

## RESEARCH

# A novel activity space approach to discover displacement patterns via mobile phone data: An analysis of the 2023 Türkiye-Syria Earthquakes

Bilgeçağ Aydoğdu<sup>1\*</sup>, Didem Danış<sup>2</sup>, Özge Bilgili<sup>3</sup>, Cemil Yıldızcan<sup>4</sup>, Semiha Nur Yağcıklı<sup>5</sup>, Subhi Güneş<sup>5</sup> and Albert Ali Salah<sup>1</sup>

\*Correspondence:

[b.aydogdu@uu.nl](mailto:b.aydogdu@uu.nl)

<sup>1</sup>Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands

Full list of author information is available at the end of the article

## Abstract

This study introduces the novel Activity Space Approach (ASA) for measuring disaster-induced displacement patterns using mobile call detail records (CDR), where we explore shifting the focus in displacement detection from home locations to habitual living spaces. We apply our method to analyze the February 2023 Türkiye-Syria earthquakes, which affected over 14 million Turkish citizens and 1.7 million Syrian refugees within Türkiye. Using anonymized and hourly aggregated CDR data from 127,700 individuals, complemented with insights from qualitative fieldwork conducted in the regions affected by the earthquakes, we show that the proposed approach overcomes the main limitations of traditional home location methods and provides more granular spatial insights into displacement patterns. By incorporating measurements of urbanization and infrastructure damage, we illustrate how post-disaster mobility shows variation among locals and refugees, given their pre-existing socioeconomic vulnerabilities and unequal capacities to respond. Our findings demonstrate that, while Turkish citizens were able to evacuate more swiftly, Syrian refugees experienced slower and more spatially constrained displacements, often toward institutional settings such as camps, reflecting legal precarity and constrained mobility options. This comparative perspective underscores the importance of recognizing and mapping variations in displacement experiences across different population segments. Consequently, ASA can inform more targeted short-term policies and support more inclusive long-term recovery planning.

**Keywords:** Mobile phone data; Call Detail Records; Earthquake; Displacement; Türkiye; Syria

## 1 Introduction

Natural disasters like earthquakes often cause various forms of mobility, such as evacuation, displacement, and resettlement, which unfold differently across time and space. In its essence, the ability to react to a natural disaster by moving elsewhere to seek refuge and safety can be framed as an immediate coping strategy or a form of adaptation [1]. However, how individuals experience this form of mobility depends on contextual factors as well as their diverse capacities and resources (or the lack thereof) – social, cultural, material, or political. In fact, when a crisis hits an area, pre-existing vulnerabilities, due to unequal distribution of resources and access to rights and services, may come to the fore among different segments of the population. These vulnerabilities may significantly influence how individuals

and communities respond to disaster and how resilient they are in the post-crisis period. Consequently, it is to be expected that individuals and communities will experience different types of displacement, often leading to divergent outcomes for affected groups in the short to long term.

February 2023 earthquakes, with magnitudes of 7.8 and 7.5, caused extensive destruction in the southeastern Türkiye and northern Syria, profoundly altering the lives of millions. Within Türkiye, the earthquakes impacted both the local resident population and the Syrian refugee community. Official data from the Ministry of Environment, Urbanization and Climate Change (MoEUCC) revealed that in the 11 most affected provinces, around 676,000 residential buildings and 115,000 workplaces sustained moderate to severe damage, with more than half of these located in the epicentral provinces of Hatay and Kahramanmaraş [2]. According to 2022 population data, these provinces were home to more than 14 million individuals registered in the Address-Based Population Registration System (ABPRS), while an additional 1.7 million Syrians under temporary protection were also residing in the region, as Presidency of Migration Management (PMM) reported shortly before the disaster (February 2, 2023) [3]. This means that 16.43 % of Türkiye's population and nearly half (49.64 %) of the country's temporary protection population were concentrated in the affected area.

The unprecedented scale of destruction, combined with the diverse demographic structure of the region, shaped by its sociocultural and ethnic plurality, and socio-economic disparities, has created tremendously complex conditions for displacement in the aftermath of the disaster. These conditions were especially challenging for Syrian refugees whose experiences were likely to be aggravated due to pre-existing intersectional vulnerabilities, including their more precarious legal, social and economic status. In this context, a systematic analysis of post-disaster displacement and a comparison between the local population and the Syrian refugee population become warranted.

Mobile phone data (MPD) can be used for real-time assessment and monitoring of disaster-induced displacements [4, 5, 6]. However, its potential for understanding displacement nuances in relation to pre-disaster vulnerabilities remains understudied. This study employs a novel approach to measure disaster-induced displacements using mobile call detail records (CDR) data. Mainstream approaches for measuring displacement using mobile data in the literature assign home locations to individuals, and detect shifts in those locations [7, 8, 9, 10]. Our proposed Activity Space Approach (ASA) uses shifts in activity spaces before and after the disaster. It calculates the origins and destinations not as single locations, but as spatial distributions, which enables detection of origin and destination hotspots that are not possible to identify with the standard home location-based methods. We make the code for ASA publicly available.

We use the 2023 Türkiye-Syria Earthquake as a case study to show how the proposed method can identify granular spatial and temporal displacement patterns that reveal distinct mobility outcomes in Türkiye among local and refugee populations following the earthquake. Namely, our findings indicate that internally displaced Syrian refugees exhibited more constrained displacement patterns concentrated around pre-existing Temporary Accommodation Centers (TACs), which

are often referred to as refugee camps. Moreover, compared to the local population, Syrian refugees were able to evacuate the region more slowly. Finally, our results also suggest that Syrian refugees, who were displaced to other cities, were drawn into neighborhoods with high Syrian populations, exhibiting reliance on their social networks. On the other hand, Turkish people were displaced faster, and largely to rural areas.

## 2 Measuring displacements with mobile phone data

Displacements refer to involuntary movements of individuals from their habitual living spaces to safer areas due to disruptions caused by disasters or conflicts [11]. These movements are characterized by spatial attributes (origin, destination, distance), temporal attributes (start date, end date, duration), and demographic attributes (age, gender, ethnicity) of the displaced people (DPs). The definition of displacement varies based on context. Cross-border displacements follow legal frameworks for refugees and asylum seekers, while internal displacements often involve needs-based frameworks. Internal displacements are typically measured across administrative boundaries. Internal Displacement Monitoring Centre (IDMC) has one of the most comprehensive databases on disaster induced-displacements [1]. IDMC follows event-based tracking to calculate the number of displaced people for each disaster event [12]. Their calculations rely on data triangulation through multiple sources, including government agencies, United Nations (UN) agencies, as well as other international and non-governmental organizations. Despite great efforts in event-based data collection, data gaps remain a big issue. It is difficult to estimate the number of voluntarily evacuated people, and the duration of tracking remains limited with little known regarding returns, local integration, and relocations[12]. Mobile phone data (MPD) offers a unique opportunity to enhance the measurements and provide more comprehensive displacement data potentially closing some of the data gaps in disaster-induced migration and mobility.

Despite a growing body of work using MPD as a tool for monitoring and managing the disaster-induced mobility [13], how to measure displacements remains an open question. The statistics collected on migrants tend to focus on two aspects of mobility; flows and stocks, respectively [14]. Flow indicators show movements between certain points within a time frame, whereas stocks are snapshot metrics that capture the number or density of the population in an area. MDP can be used to compute indicators of stocks and flows, as well as provide features of individual-level mobility.

There are two approaches for measuring the flows via MPD; the first is to show all types of movements between regions within a time period, and the second is to detect DPs first and then show their movements in an aggregated fashion. In the context of sudden-onset disasters, the individuals in affected areas tend to evacuate in large numbers. In such a context, it is possible to measure the exodus of people without detecting displacements at the individual level but by contrasting aggregated measures [4, 5, 8, 15]. The aggregated flow indicators during disasters are often compared to a baseline level to understand the displacements [4, 5, 6]. Despite their

---

<sup>[1]</sup>The database on internal displacements can be accessed at: <https://www.internal-displacement.org/database/displacement-data/>

simplicity, such comparisons can provide important insights on the crisis-induced population movements, such as the magnitude of effects of the disaster on departures and/or mobility disruption across time and space, and predictability of the movements.

While these approaches can be useful to describe the crisis, questions on the impact of economic inequalities, other socioeconomic disadvantages, and pre-existing vulnerabilities on people's post-disaster mobility require more nuanced approaches to measure displacements at individual level. This is especially crucial when analyzing distinct mobility profiles between different population groups, such as refugees and local populations, where pre-existing social vulnerabilities may lead to divergent displacement experiences.

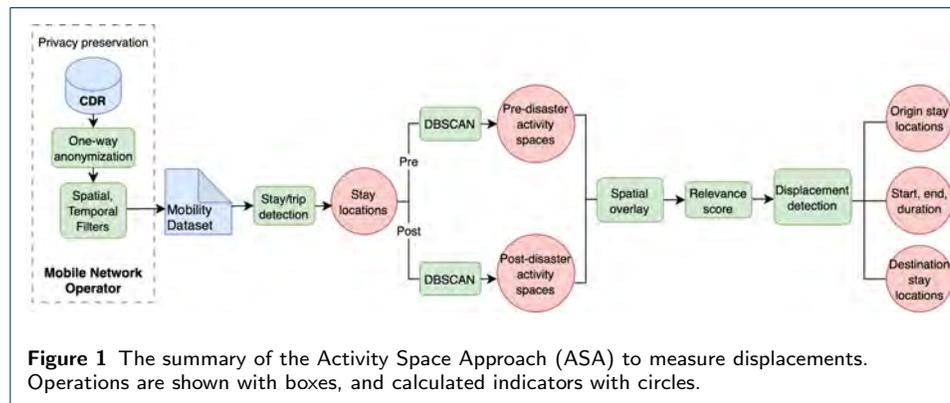
The measurement of displacements at individual level using mobile data often relies on changes in residential areas, which we call as home location-based approaches (HLA) [7, 8, 9, 10]. HLA have different ways of calculating the shifts in home locations, which we explain later in model comparisons. While HLA can reasonably estimate the number of displaced persons between administrative boundaries using CDR data, they fail to fully capture the complexity of displacement patterns.

HLA have two significant limitations. First, they overlook the broader spatial context of people's lives. Home locations alone provide no information about other places of high importance, such as residences of friends and family, which become crucial during disasters [16]. Furthermore, people's routines typically span areas beyond their homes—a phenomenon conceptualized as “activity space” in sociology and geography [17, 18]. Disasters displace people not only from their residential areas but also from these activity spaces, a dimension that HLA neglect. Second, data sparsity presents a considerable challenge to existing home location detection methods, whether they rely on the most frequent signals throughout the entire day/week, areas with concentrated signals during specific time frames (e.g., nighttime), or signals concentrated during particular days (e.g., weekends) [19]. Home location algorithms filter the limited data further and reduce the available information. One solution is to replace the home location with the modal location, i.e. the location where the individual is most frequently observed [7, 10]. However, this approach still reduces the complex spatial behavior to a single location, ignores routine and/or socially important mobility behavior. When the cell level information is used, the noise around home locations causes problems with respect to understanding the spatial extent of the displacements. HLA typically mitigate these challenges by measuring and aggregating displacements at administrative boundary levels rather than cell tower levels, which serves two purposes; (1) it distributes the spatial uncertainty around residences to larger spatial areas, (2) it facilitates the comparison with official sources.

The decision to migrate involves a complex set of considerations, this process can be even more complicated when one has to leave their place of residence due to a natural disaster and its aftershocks [20]. Moving elsewhere as a response to a natural disaster and seeking safety abstains individuals from time to prepare and to mobilize resources. In most cases, individuals will be less likely to ‘choose’ a destination, but seek asylum in their vicinity. However, depending on conditions related to the disaster (e.g. physical characteristics of the disaster, response capacity and exposure), the level of urgency to leave may show variation. In the context

of earthquake this may for example include the level of destruction caused by the disaster, which is not the same for everyone; some lose their entire homes and livelihoods, while others can continue their lives despite some damage. If preparation is at all possible, the decision for a destination can be shaped by the availability of social networks elsewhere and by perceived economic opportunities. Traditionally, migration models like the gravity and radiation models incorporate push and pull factors after spatial aggregation to explain migration patterns between regions [21]. However, individual level vulnerabilities before the disaster interact with these push and pull factors; some population segments have less capacity to leave the disaster areas due to their lower socioeconomic background [22, 23]. Spatially and demographically aggregated analyses of displacements cannot fully capture these subtle differences in push and pull factors that show how different populations cope with disasters in different ways.

### 3 Methodology



#### 3.1 Activity Space Approach

Our activity space approach (ASA) measures displacements by relying on the United Nations (UN) definition of the internal displacements in the guiding principles on displacements [24], as significant ruptures from the habitual living spaces in the aftermath of a disaster or to avoid effects of an armed conflict. In this sense, the displacements are not necessarily defined solely by the shift in home locations, but more by the post-disaster shifts in areas where the individuals spent time consistently prior to the disaster. To be able to measure the displacements on individual CDR records, we have developed ASA, which is based on the estimation of the shifts in habitual living spaces, works better when the data sparsity is a problem (which is becoming more concerning with CDR), is able to detect displacements at shorter distances, and enables more granular analyses of the displacements. Illustrated as a flowchart in Figure 1, it has four main components:

1. Measurement of stay locations
2. Determination of activity spaces
3. Computing the relevance of stay locations
4. Displacement detection

In the first component, we process CDR data into stay locations—geographical regions where individuals spend significant time. ASA distinguishes between people in transit and those lingering in specific locations. It helps to differentiate signals that are coming from people who are on the move, from people who are spending their time in a certain area. In the second step, the activity spaces of the individuals before and after the disaster are computed. This step ensures that the commonly visited areas within certain distance to one another are clustered together to represent the individuals’ use of space.

The third component measures how familiar the post-disaster activity spaces are based on their relevance scores. We measure the nighttime presence in each pre-disaster activity space, and by computing each activity space’s relative nighttime weight before the disaster. Post-disaster activity spaces are then assigned a relevance based on their spatial overlap with pre-disaster activity spaces, weighted by the relevance score of those overlapped areas.

Lastly, in the fourth component, we detect shifts in people’s stay locations based on their relevance scores falling under certain threshold and we classify everyone either as displaced or non-displaced using the migrant detection algorithm developed by Chi et al. [25]. In Section 4, we compare the measurements of ASA to that of the TMB method [10] and show how using stay locations and activity spaces can improve the measurement at close distances, enabling granular analysis of displacement patterns.

## 3.2 Data sources

### 3.2.1 CDR data

The CDR data were collected in collaboration with Turkcell Technology (TTECH) between January 1 and March 15, 2023, across Türkiye [2]. The dataset includes information on users’ outgoing call records, specifically the antenna used and the call timestamps. The CDR data sets are carefully anonymized and spatio-temporally aggregated in line with CDR processing norms established in data challenges like Data for Development (D4D) [26] and Data for Refugees (D4R) [27]. The dataset is specifically collected to enable analysis of fine-grained mobility patterns in the aftermath of the earthquake. Additionally, we incorporated demographic flags to differentiate between Syrian and Turkish users [28], following a methodology similar to the previous D4R Challenge. These nationality flags are based on registration information collected by the company; individuals who subscribed using a temporary protection card issued to Syrian refugees or a Syrian passport are flagged as Syrian in the CDR data. The precise cell tower locations are deemed sensitive information by the mobile network operator (MNO), so a small amount of noise was added to the locations within the service area of each tower.

### 3.2.2 Population sampling and biases

The CDR data comprise a small sample of the total Syrian and Turkish population living in Türkiye; we have 74,902 Turkish and 60,000 Syrian individuals in our sample, as opposed to 14,013,196 Turkish and 1,738,035 Syrian individuals in the region according to the official figures. We followed a stratified sampling technique

---

<sup>[2]</sup>See Supplementary Material for more detail about the CDR data.

while sampling users in Türkiye. For each city, we sampled considering the share of Syrian population living in the city [3] in proportion to the total Syrian population in the country. Then, we added an equal number of randomly selected Turkish users in each city for enabling a comparative study [28]. In other words, we sampled more users, Syrian or Turkish, in cities with larger Syrian populations. The different representativeness of the sample for the two populations will be important in the analysis.

At any given time, the population in a given area consists of residents, as well as a mixture of temporary residents, workers, tourists, and transit visitors [29]. During disasters, these groups will also be impacted. However, from a displacement perspective, we are mainly interested in the residents of the earthquake region, i.e., the population that was consistently present in the region before the earthquake occurred. During processing, we add a filter to exclude the people who spent less than 90% of their time in the earthquake region before the earthquake happened. After the filtering, we end up with 70,123 Turkish and 57,586 Syrians. This way, we ensure that the detected displacement patterns are not from temporary residents, tourists, or transit visitors going back to their homes after the disaster.

### 3.2.3 Other data sources

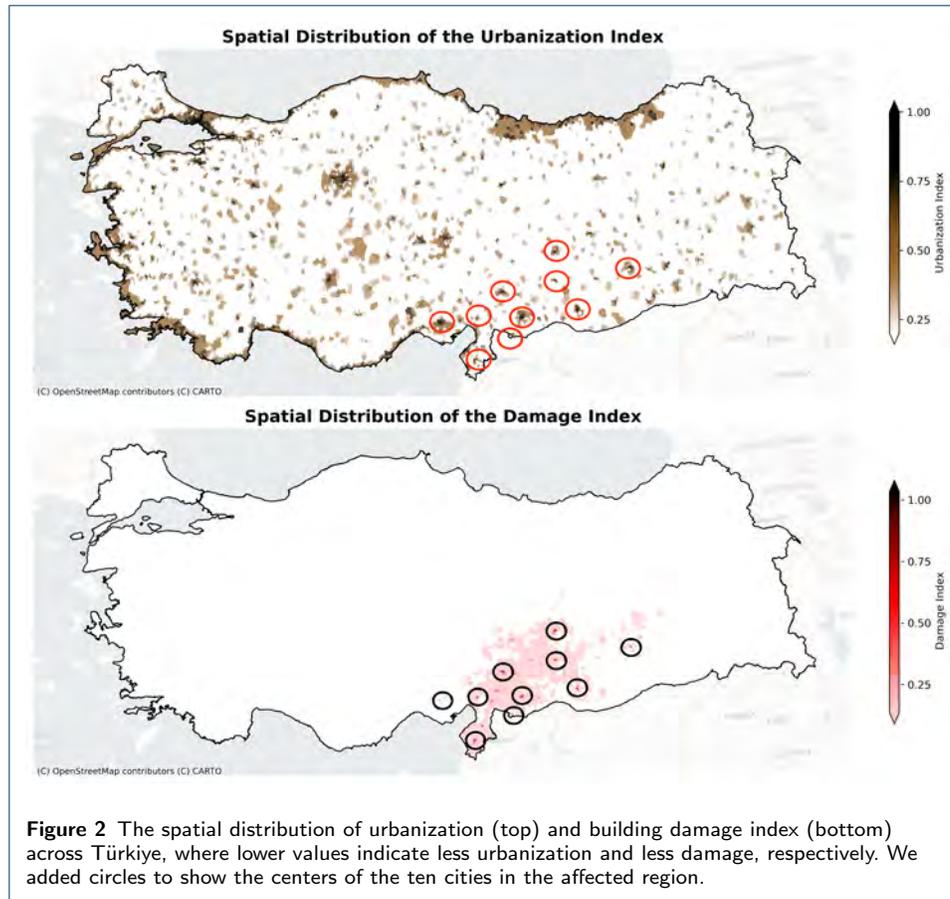
In our analysis of the displacement patterns, we further use two indices, one to indicate the amount of urbanization, and another for indicating the amount of damage [4]. The MNO internally categorizes urbanization in different categories as “densely urban”, “urban”, “suburban”, and “rural”. We base our index on this classification. The damage indicator data were collected and processed by various volunteer organizations within a month after the earthquake [30] using the official data shared by the ministry of environment and urban planning [5]. The data show damage levels for around 210,000 buildings in the earthquake area, but not all buildings are included. The buildings are classified into four damage categories: “collapsed buildings,” “buildings that need to be demolished,” “heavily damaged buildings,” and “slightly damaged buildings.” These categories represent distinct conditions rather than a strictly linear progression of damage severity. To quantify these categories, we assigned values to distinct damage categories reflecting their severity per building, and summed number of buildings per cell tower weighted by their damage level. We make both indices available to the reader online, at a  $10km \times 10km$  grid resolution. Further information including a detailed visual reconnaissance report is published by Dilsiz et al. [31]. We show the spatial distribution of both indices in Figure 2 (bottom), where darker colors indicate more intense infrastructure damage or higher level of urbanization. We also highlighted the major city centers in the earthquake region with circles.

To provide additional context to the indices and the patterns observed in the CDR-based analysis, we also incorporated on-the-ground qualitative insights. Two

[3]Data on the distribution of Syrian population across Türkiye is published, and regularly updated by the Presidency of Migration Management. See: <https://en.goc.gov.tr/>

[4]See Supplementary Material for more detail about how the damage and urbanization indices are calculated.

[5]The damage level of buildings were queried at: [hasar.cbs.gov.tr](https://hasar.cbs.gov.tr).



of the authors, a sociologist and a political scientist, conducted fieldwork across earthquake-affected provinces on three separate occasions: February 23–27, 2023 (two weeks after the disaster); June 10–15, 2023; and April 29–May 4, 2024. These field visits included semi-structured interviews with displaced individuals in cities such as Maraş, Gaziantep, and Hatay; in informal resettlement zones on the urban peripheries of Hatay; and participant observation in formal refugee camps such as Altınözü (Boynuyoğun) and Hilalkent. These qualitative observations were instrumental in contextualizing the spatial and temporal patterns identified in the ASA analysis—such as the clustering of Syrian displaced persons (DPs) around refugee camps and the differing displacement trajectories and return rhythms between Turkish citizens and Syrian refugees. Importantly, field-based insights underscored the relevance of distinguishing between citizens and refugees when analyzing post-disaster mobility, as their legal status, access to aid, and integration into national recovery systems varied considerably. This distinction also illuminated the specific vulnerabilities faced by refugee communities, including legal precarity, limited access to transport and housing, and heightened exposure to anti-refugee sentiment [32, 33, 34]. Such factors not only constrained their mobility but also deepened existing social inequalities in the aftermath of the disaster.

### 3.3 Measurement of stay locations

We use a parametric approach based on CDR data for describing the stay locations of individuals. Two thresholds, one spatial and one temporal, are used to control and adapt the algorithm to the particular context. This is similar to the algorithm proposed by [35], which however uses GPS coordinates for determining stay points. We define a “stay location” as the area covered by cell towers within the specified spatial threshold to which a person’s device connects for longer than a specified temporal threshold. We share the pseudocode for the stay detection algorithm, and describe its use, in the Supplementary Material.

In the literature, the location of a person is often associated with the Voronoi cell of the connected cell tower when processing CDR. This refers to the polygon that contains all points for which the corresponding cell tower is the closest one. However, the service area of a cell tower can be small (e.g., in city centers) or large (e.g., rural areas) depending on the population density. In dense areas, a person can receive signals from different towers despite being in the same location. Furthermore, infrequent calls by a person will result in a data sparsity problem in the CDR. Using stay locations instead of the cell tower footprints helps us deal with both data sparsity and noise in cell tower assignments.

We determine the area of a stay location by the convex hull of the Voronoi cells used, which is the smallest convex shape that fully encloses all the polygons. (See Figures S1 and S2 in the Supplementary Material for the stay location area calculations in rural and urban areas, respectively). Lastly, we assume that the individual stays in the last calculated stay location, unless a new one is calculated. Individuals with a low signal frequency from a single location are not considered as displaced by the algorithm.

### 3.4 Activity spaces

Assessing displacement is frequently achieved by determining home locations of people and detecting changes in these. Home location-based approaches (HLA) focus on the shifts in the district of the residence, yet the areas where people spend their time tend to be larger than their residences. “Activity space” is a widely accepted concept in sociology and geography and stipulates that the space in which people spend their time tends to be larger than their residential areas due to their work, routines, and social connections [17, 18]. The way people interact with their environment shows not only their personal values and preferences, but also their socioeconomic position and ethnic background [36, 37]. Incorporating activity spaces is an important component for our displacement measurement approach, as we expect loss of livelihood in the aftermath of an earthquake to result in substantial shifts in activity spaces.

There are different established approaches to measure activity spaces of individuals using MPD [38, 39], and new approaches are still emerging [40, 41]. Many studies first identify the so-called “anchor points,” such as home, work, and other significant locations, where the individuals tend to spend most of their time [38, 39]. The second step is to measure the spatial extent of anchor points by fitting ellipse-shaped regions, or by using minimum convex hulls (MCH) [40]. MCHs are easier to compute and fully enclose all the stay location polygons, but tend to overestimate the activity spaces compared to ellipse-based approaches.

In our method, we do not define work and home locations as anchor points, but partially account for them, as our stay locations identify stationary areas, which are expected to include residential and work places. Due to the way they are calculated, stay locations tend to overlap with each other (see Figures S1 and S2 in the Supplementary Material for examples). If an individual leaves and returns to the same location, it is recorded as two different stay locations. To discover activity spaces from stay locations, we use the DBSCAN clustering algorithm [42], using the coordinates of the centroids of each stay location. DBSCAN performs density estimation and clustering using two hyperparameters; the maximum distance between clusters and the minimum number of points per cluster, respectively. For the purposes of our method, we fixed the minimum cluster size to a single point so that no stay location is treated as noise. As a result of DBSCAN clustering, each stay location is assigned to an activity space.

As stay location areas are calculated using a Voronoi tessellation, we prefer the MCH approach to calculate the areas of the activity spaces as well. Ellipse-based approaches create the ellipses based on the exact cell tower locations, and do not incorporate their actual service areas. We discuss how to set the spatial parameters of ASA in the Supplementary Material, as the spatial threshold and maximum distance are important to ensure that the habitual areas are not artificially enlarged. Figures S3 and S4 there illustrate how activity spaces are calculated given a set of stay locations and maximum distance parameters and show that our approach focuses on geographic areas, where the service coverage of the cell towers plays a significant role in defining the boundaries of those areas.

Our proposed approach differs from the standard method of calculating activity spaces with mobile data by anchoring them to work and home locations [38, 39]. The distance parameter may determine that home and work locations belong to two different activity spaces, potentially classifying work locations as “unfamiliar” places due to low relevance scores. Even if this occurs, the fact that people begin spending nights in areas previously designated as work areas could signal a form of displacement.

### 3.5 Relevance of the stay locations

The destinations people relocate to after disasters can often be predicted based on their social connections or previously visited locations [16]. Nighttime visits are used to determine the residential areas of individuals, but they can also reveal close social connections. We define nighttime as the time between 10 PM and 7 AM. These hours are chosen as limited as possible to ensure that the observed patterns are either residential or related to close social connections. Specifically, for the 2023 earthquake that we study, nighttime leisure activities are very limited in the areas hit by the earthquake, compared to metropolitan areas in Türkiye, such as Istanbul or Ankara. If an individual relocates to an area where they maintain close social connections or where they have second residences, we do not want ASA to classify the movements as displacement. Therefore, we introduce the concept of a relevance score for each activity space that quantifies familiarity with the visited area according to pre-earthquake nighttime visit frequency.

$$R_{post,k} = \sum_j \left( R_{pre,j} \times \frac{A_{pre,j} \cap A_{post,k}}{A_{pre,j}} \right) \quad (1)$$

In Eq. 1,  $j$  represents a pre-disaster activity space,  $R_{pre,j}$  is the relevance score for pre-disaster activity space  $j$ ,  $A_{pre,j}$  is the area of pre-disaster activity space  $j$ , and  $A_{post,k}$  is the area of post-disaster activity space  $k$ . The calculated  $R_{post,k}$  is also assigned to all clustered stay locations in the post-disaster activity space  $k$ .

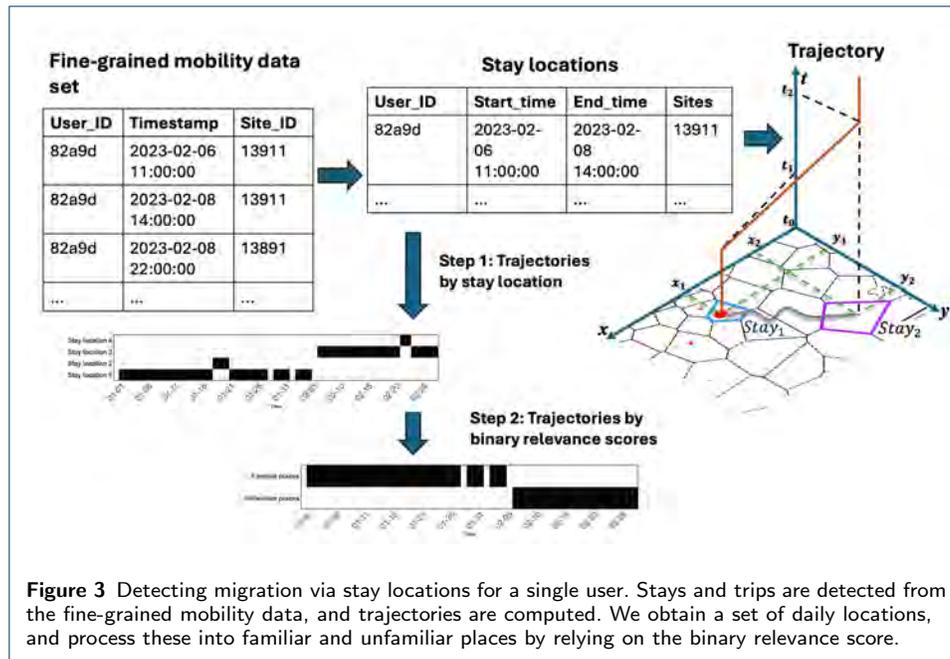
For clarity of explanations, we determine whether an activity space  $k$  is relevant or not as a Boolean variable, calculated as:

$$R_k = \begin{cases} 1, & \text{if } R_{post,k} \geq R_{threshold} \\ 0, & \text{if } R_{post,k} < R_{threshold} \end{cases} \quad (2)$$

The relevance score threshold  $R_{threshold}$  is the final hyperparameter of the ASA. This parameter controls the overlap between the pre- and post-disaster activity spaces  $A_{pre,j}$  and  $A_{post,k}$ .

Our proposed approach relies on detecting maintained drops in the binary relevance scores of stay locations after the disaster. This means that the individual has been displaced to a new set of stay locations. For this reason, each stay location of the individual will inherit the binary relevance value of its activity space cluster, as determined by DBSCAN.

### 3.6 Displacement detection



We detect displacements with a migration detection algorithm that works on a daily time series of location identifiers, which can represent various geographic units

ranging from administrative boundaries to individual cell towers [25], but adapt it to detect disaster-induced displacements. This algorithm detects continuous segments of locations for each individual, allowing for gaps and deviations, to distinguish genuine residential movements from temporary travels, while accommodating data sparsity. In Figure 3, we give an overview of the data processing for calculating the trajectories, which are the input for our displacement detection algorithm (see the Supplementary Material for more detail).

The key difference in detecting displacements and voluntary migration on MPD is related to the definitions. While there is no universally agreed definition of migration [43], the voluntary migration is often more clearly defined temporally and spatially than involuntary migration. For instance, according to the United Nations (UN) [44] a movement between two countries is considered as migration if the person spends at least 12 months at the destination, and as short-term migration if the time spent is between 3 and 12 months. Similar temporal thresholds are employed for collecting statistics on internal migration by National Statistical Offices (NSOs), although specific thresholds vary by country [45]. Clear spatial and temporal thresholds for defining what migration constitutes align well with the algorithm's segment-based approach. Involuntary movements as a result of disasters can be more sudden, irregular, include multiple moves, and may not follow the clear change between two stable locations that the algorithm expects to see. In addition, due to the crisis context definitions of displacements do not necessarily employ temporal or spatial thresholds except the distinction between cross-border and internal movements. We relied on the definition of displacements as loss of habitual spaces, and loss of familiarity, instead of spatial or temporal thresholds. Thus in ASA, we detect displacements on two types of segments; familiar and unfamiliar locations.

Given a time series of stay locations, the migration detection algorithm will detect the migration events, origins, and destinations, as well as an estimate of the date of migration. Its most important hyperparameters are the minimum number of days stayed at a location to retain a segment, which is denoted by  $k$ , and the maximum gap allowed between consecutive observations to consider them within the same location segment, denoted by  $\epsilon$ . We focus only on the migration events that start after the disaster, within the disaster affected region. When administrative boundaries are used as the location identifier, the algorithm detects the origin and destination as single locations. Since we use the binary relevance score as the location identifier, the origins refer to all stay locations with relevance scores above the relevance threshold, and the destinations are those below it. The idea behind this is that the displacements, unlike regular migration events, may be associated with loss of livelihood, which reflects on the decrease in relevance scores of stay locations, thereby indicating substantial shifts in activity spaces.

### 3.7 Model comparison

Home location-based approaches (HLA) are the most popular methods for detecting displacements on individual mobile records. HLA measure displacements through shifts in residential areas. There are different methods used for identifying shifts in home locations, including frequency-based approaches [7, 8, 9], and more recently, segment-based approaches [10]. Frequency-based approaches assign

users to home locations (cell towers or administrative areas) in varying time frames (weekly, monthly, yearly), and then compare consecutive periods to see if the individual has shifted their home location. If there is a shift, the origin and destination often refer to administrative areas (neighborhood, district, city) where individuals are most frequently found to reside before and after the disaster, respectively. On the other hand, segment-based approaches are designed explicitly to accommodate noisy mobility data, and use segments to detect the shifts. They define the origins and destinations by contiguous segments before and after the migration event, respectively. There are different ways to represent the main residential area; it can be the daily modal location (i.e. the area where the individual receives most of their signals) [10, 7], or the location where most signals are received during nighttime or weekends.

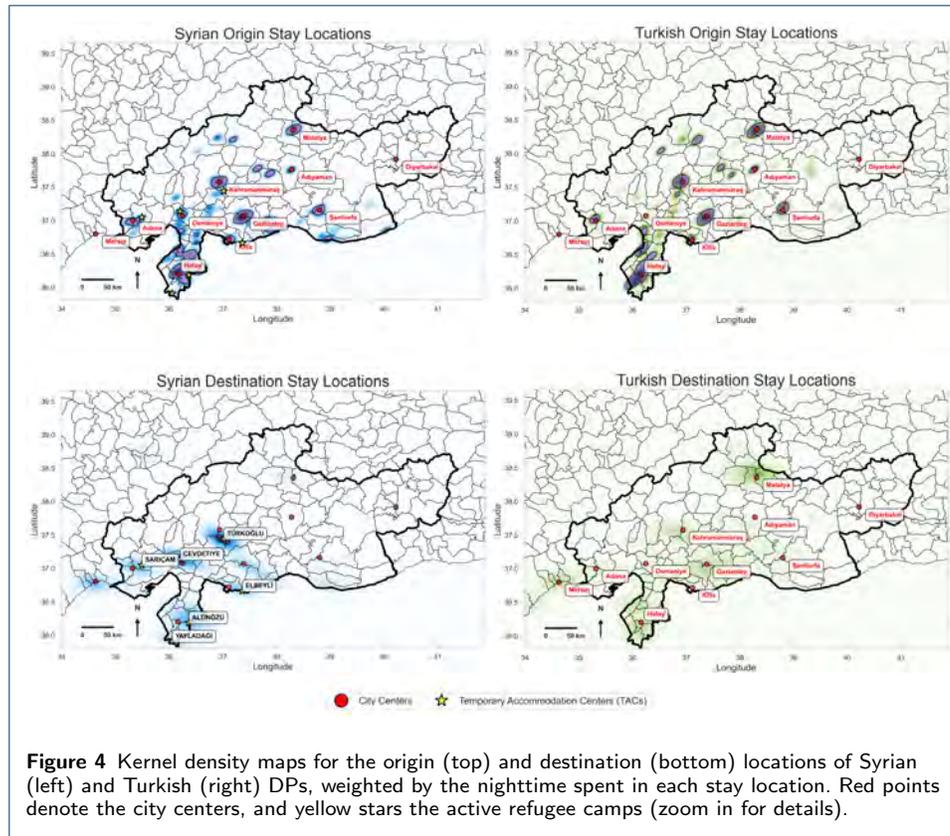
In ASA, we adopt the segment-based approach to measure displacements, but we focus on shifts from familiar activity spaces to unfamiliar ones, rather shifts between residential areas. HLA filters for home locations to trace residential movements, whereas ASA analyzes mobility patterns as a whole, through stay locations. This enables ASA to have a more individualized spatial representation of origin and destination areas, capturing not only residential areas, but also how individuals use the space more broadly, including work, leisure, and other important places. In other words, for ASA, all familiar locations before the disaster constitute the origin, and all unfamiliar locations after the disaster are the destinations. Consequently, ASA uses a segment-based approach to measure displacements, and proposes a novel way of spatially representing origins and destinations.

To evaluate ASA, we compare it to HLA proposed by Chi and others [25], which was applied in TMB [10]. TMB is successful in measuring displacements, and it can help establishing causal links between push factors such as armed violence and migration. However, it is difficult to understand what types of shifts are occurring for DP's beyond their residences. As we demonstrate with examples in Section 4, seeing such shifts at high spatial resolution can give additional insights on different experiences of displacements. We present additional comparisons with a simpler frequency-based HLA in the Supplementary Material.

## 4 Experimental results

### 4.1 Origins and destinations

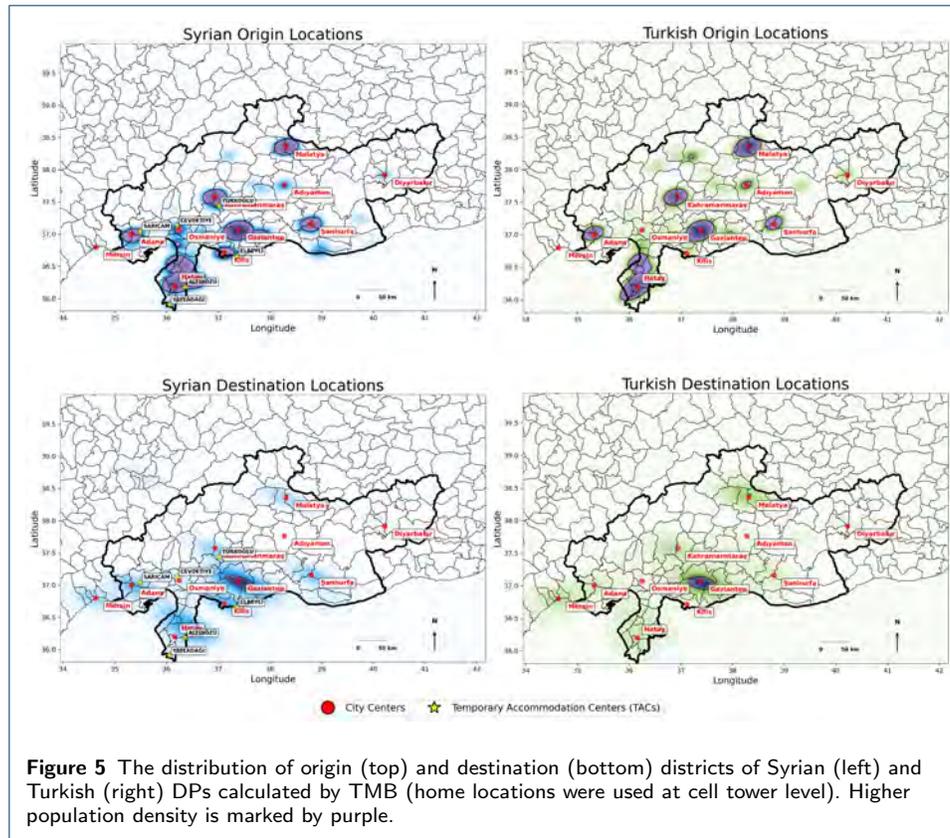
Using the ASA, we detected displacements and obtained the origins, the destinations, and the timings of the displacements. We chose the spatiotemporal thresholds as 2km and 2h, DBSCAN distance threshold at 5km and relevance score at 5%. We selected stay location thresholds to balance spatial granularity while reducing noise. In our data set, the median distance between the cell towers within the earthquake area is around 3km. By choosing a spatial threshold at 2km, we effectively cluster the towers at city centers together. We chose a 2h temporal threshold, because the CDR data is aggregated into hourly bins, following earlier mobile dataset preparation efforts [26, 27]. In the migration detection algorithm, we fix the parameter controlling the number of days needed to be spent at a location to consider it a location segment to 14 days. A sensitivity analysis of selected parameters is conducted and reported in Section 4.2.



**Figure 4** Kernel density maps for the origin (top) and destination (bottom) locations of Syrian (left) and Turkish (right) DPs, weighted by the nighttime spent in each stay location. Red points denote the city centers, and yellow stars the active refugee camps (zoom in for details).

In ASA, the origins and the destinations are calculated as distributions over space representing overall mobility of individuals. In Figure 4, we show the spatial density of the origin and destination stay locations for Syrians and Turkish people separately in the earthquake affected regions of Türkiye, using kernel density estimation [46, 47]. These are weighted by the nighttime spent in each stay location to emphasize the nighttime durations. Weighing by the night time slightly increases the density around refugee camps, but the patterns remain similar without the weights. The overall density of the stays gets reduced in the aftermath of the disaster due to the fact that many DPs left their home areas. Depending on spatiotemporal thresholds, increasing the size of stay locations lowers the density of these patterns, but the main visual insights remain very similar.

In TMB, origin and destination each refer to a single cell tower, but they can be analyzed at district level as well (see Figure S9 in Supplementary Material). In Figure 5, we show the distribution of the origins and the destinations as calculated by TMB at the cell tower level. Both ASA and TMB detect similar distributions of origin and destination areas in the earthquake region. Notably, ASA detected around 10,500 DPs, whereas TMB detected 3,700 DPs at cell tower level. TMB identifies less displaced people at cell tower level, as the segment-based algorithm struggles more to identify contiguous segments in granular space representations (see the Supplementary Material for details). Both Figure 4 and 5 show that the origin locations of both displaced Syrian and Turkish populations were largely concentrated in city centers, which are marked by the red points on the maps. Figure 4



uses the density of stay locations instead of the density of residences, so it is able to show the origins in great detail, highlighting areas with frequent visits that get lost in simple residency-based measurements.

In Hatay, which was particularly impacted by the earthquake, the distribution of origin stay locations is spatially more extensive compared to other regions (see Figure 4). Both Figure 4 and 5 show that some people were displaced to unaffected cities close to the earthquake area, such as Mersin. The destination distributions of Syrians calculated using ASA shows that there was an increased population density around active Temporary Accommodation Centers (TACs), which are marked by yellow stars on the maps.<sup>[6]</sup> TACs are commonly referred to as “camps” by locals and they were previously housing Syrian refugees under Temporary Protection. We could not find the origins and destinations at this level of detail, neither with TMB, nor with the frequency-based approaches<sup>[7]</sup>.

#### 4.2 Sensitivity analysis

We investigated the effect of different hyperparameter settings on the proposed method and the TMB baseline. We firmly believe that in complex settings, such

<sup>[6]</sup>For the locations of TACs, we used the data found at: <https://data.humdata.org/dataset/turkey-refugee-camps>, Accessed June 10, 2025. We validated the locations of the TACs, but had to manually correct the location of the Türkođlu camp, which was incorrect.

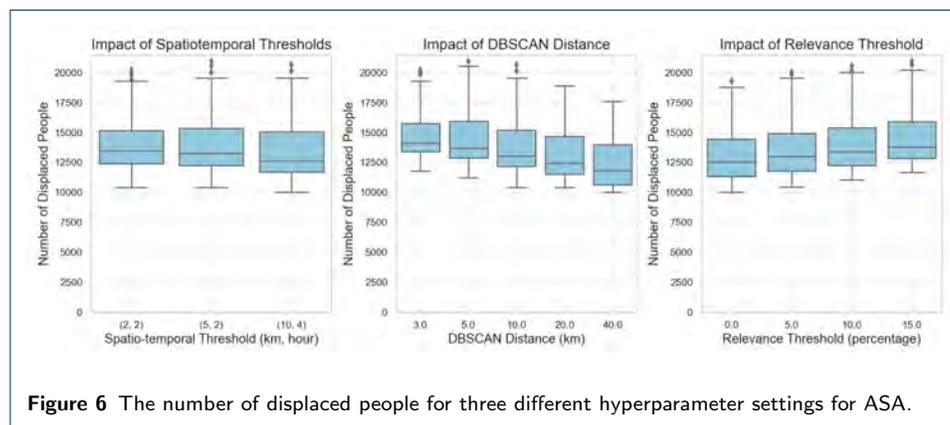
<sup>[7]</sup>See the Supplementary Material for origin and destination maps based on frequency-based approach and TMB.

as post-disaster mobility, there is no single parameter set that will work uniformly well. Instead, setting different parameters will function like changing the zoom rate of a microscope, and one should investigate broadly in collaboration with domain experts, informed by further qualitative approaches.

We start by assessing the number of displaced people detected in ASA and TMB approaches. We test three spatiotemporal threshold settings for the ASA: 2km/2h; 5km/2h; and 10km/4h, respectively. Then for each of the stay location measurements, we investigate the relevance score levels of {0%, 5%, 10%, 15%}. The DBSCAN distances are chosen separately for each spatiotemporal threshold; {3km, 5km, 10km} for 2km/2h, {5km, 10km, 20km} for 5km/2h and {10km, 20km, 40km} for 10km/4h. The minimum sample parameter of DBSCAN is fixed to one sample, as explained previously.

The TMB method can work with different levels of geographical boundaries. We contrast here three alternatives; the city, the district, and cell boundaries (defined as Voronoi cells), respectively. We set the hyperparameters of migration detection algorithm  $k$  and  $minDays$  ( $\epsilon$ ) to {3, 5, 7, 10, 14} days for both TMB and ASA.

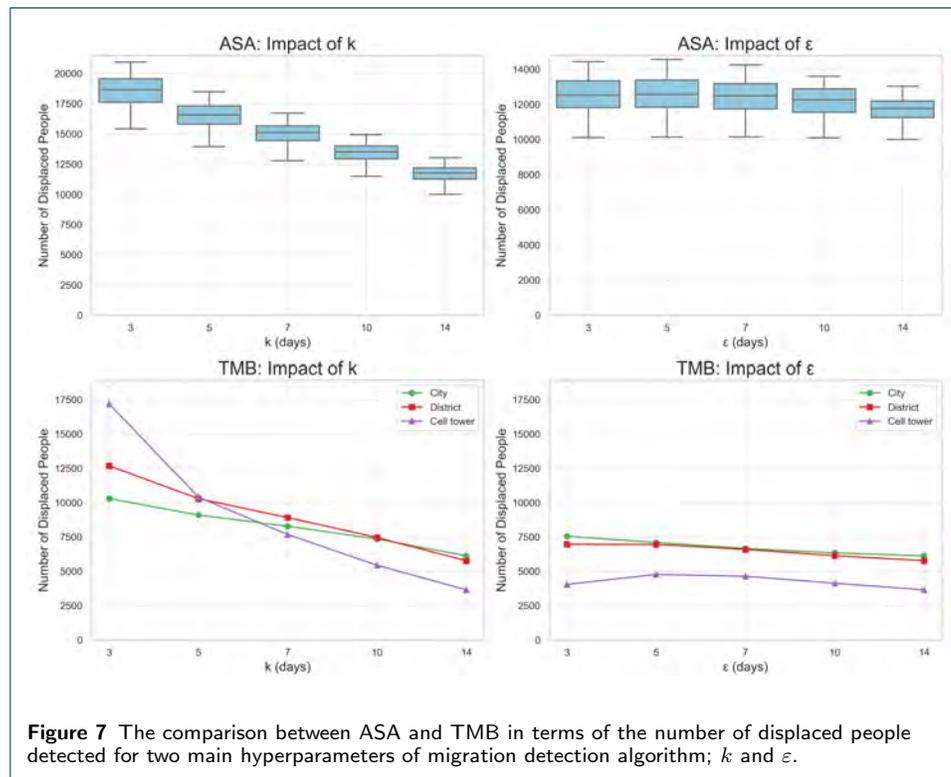
With this range of hyperparameters, ASA detects between 10,000 and 20,900 DPs, whereas TMB detects between 3,700 and 17,500 DPs at cell tower level, between 5,700 and 12,500 DPs at district level, and between 6,100 and 10,000 DPs at city level. We calculate an overlap coefficient, which measures the overlap between two sets divided by the size of the smaller set. We establish that ASA detects between 80% and 94% of DPs detected by TMB, in addition to finding other displacement groups.



**Figure 6** The number of displaced people for three different hyperparameter settings for ASA.

Figure 6 illustrates how the three hyperparameters affect the displacement detection in ASA. The left panel shows that increasing the spatiotemporal thresholds (from 2km/2h to 10km/4h) do not substantially impact the number of detected displacements. The middle panel demonstrates that the DBSCAN distance parameter, which determines activity space size, has an inverse relationship with displacement detection - larger distances (20km) generally result in fewer detected displacements compared to medium distances (3-5km). The impact of DBSCAN distances is dependent on the spatial threshold to determine stay locations. Increasing the DBSCAN distance parameter decreases the number of detected DPs, as this way, more movements remain within the activity spaces and hence not classified as displacements.

Lastly, the right panel confirms that increasing the nighttime relevance threshold (from 0% to 15%) directly increases displacement detection. Higher thresholds mean areas must have a greater pre-disaster familiarity to be considered ‘familiar’, making it more likely for post-disaster locations to be classified as unfamiliar, and thus indicating displacement.



We also test the parameters of the migration detection algorithm, used both in ASA and TMB. We sampled the number of displaced people by running the migration detection algorithm for different values of  $k$ , the minimum days stayed to consider the area as residence area, and  $\epsilon$ , the maximum gap allowed without breaking the segments. In Figure 7, we show the impact of these parameters on the number of detected displaced people.  $k$  is the most influential parameter in determining the number of DPs both for TMB and ASA. As expected, increasing  $k$  yields less displaced people, as we are more stringent on how much time needs to be spent at the destination to consider the person as displaced. Decreasing  $k$  from 14 to 3 days almost doubles the number of detected DPs. Chi et al. [25] previously showed that decreasing  $k$  has the same impact on longer periods of migration. For  $\epsilon$ , we do not see a large impact for TMB or for ASA.

## 5 Displacement patterns

After ASA is executed, we obtain origins, destinations, and dates of displacements<sup>[8]</sup>. In this section, we analyze our findings.

<sup>[8]</sup>We included further information on how dates of displacements are calculated in the Supplementary Material.

### 5.1 Stay locations

To characterize the origin and destination stay locations, we used multiple data sources. We applied the previously explained urbanization index to measure the level of urban development at both origins and destinations, and the damage index to quantify destruction levels. To account for the broader context, we considered the total population of the cities where these stay locations are situated. We also introduced a dummy Boolean variable called “Border Index” that equals 1 if a stay location falls within a 20 km radius of Türkiye’s border with Syria and 0 if it falls outside this range. As the earthquake was highly destructive in Syria too, we suspect this could create a strong pull effect for Syrian refugees living in Türkiye, especially in the immediate aftermath of the disaster. This variable helps to identify movements towards Syria.

There are multiple stay locations categorized as the origin and the destination for each individual. The urbanization, damage, border and population indices are calculated separately for each stay location of each individual, i.e. we know how urbanized the area is, how much damage is recorded within the area, the population count of the city where the stay location is situated, and whether or not the area is within 20 km of Syrian borders. The indices are first calculated at cell level, then the stay location polygons are intersected with the cells to calculate their specific values. To summarize the indices for the origin and destination stay locations, we weigh them with the nighttime duration spent in each. For a displaced person  $c$ , let  $S_{c,o}$  be the set of stay locations that are identified to be in the origin areas by our algorithm, and  $S_{c,d}$  to be the stay locations related to the destination. Then:

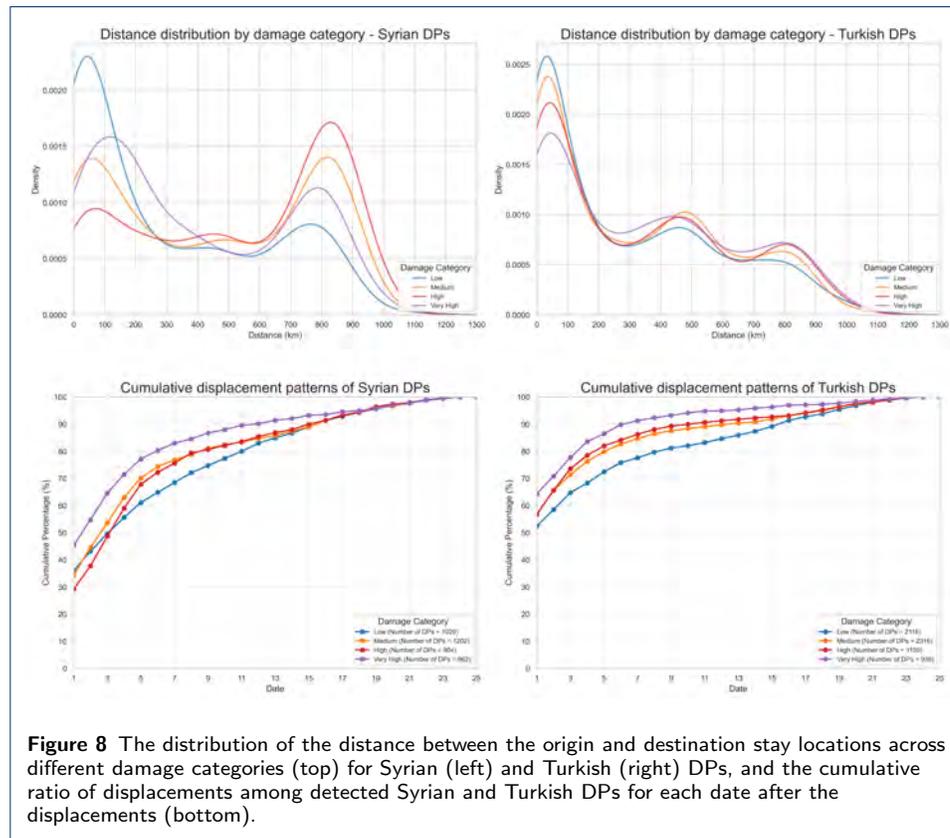
$$\text{Scaled index}_{c,l} = \frac{\sum_{s \in S_{c,l}} \text{index}_s \cdot \text{duration}_{c,s}}{\sum_{s \in S_{c,l}} \text{duration}_{c,s}}, \quad \text{where } l \in \{\text{origin, destination}\} \quad (3)$$

The indices calculated via Eq. 3 are showing the levels at the individual level, giving a summary of all locations that the person stayed in the origin and destination. We calculated the difference of these indices between the destination and the origin to quantify the changes in people’s stay locations after the disaster.

First, we look at the distribution of the distances and displacement dates. Since we have multiple origin and destination stay locations per DP, we look at the distance between the weighted centroids of origin and destination locations. In Figure 8, the upper plots show the distribution of distance for Syrians and Turkish DPs. We stratified the distributions by the damage categories, which are calculated using the damage index explained previously.

We see that for the Turkish group, DPs whose origin areas received lower levels of damage shifted their stay locations at shorter distances compared to DPs who were more impacted by the earthquake. The distribution of distances are similar across all damage categories, where DPs generally traveled shorter distances. On the other hand, for Syrians, we see a bimodal distribution, where DPs who experienced lower damage traveled shorter distances, and DPs who experienced medium or high damage traveled greater distances. When we look at the origin areas of the Syrians who were in the second peak emerging around 800–900 km, we saw that the origins

notably include the cities of Gaziantep and Şanlıurfa. One plausible explanation for this is that these cities traditionally send large numbers of seasonal agricultural workers to western Türkiye, and long-standing migration pathways may have been reactivated after the disaster. This indicates that once displacement occurs, Syrian DPs are more likely to engage in long-distance relocations, shaped by long-standing migratory practices and collective familiarity with specific routes or destinations.



## 5.2 Temporal patterns and mobility factors

A critical consideration in disaster response is the temporal pattern of population displacement. Research consistently demonstrates socioeconomic disparities in evacuation behavior; individuals from lower socioeconomic backgrounds typically evacuate later and in fewer numbers compared to those from higher socioeconomic strata [23, 22]. In addition, race and ethnicity was shown to impact the evacuation rates where the privileged demographic groups are overrepresented in early evacuations [48]. To understand the speed of displacements, we looked at the date of displacements after the earthquake occurred. In Figure 8, the lower plots show the displacement date for Turkish and Syrian communities for different levels of damage they experienced during the earthquake. On the first day 56% of Turkish detected DPs were already displaced, whereas only 36% of Syrians DPs were displaced on that day. Temporal patterns in Figure 8 highlight a stark contrast in mobility capacity. Turkish DPs, particularly from heavily damaged areas, tend to evacuate rapidly—peaking on the first or second day. In contrast, Syrian DPs

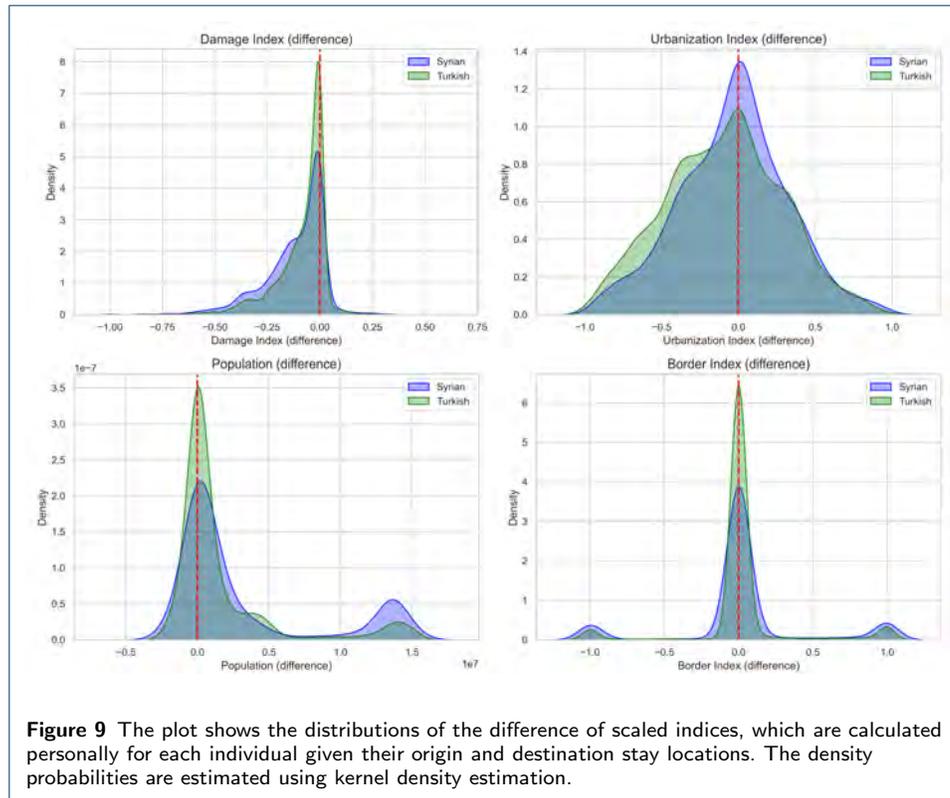
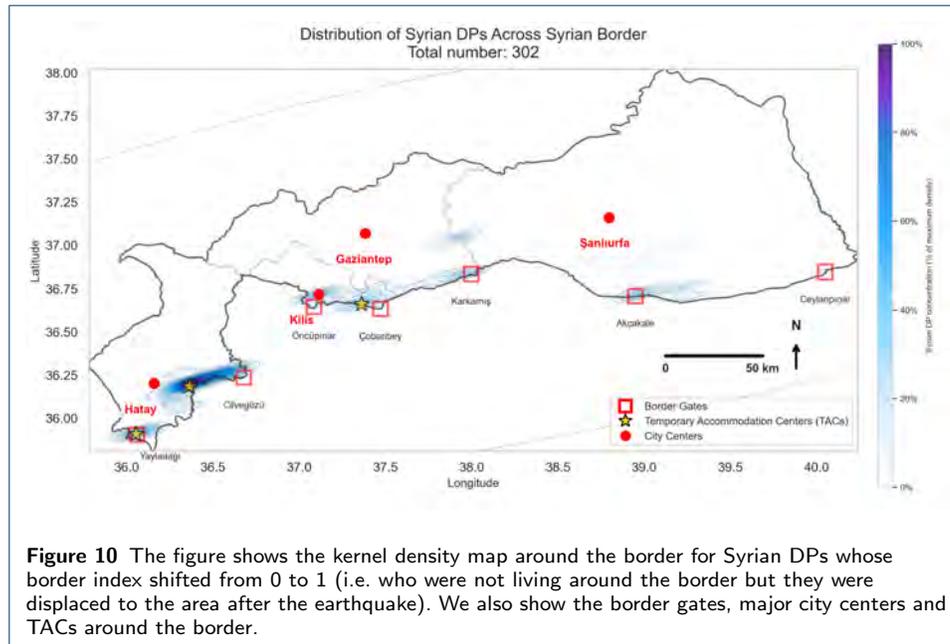


exhibit delayed displacement even from high-damage zones, with a noticeably long-tailed distribution. This temporal lag signals that exposure alone does not determine displacement timing; rather, the capacity to act on that exposure—influenced by access to transportation, legal protection, and social ties—plays a pivotal role. Vulnerable populations—such as the elderly, persons with disabilities, minorities, and undocumented migrants—face significant barriers in accessing timely information, transportation, and support services during disasters [49]. Their evacuation processes tend to be slower due to structural inequalities, limited social networks, and a lack of resources or legal protections [50]. As a result, disaster impacts are disproportionately severe for these groups, reinforcing existing vulnerabilities and reducing their chances of survival and recovery.

In Figure 9 top left, we see that most people shifted their stay locations to less damaged areas. The distributions center at a slightly negative value for both Turkish and Syrian populations, with a pronounced long tail extending toward the negative values, clearly indicating movement away from more severely damaged areas. Interestingly, Turkish DPs were more likely to relocate to less urbanized, often rural areas in the aftermath of the disaster (top right plot in Figure 9). This pattern reflects more than a retreat from risk; it reveals an adaptive capacity rooted in rural social embeddedness, including kinship ties, familiarity with land, and—in many cases—access to secondary housing or property. These resources enabled relatively autonomous and rapid evacuation, drawing on longstanding socio-spatial ties and informal support structures. Notably, official statements at the time confirm this trend: on 1 March 2023, the Turkish President announced that while 3.3 million people had left the disaster zone, approximately 800,000 had returned to their villages,



highlighting the scale of rural-directed displacement during the early post-disaster period.

For population differences between the origins and destinations (see lower left plot in Figure 9), we see that the DPs were attracted to larger population areas, as there is more density in the right hand side. This aligns well with the predictions of traditional migration models, which expect more migration towards higher population areas. The smaller peak (around 14 million) reflects DPs who traveled to Istanbul, which has an official populace of 16.5 million. Most DPs who migrated to Istanbul were Syrians. Turkish DPs show a more dispersed pattern of displacement across various metropolitan cities in the western regions of Türkiye. When we look at the destination hotspots in Istanbul, we see a similar pattern that the Syrians DPs are concentrated in a couple of districts in the European side of the city, whereas Turkish DPs are spatially more dispersed. Notably, Syrian DPs are concentrated in Esenyurt, Bağcılar, Esenler and Fatih, which are all districts with high concentration of Syrian refugees [51]. This suggests that Syrians relied on their social networks in Istanbul when they choose their destinations. It is possible that Turkish DPs have done the same, yet due to the larger spatial variation in their social networks in Istanbul such concentrations are not possible to observe.

There are around 302 Syrians, and 250 Turkish DPs whose destination stay locations got shifted towards the Syrian border compared to their origin stay locations measured by the border index difference in Figure 9. These people were generally residing in the central areas of Hatay, Gaziantep and Şanlıurfa. We gave a visual summary of the destinations of this group around the Syrian border in Figure 10. We see that the concentration of people were around various hotspots, most notably border gates, TACs, and city center of Kilis. We cannot definitively know if the Syrian DPs were attracted to this region to cross the border and possibly help

their social networks in Syria<sup>[9]</sup>, or they were coming near the TACs or their social networks near the border residing in Türkiye. The destinations for Turkish people who get closer to the border region were not clustered as close to the border gates, rather were around the rural areas of Gaziantep, Hatay and Şanlıurfa.

### 5.3 TACs

The finding in Section 4.1 that Syrian refugees were displaced to TACs align with the qualitative observations from our field visits. Notably, none of the methods we tried (except ASA) could identify that the Syrian refugees chose TACs as one of their main destination areas in the earthquake zone. Following our initial visit to the earthquake-affected region, we observed that some TACs had become post-disaster shelters not only for Syrians, but also for Turkish citizens. For instance, in the Kahramanmaraş Türkoğlu (Sivricehüyük) TAC, which previously hosted approximately 10,000 Syrians, the population reportedly rose to 16,500 after the earthquake [52]. This increase reflects the arrival of displaced Turkish citizens seeking shelter in spaces with already functioning infrastructure, including heating, sanitation, and medical services. While ASA-based measures can capture a rise in Syrian presence in Türkoğlu TAC, the mobility of Turkish citizens does not appear with the same clarity, which is likely due to limitations in sampling and the absence of concentrated movement patterns among them.

Our findings suggest that refugee camps have functioned not only as pre-disaster residences, but also as post-disaster spatial anchors – limiting the range of dispersal and reinforcing institutional dependency. These patterns may be interpreted as both protective and limiting, shaping uneven geographies of recovery. This spatial gathering observed among Syrians, especially the concentration in post-disaster camp zones, may reflect multiple overlapping dynamics. First, information asymmetries and legal precarity may have restricted alternative destination options. Second, existing infrastructure in and around the camps (despite being established before the earthquake for different reasons) offered familiar settings and perceived safety. Third, in the immediate aftermath of the disaster, the intensification of anti-refugee sentiment and public discrimination may have led some Syrians to seek invisibility or social shielding by retreating into more enclosed and institutionally managed spaces such as camps—places where their presence, while often marginalized, was at least formally recognized [32, 53].

Despite the co-settlement in TACs such as Türkoğlu, the spatial convergence did not necessarily translate into social integration in these areas. Instead, forms of micro-level segregation emerged, highlighting the enduring nature of group boundaries even in contexts of shared vulnerability and displacement. One of the additional sites we visited during the qualitative fieldwork was the Hilalkent camp, established nearly a year after the earthquake. Unlike Türkoğlu TAC, Hilalkent was not a spontaneous shelter, but a state-initiated response to escalating tensions in the area. Its establishment highlights that the post-disaster period did not necessarily reduce social divisions; in some cases, it reinforced them—supporting our argument

---

<sup>[9]</sup>There are news sources documenting the increased queues at the border gates in Syrian border after the earthquake: <https://www.bbc.com/turkce/articles/crgz584x9gro>, Accessed April 8, 2025

that co-settlement did not foster cohesion and, at times, even deepened spatial and social segregation.

## 6 Validation

The validation of MPD-based indicators of displacement in the context of disasters remains a challenge, as usually ground truth data do not exist and existing data sources focus on different aspects of the displacements. Nevertheless, we used the Turkish Statistical Institute (TURKSTAT) annual migration inflows and flows per city to validate our results further. These are based on the changes in address registrations recorded at the end of each year compared to the previous year’s end [10], and are disaggregated by the reason of migration. Earthquake-induced migration is included in a category labeled “other”, and migration outflows with this label highly correlate with earthquake-induced migration outflows from the cities affected by earthquake (approximately 70 % of the people migrated out from the earthquake affected cities are in the “other” category).

TURKSTAT data are challenging to compare to our own measurements for three reasons. Firstly, the data show the migration figures as of December 2023 whereas our measurements are limited to the 26 days after the earthquake. Secondly, TURKSTAT data show only outflows and inflows and it is not disaggregated by origins and destinations, lastly there is no information on the flows of Syrian DPs. In addition, the total number of people migrated with the “other” reason in 2023 is somewhere around 500,000 and 600,000 (the statistics changes for inflows and outflows). IDMC, on the other hand, claims that there were 4 million internally displaced people in Türkiye as a result of 2023 Syria-Türkiye earthquake[11]. We keep these limitations in mind while comparing our results to TURKSTAT figures.

In Eq. 4, we debias the DP outflows from each affected city  $f \in F$  for each nationality group  $n \in \{\text{Turkish, Syrian}\}$  by correcting for sampling biases in the CDR data relative to census population figures. This balances factors like the market share of the MNO, and provides interpolation at city level, but assumes similar customer behaviors across the three Turkish MNOs.

$$\text{out}_{f,n}^d = \frac{p_{f,n}^{CDR}}{p_{f,n}^{census}(1 - \delta_n)} \times \text{out}_{f,n}^r \quad (4)$$

In Eq. 4,  $\text{out}_{f,n}^d$  represents the debiased DP outflows from affected cities, whereas  $\text{out}_{f,n}^r$  represents the raw calculations. For debiasing, we multiply CDR estimated DP outflow numbers with the CDR population residing in each city and divide it by the census population. Secondly, as children and elderly (represented by delta for each group) are not well represented in CDR data set, we remove them from the census population. According to TURKSTAT, population under 15 constitute

[10] <https://data.tuik.gov.tr/Bulten/Index?p=Uluslararası-Göç-İstatistikleri-2023-53544>, Accessed April 7, 2025.

[11] The numbers on the internally displaced people per country and disaster can be queried at: <https://www.internal-displacement.org/database/displacement-data/>

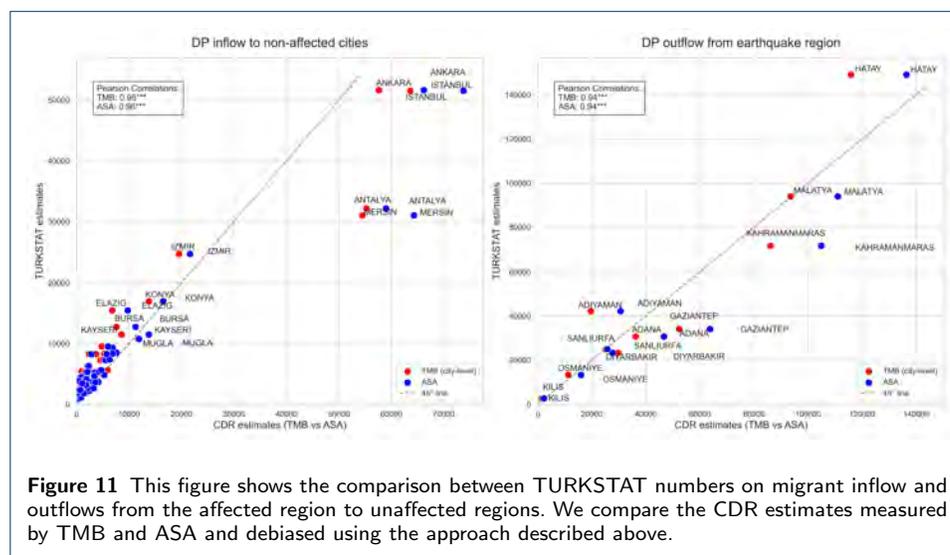
21 %, whereas population over 85 constitute 1 % of the population [12]. For Syrians, in January 2023 approximately 40 % of population was under 15, and a negligible portion was over 85 according to Presidency of Migration Management (PMM)[13]. As a result, Eq. 4 calculates the ratio only for the eligible census population.

To calculate the DP inflows, first we calculate the share of each flow from affected cities  $f \in F$  to unaffected cities  $u \in U$  as a ratio to all flows from the affected city  $f$  as in Eq. 5:

$$\text{share}_{f \rightarrow u}^{CDR} = \frac{\text{flow}_{f \rightarrow u}^{CDR}}{\sum_{u \in U} \text{flow}_{f \rightarrow u}^{CDR}} \quad (5)$$

Equation 5 gives us the ratio of expected flows from each affected city to unaffected city. Then, in Eq. 6, we calculate the total DP inflows to unaffected city  $u$ , by multiplying the flow share from each affected city  $f$  with debiased DP outflows and summing over all affected cities  $F$ .

$$\text{in}_{u,n}^d = \sum_{f \in F} \text{out}_{f,n}^d \times \text{share}_{f \rightarrow u}^{CDR} \quad (6)$$



In Figure 11, we compare the estimated number of Turkish DP inflows and Turkish DP outflows for TMB and ASA. To be able to compare the measurements of ASA to TURKSTAT figures and TMB estimates, we needed to assign each person to

[12] There are significant differences in age distribution among cities as well, however for simplicity we will assume a constant number for our calculations. More statistics on the demographics of Turkish citizens is available at: <https://data.tuik.gov.tr/>

[13] Age and gender distribution of Syrians are sporadically shared by PMM. We used the infographics shared on 19th of January, 2023 on official PMM website: <https://en.goc.gov.tr/>

a single city before and after the disaster. For each DP, we identified their origin city by determining which city contained the origin stay locations where they spent the most cumulative nighttime hours. Similarly, we designated their destination city as the one containing the destination stay locations where they spent the most cumulative nighttime hours.

The left plot shows the correlation between the CDR estimated DP outflows and the figures reported by TURKSTAT. Importantly, 45 degree line represents the perfect CDR-estimation of DP outflows. The Pearson correlation coefficient between TURKSTAT and CDR estimates for DP outflows is calculated between 0.94 both for TMB and ASA (0.01 significance level). The difference between TMB and ASA measurements are due to the fact that ASA detects more DPs than TMB, hence the debiased CDR estimates are higher for ASA. In Şanlıurfa, Osmaniye and Kilis, the CDR-based estimates are very close to the TURKSTAT figures, whereas in Hatay and Adıyaman, there is an underestimation, and in Kahramanmaraş and Gaziantep, there is an overestimation both compared to the TURKSTAT figures. TURKSTAT figures are calculated throughout the whole year, and potentially capture the displacements that were more persistent. For instance, the difference between TURKSTAT and MPD-based estimates in Hatay might be pointing to the persistence of the displacements from this region due to larger infrastructure damage experienced. Possibly, the underestimation might be stemming from the larger infrastructure damage experienced in Hatay.

**Table 1** Syrian Displaced Persons Inflows and Outflows by City

Destination City	DP Inflows	Origin City	DP Outflows
İSTANBUL	13,379	HATAY	18,598
MERSİN	5,492	GAZİANTEP	13,790
BURSA	5,004	KAHRAMANMARAŞ	8,043
ADANA	3,026	ŞANLIURFA	5,514
ANKARA	2,917	MALATYA	4,322
KONYA	2,784	ADANA	2,345
İZMİR	2,502	ADİYAMAN	2,052
GAZİANTEP	2,416	OSMANİYE	923
ŞANLIURFA	2,052	KİLİS	488
ANTALYA	1,873	DİYARBAKIR	434

Note: The numbers are calculated from CDR as de-biased migrant inflow and outflow estimates using ASA.

In the right plot of Figure 11, we see the DP inflow calculations from affected cities to unaffected cities. TMB, ASA, and TURKSTAT all rank the number of DP inflows and outflows per unaffected and affected city in the same order. However, in some cases, CDR estimates are significantly above the DP inflows measured by TURKSTAT. This pattern is especially noticeable in Ankara, Istanbul, Antalya, and Mersin, which might be a signal that the DPs from these regions returned at a higher rate than other places.

In total, with ASA we estimate around 560,000 Turkish DPs migrated to other cities. We estimate around 300,000 Turkish were displaced within the city boundaries of their origins within the time period we have investigated. We recognize that these numbers are well below the 4 million declared by IDMC, but align more closely with TURKSTAT statistics. For Syrians, there is no official statistics to compare the estimated number of earthquake-induced displacements. We estimate that around 55,000 Syrians were displaced to other cities, and around 15,000 Syrians

were displaced within the city boundaries. In Table 1, we share the statistics on the main origin and destination cities of Syrian DPs as estimated.

There are a couple of limitations that might be affecting our results. Firstly, the earthquake caused high levels of damage to the cell tower infrastructure, especially in the most impacted regions. At any given moment, a proportion of cell towers registered to the MNO will not receive signals, and this proportion is somewhat higher in the earthquake region after the earthquake, particularly in Hatay. Since nearby towers often compensate this, we relied on towers that continued to provide service throughout the whole period. We recognize this will cause a bias in how well we are able to assess the damage per area, and there may be groups of DPs in very highly damaged areas missing from our CDR analysis.

## 7 Conclusion

In this paper, we introduced the novel ASA to predict and monitor forced displacements caused by natural disasters via CDR. We tested ASA using the Türkiye and Syria earthquakes as a case study, where the disaster induced widespread displacements occurring within a region characterized by coexisting local and refugee populations. ASA is built on the concept of activity spaces, and overcomes some limitations of HLA in terms of accurately describing the living spaces of individuals at spatially granular level. ASA detects multiple origins and destinations per person as stay areas, which also enable the detection of origin and destination hotspots. This flexible approach allows us to see more detailed mobility patterns at different levels. For example, our experiments illustrated that the alternative home location based TMB approach could not identify the main destination hotspots for Syrian refugees, such as the Temporary Accommodation Centers (TACs). In addition, ASA enables building more specific spatial indices, which can show the impact of the disaster on DPs across important factors. Our findings underscore the value of spatially disaggregated analyses in understanding displacement and allow for comparisons among different segments of the population.

Our analyses have shown that local and Syrian DPs exhibit divergent displacement patterns across spatial and temporal dimensions. Syrian refugees were able to leave the disaster areas considerably later than locals and were primarily displaced to TACs within the region, whereas Turkish population showed more diverse spatial displacement patterns both within the earthquake-affected areas and across other cities, with a notable pattern of urban-to-rural displacement. Syrian displacement appears to be shaped by constrained agency due to various conditions. For example, Syrian refugees are subject to travel permit obligations, have more limited social networks, and are delimited by institutionally defined geographies as they are more visibly concentrated around TACs. Their limited mobility patterns demonstrated spatial clustering around pre-existing TACs, which acted as stabilizing infrastructures in the immediate aftermath. Conversely, local DPs exhibit broader dispersal patterns, more rural transitions, and earlier displacements and higher mobility capacity.

All in all, our research shows systematically how coping strategies as a response to a natural disaster are conditional upon various economic, social and legal circumstances of individuals and show considerable variation across different segments

of the population. In line with Doreen Massey's concept of "power geometries", we see how socio-spatial mobilities are unevenly produced, regulated, and constrained [54]. ASA allowed us to bring attention to the ways in which specific forms of movement are shaped by structural inequalities, institutional regulations, and disciplinary forces that restrict or enable mobility, though the specific mechanisms behind these patterns remain open questions for future research.

Field observations suggest that concentrated camp-based populations may be channeled into precarious labor roles, constraining their bargaining power and prospects for sustainable recovery. Similarly, local populations which have moved towards rural areas may not be able to have a sustainable livelihood considering the rather limited absorptive capacity of rural systems. In the face of such differences, a one-size-fits-all solution is not possible, and flexible computational tools are needed.

In conclusion, these findings highlight the importance of examining varying mobility patterns among different population segments, calling for more targeted approaches that address the distinct needs of each group. Our parametric approach enables detailed MPD analyses at different resolutions, and highlights interactions between the vulnerabilities of certain communities and their displacements in the face of a natural disaster. We emphasize that MPD analysis must be further supported by local knowledge and long term reflections with regards to the needs of the communities for building resilience in face of future disasters.

#### Funding

This study is supported by European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 870661 and by the Scientific Research Projects Commission of Galatasaray University under grant number FCD-2022-1083.

#### Abbreviations

**ABPRS:** Address-Based Population Registration System  
**ASA:** Activity Space Approach  
**CDR:** Call detail records  
**D4D:** Data for Development  
**D4R:** Data for Refugees  
**DBSCAN:** Density-Based Spatial Clustering of Applications with Noise  
**DP:** Displaced person  
**GPS:** Global positioning system  
**HLA:** Home location approach  
**IDMC:** Internal Displacement Monitoring Centre  
**MCH:** Minimum convex hull  
**MNO:** Mobile network operator  
**MoEUCC:** Ministry of Environment, Urbanization and Climate Change  
**MPD:** Mobile phone data  
**PMM:** Presidency of Migration Management  
**TAC:** Temporary accommodation center  
**TMB:** Tai, Mehra, Blumenstock  
**TTECH:** Turkcell Technology  
**TURKSTAT:** Turkish Statistical Institute  
**UN:** United Nations

#### Availability of data and materials

The code for Activity Space approach is extensively documented at Github repository, and can be accessed at: [https://github.com/bilgecag/activity\\_space\\_based\\_displacement\\_detection](https://github.com/bilgecag/activity_space_based_displacement_detection). The damage and urbanization indices are re-calculated at 10km grids and made accessible. Final CDR-based estimates on migrant inflows and outflows in combination with official data sources will be made available.

#### Ethics approval and consent to participate

This study uses anonymized and aggregated mobile phone data from Turkcell Technology, with customers providing consent for research purposes at subscription time. The data were prepared following strict ethical protocols, privacy-by-design principles, and legal frameworks, including the Turkish Law on Protection of Personal Data, and EU's General Data Protection Regulation. The data are processed to prevent identification of individuals and is shared only under limited access conditions as documented in the HumMingBird EU project report of the ethics committee [55].

### Competing interests

The authors declare that they have no competing interests.

### Authors' contributions

B.A., S.N.Y., and S.G. collected, anonymized and aggregated the call detail records at TTECH. B.A. conceptualized the methodology, analyzed the data, and prepared the first draft. A.A.S. provided guidance in methodology and data analysis. D.D. and C.Y. conducted the fieldwork. B.A., A.A.S., D.D., C.Y. and Ö.B. wrote the paper. All authors reviewed and approved the manuscript.

### Author details

<sup>1</sup>Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands. <sup>2</sup>Department of Sociology, Galatasaray University, Istanbul, Türkiye. <sup>3</sup>Department of Interdisciplinary Social Science, Utrecht University, Utrecht, The Netherlands. <sup>4</sup>Department of Political Science, Galatasaray University, Istanbul, Türkiye. <sup>5</sup>Turkcell Technology, Istanbul, Türkiye.

### References

- Adger, W.N., De Campos, R.S., Mortreux, C.: Mobility, displacement and migration, and their interactions with vulnerability and adaptation to environmental risks. In: Routledge Handbook of Environmental Displacement and Migration, pp. 29–41. Routledge, London (2018)
- Strategy and Budget Office: Kahramanmaraş Ve Hatay Depremleri Yeniden İmar Ve Gelişme Raporu 2024. T.C. Strateji ve Bütçe Başkanlığı, Ankara (2024)
- Sağiroğlu, A.Z., Ünsal, R., Özenci, F.: Deprem sonrası göç ve insan hareketlilikleri: Durum değerlendirme raporu. Technical report AYBÜ-GPM Rapor Serisi-15, Ankara Yıldırım Beyazıt Üniversitesi Göç Politikaları Uygulama ve Araştırma Merkezi, Ankara (2023). Updated 2nd Edition: 15 April 2023
- Wilson, R., Erbach-Schoenberg, E.Z., Albert, M., Power, D., Tudge, S., Gonzalez, M., Guthrie, S., Chamberlain, H., Brooks, C., Hughes, C., Pitonakova, L., Buckee, C., Lu, X., Wetter, E., Tatem, A., Bengtsson, L.: Rapid and near real-time assessments of population displacement using mobile phone data following disasters: The 2015 Nepal earthquake. *PLoS Currents* **8**(Disasters) (2016). doi:[10.1371/currents.dis.d073fbec328e4c39087bc086d694b5c](https://doi.org/10.1371/currents.dis.d073fbec328e4c39087bc086d694b5c)
- Acosta, R.J., Kishore, N., Irizarry, R.A., Buckee, C.O.: Quantifying the dynamics of migration after Hurricane Maria in Puerto Rico. *Proceedings of the National Academy of Sciences of the United States of America* **117**(51), 32772–32778 (2020). doi:[10.1073/pnas.2001671117](https://doi.org/10.1073/pnas.2001671117)
- Lu, X., Bengtsson, L., Holme, P.: Predictability of population displacement after the 2010 Haiti earthquake. *Proceedings of the National Academy of Sciences of the United States of America* **109**(29), 11576–11581 (2012). doi:[10.1073/pnas.1203882109](https://doi.org/10.1073/pnas.1203882109)
- Li, T., Dejby, J., Albert, M., Bengtsson, L., Lefebvre, V.: Detecting individual internal displacements following a sudden-onset disaster using time series analysis of call detail records. arXiv preprint arXiv:1908.02377 (2019)
- Cumbane, S.P., Gidófalvi, G.: Spatial distribution of displaced population estimated using mobile phone data to support disaster response activities. *ISPRS International Journal of Geo-Information* **10**(6), 421 (2021)
- Isaacman, S., Frias-Martinez, V., Frias-Martinez, E.: Modeling human migration patterns during drought conditions in La Guajira, Colombia. *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies, COMPASS 2018* (2018). doi:[10.1145/3209811.3209861](https://doi.org/10.1145/3209811.3209861)
- Tai, X.H., Mehra, S., Blumenstock, J.E.: Mobile phone data reveal the effects of violence on internal displacement in Afghanistan. *Nature Human Behaviour* **6**(5), 624–634 (2022)
- Barrett, S., Steinbach, D., Addison, S.: Assessing vulnerabilities to disaster displacement: A good practice review. IIED Working Paper (2021)
- Internal Displacement Monitoring Centre (IDMC): Systematic data collection and monitoring of displacement and its impacts at local, national, regional and international level to inform comprehensive needs and risk assessments for the formulation of policy and plans. Technical report, The Warsaw International Mechanism for Loss and Damage Associated with Climate Change Impacts (August 2018). Activity III.1-3 of the workplan the Task Force on Displacement
- Aydoğdu, B., Bilgili, Ö., Güneş, S., Salah, A.A.: Mobile phone data for anticipating displacements: practices, opportunities, and challenges. *Data & Policy* **7**, 5 (2025)
- Black, J.: Global Migration Indicators 2021. International Organization for Migration (IOM), Geneva (2021). <https://publications.iom.int/books/global-migration-indicators-2021>
- Marzuoli, A., Liu, F.: Monitoring of natural disasters through anomaly detection on mobile phone data. In: 2019 IEEE International Conference on Big Data (Big Data), pp. 4089–4098 (2019). IEEE
- Song, X., Zhang, Q., Sekimoto, Y., Shibasaki, R.: Prediction of human emergency behavior and their mobility following large-scale disaster. *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 5–14 (2014). doi:[10.1145/2623330.2623628](https://doi.org/10.1145/2623330.2623628)
- Golledge, R.G., Stimson, R.J.: *Spatial Behavior: A Geographic Perspective*. Guilford Press, New York (1997)
- Horton, F.E., Reynolds, D.R.: Effects of urban spatial structure on individual behavior. *Economic geography* **47**(1), 36–48 (1971)
- Pappalardo, L., Ferres, L., Sacasa, M., Cattuto, C., Bravo, L.: Evaluation of home detection algorithms on mobile phone data using individual-level ground truth. *EPJ data science* **10**(1), 29 (2021)
- Gemenne, F., Zickgraf, C., Hut, E., Castillo Betancourt, T.: Forced displacement related to the impacts of climate change and disasters. Reference Paper for the 70th Anniversary of the 1951 Refugee Convention (2022)
- Zipf, G.K.: The P 1 P 2/D hypothesis: on the intercity movement of persons. *American Sociological Review* **11**(6), 677–686 (1946)
- Naushirvanov, T., Elejalde, E., Kalimeri, K., Omodei, E., Karsai, M., Ferres, L.: Evacuation patterns and socioeconomic stratification in the context of wildfires. *EPJ Data Science* **14**(1), 23 (2025)

23. Yabe, T., Ukkusuri, S.V.: Effects of income inequality on evacuation, reentry and segregation after disasters. *Transportation research part D: transport and environment* **82**, 102260 (2020)
24. de Mello, S.V., Deng, F.M.: Guiding principles on internal displacement. United Nations (1998)
25. Chi, G., Lin, F., Chi, G., Blumenstock, J.: A general approach to detecting migration events in digital trace data. *PloS one* **15**(10), 0239408 (2020)
26. Blondel, V.D., Esch, M., Chan, C., Clérot, F., Deville, P., Huens, E., Morlot, F., Smoreda, Z., Ziemlicki, C.: Data for development: the D4D challenge on mobile phone data. *arXiv preprint arXiv:1210.0137* (2012)
27. Salah, A., Pentland, A., Lepri, B., Letouzé, E.: *Guide to Mobile Data Analytics in Refugee Scenarios*. Springer, Cham, Switzerland (2019). doi:[10.1007/978-3-030-12554-7](https://doi.org/10.1007/978-3-030-12554-7)
28. Aydogdu, B., Salah, A.A., Ones, O., Gurbuz, B.: Description of the mobile CDR database. *HumMingBird Project Deliverable* **6**(1) (2021)
29. Tominga, A., Silm, S., Orru, K., Vent, K., Klaos, M., Vöik, E.J., Saluveer, E.: Mobile positioning-based population statistics in crisis management: An Estonian case study. *International Journal of Disaster Risk Reduction* **96**(September 2022) (2023). doi:[10.1016/j.ijdr.2023.103887](https://doi.org/10.1016/j.ijdr.2023.103887)
30. 2023 Turkey Earthquakes - Building Damage Assessment Map Accessed: 2024-02-07 (2023). <https://hasar.6subatdepremi.org/>
31. Dilsiz, A., Günay, S., Mosalam, K.M., Miranda, E., Arteta, C., Sezen, H., Fischer, E., Hakhamaneshi, M., Hassan, W.M., Alhawamdeh, B., et al.: StEER-EERI: 2023 mw 7.8 Kahramanmaraş, Türkiye earthquake sequence joint preliminary virtual reconnaissance report (PVRR). Technical report, ETH Zurich (2023)
32. Sevinin, E., Daniş, D., Sert, D.: Double displacement of refugees in the context of the 2023 Turkey-Syria earthquake. *International Migration* **61**(4), 341–344 (2023)
33. Çakın, B., AlMajdalawi, R.: Unsettled ground, unsettled lives: The ontological (in)securities of Syrian refugees in the shadow of Turkey's 2023 earthquake. *Journal on Migration and Human Security* **13**(2), 257–275 (2025)
34. Nasır, A., Safi, G., Ms, F.B., Tekin, S., Drury, J.: Understanding the socio-economic and wellbeing impacts of the 2023 earthquakes in turkey on syrian refugees. *International Journal of Disaster Risk Reduction* (2025). doi:[10.1016/j.ijdr.2025.105548](https://doi.org/10.1016/j.ijdr.2025.105548)
35. Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., Ma, W.-Y.: Mining user similarity based on location history. In: *Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, pp. 1–10 (2008)
36. Silm, S., Ahas, R.: Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data. *Annals of the Association of American Geographers* **104**(3), 542–559 (2014)
37. Järv, O., Ahas, R., Witlox, F.: Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies* **38**, 122–135 (2014)
38. Ahas, R., Aasa, A., Silm, S., Tiru, M.: Daily rhythms of suburban commuters' movements in the Tallinn metropolitan area: Case study with mobile positioning data. *Transportation Research Part C: Emerging Technologies* **18**(1), 45–54 (2010)
39. Ahas, R., Silm, S., Järv, O., Saluveer, E., Tiru, M.: Using mobile positioning data to model locations meaningful to users of mobile phones. *Journal of urban technology* **17**(1), 3–27 (2010)
40. Xiong, Q., Liu, Y., Xing, L., Xie, P., Liu, Y.: Exploring temporal and spatial reliability of individual activity space measurements using massive mobile phone data. *Transactions in GIS* **26**(6), 2669–2690 (2022)
41. Sila-Nowicka, K., Fotheringham, A.S., Demšar, U.: Activity triangles: a new approach to measure activity spaces. *Journal of Geographical Systems* **25**(4), 489–517 (2023)
42. Ester, M., Kriegel, H.-P., Sander, J., Xu, X., et al.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: *Kdd*, vol. 96, pp. 226–231 (1996)
43. McAuliffe, M., Oucho, L.A. (eds.): *World Migration Report 2024*. International Organization for Migration (IOM), Geneva (2024). <https://www.iom.int>
44. United Nations: Recommendations on statistics of international migration. *Statistical Papers Series M*, No. 58, Rev. 1. Statistics Division, Department of Economic and Social Affairs (1998)
45. Ahmad Yar, A.W., Bircan, T.: Big data for official migration statistics: Evidence from 29 national statistical institutions. *Big data & Society* **10**(2), 20539517231210244 (2023)
46. Parzen, E.: On estimation of a probability density function and mode. *The Annals of Mathematical Statistics* **33**(3), 1065–1076 (1962)
47. Worton, B.J.: Kernel methods for estimating the utilization distribution in home-range studies. *Ecology* **70**(1), 164–168 (1989)
48. Deng, H., Aldrich, D.P., Danziger, M.M., Gao, J., Phillips, N.E., Cornelius, S.P., Wang, Q.R.: High-resolution human mobility data reveal race and wealth disparities in disaster evacuation patterns. *Humanities and Social Sciences Communications* **8**(1), 1–8 (2021)
49. Myers, C.A., Slack, T., Singelmann, J.: Social vulnerability and migration in the wake of disaster: The case of Hurricanes Katrina and Rita. *Population and Environment* **29**, 271–291 (2008)
50. Uekusa, S., Matthewman, S.: Vulnerable and resilient? immigrants and refugees in the 2010-2011 Canterbury and Tohoku disasters. *International Journal of Disaster Risk Reduction* **22**, 355–361 (2017)
51. Aydođdu, B., Ahat, B., Salah, A.A., Bircan, T.: Temporal and spatial analysis of indicators on segregation of Syrian refugees in Turkey with mobile phone data. In: *IEEE 30th Signal Processing and Communications Applications Conference (SIU)*, pp. 1–4 (2022)
52. Sert, D., Daniş, D., Sevinin, E.: Göç ve deprem: Durum tespit raporu. Research report, Göç Araştırmaları Derneği (GAR) (2023). Accessed: 2025-06-04. <https://gocarastirmalaridernegi.org/tr/calismalar/arastirmalar/goc-ve-deprem-durum-tespit-raporu/311-goc-ve-deprem-durum-tespit-raporu>
53. Halilođlu Kahraman, Z.E.: Reflections on Kahramanmaraş and Hatay earthquakes: Evaluating Syrian refugees' location choices in terms of earthquake hazards and risks. *Journal of Planning* **33**(3), 506–514 (2023)
54. Massey, D.: *Space, Place and Gender*. University of Minnesota Press, Minneapolis (1994)
55. Salah, A.A., Aydođdu, B.: Report of the ethics committee. Technical report, Utrecht University (2023).

Deliverable 6.6, Work Package 6, HumMingBird Project

# Supplementary Material for “A novel activity space approach to discover displacement patterns via mobile phone data: An analysis of the 2023 Türkiye-Syria Earthquakes”

Bilgeçağ Aydoğdu<sup>1</sup>, Didem Danış<sup>2</sup>, Özge Bilgili<sup>3</sup>,  
Cemil Yıldızcan<sup>4</sup>, Semiha Nur Yağcıklı<sup>5</sup>, Subhi Güneş<sup>5</sup>, Albert Ali Salah<sup>1</sup>

<sup>1</sup> Department of Information and Computing Sciences, Utrecht University, The Netherlands

<sup>2</sup> Department of Sociology, Galatasaray University, Türkiye

<sup>3</sup> Department of Interdisciplinary Social Science, Utrecht University, The Netherlands

<sup>4</sup> Department of Political Science, Galatasaray University, Türkiye

<sup>5</sup> Turkcell Technology, Türkiye

July 10, 2025

## Introduction

In this document, we provide a range of supplementary materials, starting with methodological details on call detail records, discussion of noise in cell tower data, urbanization and damage indices, measurement of stay locations, activity spaces, parameter selection, and displacement detection. We provide additional model comparisons and experimental results, particularly, figures that provide more detailed insights on alternative models.

## Additional Details for Methodology

### CDR preparation details

The data collaborative with Turkcell Technology started in 2020 within the framework of the HumMingBird EU project<sup>1</sup>. Initially, the subscriber base of Turkcell was subsampled for around 2 million foreigners, essentially using the information collected on users’ nationality at the time of registration, and for another 2 million Turkish users. We used stratified sampling for the Turkish users considering the number of foreigners we sampled in each of the 81 cities of Türkiye. Therefore, the sampling of Turkish users follows the distribution of foreigners across different parts of the country; more foreigners meant more Turkish users. We created five broad nationality-based flags

---

<sup>1</sup><https://hummingbird-h2020.eu/>, Accessed June 13, 2025.

for each user based on this information for comparative analysis: Turkish, Syrian, Middle Eastern or Northern African (except Syria and Türkiye), Central Asian or Afghan.

We collected the Call Detail Records (CDR) from this user base between November, 2022 and September, 2023. No personal information was stored in CDR, and we ensured one-way anonymization. For the earthquake study, we have processed this CDR into a fine-grained mobility data set format, where we had a timestamp, anonymised ID for caller, nationality segment of the caller and the callee, cell tower ID used by the caller and the callee, respectively. For this study, we only used outgoing call records, and not incoming call records. We sampled around 500,000 users across Türkiye (around 25% of the user base), which resulted in approximately 125,000 people in the earthquake region. This region is overrepresented in the data we have prepared, as a high proportion of Syrians lived in this region, and we subsampled more Turkish users there.

Our sampling strategy was explicitly designed for understanding migrant mobility. Therefore, in our study we were better able to identify the displacement patterns of Syrians, compared to Turkish users. Among migrant groups, we only focused on Syrians in this study, because the other groups were too small to enable a meaningful analysis.

## Noise in cell tower data

The precise locations of cell towers is considered as sensitive information. Consequently, an initial Voronoi tessellation was computed using the actual cell tower locations inside TTECH. A small portion of the towers were unused throughout the whole data collection period (i.e. no data acquired from the Voronoi cell), and these were excluded from the analyses. Then, we clustered the active cell towers using agglomerative hierarchical clustering with Ward’s method and a 1 km distance threshold, which helped us group towers into spatially cohesive clusters, merging small cells into larger areas. A new Voronoi tessellation was created using the centers of the clustered towers, and the centroids of cells were used in our analysis. This adds some minor noise to our location estimation, but prevents potential privacy risks due to too-small Voronoi cells.

## Urbanization index

The inactive cells required additional processing for the computation of the urbanization index. We redistributed the urbanization categories of these cells to the adjusted service areas of functioning cells probabilistically. For example, a single cell tower might cover an area that is 60% urban, 30% suburban, and 10% rural. To quantify the overall urbanization level for each cell tower’s service area, we assigned progressively higher weights to more urbanized categories (with “densely urban” receiving the highest weight and “rural” the lowest). We then calculated a weighted sum of these percentages to create a normalized urbanization index between 0 and 1, where higher values indicate more urbanized areas.

The spatial distribution of urbanization is visually very similar to nighttime satellite images of Türkiye<sup>2</sup>, with the notable difference that the northeastern parts of the country show a larger spread of high levels of urbanization index compared to what the nighttime satellite images suggest. This is related to high levels of seasonal mobility in and out of this area. The MNO categorizes urbanization levels of areas that receive high signals seasonally in the western, southern and northern coastal areas as “seasonal densely urban”, “seasonal urban”, and “seasonal suburban”. While creating the

---

<sup>2</sup>See <https://www.sciencephoto.com/media/662978/view/turkey-at-night-satellite-image>, Accessed June 16, 2025.

urbanization index we ignored the seasonality aspect, and only used the urbanization descriptors. In the earthquake region (see the red circles in the upper plot of Figure 2 in the main paper) the index aligns well with the expected urbanization levels. However, for the Northeastern regions of Türkiye, this resulted in an index that overestimates the urbanization levels due to high seasonal mobility to rural areas, which were classified as “seasonal suburban” by the MNO. Since the mobility related to this region is very limited in our dataset, this mismatch does not constitute an issue for our analysis.

## Damage index

To compute the damage index, we assigned a value of 1 to collapsed buildings and discounted values to the other categories (buildings that need to be demolished, heavily damaged buildings, and slightly damaged buildings), with each value weighted to reflect its relative severity compared to collapsed buildings. Then we calculated the weighted sum of all damaged buildings, and divided it by the area spanned by each cell to have a density-based measure of the damaged buildings. Lastly, we applied a min-max normalization, which produced a normalized measure between 0 and 1, where higher values indicate areas with more severe infrastructure damage. This approach allowed us to create a single comparable measure that accounts for the varying degrees of damage across the affected regions, while acknowledging the qualitative differences between damage categories. We recognize that choosing the weights for these categories requires simplifying assumptions, but we show empirically that they are still useful to understand the patterns of displacement. The resulting index varies between 0 and 1, but most of the distribution is between low values of 0 and 0.2. Figure 2 in the main paper clearly shows that there were higher levels of damage in the city centers.

## Measurement of stay locations

In Algorithm 1, we share the pseudo-code of the stay location detection approach, inspired by [1], which is a differential based strategy that tries to identify stationary areas by focusing on the temporal and spatial differences between consecutive GPS traces in individual movements [2]. Our measurement of stay locations differs from [1] in one aspect; they use a single distance check between consecutive points until a threshold is exceeded, whereas we use a double distance check both between consecutive points and between the initial point and the current point. The original algorithm was written for processing GPS data, where the data frequency is higher than CDR. In addition, GPS works on smaller distance thresholds compared to CDR, as the former relies on satellites for data collection, whereas the latter on cell towers. Because of the lower resolution in CDR, there is a chance that consecutive data points will be below the distance threshold even if the individual is actually moving. This issue can potentially exist for GPS data as well, but it has not emerged as an issue in the literature of stay location detections, possibly because the data frequency and accuracy of GPS compensate for this concern. In CDR, without adding a second distance check, some stay locations might capture movements across neighboring cells instead of actual stationary areas. Instead, we required each point to remain within the distance threshold of both the previous point and the stay’s starting location.

The algorithm groups cell towers together as stay locations, but it does not define their shape. For this, we have calculated the convex hull of the Voronoi cells, using the minimal convex polygon that fully encloses all the exterior boundaries. In Figures S1 and S2, we demonstrate how the convex hulls are fit to the same individual CDR trace for different temporal and spatial thresholds. While

---

**Algorithm 1** Stay Detection( $D, distThresh, durThresh$ )

---

**Input:** An MPD dataset  $D$ , distance threshold  $distThresh$ , duration threshold  $durThresh$

**Output:** A set of stays for each user  $AllStays$

```
1:  $AllStays \leftarrow \emptyset$ ;
2: for each unique user  $u$  in  $D$  do
3:    $G \leftarrow \{p \in D : p.userId = u\}$ 
4:   Sort  $G$  by timestamp;
5:    $pointNum \leftarrow |G|, i \leftarrow 0$ ;
6:    $stays \leftarrow \emptyset$ ;
7:    $currentStay \leftarrow null, cumulativeDist \leftarrow 0$ ;
8:    $stayStartPoint \leftarrow null$ ;
9:   while  $i < pointNum$  do
10:     $p_i \leftarrow G[i]$ ;
11:    if  $i = 0$  then
12:      Initialize  $currentStay$  with  $p_i$  properties;
13:       $stayStartPoint \leftarrow p_i.geometry$ ;
14:    else
15:       $distFromPrev \leftarrow Distance(p_i.geometry, p_{i-1}.geometry)$ 
16:       $distFromStart \leftarrow Distance(p_i.geometry, stayStartPoint)$ 
17:       $\Delta T \leftarrow p_i.time - currentStay.startTime$ 
18:      if  $\max(distFromPrev, distFromStart) \leq distThresh$  and  $\Delta T \geq durThresh$  then
19:        Update  $currentStay$  with  $p_i$ ;
20:         $cumulativeDist \leftarrow cumulativeDist + distFromPrev$ ;
21:      else
22:        Finalize  $currentStay$  and add to  $stays$ ;
23:        Initialize new  $currentStay$  with  $p_i$ ;
24:         $stayStartPoint \leftarrow p_i.geometry; cumulativeDist \leftarrow 0$ ;
25:      end if
26:    end if
27:     $i \leftarrow i + 1$ ;
28:  end while
29:  if  $currentStay \neq null$  then
30:    Finalize  $currentStay$  and add to  $stays$ ;
31:  end if
32:   $AllStays \leftarrow AllStays \cup stays$ ;
33: end for
34: return  $AllStays$ ;
```

---

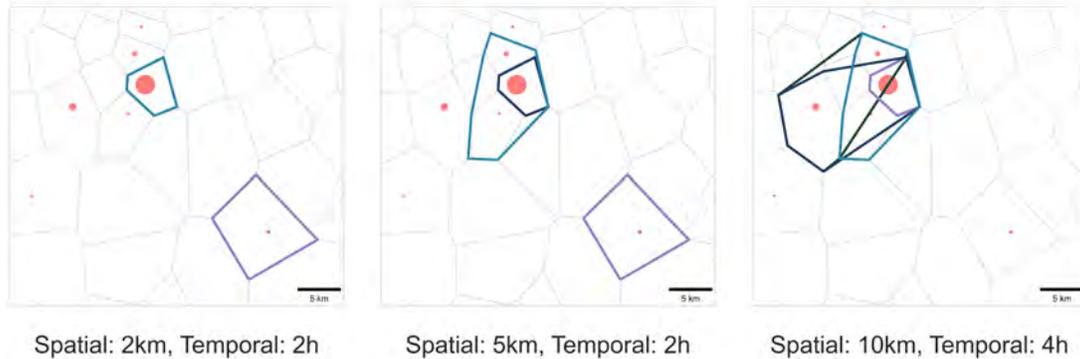


Figure S1: This is an example of fitting stay locations on the individual call records in rural areas depending on different spatiotemporal thresholds. The calls are represented with red dots; dots size indicates number of calls.

choosing the spatial thresholds, considering distances between towers is important when CDR traces are used. For instance, in both examples presented in these figures, increasing spatial thresholds progressively from 2 km to 10 km substantially changes the shape of detected stay locations. In Figure S1, when the temporal threshold increases from 2 hours to 4 hours, one of the stay location disappears, as the individual spent there more than 2 hours, but less than 4 hours.

## Activity spaces

Figures S1 and S2 illustrate how stay locations can overlap in our algorithm. It is also possible that they do not overlap but remain in close proximity to one another. Since our analysis focuses on calculating shifts in people’s locations over time, we define activity spaces as larger areas that group together related stay locations. The spatial distribution of cell towers tends to be non-uniform, which could be a concern for clustering using DBSCAN. We have considered HDBSCAN as an alternative, which builds a hierarchical tree of potential clusters at all density levels and then selects the most stable clusters from this hierarchy by identifying cluster formations that persist across the widest range of density threshold [3]. However, in ASA, we focus on clustering the centroids’ stay locations (not raw tower locations) for individual users. Since stay locations represent areas where users spent significant time, they tend to be more spatially coherent than raw tower locations, making the uniform density assumption of DBSCAN more reasonable at this individual level.

Figures S3 and S4 demonstrate how different combinations of spatial and temporal thresholds in the stay detection algorithm, along with DBSCAN’s maximum distance parameter, result in different activity space representations. In Figure S3, the first and second rows, we observe two distinct activity spaces. If the maximum distance parameter is increased further, these two activity spaces would merge into a single representation. The bottom row illustrates this effect more clearly: in the leftmost plot, setting the distance threshold to 6 km yields two distinct activity spaces despite substantial overlap between them. These two activity spaces merge when larger distance values are applied, as shown in the remaining plots of the bottom row.

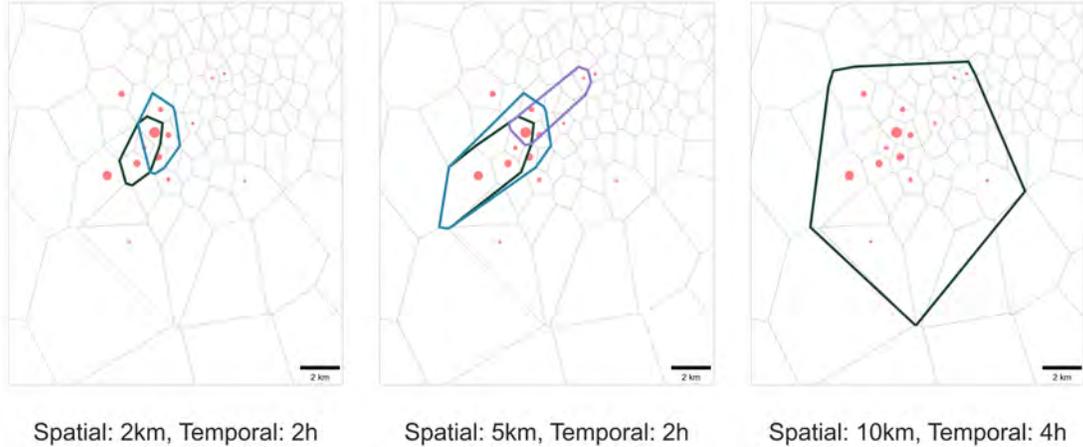


Figure S2: This is an example of how stay locations are fitted on the individual call records in highly urbanized areas. The calls are represented with red dots; larger the dots more calls were made from the associated cell.

Figure S4 demonstrates how parameters of DBSCAN and stay location algorithm interact within a dense urban area. The metropolitan area covered by the city has approximately 10 to 12 km diameter, so using 10 km as spatial threshold creates a single stay location, and a single activity space that covers all urban mobility, whereas using 2 km as the spatial threshold yields a compact activity space with several stay locations. These examples (and many other examples we have inspected) show that generally, the spatial threshold of stay locations is the determining factor for the size and division of activity spaces, whereas maximum distance can play a role in merging activity spaces together when applied at the right level.

## Setting the spatial parameters

The spatial threshold for stay locations should be chosen based on the inherent spatial resolution and characteristics of the mobile data type at hand. For telco data (CDR, xDR, etc.), the threshold must consider the typical distances between cell towers, as setting it too small (e.g., less than 500 meters) may result in overly fragmented single-cell stay locations, while setting it too large loses spatial granularity. GPS data, with its higher precision and frequency, allows for much smaller spatial thresholds, since measurements are more accurate and frequent. The key is to balance capturing meaningful stationary behavior while avoiding both over-fragmentation and loss of spatial detail, with the optimal threshold varying significantly between urban areas (where towers are closer) and rural regions (where towers are more dispersed).

The DBSCAN maximum distance parameter should be set by considering two key factors: the size of stay locations and the urban context of the studied area. First, the parameter must be large enough to effectively cluster related stay locations together, especially when they are in close proximity or have some overlap. Second, it should reflect the urban structure and mobility patterns of the region; in cities where residential areas are in the suburbs and commercial centers are downtown, the maximum distance should be sufficient to capture these cross-city movements. For

most metropolitan areas, setting the parameter to match the city’s diameter (typically 10-12 km for medium-sized cities) ensures that activity spaces capture meaningful intra-city mobility patterns. However, if the parameter is set too large, distinct activity spaces may merge inappropriately, losing important spatial granularity for the analysis. Checking a range of thresholds will provide a more comprehensive view.

In disasters covering very large areas, as was the case with the 2023 Türkiye–Syria earthquakes, setting appropriate spatial thresholds for stay locations and maximum distances for DBSCAN was a challenge. The area affected by the earthquake covered both rural and urban places where cell towers can be as far as 10 km from their neighbors, or as close as 20 meters. Choosing 2 km as spatial threshold creates single cell stay locations in rural areas (see Figure S1 as an example), but merges multiple cells in highly urbanized areas (see Figure S2).

Activity spaces are expected to account for movements that are beyond residential areas, ideally capturing the urban mobility including the regular movements between work and home, but also other frequented trips to urban centers for leisure activities. The MPD sources are noisy, and they might include signals that are not essential to the measurement of activity spaces. An individual might be receiving signals here and there, from places that are not essential to their lives, or they might be passing through these areas without meaningful stays. Therefore, researchers working with MPD try to mitigate these issues while estimating the activity spaces with techniques such as anchoring the meaningful locations [4, 5, 6]. ASA differs from this literature, as it does not employ spatial anchoring to predetermined home and work locations. Instead, our approach uses a data-driven method where stay locations are identified based on spatial and temporal thresholds, then clustered using DBSCAN to form activity spaces. In this way, we hope to account for regions that are of high importance to individuals.

Depending on the spatiotemporal thresholds of stay locations and DBSCAN maximum distance, the activity spaces might be over or underestimated for some individuals. For instance, in Figure S4, the top plots showcase a minimalist approach to the measurement of activity spaces, whereas bottom plots demonstrate a rather maximalist approach. Here, what is more important for us is not the most correct representation of the activity spaces, but to correctly account for their shifts after the disaster. The approach we follow is adequate for measuring the activity space shifts, which remains as the key function of activity spaces in our methodology.

## Relevance of the stay locations

In Figure S5, we show how activity spaces before and after the disaster are intersected to calculate the relevance scores of the post-disaster stay locations for different combinations of spatial and temporal thresholds of stay detection algorithm, and the maximum distance of DBSCAN. In the example presented in Figure S5, the person was not displaced, but the stay locations, hence the activity spaces, have shifted after the earthquake. These subtle changes are not found when the stay locations are calculated at a 10 km level (rightmost plot in Figure S5). After we calculate the relevance score at the activity space level, the stay locations that made up the post-disaster activity spaces inherit this relevance score. For the leftmost and rightmost plots in Figure S5, the whole area of pre-disaster activity space overlaps with the post-disaster activity space, so at these configurations, the person will be marked as staying in familiar places after the disaster, whereas in the middle plot, the intersection is less than a hundred percent, so there will be some loss of familiarity. However, as relevance scores are calculated in a binary fashion, this level of loss of familiarity is tolerated by our measurements.

## Displacement detection

For detecting displacements, we adapted the migration detection algorithm of Chi et al. [7]<sup>3</sup>. The algorithm uses a segment-based clustering approach similar to DBSCAN, where contiguous periods in which individuals remain in the same location are grouped together, allowing for some gaps to occur in between. The migration event is defined as the movement between two stable location segments. The algorithm also infers the date of migration, and calculates an associated uncertainty value for it. The segment-based approach was shown to outperform frequency-based methods [7] in accurately detecting migration events, and was used to detect conflict-induced displacements in Afghanistan by TMB [9]. In Figures S6 and S7, we explain how  $\varepsilon$  controls the segment formation, both in the context of ASA and TMB, respectively. The migration detection algorithm also infers the migration date, alongside with the migration event. In Figure S6, we show the inferred migration date, the date of departure from the origin, and the date of arrival to destination.

The original algorithm is meant to detect regular migration between administrative areas (districts, cities, etc.). TMB uses the daily modal location for these calculations. This way, individuals are assigned to their modal locations for each day. At district level, using a single daily modal location to represent individual’s location is a similar idea to using their home locations. We measured displacements with TMB at cell-tower for discovering origins and the destinations hotspots, but as we showed also in sensitivity analysis the algorithm struggles to find stable segments (demonstrated in Figure S7) especially for higher values of  $k$ .

In ASA, individuals can have multiple stay locations within the same night. Each of these stay locations can have different levels of relevance scores too. Therefore we took the average relevance score of detected stay locations for each night to have something similar to daily modal locations. However, this score is a summary of familiarity of multiple locations. In Figure S8, we demonstrate the impact of using relevance scores instead of the stay locations in the segment-based algorithm. It serves a very similar function to using the administrative boundaries (instead of cell towers) in the segment-based algorithm.

## Model comparison

### TMB at district level

In the main article, we compared ASA to TMB calculated at cell tower level. We chose this model because of its similarity in terms of segment-based calculations, and its ability to visualize origin and destination at cell tower level. However, as we previously noted, TMB is meant to measure displacements between administrative boundaries. Therefore, here we present how origins and destinations look like when they are calculated with TMB at district level. When we run the migration detection algorithm at district level by setting  $\varepsilon$  and  $k$  at 14 days, we detect 5,700 DPs. In Figure S9, we show the number of DPs per district by their origins and destinations. We see the general pattern of exodus from the earthquake region, and nearby cities (like Mersin) appearing as a prominent destination, but the origin and destination hotspots are not observable.

---

<sup>3</sup>The algorithm was built using the data structure of Turi Create, which was specifically used for speeding up calculations with parallel computing. The library can be downloaded here: <https://github.com/apple/turicreate/>. Turi Create is not maintained since 2023, so it was not possible to use it within our programming framework. We replaced it with a pandas dataframe [8].

## Frequency-based approach

One of the reasons why the subtle displacement patterns do not emerge on origin destination density plots with TMB at cell tower level is due to the low density of residential areas compared to the stay locations. We suspected that the segment-based approach might be too demanding on the definition of displacements on location data, so in addition to TMB, we also measured displacements with a simple frequency-based approach.

For frequency-based measures, we calculated the home locations before and after the disaster for all users. The home locations are defined as the most frequent cell towers used at night (between 10 PM and 7 AM). We filtered the people who were residing in the disaster area before the earthquake, and have changed their most frequent tower after the earthquake.

We tried three ways of defining displacements:

1. If the most frequent cell tower has changed after the earthquake
2. If the district of the most frequent cell tower has changed
3. If the city of the most frequent tower has changed

According to the first definition, we have 54,000 displaced people (DPs), which is five times more than ASA and TMB. This is likely because such a simple definition (change in the most common night tower) captures movements that are not related to displacements. According to the second definition, there are 27,000 DPs and according to the last one there are 15,000 DPs. In Figure S10, we show the detected origin and destination hotspots following the first definition, in Figure S11 following the second definition and in Figure S12 following the last definition. The first definition emphasizes people who have changed their locations within the same city. Therefore the density of destinations are emphasized in and around the city centers. The origins hotspots in Figure S10 are remarkably similar to the origins calculated by ASA as we shared in Figure 4. The destinations have similar patterns, but any type of agglomeration of Syrian refugees around Temporary Accommodation Centers (TACs), are not as clearly visible as ASA could demonstrate.

To deal with issues of sparse and noisy data, we also defined displacement as the shift of district and city of the most frequently used tower at night. In Figure S11, we show the distribution of origins and destinations for people who changed districts. This definition reduces the emphasis of the destinations around the city centers. Notably for the destinations of the Syrian DPs, there is still some emphasis around the border area, which includes a handful of TACs. But again the displacement patterns of Syrians do not as clearly indicate that the Syrians chose TACs as one of the primary destinations. Lastly, when we focus on cross-city movements in Figure S11, we see that only Mersin, Adana, and Kilis appear as main destination areas for Syrian DPs, and only Mersin for Turkish DPs. In short, we could not replicate the findings of ASA as clearly using any of the frequency-based approaches.

## Additional Experimental results

### Displacement dates and returnee rates

The migration detection algorithm detects the start and end dates at the origins and destinations, as well as estimated migration dates for each migration event. Due to noise and data sparsity,

finding the exact date of migration events remains a challenging task. The migration detection algorithm decides on the migration date as the day between the end of the origin segment and the start of the destination segment “that minimizes the number of ‘misclassified’ days, i.e., the number of days when the migrant appears at destination before the migration date and days when the migrant appears at home after the migration date” [7, p. 6]. The origin end date and destination start date are the last day of continuous presence at the origin, and the first day of continuous presence at the destination, respectively.

In Figure S13, we compare TMB to ASA with respect to the number of displacements detected each day after the earthquake. The figure shows the differences between the number of DPs when using the migration date, origin end date or destination start date as the displacement date. As before, we use three alternatives for the TMB method; the city, the district and cell boundaries. Both TMB and ASA are calculated with the same parameters as we calculated the origins and destinations. ASA is able to detect displacements at an earlier date than TMB at district and city level. This is expected as we define displacements as loss of habitual areas, which happen faster than changing cities or districts. Such ruptures from habitual living areas are the first signals of displacements, and missed in district or city level calculations.

It is important to pay attention to how the displacement date is defined. Once an origin and a destination are determined for an individual, it is possible for that individual to occasionally receive signals from the origin after the displacement. The detection algorithm infers “migration date” as the date that minimizes such receptions, but this can be noisy [7]. We use the “origin end date” as the preferred approximation for the displacement date, as using the “destination start date” can understate the date of displacement for DPs who lost signals along the way. With this choice, ASA suggests that an overwhelming ratio of displacements happened immediately in the first day. Although, changing districts or cities may have taken a couple of more days (as suggested by TMB) people lost their habitual living spaces very rapidly in the beginning of the disaster.

Within the given period, TMB detected 5 % returnees at city, and 4 % at district level. ASA detected 4 % returnees in the same period. Similar to displacements, the selection of  $k$  is very influential on how many returnees are detected. If  $k$  is set to 3 days, ASA detects around 16 % returnees whereas this drops to 4 % at 14 days. For ASA, relevance threshold has also small but meaningful impact where the larger relevance thresholds decrease the returnee rates. We have not seen a large impact of spatiotemporal thresholds for the calculation of the returnee rates.

## References

- [1] Li, Q., Zheng, Y., Xie, X., Chen, Y., Liu, W., Ma, W.-Y.: Mining user similarity based on location history. In: Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems, pp. 1–10 (2008)
- [2] Pérez-Torres, R., Torres-Huitzil, C., Galeana-Zapién, H.: Full on-device stay points detection in smartphones for location-based mobile applications. *Sensors* **16**(10), 1693 (2016)
- [3] Campello, R.J., Moulavi, D., Sander, J.: Density-based clustering based on hierarchical density estimates. In: Pacific-Asia Conference on Knowledge Discovery and Data Mining, pp. 160–172 (2013). Springer
- [4] Ahas, R., Silm, S., Järv, O., Saluveer, E., Tiru, M.: Using mobile positioning data to model locations meaningful to users of mobile phones. *Journal of urban technology* **17**(1), 3–27 (2010)

- [5] Järv, O., Ahas, R., Witlox, F.: Understanding monthly variability in human activity spaces: A twelve-month study using mobile phone call detail records. *Transportation Research Part C: Emerging Technologies* **38**, 122–135 (2014)
- [6] Silm, S., Ahas, R.: Ethnic differences in activity spaces: A study of out-of-home nonemployment activities with mobile phone data. *Annals of the Association of American Geographers* **104**(3), 542–559 (2014)
- [7] Chi, G., Lin, F., Chi, G., Blumenstock, J.: A general approach to detecting migration events in digital trace data. *PloS one* **15**(10), 0239408 (2020)
- [8] pandas development team, T.: Pandas-dev/pandas: Pandas. doi:10.5281/zenodo.3509134. <https://doi.org/10.5281/zenodo.3509134>
- [9] Tai, X.H., Mehra, S., Blumenstock, J.E.: Mobile phone data reveal the effects of violence on internal displacement in Afghanistan. *Nature Human Behaviour* **6**(5), 624–634 (2022)

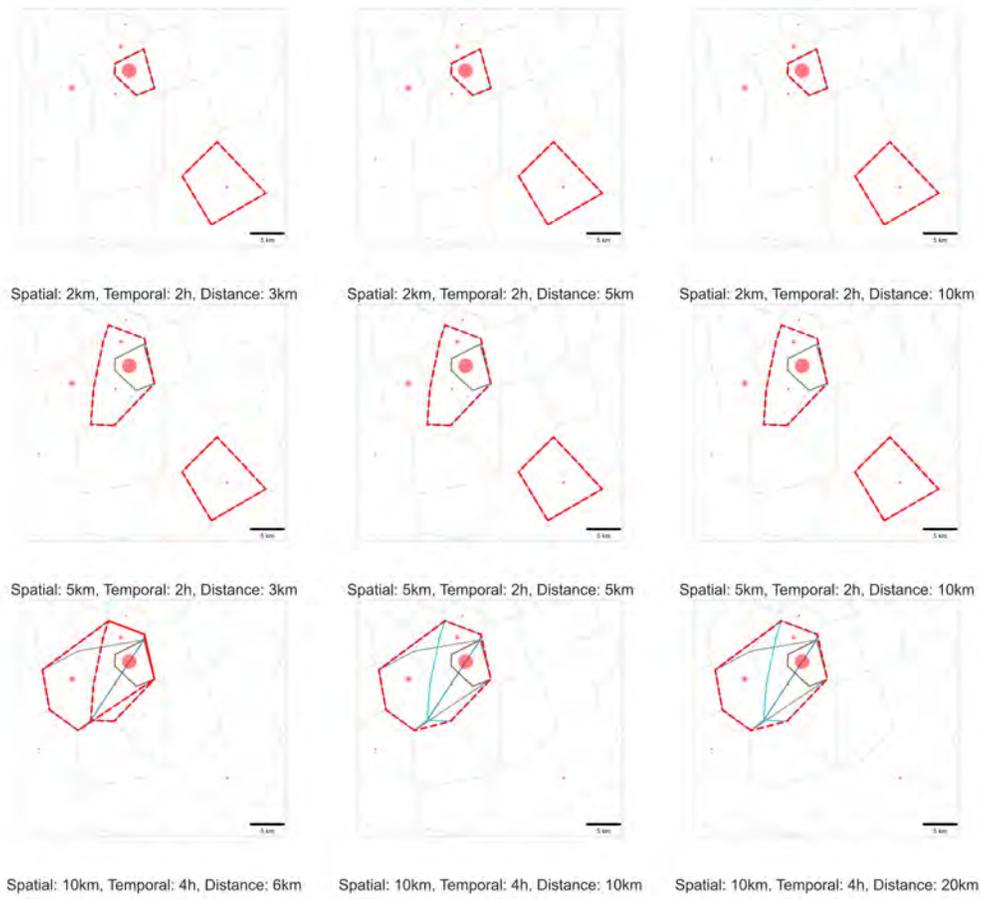


Figure S3: This is an example of how activity spaces are fitted using the stay locations in rural areas. The calls are represented with red dots. Activity spaces are represented by dashed red lines.

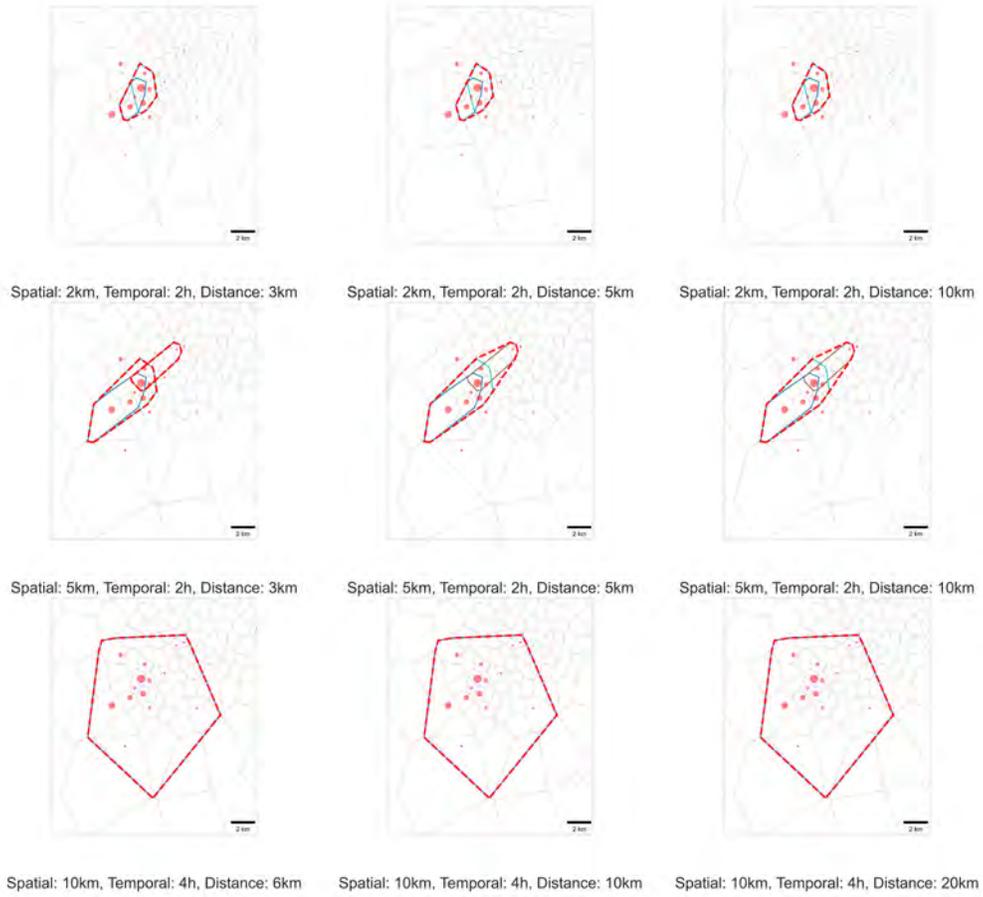


Figure S4: This is an example of how activity spaces are fitted using the stay locations in highly urbanized areas. The calls are represented with red dots. Activity spaces are represented by dashed red lines.

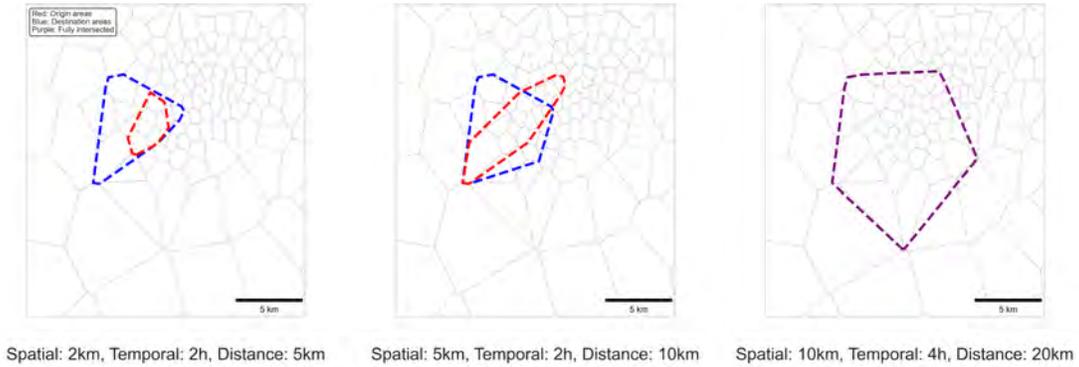


Figure S5: This is an example of how pre-disaster and post-disaster activity spaces are intersected for a non-displaced person. Red dashed polygons represent the pre-disaster activity spaces and the blue dashed polygons represent post-disaster activity spaces. When they fully intersect, we represent it with a purple color.

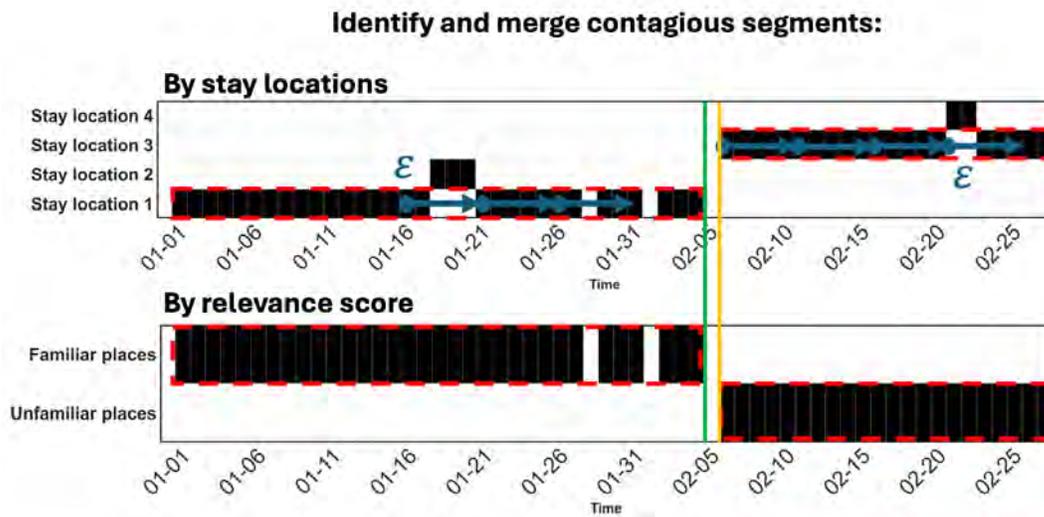


Figure S6: The red dashed boxes demonstrate the continuous segments. The upper figure shows how the parameter  $\epsilon$  controls what constitutes a continuous segment. If there are neighboring segments in the same location, they are merged by the algorithm. The green line is the date of departure from the origin, whereas the orange line is the date of arrival to destination, as well as the inferred migration date by the algorithm.

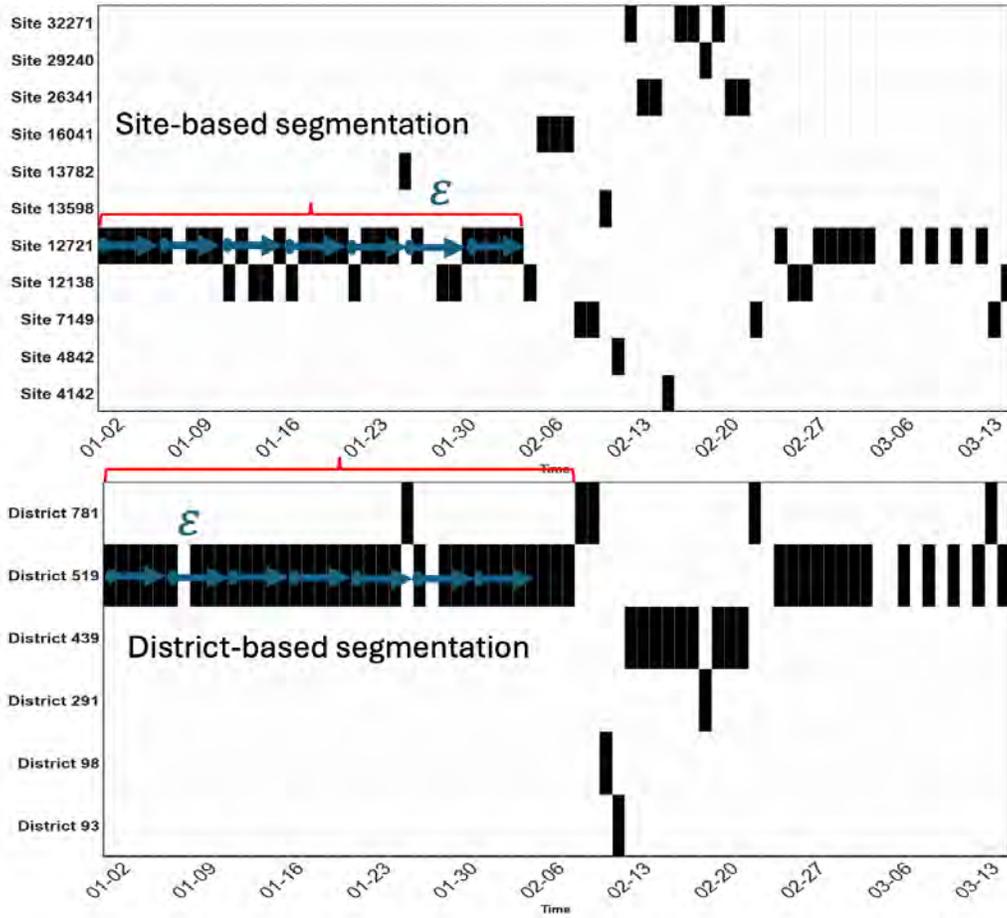


Figure S7: In this figure, we show the difference between using districts and sites in the detection of displacement in TMB approach.

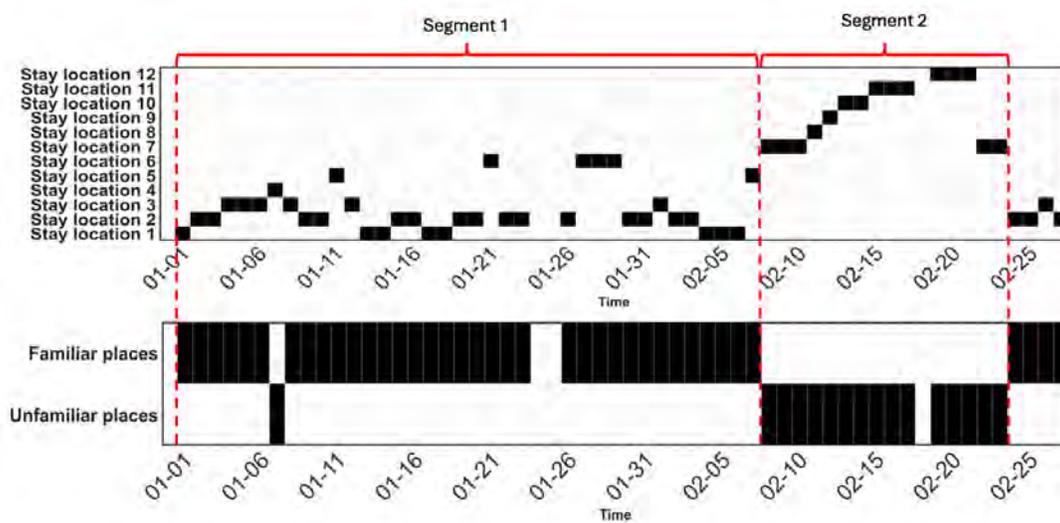


Figure S8: In this figure, we show how using binary relevance score helps detection of displacements with the segment-based algorithm.

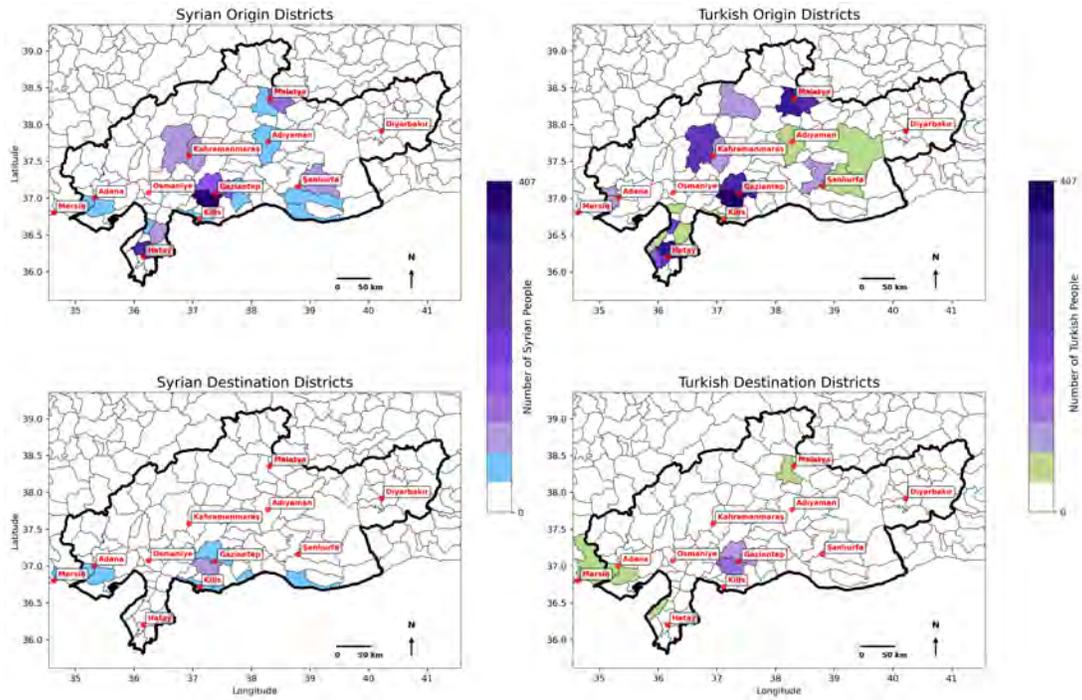


Figure S9: The distribution of origin (top) and destination (bottom) districts of Syrian (left) and Turkish (right) DPs calculated by TMB. Any value less than 37 is not colored. Any value above 37 and less than 74 DPs are colored to light blue for Syrians, and light green for Turkish. Higher than 74 DPs are marked by shades of purple

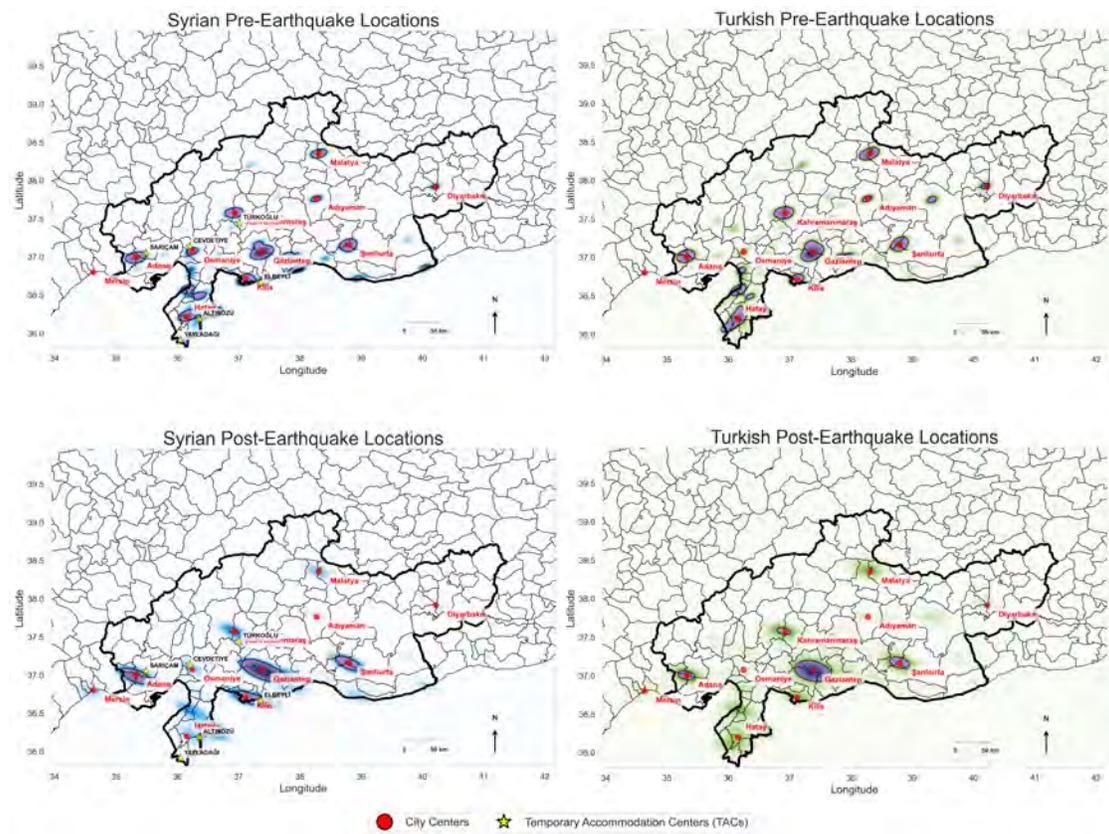


Figure S10: The distribution of origin (top) and destination (bottom) districts of Syrian (left) and Turkish (right) DPs calculated by the first definition of the frequency-based approach (anyone who has changed cell towers are displaced.). Higher values are shaded with purple.

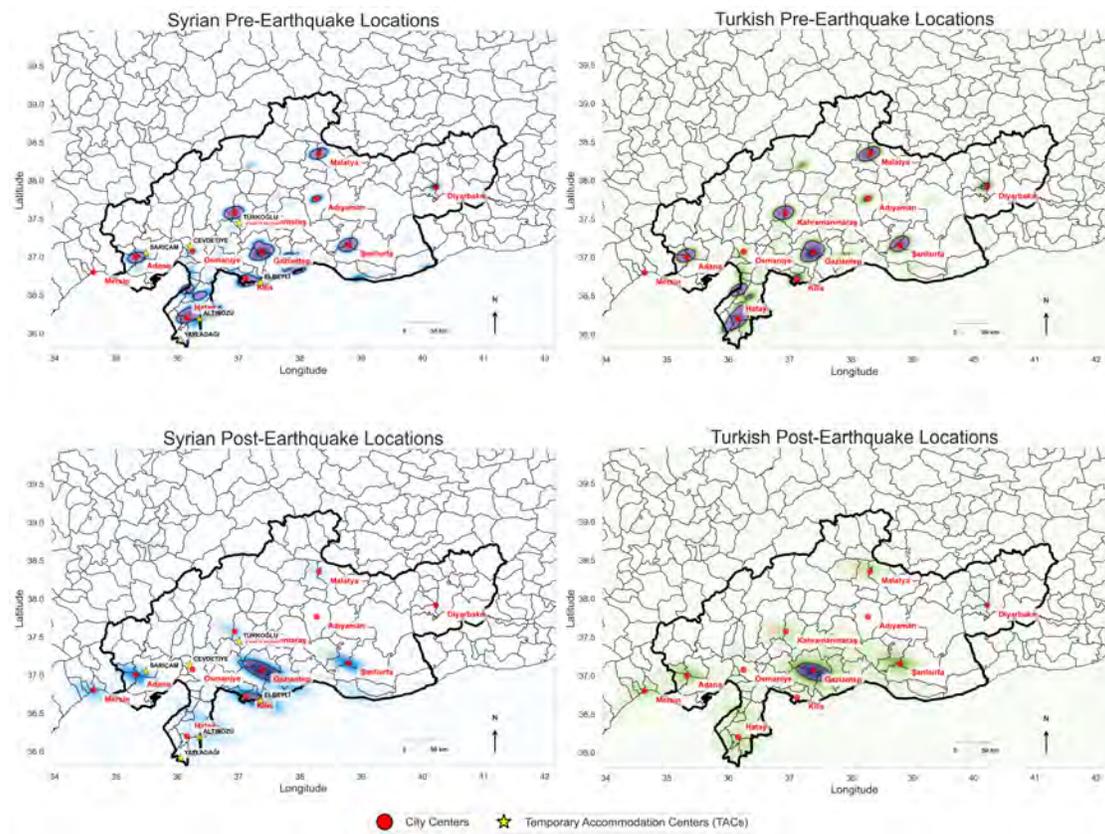


Figure S11: The distribution of origin (top) and destination (bottom) districts of Syrian (left) and Turkish (right) DPs calculated by the second definition of the frequency-based approach (anyone who has changed districts are displaced.). Higher values are shaded with purple.

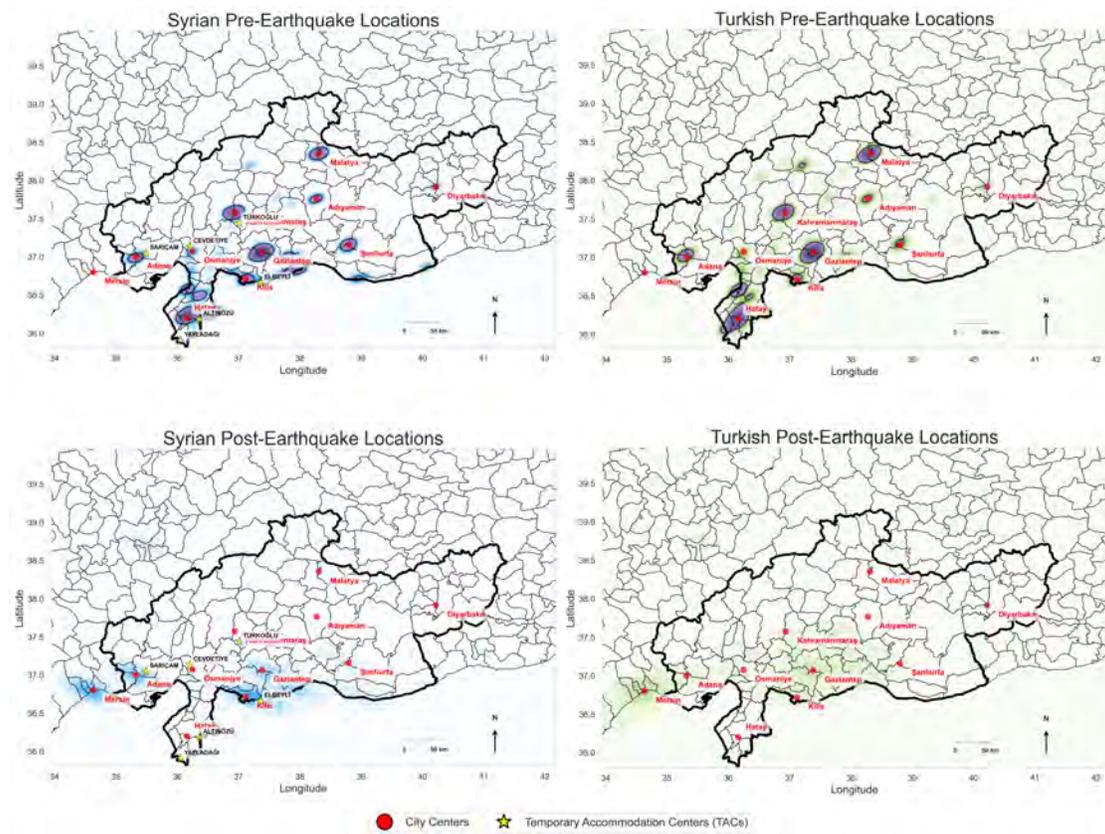


Figure S12: The distribution of origin (top) and destination (bottom) districts of Syrian (left) and Turkish (right) DPs calculated by the third definition of the frequency-based approach (i.e. anyone who has changed cities is displaced.) Higher values are shaded with purple.

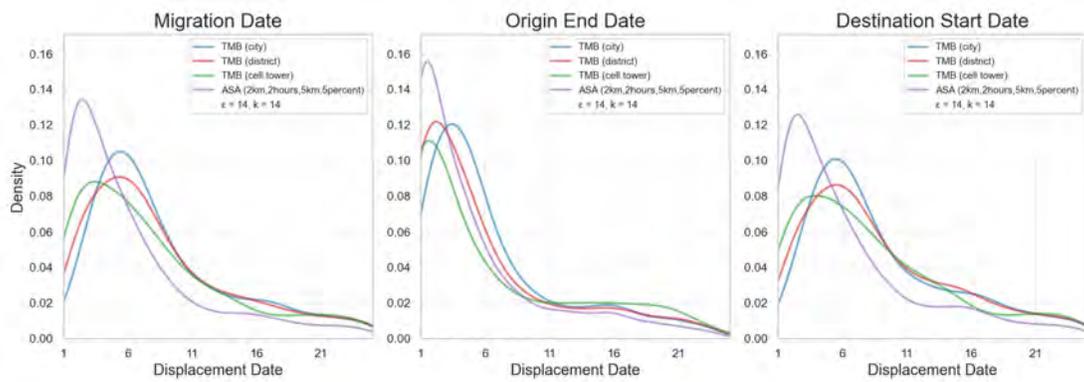


Figure S13: The number of DPs for each day after the earthquake calculated via TMB and ASA for three different candidates for displacement date calculated by the migration detection algorithm; migration date, origin end date, and destination start date, respectively.