than half of the affected provinces—Adana, Şanlıurfa, Osmaniye, Hatay, Gaziantep, and Kahramanmaraş—align with pre-event mobility patterns by the third week for the no movement category. Minor variations remain but are within a margin of 0.5%. Malatya, Adıyaman, and Diyarbakır also show progress toward normalization, though some deviations persist, such as a 2.3% higher stationary population in Malatya on Saturday, a 1% increase in Diyarbakır on Sunday, and significant differences in Adıyaman on Monday (2.3%) and

Thursday (1.8%). Kilis and Elazığ continue to exhibit more dispersed patterns. By Wednesday of the third week, Kilis approaches pre-event mobility levels, whereas Elazığ deviates from modeled trends with a decline on Thursday followed by an increase on Friday.

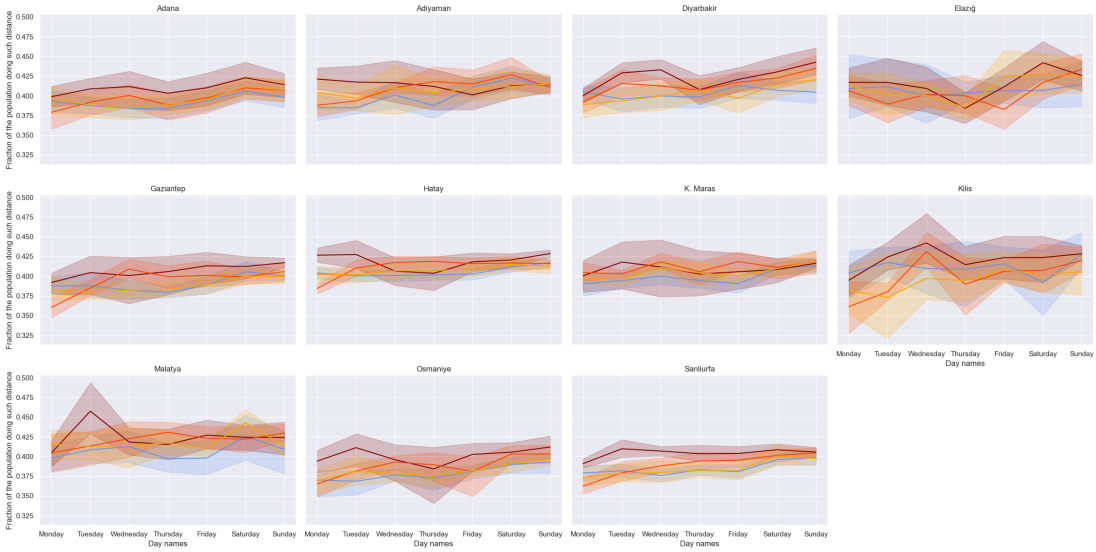
For individuals travelling 0 - 10 km a day, except in Elazığ and Kilis, the EQ's impact is evident as an average decrease in mobility. The first week is the most disrupted, with mobility patterns gradually returning to normal in subsequent weeks. This analysis focuses on week one disruptions, irregularities in weeks two and three, and anomalies in Elazığ and Kilis. Despite the overall decrease, data reveal limited impacts on short-distance mobility in Gaziantep, Malatya, Diyarbakır, Osmaniye, Adana, and Şanlıurfa, ranging from a 1.5% decrease in Şanlıurfa to 4.0% in Gaziantep. These provinces show low peaks on Tuesday/Wednesday and Thursday/Friday, except for Şanlıurfa, which peaks only on Wednesday, and Diyarbakır, where mobility decreases by 2% starting Wednesday.

Kahramanmaraş, Hatay, and Adıyaman display prominent peaks on Tuesday or Wednesday, with stationary population rates reaching 47.5% (down from 55.7%) in Kahramanmaraş, 48% (down from 55%) in Adıyaman, and 50% (down from 55%) in Hatay. Weeks two and three show widespread stabilization, excluding Elazığ and Kilis.

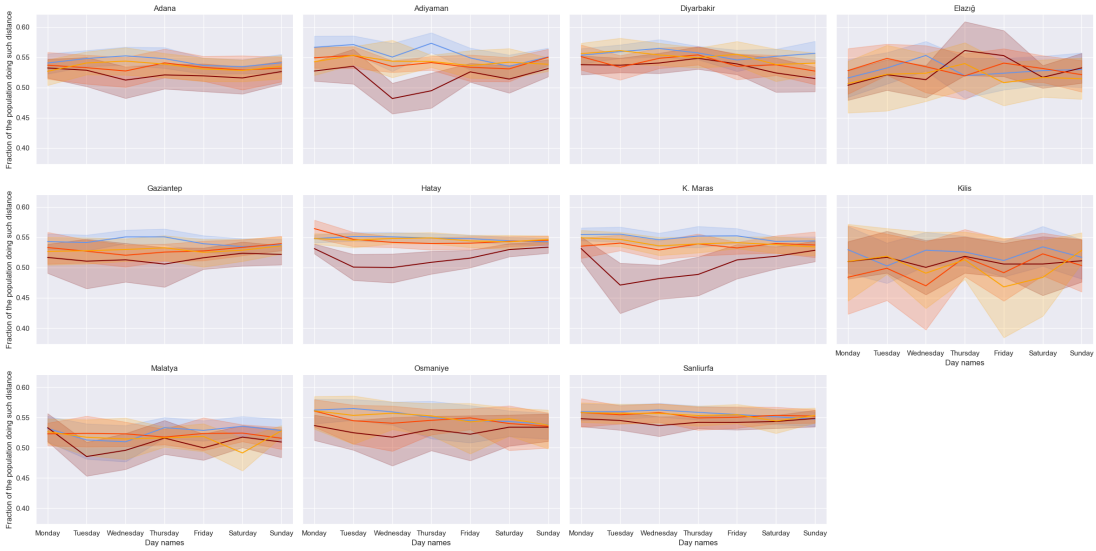
Exceptions include a significant dip in Malatya's stationary population on the second Saturday (49.5%, down from 52.6%) and an initial increase in Hatay's stationary population at the start of the second week. Diyarbakır continues to exhibit daily fluctuations with notable peaks and troughs.

Elazığ and Kilis stand out with pronounced and persistent disruptions. Elazığ's first week sees a midweek increase in mobility, followed by a sharp decline, while Kilis shows stability during week one but experiences peaks in weeks two and three.

Figure 26 illustrates the continuation of Figure 25, focusing on mobility bands of 10–100 km and over 100 km. Mobility in the 10–100 km range shows minimal earthquake impact, with slight peaks below 1%. Across most provinces, this band experiences an average decline on Monday and Tuesday, with peaks on Wednesday and Thursday in Adıyaman, Hatay, and Kahramanmaraş. Post-earthquake, Malatya exhibits smoother curves, indicating reduced variability compared to pre-event patterns



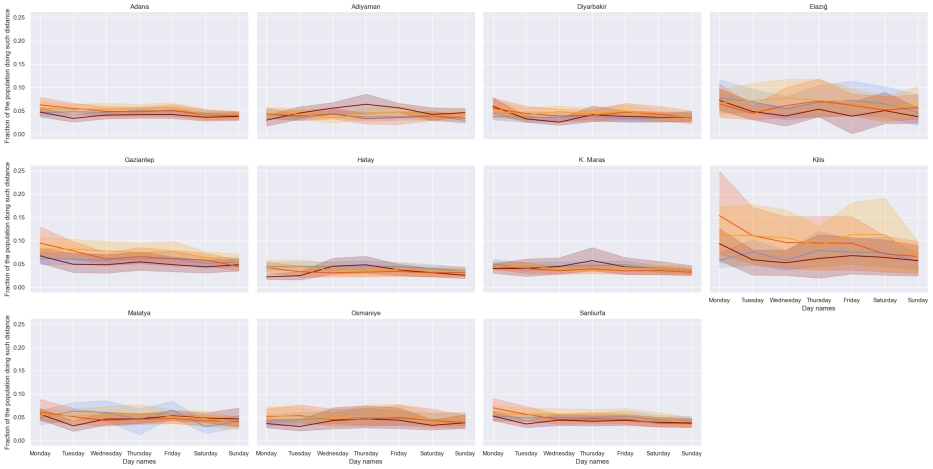
(a) Fraction of the population no moving



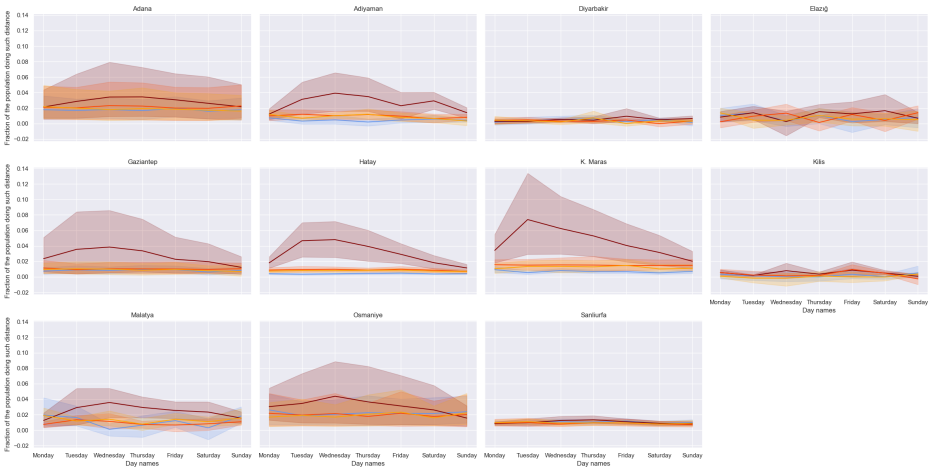
(b) Fraction of the population doing between 0km and 10kms

(c) Legend

Figure 25: Time Series by days of the week and the three week (and during) the disaster and by the provinces touched by it.



(a) Fraction of the population doing t between 10km and 100km



(b) Fraction of the population doing from 100km and over

(c) Legend

Figure 26: Time Series through days of the week and the three week (and during) the disaster and by the provinces touched by it.

Elazığ and Kilis exhibit notable dissimilarities compared to other provinces in the 10 to 100 km distance category. In Elazığ, low peaks are observed during the first week, followed by a trend that aligns more closely with the regular curve in subsequent weeks. Conversely, Kilis stands out as the province most impacted in terms of mid-distance mobility. While the first week shows a decrease similar to other provinces, weeks two and three are particularly remarkable. In week three, up to 15.1% of Kilis's population engaged in mid-distance travel on Monday, followed by a gradual decline through Sunday. During week two, the proportion of the population travelling such distances increases by 5% compared to regular week levels.

For distances over 100 km, the fraction of the population travelling shows a sharp increase during the first week, with subsequent weeks approaching regular patterns. Şanlıurfa, Kilis, and Diyarbakır are less impacted than other provinces in this category. Kilis displays

two pronounced peaks during week one, with a deviation of 0.8% compared to normal levels on both Wednesday and Friday. Elazığ again stands out with unique patterns compared to other provinces. Malatya exhibits disturbances akin to Elazığ, with the first days of week one resembling pre-event levels before deviations emerge, including a 1% increase in population travelling over 100 km. Week two closely mirrors pre-event levels, except for a notable variation on Tuesday. Week three shows minor disturbances, with the largest deviation being 0.6%.

Adana shows a relatively subdued peak, spread across Wednesday and Thursday, with 3.8% of its population travelling over 100 km, compared to 2% in a regular week. Osmaniye experiences a peak of 4.2% on Wednesday, up from 2%. Adıyaman shows a significant increase, with 4% of its population traveling such distances on Wednesday, compared to 0.2% under normal conditions. Similarly, Gaziantep reaches 4% on Wednesday, up from 1%. Hatay exhibits a smoothed peak spanning Tuesday and Wednesday, with 4.3% compared to 0.2%. Malatya experiences a peak of 3.2% on Wednesday, compared to 0%. Kahramanmaraş records the most pronounced increase, with 7.8% of its population travelling over 100 km on Tuesday, compared to 0.2%.

Summary

The province-scale analysis of the case study reveals that mobility changes induced by the earthquake are largely absorbed within the first month following the event.

Interestingly, mobility changes are not confined to the directly impacted provinces. Neighboring provinces and even those not directly affected also exhibit shifts in mobility patterns. This suggests that the shock of the earthquake was felt far beyond the epicenter.

Neighboring provinces demonstrate a more pronounced change in mobility compared to provinces that were not affected, highlighting a gradient of impact based on proximity.

In terms of distance categories, two stand out as experiencing the most significant changes: the "no movement" category and the "100 km and over" category. These categories reflect the population's immediate and long-distance responses to the disaster.

Over the timeline, there are notable peaks and troughs in mobility patterns. These fluctuations may be attributed to outlier effects, such as external influences or irregularities in population movement.

VII.1.b District scale analysis

At the provincial scale, the results reveal remarkably intriguing patterns when compared to the mean PGA. This suggests that spatial scale may play a significant role in interpreting the data. To address this, this section offers insights into the earthquake's impact using a normalized difference metric[9] (see Figures Figure 31, Figure 32, and Figure 29).

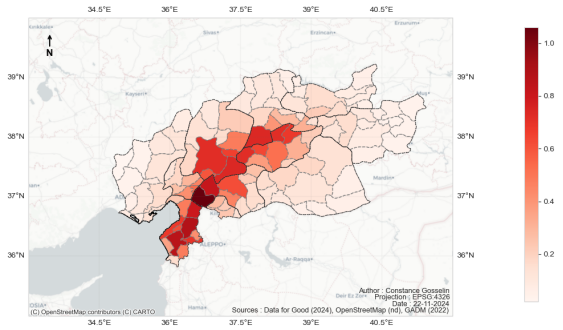


Figure 27: PGA mean per districts.

Correlation and linear regression analyses offer valuable insights into the relationship between the earthquake event and population mobility across all districts. Remarkably, not all population movements in the affected districts demonstrate statistically significant alignment with PGA values. The difference between pre-event and weekly mobility, normalized by the mean, shows a correlation with time, as illustrated in Figure 28. All statistically significant correlations indicate a low positive relationship, with no evidence of negative correlations.

Additional details regarding the relationship between PGA and population displacement are presented in Table 29. The analysis reveals that only the 100 km and over and no movement categories exhibit statistical significance during the first and third weeks following the earthquake. Specifically, there was a significant increase in the number of people travelling 100 km or more during the first week, while the third week saw a notable decrease in the proportion of people remaining stationary.

Figure 29 illustrates the impact of the EQ on mobility patterns on the day of the event. Overall, there are notable disparities in mobility reactions across districts. The first key observation is the variation in the range of differences across categories. The largest range is observed in the 100 km and over category, spanning from -24.74 to 30.46, indicating a wide diversity of movement in this group. In contrast, the smallest range is found in the 0 to 10 km category, suggesting more consistent behaviour. The second smallest range is observed in the no movement category, followed by the 10 to 100 km category.

[9]Refer to the methodology outlined in the study case for further details

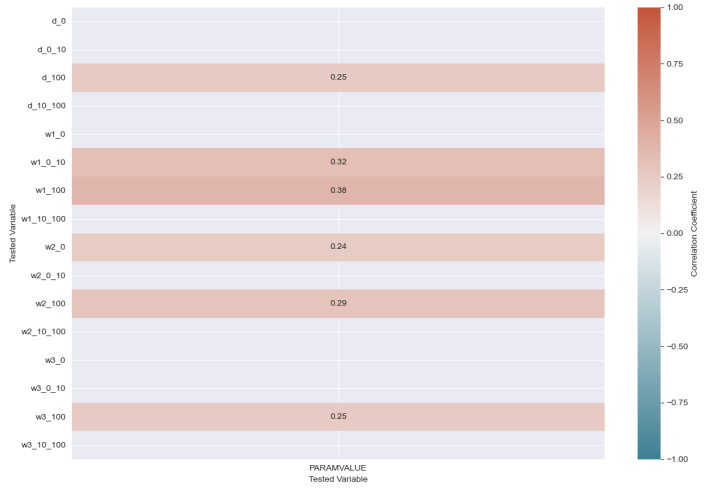


Figure 28: Spearman correlation of PGA and normalized difference of the FP.

Fraction of population	**Coefficient**	**p-value**	**Significant**
Between before and during the earthquake			
No movement	0.0122	0.053	Not significant
0 to 10 km	0.0008	0.945	Not significant
10 to 100 km	-0.0080	0.318	Not significant
100 km and over	-0.005	0.089	Not significant
Between before and first week after the earthquake			
No movement	0.0419	0.303	Not significant
0 to 10 km	-0.0065	0.903	Not significant
10 to 100 km	-0.0090	0.479	Not significant
100 km and over	0.0859	0.011	Significant
Between before and second week after the earthquake			
No movement	-0.0098	0.578	Not significant
0 to 10 km	0.0008	0.945	Not significant
10 to 100 km	-0.0016	0.322	Not significant
100 km and over	-0.0130	0.090	Not significant
Between before and third week after the earthquake			
No movement	-0.0191	0.027	Significant
0 to 10 km	0.0403	0.086	Not significant
10 to 100 km	-0.0080	0.318	Not significant
100 km and over	-0.0057	0.089	Not significant

Table 29: Linear regression between pre-EQ values and the three week following the EQ.

District-level patterns reveal that southeastern districts are more likely to exhibit stationary behaviour and display greater heterogeneity in this context. In the extreme south, populations are more likely to move, with slight deviations from the mean. Hatay stands out as a province with higher mobility overall, with few exceptions.

During the event, districts in the north show an increased proportion of the population undertaking higher-than-normal displacements. Notably, districts along the fault line in Elazığ exhibit a significant decrease in the fraction of the population travelling 0 to 10 km. Conversely, southern districts display a slight increase in this category but are among the least impacted overall. Hatay and Kilis demonstrate considerable heterogeneity, with districts showing both high negative and positive changes in mobility.

The 10 to 100 km category exhibits significant disparities. Figure 29 highlights that this disparity is driven by a few extreme differences: a pronounced negative deviation in Malatya and a marked positive deviation in Kilis. Overall, districts display heterogeneity in this category, with those along the fault line and to the north of the epicenter showing an increase in this mobility range. In contrast, districts in Kahramanmaraş along the fault line experience a significant negative impact.

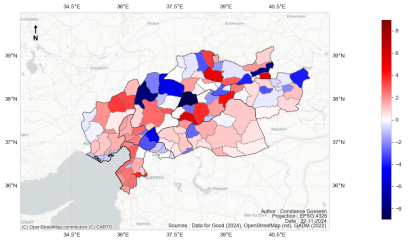
For the 100 km and over category, disparities are evident, as shown in Figure 29. These movements often contrast with those observed in the 0 to 10 km category. Populations in the "100 km and over" category tend to either slightly decrease or remain unchanged in their mobility. Northern districts along the fault show increased movement in this category, with exceptions in less-affected districts of Adana and eastern districts of Elazığ.

Lisa's cluster analysis in Figure 29 highlights outliers and districts with notable spatial correlations. Hatay and provinces north of the fault exhibit more outliers, with Kahramanmaraş's northern districts showing low outliers (LH) in the no movement category. For the 0 to 10 km and 100 km and over categories, high clusters are evident in northern districts.

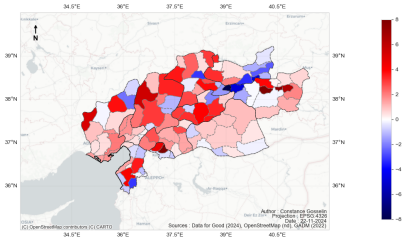
In Malatya, two districts stand out as high outliers in the same categories compared to surrounding districts. Similarly, for the 10 to 100 km distance category, high outliers are observed in Elazığ districts, consistent with earlier observations. Border districts between Malatya, Kahramanmaraş, and Adıyaman also show low outliers for the 10 to 100 km category.
Figure 31 illustrates the differences between districts for the no movement and 0 to 10 km (FP). The no movement fraction appears to be most impacted during the third week, with significant positive and negative differences observed in Malatya and Elazığ.

Districts along the fault line show only slight impacts during the first week, while southern districts exhibit a higher proportion of people remaining stationary. In contrast, districts north of the fault and those on the fault line demonstrate greater heterogeneity. During the second and third weeks, the "no movement" fraction increases on the fault and in northern districts, with the third week showing the most pronounced disparities, particularly in districts close to the epicenter.

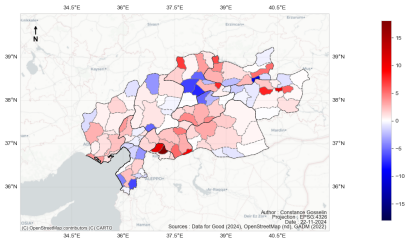
For the 0 to 10 km category, the range of differences is relatively narrow compared to other mobility categories. Overall, districts tend to show increased mobility in this range over the three weeks. However, districts bordering Malatya, Elazığ, and Diyarbakır experience a decrease in this fraction. Once again, districts south of the fault are less impacted compared to others, while slight heterogeneity is observed on the fault and north of it.



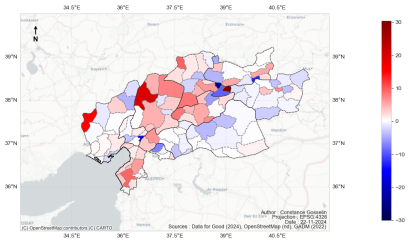
(a) No movement category



(b) 0 to 10 km distance category

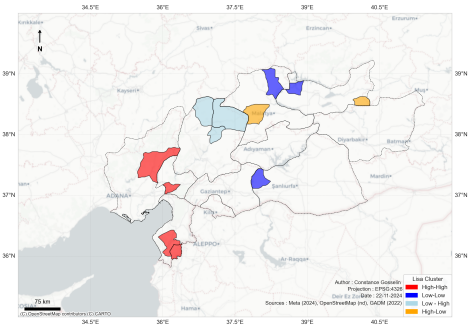


(c) 10 to 100 km distance category

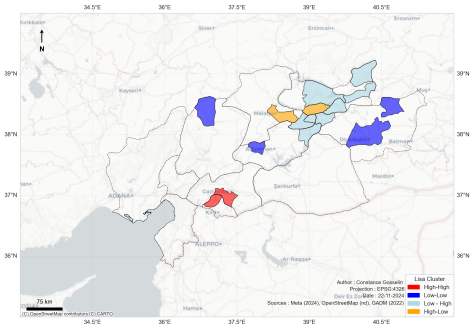


(d) 100 km and over distance category

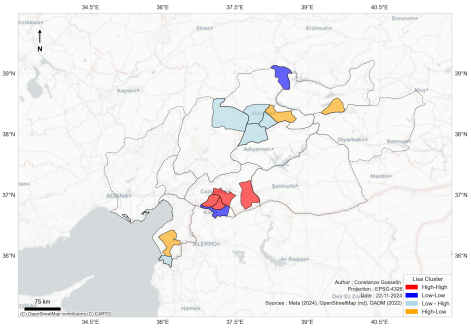
Figure 29: Normalized difference of the FP values during the EQ and per district.



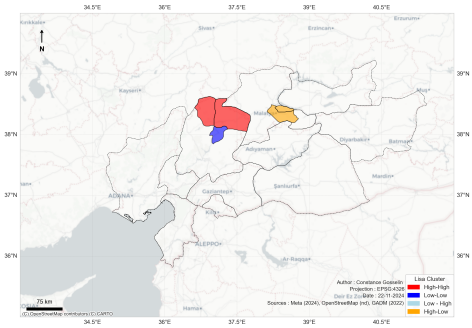
(a) No movement distance category



(b) 0 to 10 km distance category

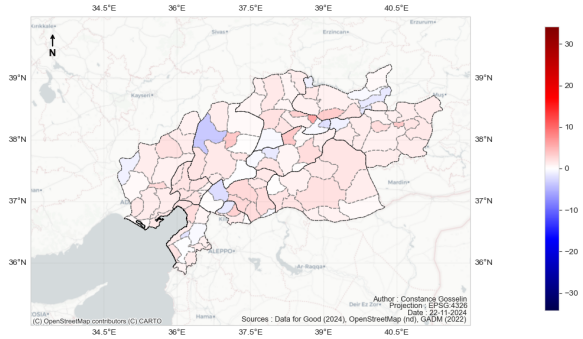


(c) 10 to 100 km distance category

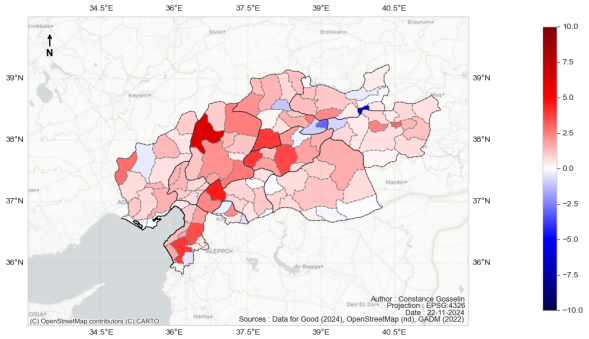


(d) 100 km and over distance category

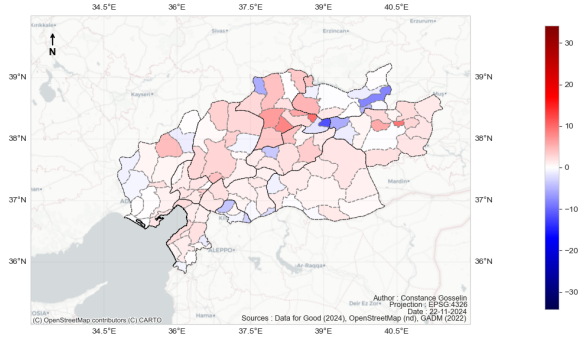
Figure 30: Local moran index cluster of the mobiloty change during the EQ



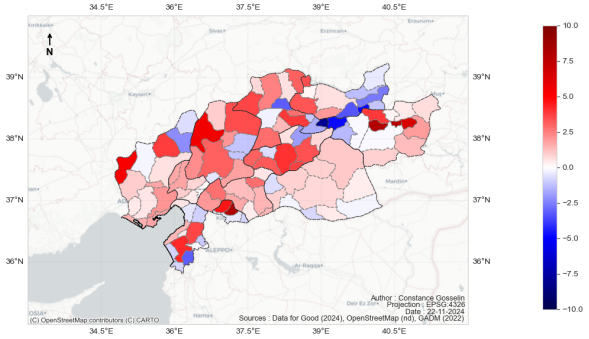
(a) First week and no movement category



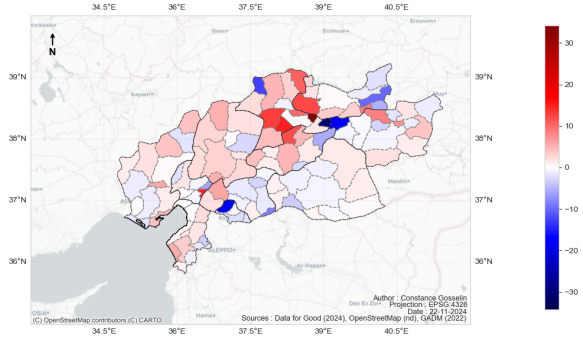
(b) First week and 0 to 10 km distance category



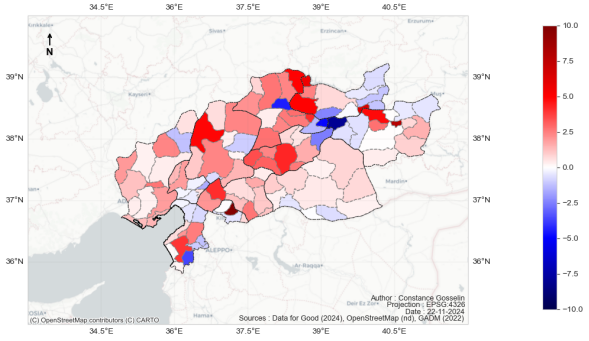
(c) Second week and no movement category



(d) Second Week and 0 to 10 km distance category

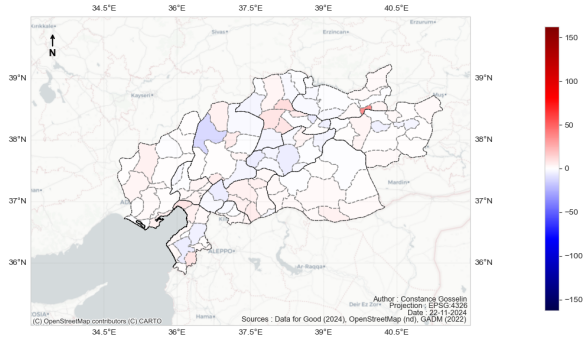


(e) Third week and no movement category

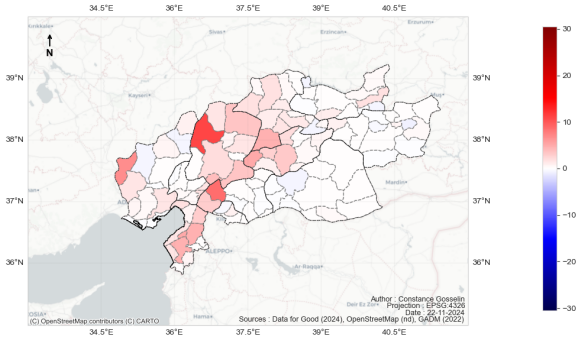


(f) Third week and 0 to 10 km distance category

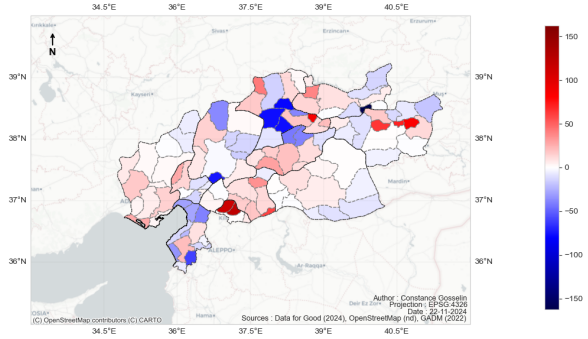
Figure 31: Normalized difference of the FP value following the three weeks after the EQ for the no movement and 0 to 10 km distance category.



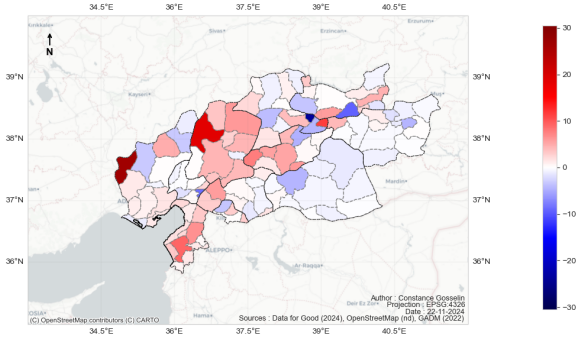
(a) First week and 10 to 100 km distance category



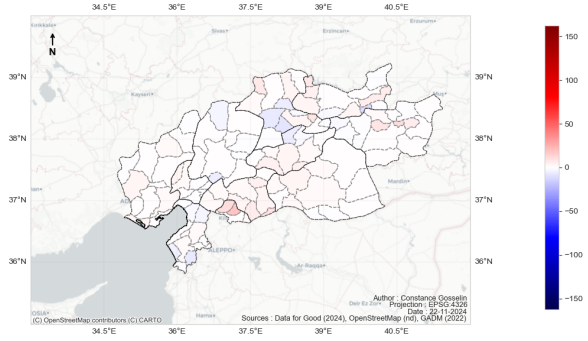
(b) First week and 100km and over distance category



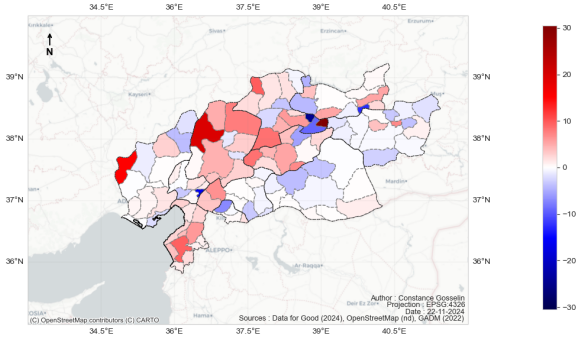
(c) Second week and 10 to 100km distance category



(d) Second Week and 100 and over distance category



(e) Third week and 10 to 100 km distance category



(f) Third week and 100 km and over distance category

Figure 32: Normalized difference of the FP value following the three weeks after the EQ for the 10 to 100 km and 100 km and over distance category.

The 10 to 100 km distance category (Figure 32) exhibits the widest range of differences, reaching up to 150. This category is most affected during the second week, with significant negative changes in Malatya and Hatay provinces. Conversely, large positive differences are evident in Kilis and Diyarbakır.

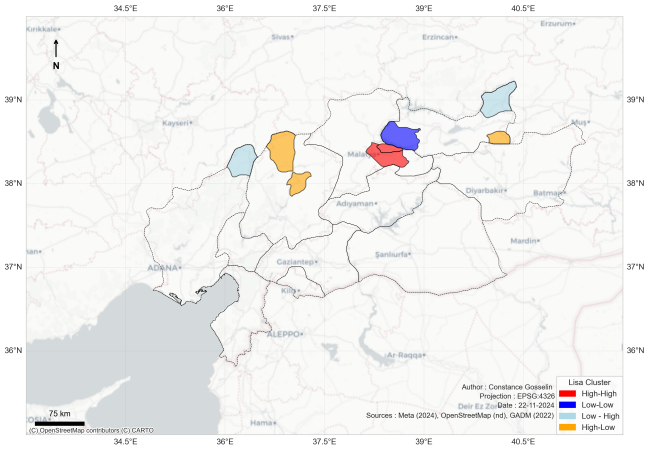
The final category, 100 km and over, exhibits a range of values similar to the no movement category. Figure 32 clearly shows the earthquake's impact, with increased mobility in the districts along and near the fault line. There is a noticeable contrast between the northern and southern parts of the fault. Districts in the south tend to show a decrease in the fraction of the population travelling more than 100 km, while the northern districts exhibit greater heterogeneity, with a marked increase in mobility, particularly in the districts of Kahramanmaraş. Mobility for distances greater than 100 km compared to the mean for all impacted districts is more significant in the second and third weeks.

Local Moran's indices for the differences per district, divided by the mean, provide valuable insights into the spatial patterns of displacement. In general, the northern part of the region shows more spatial clusters and outliers compared to the south. In the first week, districts with high autocorrelation (high-high) are found in Hatay, Kahramanmaraş, and Malatya for the 0 to 10 km and 100 km and over categories. Low autocorrelation (low-low) is observed in Elazığ and Diyarbakır for the no movement and 10 to 100 km categories.

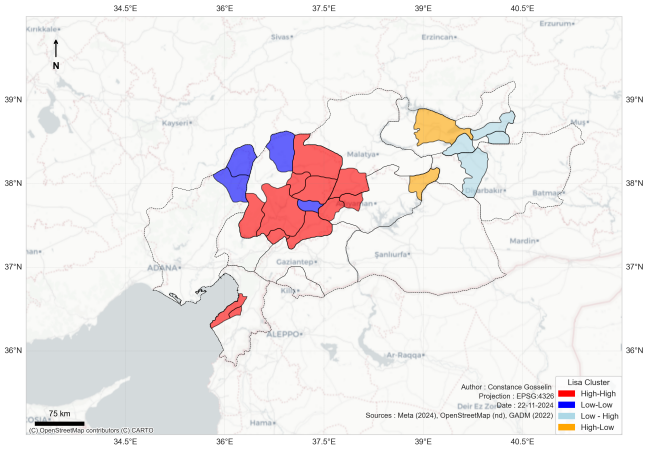
The outliers in the first week are predominantly located in the southern provinces, such as Gaziantep, Şanlıurfa, and Diyarbakır, particularly for the 100 km and over category. Both Elazığ and Diyarbakır exhibit both high and low outliers in the 0 to 10 km category. The no movement category does not show significant spatial patterns, with few spatial clusters or outliers observed.

In the second week, the northern and eastern regions display more spatial division, with clusters and outliers becoming more pronounced. Elazığ and districts along the fault line, after experiencing high PGA, show low outliers surrounded by high values in the no movement and 0 to 10 km categories. Kahramanmaraş shows predominantly high autocorrelation (high-high) in the 100 km and over category. Kilis and Gaziantep border districts exhibit high autocorrelation (high-high) in the 0 to 10 km category. Other areas in this week show low autocorrelation (low-low) clusters, particularly in Diyarbakır, Kahramanmaraş, and Adıyaman. The 10 to 100 km category predominantly shows clusters in Kilis, Gaziantep, and Diyarbakır.

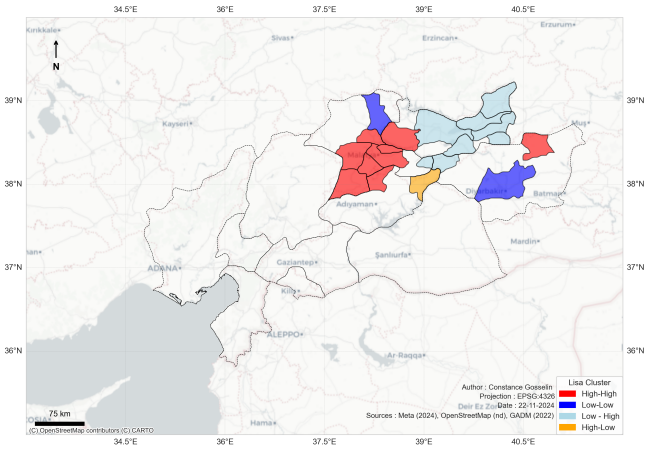
In the third week, the no movement category reveals that Malatya has formed distinct clusters of districts. For both the no movement and 0 to 10 km categories, districts sharing administrative borders in Malatya, Diyarbakır, Elazığ, and Adıyaman appear as outliers with lower values compared to surrounding districts. In the 0 to 10 km and 10 to 100 km categories, Kilis and Gaziantep districts display both high-high clusters and low-low clusters. Hatay exhibits two types of outliers within the same category. In the 100 km and over category, the northern districts of Kahramanmaraş show both high-high and low-low clusters.



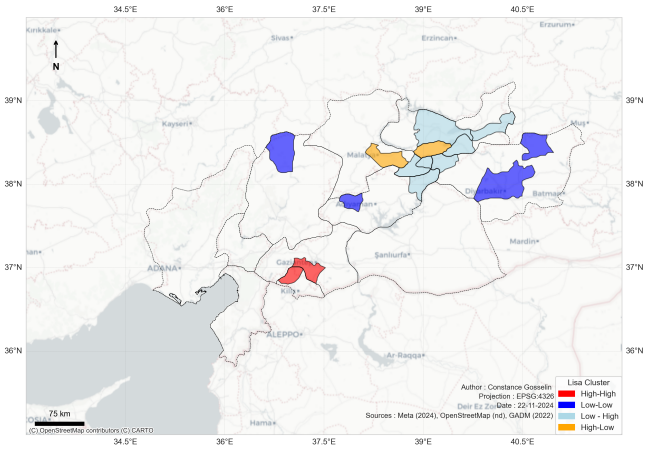
(a) First week and no movement distance category



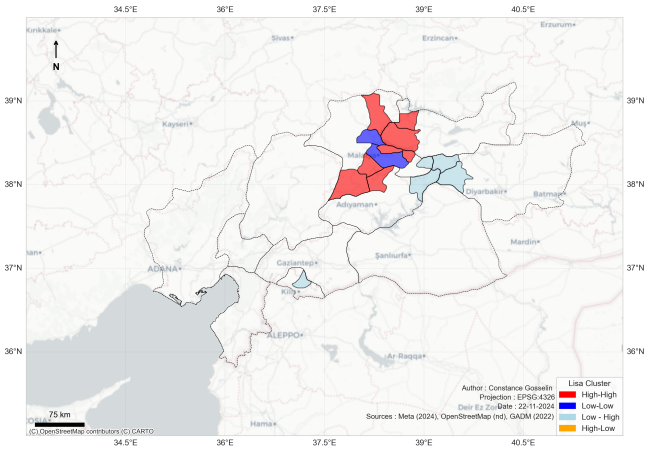
(b) First week and 0 to 10km distance category



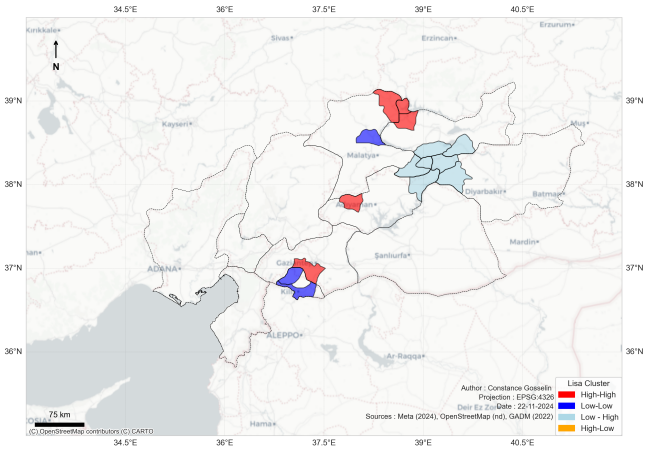
(c) Second week and no movement distance category



(d) Second Week and 0 to 10 km distance category

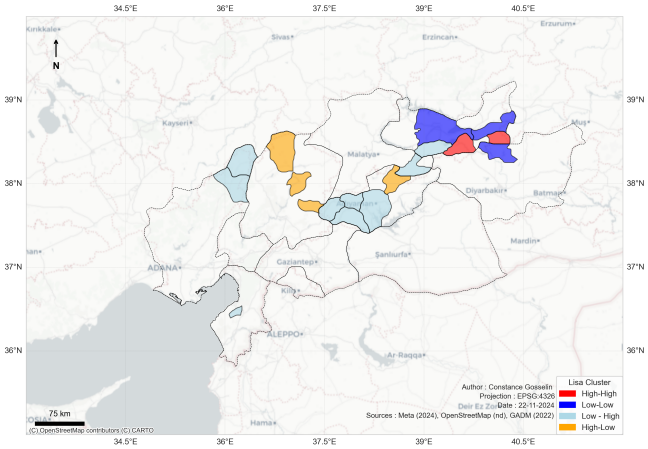


(e) Third week and no movement category

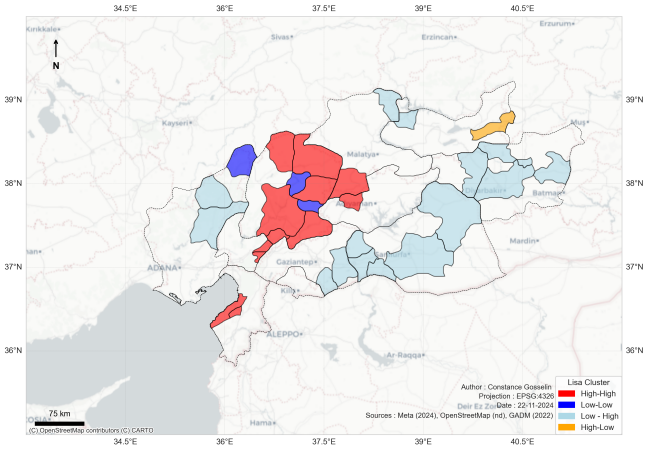


(f) Third week and 0 to 10 km distance cateogory

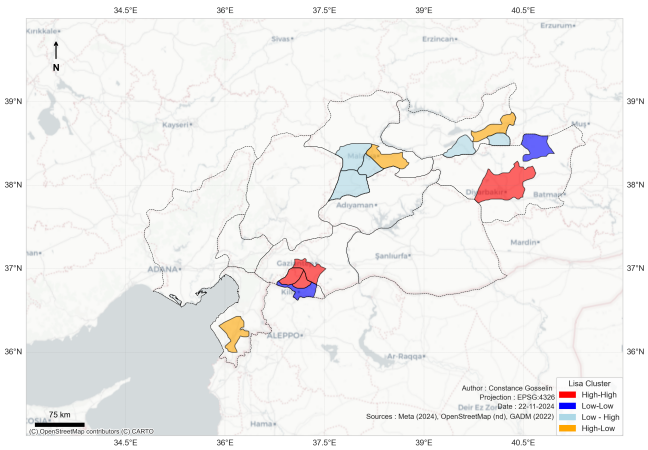
Figure 33: Local moran indicator of normalized difference of the FP value following the three weeks after the EQ for the no movement and 0 to 10 km distance categories.



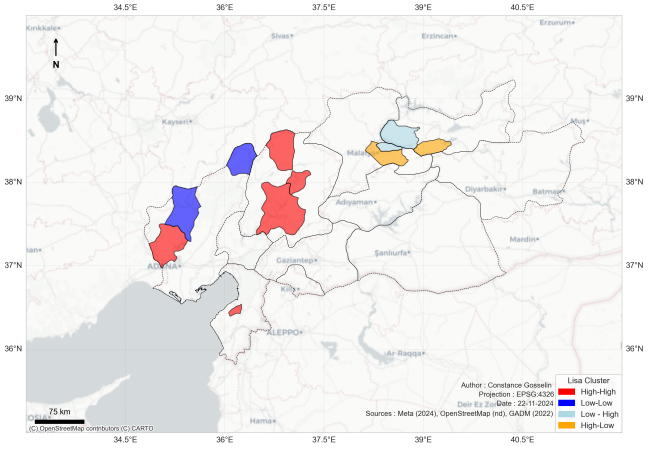
(a) First week and 10 to 100 km distance category



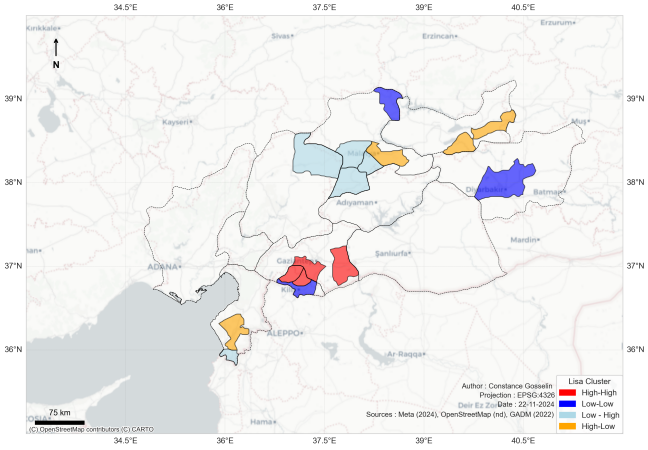
(b) First week and 100km and over distance category



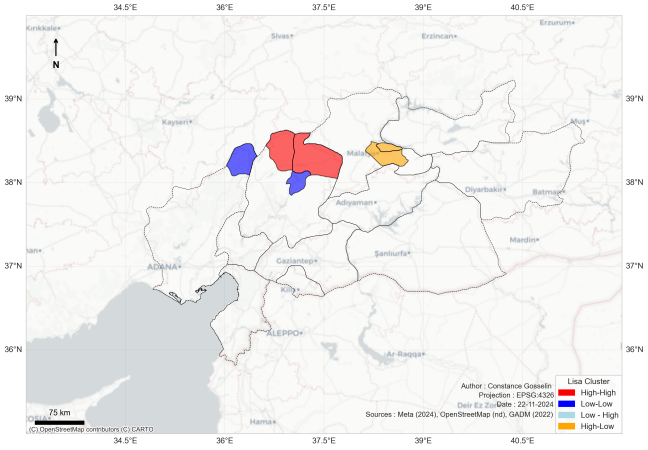
(c) Second week and 10 to 100km distance category



(d) Second Week and 100 km and over distance category



(e) Third week and 10 to 100 km distance category



(f) Third week and 100 km and over distance category

Figure 34: Local moran index cluster of normalized difference of the FP value following the three weeks after the EQ for the 10 to 100 km and 100 km and over distance categories

Summary

The normalized difference in mobility does not align with the PGA (Peak Ground Acceleration).

However, districts located north of the fault exhibit greater changes in mobility compared to those south of the fault.

Districts along the southern fault display relatively uniform changes in mobility, while districts along the northern fault show significant heterogeneity. This variability leads to the formation of spatial clusters and the identification of outliers.

In terms of the distance categories, they all show a difference in their mobility which are similar to the one shown in the Province scale analysis.

VII.2.c Rural versus Urban

In the context of a geographical study examining the impacts of the earthquake on urban versus rural areas, the following results provide insights into the districts with the highest destruction reports in urban areas and their rural equivalents in terms of PGA.

Figure 35 and Figure 36 represent the timeline for the first three weeks following the earthquake. Figure 35 presents data on the fraction of the population (FP) that remained stationary (no movement) and traveled between 0 to 10 km. Rural districts in Adıyaman, Malatya, and Kahramanmaraş show significant shifts in mobility, with notable changes in the FP doing 0 to 10 km, no movement, and 10 to 100 km throughout the three weeks. In contrast, urban districts display smoother changes, even crossing the reference lines for typical weekly mobility. The earthquake's impact does not always align with the PGA values for these districts.

At a finer level of analysis, the no movement category provides further insights into the differences between rural and urban areas. Gaziantep's rural district shows a lower increase of 4.6% in the fraction of the population not moving, compared to 5.9% and 6.2% in the urban districts. On the other hand, Hatay exhibits a significant dip in movement, with a low peak of 36.1% in the rural district, while the urban districts show smoother trends and higher proportions of the population not moving. Apart from Adıyaman, rural areas with higher shifts tend to have larger fractions of the population remaining stationary.

In Matalya and Adıyaman's rural districts, the fraction of the population not moving increases after the first week. Gaziantep and Hatay's rural districts, however, show less impact in terms of people moving between 0 to 10 km. All urban districts demonstrate smoother patterns of mobility, with a lower peak of 43% of the population not moving compared to 57.6% in Hatay. Kumlu shows a slight decrease in the FP for 0 to 10 km compared to the regular week, and after the first week, urban districts stabilize their mobility at levels below the regular weekly trend. The least affected urban districts are located in Gaziantep province.

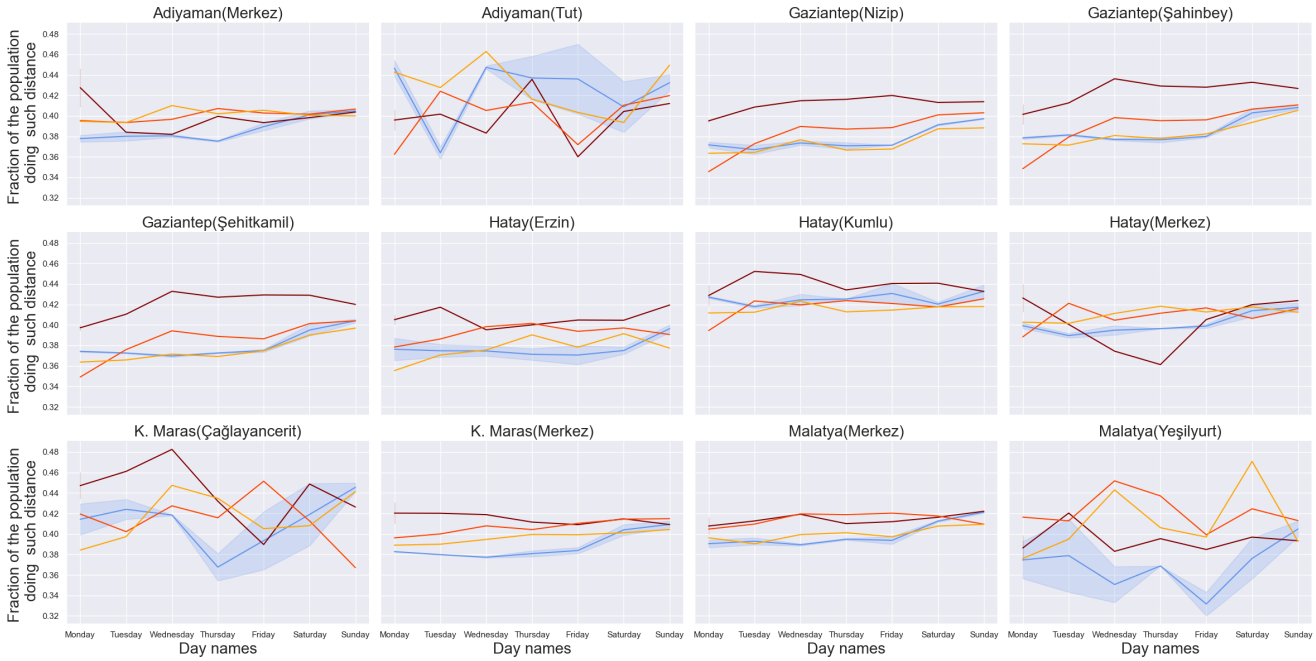
The disparities between rural and urban areas are also evident in the FP for distances between 10 to 100 km. With the exception of Tut, all rural districts show a mean decrease in the fraction of the population traveling this distance. There are no clear similarities between urban districts in this category, even when compared with the PGA. Urban districts in Matalya, Gaziantep, and Kahramanmaraş show lower FP for the 10 to 100 km category. However, these urban districts begin to stabilize in the second week (for Kahramanmaraş) and towards the end of the third week (for Gaziantep).

Matalya's FP remains elevated after the first week (4% compared to 2.8% before the earthquake). Hatay and Adıyaman's urban districts exhibit significant peaks in mobility. For example, on Thursday, Hatay's urban district sees a peak of 9.8% of the population traveling 10 to 100 km, compared to just 2.3% before the earthquake. Similarly, Adıyaman's urban district peaks at 7.1% on Wednesday, up from 2.0% before the event.

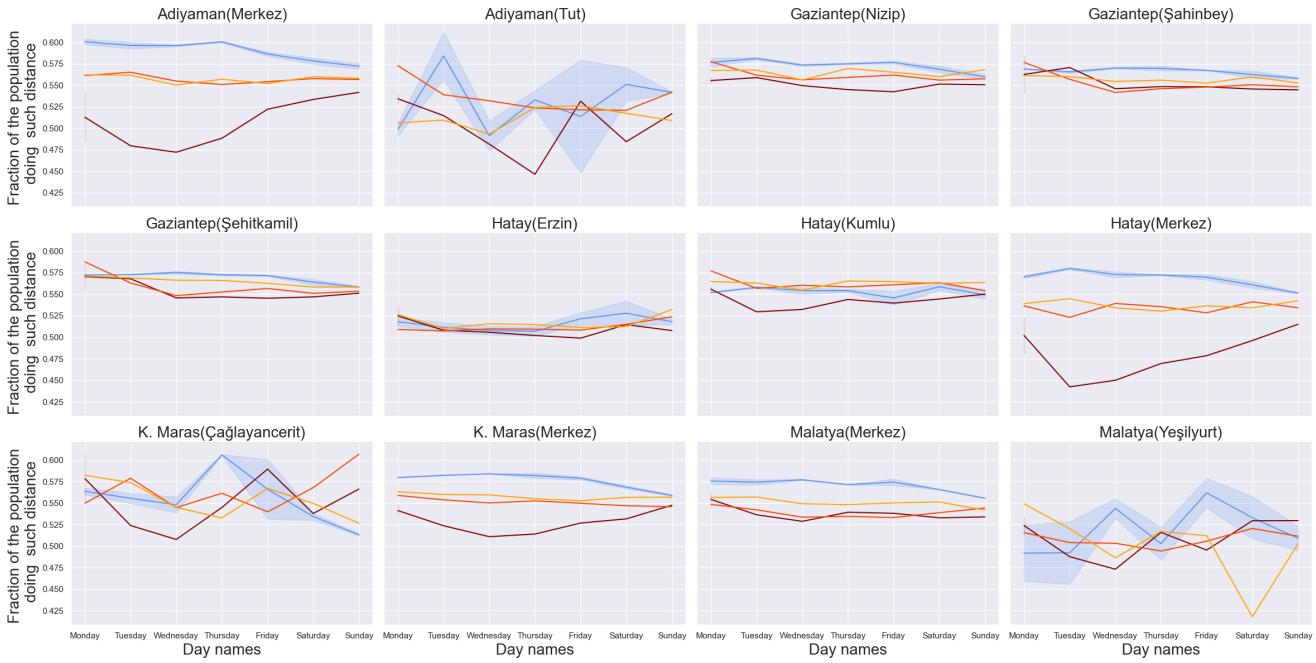
Urban areas are more significantly impacted in terms of people traveling over 100 km after the earthquake (except for Matalya's rural district). The first week shows the most significant changes, with high peaks in the fraction of the population traveling on Tuesday

in Adıyaman (8% compared to 0.3% before) and Hatay (12.8% compared to 0% before).

Matalya also shows a peak on Tuesday (2.8% compared to 0.1% before), and Kahramanmaraş experiences a peak on Wednesday (4.1% compared to 1.8% before). The rural districts of Adıyaman and Matalya experience two significant peaks on different days—Wednesday for both and Friday for Matalya, with Tut showing a peak on Saturday. After the first week, the FP for rural districts stabilizes, often returning to levels similar to the pre-earthquake week or showing an increase in the fraction of the population traveling over 100 km.



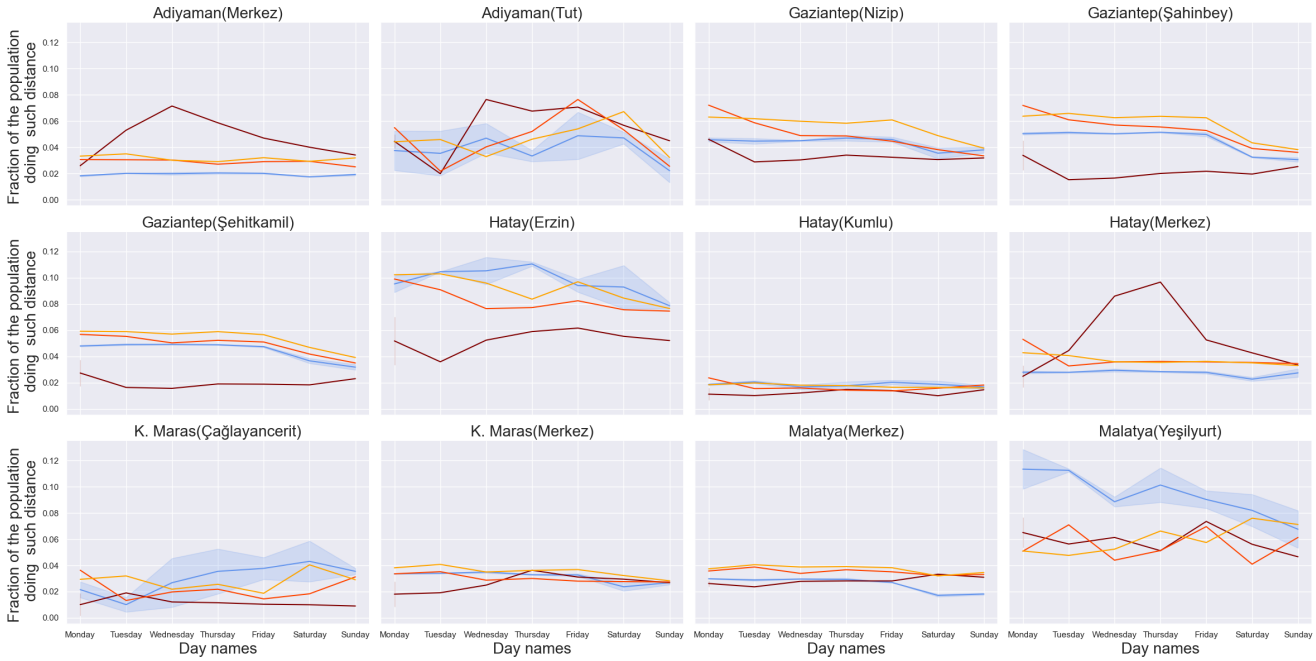
(a) Fraction of the population no moving



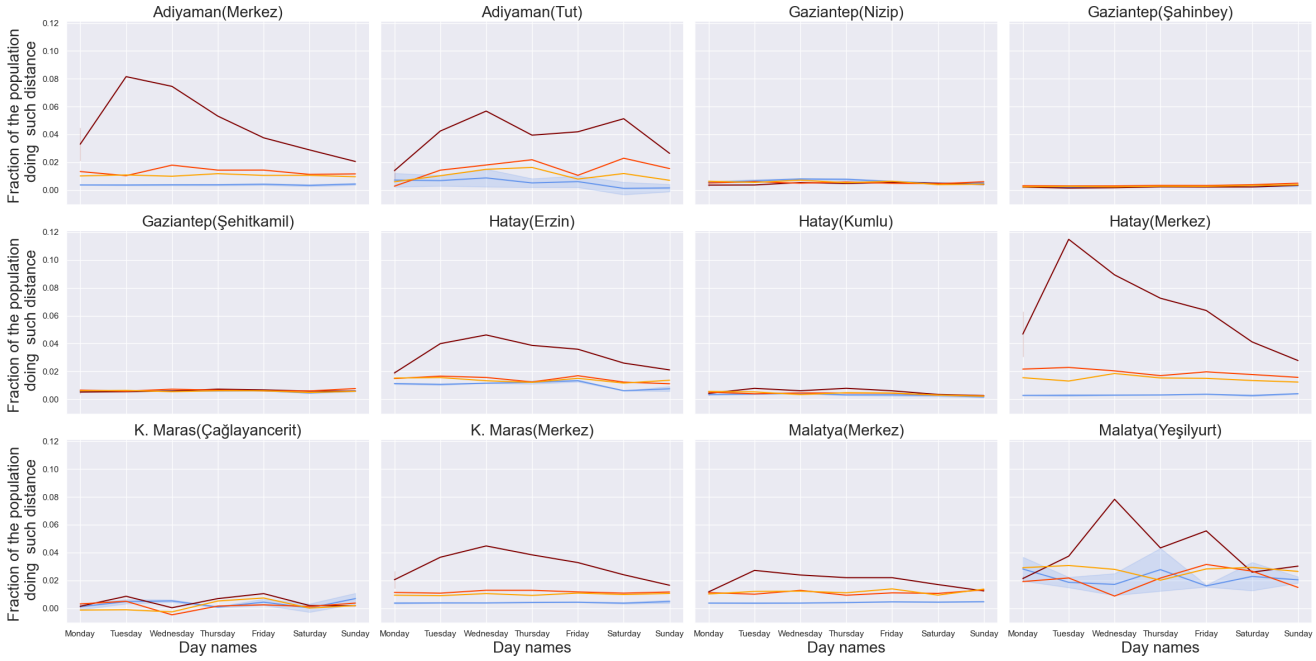
(b) Fraction of the population between 0 and 10km

(c) Legend

Figure 35: Time Series through days of the week and the three week (and during) the disaster for the no movement and 0 to 10 km distance category.



(a) Fraction of the population between 10 and 100km



(b) Fraction of the population from 100km and over

(c) Legend

Figure 36: Time Series through days of the week and the three week (and during) the disaster for the distance of 10 to 100 km and 100 km and over distance categories.

Summary

Rural and urban districts react differently. Urban areas seem to be less impacted in their mobility compared to the urban districts.

This part also informs that the PGA does not affect the proportion of the mobility.

However, two rural districts appears to be less impacted than the other rural districts.

VII.3 Interpretation

All of the results clearly demonstrate the significant impact of tectonic plate movement on population mobility. What is particularly interesting is that the effects of the earthquake were felt across the entire country, albeit with varying degrees of impact. Provinces that were not directly impacted by the earthquake showed fewer changes in mobility.

However, provinces close to, but not neighbouring, the affected areas might have experienced some effects, as people in these regions likely altered their daily routines in anticipation of potential aftershocks. This behaviour could be linked to a heightened sense of risk perception within these communities. Notably, the fact that fewer people in unaffected areas followed their usual mobility patterns in the third week suggests that concern and awareness about the earthquake may have diminished over time.

Neighboring provinces were more affected by the aftershocks (as shown in Figure 37), which likely led to further disruptions in mobility. The change in mobility on Monday of the third week, for example, coincides with an aftershock on February 27th (visible in Figure 21). This suggests that aftershocks continued to influence the behavior of the population, even in areas that were not directly impacted by the main earthquake event.

In general, the provinces that were not directly impacted by the earthquake did not show any significant increases in the fraction of the population travelling long distances, such as 100 km or more. However, neighbouring provinces did experience a slight increase in mobility within the 10 to 100 km range, likely due to efforts to assist those affected by the earthquake. Despite this, the overall trend for these provinces was a reduction in the fraction of the population travelling these distances, rather than an increase. This indicates that while there was some movement, it was not necessarily driven by a large-scale mobilization of resources or humanitarian aid, but rather by localized, smaller-scale movements of people seeking to help their neighbours.

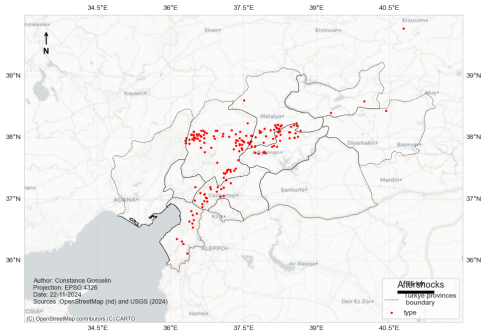


Figure 37: Aftershocks of the EQ event.

Provinces close to the epicenter are more likely to exhibit a stronger reaction to the earthquake, as shown in Figure 25 and Figure 26. These figures also highlight the significant role of aftershocks in influencing changes in population mobility. While Elazig and Malatya have relatively low PGA values, Figure 37 provides evidence of the substantial impact of aftershocks on mobility in these areas.

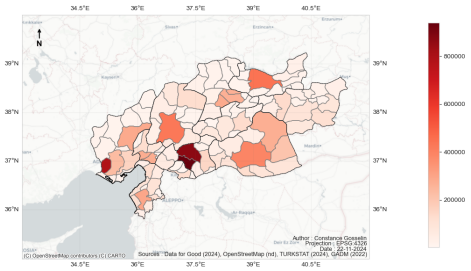
The data presented in Figure 24, Figure 25, and Figure 26 show notable similarities in terms of population displacement. In these provinces, a high proportion of the population is traveling more than 100 km, while at the same time, there is a significant fraction of the population remaining stationary. This raises the question: Are there inequalities within the provinces that could explain such mobility patterns? If the population in each province had equal access to resources and capabilities, one would expect similar mobility responses across regions. However, the contrasting mobility patterns may be influenced by the varying levels of PGA across the provinces. If this hypothesis holds, neighboring provinces should exhibit distinct mobility responses as well.

One potential factor influencing these disparities is the level of destruction and access to transportation infrastructure. Large road networks are concentrated in provinces such as Gaziantep, Sanliurfa, Diyarbakir, Adana, and Osmaniye (Figure 38). Despite these robust infrastructures, these provinces do not exhibit a notably higher fraction of the population traveling longer distances compared to others.

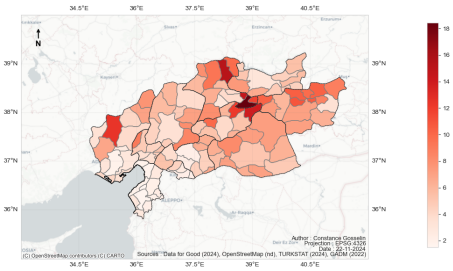
The increase in the population traveling over 100 km is documented by Yılmaz (2023) during the 24 hours preceding the first earthquake. Adana and Diyarbakir were identified as key hubs, which could explain the higher fraction of the population traveling between 10 and 100 km during the second and third weeks. According to Şenol Balaban et al. (2024), 72 hours after the first earthquake, evacuation efforts intensified, leading to a significant increase in the number of people moving on Wednesday. The subsequent reduction in population mobility could be attributed to the institutionalized displacement of affected individuals.

In the case of Kilis and Hatay, the disruption in their time series might be linked to increased migration from disaster-stricken areas in Syria. Although the MDM does not provide data from Syria, the proximity to Aleppo and the short distance to the border could have led to the displacement of Syrian refugees into the area (Dondaine, 2021). This influx of refugees, combined with the existing challenges faced by the local population, may have influenced the mobility patterns in these regions. In contrast, provinces like Sanliurfa and Gaziantep, which were less impacted by the earthquake, likely did not experience as significant a shift in mobility due to the limited need for Syrian refugees to move into these areas (Figure 37).

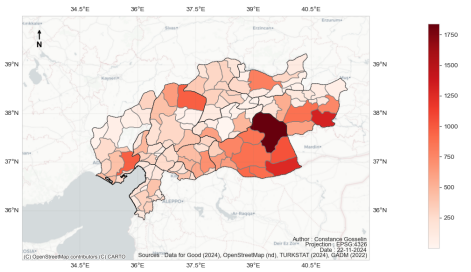
The level of destruction across different provinces does not seem to explain the disparities in mobility. Despite high levels of destruction in provinces such as Kilis, Adiyaman, and Hatay, there are no clear patterns of increased mobility that could be linked to these damages. Furthermore, poverty indicators, such as the illiteracy rate, do not appear to provide meaningful insights into the differences in population movement.



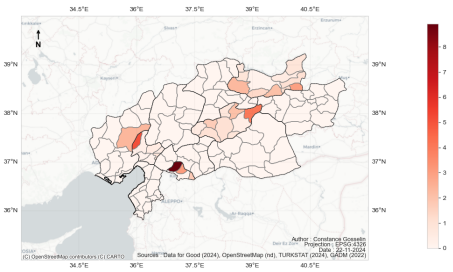
(a) Total of population



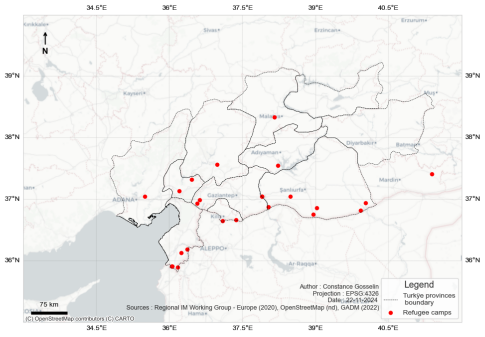
(b) Percentage of Illetrate



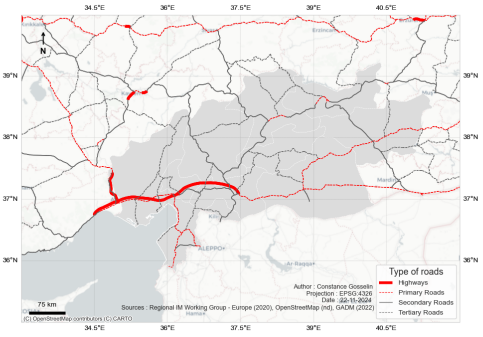
(c) Agriculture area



(d) Ratio of the total of population living in a city and in a rural area



(e) Refugee camps



(f) Ratio of the total of population living in a city and in a rural area

Figure 38: socio-demographic indicators of the districts impacted by the EQ.

As shown in the study of natural disasters, there are various noise factors at the provincial scale. For this analysis, focusing on the district level was essential. The correlations indicate a low positive relationship between PGA and mobility in some districts, suggesting that other factors are influencing mobility during the disaster. Additionally, the linear regression results show that, for this particular earthquake, PGA does not appear to correlate strongly with mobility, except for the movement of populations traveling over 100 km.

The heterogeneity observed in the northern part of the fault seems to be linked to aftershocks (Figure 29, Figure 31, Figure 32, and Figure 37). However, this explanation does not account for all the observed results. For example, districts in Elazig and parts of Adiyaman do not exhibit significant changes in mobility compared to the pre-event period. The Local

Moran's I indicator provides insights into the socio-demographic characteristics of districts in Sanliurfa (Figure 38, Figure 38). This province appears to be more rural and less affected by the earthquake (Figure 37), which could explain why mobility there was less impacted than in districts to the north of the fault.

The Local Moran's I indicator also reveals high-high clusters (areas with significantly higher mobility compared to surrounding districts) and low-low clusters (areas with lower mobility compared to neighboring districts) in Kilis. This could be related to the presence of refugee camps near the Syrian border. The outliers in the southern part of Hatay might also be linked to changes in population displacement caused by Syrian mobility, as refugees from Syria may have influenced the population dynamics in these districts (Dondaine, 2021). Note: This hypothesis is not statistically proven nor supported by the literature.

Elazig's districts often emerge as low spatial outliers (Figure 33, Figure 34, and Figure 30). However, like other districts, socio-demographic indicators do not provide clear insights into the spatial clusters or outliers. Some districts repeatedly show up as outliers and clusters, suggesting there may be underlying factors that have not been fully captured in the analysis.

The comparison between urban and rural areas reveals interesting patterns in mobility changes over time. Urban areas tend to exhibit smoother mobility patterns, with the population generally moving within a more predictable range. In contrast, rural districts show more variability in how populations react to the disaster. Adiyaman and Malatya's rural districts, for example, may be more affected due to their mountainous locations and the higher frequency of aftershocks in these regions.

Summary

The interpretation provides valuable insights into the mobility changes observed in the last three sections of this research.

The time series reveal two main patterns of response: self-evacuation by individuals and organized evacuations.

Additionally, it appears that districts north of the fault are affected by other earthquakes, contributing to the observed heterogeneity. This variability in mobility changes could be linked to the compounded effects of these events.

Currently, there is no visible evidence of the influence of socio-demographic factors on mobility changes. Further investigation into these factors may provide a more comprehensive understanding of the observed patterns.

VIII Discussion

This study has thoroughly validated and applied the MDM dataset in the context of the February 6 earthquake event, utilizing two distinct methodologies to assess its utility and reliability. These methodologies enable a comprehensive evaluation of the dataset by addressing several key aspects:

- The validity of the MDM dataset for practical use.
- Insights into daily population mobility patterns.
- Interactions between mobility data and significant national events.
- The dataset's sensitivity to disaster scenarios.

Furthermore, the case study highlights the practical applicability of the dataset in real-world contexts.

VIII.1 Validation of the *Movement Distribution Map*

The validation process employs advanced Big Data techniques, including visualization and statistical analyses. The findings demonstrate that the dataset captures meaningful mobility patterns, with 54% of the population travelling 0–10 km daily and 39% not moving at all. Notably, mobility patterns shift during weekends, starting from Friday.

Clustering of the local Moran index, which reveals spatial disparities. These observations align with previous studies (Arjona and Palomares, 2020; Carvalho et al., 2021; Huang and Wong, 2015). For instance, Arjona and Palomares (2020) identify differences in mobility between workdays and weekends, while Carvalho et al. (2021) highlight contrasting mobility patterns between locals and tourists, with locals tending to stay home on weekends.

Interestingly, the dataset also reflects higher mobility (over 100 km) along major highways, particularly from Istanbul to the capital, potentially capturing distinct behaviours among Facebook users.

The application of Laplace noise tests confirms the dataset's statistical robustness, though its sensitivity to the demographic characteristics of Facebook users introduces certain limitations. For instance, Meta (2022) note that the dataset requires at least ten Facebook users per district, meaning sparsely populated areas may experience greater noise disturbances. Nevertheless, the dataset's accuracy remains reliable, as demonstrated by its relative error. The predominance of users aged 45–55 years (Singh et al., 2019) is not impacting the noise. The total of the population in the district does.

Sensitivity tests further reveal notable variations in mobility. Workday versus weekend comparisons highlight consistent weekday mobility patterns, with reduced movement during weekends (Arjona and Palomares, 2020; Carvalho et al., 2021; Huang and Wong, 2015). Similarly, national holidays mirror these weekend mobility trends. Seasonal tests suggest increased movement over distances of 10–100 km during summer, consistent with holiday

travel patterns observed by Carvalho et al. (2021). Despite these variations, daily mobility for shorter distances (0–10 km) remains stable, reflecting continued work-related activities (Arjona and Palomares, 2020).

Disaster-event tests underscore the dataset's limitations in capturing mobility changes during disasters, particularly in understanding population vulnerabilities and responses. Differences in mobility across disasters may relate to varying levels of community preparedness and risk perception (Ozer, 2019; Yu et al., 2021; Schwaller et al., 2024).

VIII.2 Case Study: The February 6 Earthquake

The earthquake analysis reveals a profound impact on population mobility, characterized by both spatial and temporal disparities.

Nationwide, the earthquake disrupted daily mobility, with closer provinces experiencing prolonged disruptions compared to those farther from the epicentre. This raises an important question: how do high-intensity earthquakes disrupt national mobility, and what factors might mitigate such impacts? Current literature lacks comprehensive studies on these dynamics, underscoring the novelty of this research.

Within affected provinces, two primary responses emerged: evacuation and remaining in place. These responses align with disparities in vulnerability, as discussed by Yu et al. (2021); W Sharp and W Beadling (2013); Castillo Betancourt and Zickgraf (2024); Courtoy (9 01). Intra-provincial differences likely reflect underlying social, economic, and demographic inequalities.

District-level analyses highlight the role of aftershocks, which appear to significantly influence mobility patterns. Surprisingly, the primary ground acceleration (PGA) of the main shocks correlates weakly with mobility changes, suggesting that aftershocks warrant closer examination in future studies. Literature on aftershock-induced mobility remains scarce, presenting an opportunity for further investigation.

Spatial clustering and outliers in mobility responses also suggest links to community resilience and shared risk perceptions (Ozer, 2019; Yu et al., 2021). Temporal analyses further reveal that emergency evacuations are evident in MDM data, consistent with the complexities of post-disaster evacuation as noted in the literature (W Sharp and W Beadling, 2013; Castillo Betancourt and Zickgraf, 2024).

Rural versus urban analyses indicate distinct mobility patterns, with rural populations tending to remain stationary, while urban populations exhibit higher mobility, often evacuating. This aligns with differing risk perceptions and adaptive capacities (Karasözen et al., 2023; Karampotsis et al., 2024). High-density urban areas exhibit greater mobility, potentially reflecting the interplay between urban infrastructure limitations and vulnerabilities (Komatsuzaki et al., 2022; Aguilera and Villagra, 2023; Wang et al., 2024).

VIII.3 Perspectives and Limitations

VIII.3.a State of the Art

Big Data and human mobility are expansive fields requiring careful delineation. This study, focused on the MDM dataset within a disaster context, necessitated a synthesis of interdisciplinary themes. While the literature review provided a targeted exploration of Big Data and disaster-related mobility, it inevitably excluded broader discussions on migration, urban destruction, and emergency management. Future research could integrate these dimensions to offer a more holistic perspective.

The omission of traditional mobility studies unrelated to Big Data represents another limitation. Incorporating these studies could deepen insights into observed mobility trends and complement the dataset's findings.

VIII.3.b Validation of the MDM Dataset

The validation of the MDM dataset underscores both its utility and limitations. Unlike geolocated datasets, MDM lacks established validation methodologies, necessitating reliance on clustering, regression analyses, and descriptive statistics.

While effective, these methods could be enhanced by incorporating sensitivity tests that adjust Laplace noise levels. Unfortunately, such tests are currently infeasible due to insufficient information on user counts within districts. Advanced predictive modelling approaches, such as Prophet or AROMA, could provide additional insights into dataset disturbances and enhance its applicability for climate-change-related studies.

Knowing such impacts from test can also give an overview of the daily displacements. This could be a part of other type of studies more based to local scale like, peri-urban area habitat of mobility compared to people living in the city.

VIII.3.c Case Study

The case study demonstrates the MDM dataset's potential to quantify population mobility changes during disasters. Its ability to provide granular insights into mobility patterns offers a significant advantage over traditional datasets. However, its inability to track destination locations remains a notable limitation.

Moreover, the study highlights the importance of integrating Big Data with socio-demographic indicators to uncover non-visible factors influencing mobility. Future research should prioritize collecting detailed data at finer spatial scales to reduce Modifiable Areal Unit Problem errors. This could involve combining MDM data with other geo-databases and applying methodologies like geographically weighted regressions to better understand mobility dynamics.

Furthermore, this master thesis is showing the importance of being able to do interdisciplinarity. The use of Big Data in order to collect data on population's mobilities is necessary to understand the layers of displacements of people. However, there are limits.

The main limit is to not be able to collect enough data in order to understand the reason of the characteristics of mobility. As an example here, it was not possible to find data at the district's scale on business, economic loss, destructions, ects... This part shows the need of statistical relevance with socio-demographic indicators or social research methodology. These could have help to highlight non visible aspects.

Finally, longitudinal studies examining population "return" trends could provide valuable insights into recovery patterns. For instance, comparing January 2023 and January 2024 mobility data could reveal the long-term impacts of the earthquake, accounting for seasonal and exceptional variations.

IX Conclusion

This study offers valuable insights into the utility of the MDM dataset in analyzing mobility patterns, particularly in the context of natural disasters. While Big Data and social media datasets have been extensively studied, MDM represents a relatively untapped resource. It provides unique advantages, such as the ability to examine mobility on finer spatial and temporal scales, including daily variations.

Throughout the validation process, the MDM dataset demonstrated robust accuracy but exhibited sensitivity to noise, primarily influenced by the volume of users active on the Facebook platform. Notably, the dataset revealed discernible differences in mobility patterns between regular days and test scenarios, underscoring its sensitivity to national-scale population movements. The observed variability among districts further highlights the influence of socio-demographic factors on mobility responses. Compared to other Big Data studies, MDM delivers comparable insights, such as higher mobility during workdays (Arjona and Palomares, 2020; Carvalho et al., 2021; Huang and Wong, 2015).

In the context of natural disasters, the findings underscore the significance of disaster type and noise levels in interpreting mobility changes. Smaller-scale events, such as wildfires, or those with limited impact on the population, present challenges in detecting mobility shifts. Despite these limitations, the study successfully illustrated MDM's capacity to quantify mobility dynamics during significant events, capturing impacts of evacuation protocols, aftershocks, and district-level typologies.

The analysis also revealed critical insights into population behaviour following disasters. Contrary to the initial hypothesis that populations would remain stationary immediately post-disaster, the data indicated two distinct reactions: individuals either stayed within hazard zones or relocated swiftly to safer areas. These behavioural patterns likely reflect varying risk perceptions, vulnerabilities, and capacities for mobility.

While the study demonstrates the relevance of MDM in disaster mobility research, several limitations remain. First, the validation process could be strengthened through model-based approaches and by testing noise resilience with increased Laplace coefficients. Additionally, the analysis focused solely on mobility distances, without capturing the underlying motivations, purposes, or final destinations of movements. Incorporating these dimensions would significantly enhance the interpretive power of MDM analyses.

Despite these limitations, this research provides a strong foundation for leveraging MDM in disaster mobility studies. Future investigations could extend this approach to diverse types of natural disasters, urban studies, or broader explorations of community mobility habits. The dataset's potential to uncover nuanced patterns of human behavior in response to varying contexts offers promising avenues for advancing mobility research.

References

United Nations Office for the Coordination of Humanitarian Affairs (2023). Humanitarian transition overview - türkiye earthquake response (august 2023). https://www.unocha.org/publications/report/turkiye/humanitarian-transition-overview-turkiye-earthquake-response-august-2023. Accessed: 18-12-2024.

Aguilera, N. and Villagra, P. (2023). Contrastes multidimensionales y territoriales en resiliencia comunitaria ante el desastre entre zonas urbanas y rurales de la comuna de corral. *Revista de Urbanismo*, 49:66–92.

Alvarez-Dionisi, L. E. (2016). Toward grasping the dynamic concept of big data. *International journal of information technology and computer science*, 8(7):8–15.

American Association of Collegiate Registrars and Admissions Officers (2023). Emergent news. https://www.aacrao.org/edge/emergent-news/meb-2023-2024-academic-year-calendar. Accessed : 21-12-2024.

Arjona, J. O. and Palomares, J. C. G. (2020). Spatio-temporal mobility and twitter: 3d visualisation of mobility flows. *Journal of Maps*, 16(1):153–160.

Bao, J. (2023). Multidomain big data modeling: Concepts and applications. In *Proceedings of the 2023 International Conference on Image, Algorithms and Artificial Intelligence (ICIAAI 2023)*, pages 419–427. Atlantis Press.

Borko Furht, F. V. (2016). *Big Data Technologies and Applications*. Cham : Springer International Publishing AG.

British Red Cross (2024). The road to recovery is long. https://www.redcross.org.uk/stories/disasters-and-emergencies/world/turkey-syria-earthquake. Accessed: 17-11-2024.

Carvalho, A. M., Ferreira, M. C., and Dias, T. G. (2021). Understanding mobility patterns and user activities from geo-tagged social networks data. *Transportation Research Procedia*, 52:493–500. 23rd EURO Working Group on Transportation Meeting, EWGT 2020, 16-18 September 2020, Paphos, Cyprus.

Castillo Betancourt, T. and Zickgraf, C. (2024). It's not just about women: broadening perspectives in gendered environmental mobilities research. *Climate and Development*.

Centre for Research on the Epidemiology of Disasters (2024). Learn more about em-dat. https://doc.emdat.be/. Accessed: 19-11-2024.

Chen, S.-H. and Yu, T. (2018). *Big Data in Computational Social Sciences and Humanities: An Introduction*, pages 1–25. Springer International Publishing, Cham.

Chuan Liao, Daniel Brown, D. F. X. L. D. C. S. C. (2018). Big data-enabled social sensing in spatial analysis: Potentials and pitfalls. *Transactions in GIS*, 22(1):1–21.

Chun, Y., Kwan, M.-P., and Griffith, D. A. (2019). Uncertainty and context in giscience and geography: challenges in the era of geospatial big data. *International journal of geographical information science*, 33(6):1131–1134.

Courtoy, M. (2022-09-01). "to leave is to die": States' use of mobility in anticipation of land uninhabitability.

Demir Aydin, Celebi Erkan, O. H. O. Z. O. A. B. E. S. S. S. F. Z. A. E. D. Y. Z. U. M. M. N. (2024). Destructive impact of successive high magnitude earthquakes occurred in türkiye's kahramanmaraş on february 6, 2023. *Bulletin of Earthquake Engineering*, 608.

DEMIRALP, S. (2023). The economic impact of the turkish earthquakes and policy options.

Dilek Yildiz, Tugba Adali, C. (2023). La république turque a 100 ans. qu'en est-il de sa population ? *Population et Sociétés*, 608.

Dondaine, M. (2021). Camp de rÉfugiÉ-e-s de ÖncÜpinar. page 21. L'observatoire des camps de réfugiés.

ESPG (nd). `https://epsg.io/?q=Turkey`. Accessed : 21-12-2024.

European Commission (nd). Big data. `https://digital-strategy.ec.europa.eu/en/policies/big-data`. Accessed: 17-11-2024.

Feiertagskalender (nd). Archiv - schulferien 2023. `https://www.feiertagskalender.ch/ferien.php?geo=3539&jahr=2023`. Accessed : 21-12-2024.

Feizizadeh, B., Omarzadeh, D., and Blaschke, T. (2024). Spatiotemporal mapping of urban trade and shopping patterns: A geospatial big data approach. *International Journal of Applied Earth Observation and Geoinformation*, 128:103764.

for Migration (IOM), I. O. (2018). Implementation of the workplan of the task force on displacement under the warsaw international mechanism for loss and damage (wim). Technical Report Pillar II: Policy – International/Regional Activity II.2, United Nations Framework Convention on Climate Change (UNFCCC). Analysis Report, August 2018.

Ginzarly, M., Pereira Roders, A., and Teller, J. (2018). Mapping historic urban landscape values through social media. *Journal of Cultural Heritage*.

Global Administrative Area organisation (2022). Gadm data. `https://gadm.org/data.html`. Accessed: 19-11-2024.

Han, J., Zheng, Z., Lu, X.-Z., Chen, K.-Y., and Lin, J.-R. (2024). Enhanced earthquake impact analysis based on social media texts via large language model. *International Journal of Disaster Risk Reduction*, 109:104574.

Heydari, S., Huang, Z., Hiraoka, T., de León Chávez, A. P., Ala-Nissila, T., Leskelä, L., Kivelä, M., and Saramäki, J. (2023). Estimating inter-regional mobility during disruption: Comparing and combining different data sources. *Travel Behaviour and Society*, 31:93–105.

Hrehova, S. (2018). Brief overview of the concept of big data.

Huang, Q. and Wong, D. W. S. (2015). Modeling and visualizing regular human mobility patterns with uncertainty: An example using twitter data. *Annals of the Association of American Geographers*, 105(6):1179–1197.

Humanitarian OpenstreetMap Team (nd). Humanitarian openstreetmap team (hot). `https://www.hotosm.org/`. Accessed : 21-12-2024.

Huo, Z., YANG, X., LIU, X., and YAN, X. (2024). Spatio-temporal analysis on online designated driving based on empirical data. *Transportation Research Part A: Policy and Practice*, 183:104047.

Internal Displacement Monitoring Center (nd). Idmc data portal. `https://www.internal-displacement.org/database/displacement-data/`. Accessed : 21-12-2024.

Ivan, I., Singleton, A., Horák, J., and Inspektor, T. (2017). *The Rise of Big Spatial Data*. Lecture Notes in Geoinformation and Cartography. Springer International Publishing, Cham, 1st ed. 2017. edition.

Karampotsis, E., Kioskli, K., Tsirimpa, A., and et al. (2024). Understanding evacuation behavior for effective disaster preparedness: a hybrid machine learning approach. *Natural Hazards*, 120:13627–13665.

Karasözen, E., Büyükakpınar, P., Ertuncay, D., Havazlı, E., and Oral, E. (2023). A call from early-career turkish scientists: seismic resilience is only feasible with "earthquake culture". *Seismica*, 2(3).

Komatsuzaki, N., Otsuyama, K., and Hiroi, U. (2022). How the choice of temporary housing impacts on widespread displacement after large-scale flooding? a disaster recovery simulation in tokyo metropolitan area. *International Journal of Disaster Risk Reduction*, 81:103243.

Lam, N. S. N., Meyer, M., Reams, M., Yang, S., Lee, K., Zou, L., Mihunov, V., Wang, K., Kirby, R., and Cai, H. (2023). Improving social media use for disaster resilience: challenges and strategies. *International Journal of Digital Earth*, 16(1):3023–3044.

Li, H., Han, Y., Wang, X., and Li, Z. (2024). Risk perception and resilience assessment of flood disasters based on social media big data. *International Journal of Disaster Risk Reduction*, 101:104249.

Li Cai, Y. Z. (2015). The challenges of data quality and data quality assessment in the big data era. *Data Science Journal*, 14(2):1–10.

Longley, P. A., Adnan, M., and Lansley, G. (2015). The geotemporal demographics of twitter usage. *Environment and Planning A: Economy and Space*, 47(2):465–484.

Macnish, K. and Galliott, J. (2020). *Big data and democracy*. Edinburgh scholarship online. Edinburgh University Press, Edinburgh.

Maddalena Favaretto, Eva De Clercq, C. O. S. B. S. E. (2020). What is your definition of big data? researchers' understanding of the phenomenon of the decade. *PLoS ONE*, 15(2).

Mark Graham, T. S. (2013). Geography and the future of big data, big data and the future of geography. *Dialogues in Human Geography*, 2(3):255–261.

McKitrick, M., Schuurman, N., and Crooks, V. (2023). Collecting, analyzing, and visualizing location-based social media data: review of methods in gis-social media analysis. *GeoJournal*, 88:1035–1057. Accepted 08 December 2021, Published 19 January 2022, Issue Date February 2023.

Meta (2022). *Movement Distribution*.

Meta (2024). À propos de data for good chez meta. https://dataforgood.facebook.com/dfg/about. Accessed: 18-11-2024.

Ministre de la culture et du tourisme de Turquie (2024). Jours fériés. https://turquietourisme.ktb.gov.tr/FR-135775/jours-fris.html. Accessed: 19-11-2024.

Muniz-Rodriguez, K., Ofori, S. K., Bayliss, L. C., Schwind, J. S., Diallo, K., Liu, M., Yin, J., Chowell, G., and Fung, I. C. (2020). Social media use in emergency response to natural disasters: A systematic review with a public health perspective. *Disaster Medicine and Public Health Preparedness*, 14(1):139–149.

Ozer, P. (2019). Les risques naturels. In *Risques, planification d'urgence et gestion de crise*. La Charte.

Poullet, Y. (2020). *La propriété des données: balade au pays des merveilles " à l'heure du Big Data "*, pages 191–204. Dalloz.

Riahi, Y. and Riahi, S. (2018). Big data and big data analytics: concepts, types and technologies. *International journal of research and engineering*, 5(9):524–528.

Rob Kitchin, G. M. (2016). What makes big data, big data? exploring the ontological characteristics of 26 datasets. *Big data society*, 3(1):1–10.

Sarı, H., Özel, M., Akkoç, M. F., and Şen, A. (2023). First-week analysis after the turkey earthquakes: Demographic and clinical outcomes of victims. *Prehospital and Disaster Medicine*, 38(3):294–300.

Schwaller, N. L., El-Khattabi, A. R., and Nguyen, M. T. (2024). Networks matter: Examining migration networks using cellphone data following hurricane maria. *International Journal of Disaster Risk Reduction*, 100:104139.

Shukla, S., Kukade, V., and Mujawar, S. (2015). Big data: Concept, handling and challenges: An overview. *International journal of computer applications*, 114(11):6–9.

Singh, A., Halgamuge, M. N., and Moses, B. (2019). An analysis of demographic and behavior trends using social media: Facebook, twitter, and instagram. In *Social Network Analytics*, pages 87–108. Elsevier.

Tayfur, , Bayramoğlu, B., Şimşek, P., and Gunduz, A. (2024). Medical response to the february 6, 2023, earthquakes in hatay: Challenges faced in the deadliest disaster in the history of türkiye. *Disaster Medicine and Public Health Preparedness*, 18:e45.

Tiwari, N. K. (2019). Big data analytics: A concept, challenges and research issue. *International journal for research in applied science and engineering technology*, 7(10):476–478.

Turkish statistical institue (nda). Central dissemination system. https://biruni.tuik.gov.tr/medas/?kn=130&locale=en. Accessed: 31-12-2024.

Turkish statistical institue (ndb). Popular statistics. https://www.tuik.gov.tr/. Accessed : 05-12-2024.

United Nations (2022). Handbook on management and organization of national statistical systems. https://unstats.un.org/capacity-development/handbook/html/topic.htm#t=Handbook%2FC14%2FReview_of_changes_since_last_edition_and_current_trends.htm%23XREF_57804__14_2_7_Big_Data114. Accessed: 17-11-2024.

United Nations Office for the Coordination of Humanitarian Affairs (nd). Humanitarian data exchange. `https://data.humdata.org/`. Accessed : 21-12-2024.

United States Geological Survey (nd). 2.1. philosophy of estimating and interpolating ground motions. `https://cbworden.github.io/shakemap/manual4_0/tg_philosophy.html`. Accessed: 31-12-2024.

Venkatram, K. and Geetha, M. A. (2017). Review on big data analytics – concepts, philosophy, process and applications. *Cybernetics and Information Technologies*, 17(2):3–27.

W Sharp, T. and W Beadling, C. (2013). 18 - disasters, complex emergencies, and population displacement. In Magill, A. J., Hill, D. R., Solomon, T., and Ryan, E. T., editors, *Hunter's Tropical Medicine and Emerging Infectious Disease (Ninth Edition)*, pages 148–156. W.B. Saunders, London, ninth edition edition.

Wang, F., Tian, J., Shi, C., Ling, J., Chen, Z., and Xu, Z. (2024). A multi-stage quantitative resilience analysis and optimization framework considering dynamic decisions for urban infrastructure systems. *Reliability Engineering System Safety*, 243:109851.

Watson, H. J. (2014). Tutorial: Big data analytics: Concepts, technologies, and applications. *Communications of the Association for Information Systems*, 34(1):1247–1268.

Wisner, B., Gaillard, J., and Kelman, I., editors (2012). *Handbook of Hazards and Disaster Risk Reduction*. Routledge, 1st ed. edition.

Wu, F. and Ma, W. (2022). Clustering analysis of the spatio-temporal on-street parking occupancy data: A case study in hong kong. *Sustainability*, 14(13):7957.

Xiaoyao Han, Oskar Josef Gstrein, V. A. (2024). When we talk about big data, what do we really mean? toward a more precise definition of big data. *Frontiers in Big Data*, 7:1441869.

Yan, Z., Guo, X., Zhao, Z., and Tang, L. (2024). Achieving fine-grained urban flood perception and spatio-temporal evolution analysis based on social media. *Sustainable Cities and Society*, 101:105077.

Yu, J., Castellani, K., Forysinski, K., and et al. (2021). Geospatial indicators of exposure, sensitivity, and adaptive capacity to assess neighbourhood variation in vulnerability to climate change-related health hazards. *Environmental Health*, 20.

Yılmaz, S., K. O. Y. S. M. E. S. E. D. O. A. O. M. Y. K. A. F. G. E. M. E. B. . A. H. (2023). Emergency medicine association of turkey disaster committee summary of field observations of february 6th kahramanmaraş earthquakes. *Prehospital and disaster medicine*, 38(3):415–418.

Şenol Balaban, M., Doğulu, C., Akdede, N., and et al. (2024). Emergency response, and community impact after february 6, 2023 kahramanmaraş pazarcık and elbistan earthquakes: reconnaissance findings and observations on affected region in türkiye. *Bulletin of Earthquake Engineering*.