

Research Article

Comparative Analysis of Travel Patterns from Cellular Network Data and an Urban Travel Demand Model

Nils Breyer , Clas Rydergren , and David Gundlegård

Department of Science and Technology, Linköping University, Linköping, Sweden

Correspondence should be addressed to Nils Breyer; nilsbreyer@liu.se

Received 6 June 2019; Accepted 16 January 2020; Published 13 February 2020

Academic Editor: Rakesh Mishra

Copyright © 2020 Nils Breyer et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Data on travel patterns and travel demand are an important input to today's traffic models used for traffic planning. Traditionally, travel demand is modelled using census data, travel surveys, and traffic counts. Problems arise from the fact that the sample sizes are rather limited and that they are expensive to collect and update the data. Cellular network data are a promising large-scale data source to obtain a better understanding of human mobility. To infer travel demand, we propose a method that starts by extracting trips from cellular network data. To find out which types of trips can be extracted, we use a small-scale cellular network dataset collected from 20 mobile phones together with GPS tracks collected on the same device. Using a large-scale dataset of cellular network data from a Swedish operator for the municipality of Norrköping, we compare the travel demand inferred from cellular network data to the municipality's existing urban travel demand model as well as public transit tap-ins. The results for the small-scale dataset show that, with the proposed trip extraction methods, the recall (trip detection rate) is about 50% for short trips of 1–2 km, while it is 75–80% for trips of more than 5 km. Similarly, the recall also differs by a travel mode with more than 80% for public transit, 74% for car, but only 53% for bicycle and walking. After aggregating trips into an origin-destination matrix, the correlation is weak ($R^2 < 0.2$) using the original zoning used in the travel demand model with 189 zones, while it is significant with $R^2 = 0.82$ when aggregating to 24 zones. We find that the choice of the trip extraction method is crucial for the travel demand estimation as we find systematic differences in the resulting travel demand matrices using two different methods.

1. Introduction

In order to meet an increasing travel demand and the need to reduce environmental impacts, today's traffic system needs to become more efficient. To make well-informed decisions when improving the traffic system, a detailed understanding of human mobility is needed. This calls for comprehensive information on travel patterns and the actual travel demand, which today is difficult to obtain [1, 2].

Cellular network data are seen as a promising data source which can be used to augment both traffic management [3] and traffic planning [4, 5]. As a large-scale data source, it can give new insights on mobility with all travel modes. It is also easier to keep up to date than travel surveys. Several studies have investigated the possibilities to infer travel demand from cellular network data. Estimating travel demand from cellular network data involves a number of

processing steps. Few studies have used real-world cellular network data and compared all outputs generated in these processing steps to other existing data. Therefore, there is no comprehensive understanding of the quality and potential problems that can arise in the data processing steps.

This paper aims to analyse the potential and limitations of travel demand inference from large-scale cellular network data. We propose a process to obtain travel patterns from cellular network data, consisting of two alternative algorithms to extract trips and a method to infer time-sliced travel demand. In order to evaluate the trip extraction performance in terms of recall and precision, we compare trips extracted from cellular network data to trips obtained from GPS tracks collected on the same mobile device. Using this method, we can analyse which types of trips can be detected from cellular network data by applying the two trip extraction methods. We infer the time-sliced travel demand

for a Swedish municipality using a large-scale cellular network dataset and investigate the impact of the trip extraction method used on the resulting travel demand. In a comparative analysis of the estimated travel demand and the municipality's existing urban travel demand model and other data sources, we evaluate the travel demand inferred using cellular network data.

The remainder of the paper consists of an overview of previous research in the field in Section 2, followed by a description of the methods used for trip extraction as well as travel demand inference in Sections 3 and 4. The methods have been applied to the datasets presented in Section 5. The results of the analysis are presented in Section 6 and discussed in Section 7.

2. Previous Research

The common way in traffic planning to model the traffic system follows the structure of the four-step model [6]. The four steps are modelled based on census data, surveys, and calibrated against traffic counts. Travel demand in these models is modelled using a gravity model [7]. Shortcomings arise from the underlying behavioural assumptions, which are a simplification of reality. A major issue is that the quality and the level of detail of these traffic models depends on the input data which are expensive to collect.

Research studies investigating the potential uses of cellular network data as a new data source for traffic analysis have been ongoing for at least a decade. Caceres et al. [8] have been among the early studies showing the potential for travel demand estimation. Several metastudies list a large number of case-studies with focus on travel behaviour [5], specific methods like machine learning [9], and cellular network data in context of other new data sources like global positioning system (GPS) tracks and smartcard data from public transit systems [4]. A summary of the potential and obstacles of using cellular network data for traffic analysis has been given by von Mörner [10].

Cellular network data are collected also when a user is not travelling. In a traffic analysis context, however, physical movements are of interest. A change of antenna can be caused by actual, physical movements but also happens for other reasons. Therefore, most studies in the area involve some kind of trip detection step which aims to distinguish moving periods (trips) and standstill periods (stops). A common approach which is used by Alexander et al. [11], Calabrese et al. [12], and Graells-Garrido et al. [13], among others, is to detect stops by scanning the data of one user and find periods that fulfill certain criteria indicating that the user has not moved for some time. Some studies make use of the fact that people often visit the same places several times [1] by identifying important places [11], specifically home and work [14]. Recently, more advanced trip extraction methods using clustering and other methods commonly referred to as machine learning have been used [15]. Breyer et al. [14] presented a trip extraction method that instead of stopping identifies movements using different indicators. The impact of the choice of the trip extraction method when inferring travel demand is discussed by Gundlegård et al. [16].

Trips extracted from cellular network data can be used as a standalone source of travel demand data [8] or as a way to augment existing travel demand models with additional input data [17]. With today's ubiquity of mobile devices, cellular network data are a natural source to understand large-scale travel patterns. It is challenging, on the contrary, to use cellular network data to obtain additional metadata about the travel patterns besides their description as flows in time and space. Alexander et al. [11] and Widhalm et al. [15] have made some attempts to classify trip purposes and activities, and Bachir et al. [18] and Graells-Garrido et al. [13] have investigated possibilities to infer travel demand for each travel mode. Socioeconomic data for individuals are not available in cellular network data for privacy reasons. Hence, linking travel patterns to socioeconomic attributes is challenging and needs to be done on an aggregated level, as proposed by Calabrese et al. [19].

To judge the quality of the travel demand inference from cellular network data, different approaches and metrics are used in the literature to validate and compare results to other data sources. To validate the trip extraction, a trip-by-trip comparison to another data source like GPS tracking gives the most detailed understanding of the trip detection quality. Only a few studies like Fillekes [20] and Breyer et al. [14] have done this, as the collection of a second data source requires additional action by the user and can therefore only be done in a small scale. To verify trips extracted for a large-scale dataset, some studies rather look at aggregated statistics [21]. Such statistics could, for example, be the trip length distribution or the number of trips per day and user which can be compared to surveys.

Validating the inferred travel demand is difficult, as the ground truth is unknown and cannot be fully observed. Several studies compare the calculated travel demand to existing traffic models which have been built using other data sources and describe the correlation using R^2 values or the root mean square error (RMSE) [22, 23]. A difficulty here is that the correlation of the origin-destination flows (OD flows) depends on the setup and the size of the origin/destination zones used, as illustrated by Batran et al. [22]. Also, the correlation of OD flows disregards their spatial structure: if, for example, flow is assigned to a neighbouring destination zone, it would have the same effect on the correlation measure as if the flow is assigned to a different destination zone very far away. Pollard et al. [24] discuss new similarity measures to overcome this problem. Another approach is to use a traffic model to estimate link flows from the inferred travel demand as done by Iqbal et al. [25], which allows validating against actual traffic counts.

3. Trip Extraction

Cellular network data refer to events which are triggered in the mobile operator's network. In order to make use of the cellular network data to analyse travel patterns, the first step is to distinguish movements (trips) and stationary periods. For a given user, the raw data consist of a list of events $e_i = (t_i, c_i)$, where t_i is the time of the event and c_i is the antenna where the event has occurred. Table 1 gives further

TABLE 1: Notations used for the trip extraction methods.

Notation	Description
T	Total number of minutes of raw data for a user
$t \in 1, \dots, T$	Time (in minutes since start of the data)
$e_i = (t_i, c_i)$	Event i in the raw data at time t_i and antenna c_i
$p(t)$	Estimated position (antenna) at time t
$d(a, b)$	Euclidean distance between antennas a and b in km
(t_s, t_e)	Trip start and end time

notations used in this chapter, and Table 2 presents the parameter setup used in this paper, which is also further explained in Section 3.5.

Two methods to extract trips are presented in this section. Before running these algorithms, the raw events are preprocessed, as described in Section 3.1. The algorithm `STOP` then detects trips by first identifying stationary periods, while the algorithm `MOVEMENT` directly detects movements using movement indicators. The output of the trip extraction is a list of trips for a given user. Each trip has a start time and an end time as well as a start antenna and end antenna.

3.1. Preprocessing. We propose a preprocessing of the events which estimates the most likely position of the user for each moment in time. This is to simplify the trip extraction process and to get a more accurate time estimate when extracting trips. Events caused by periodic updates, for example, occur with constant time intervals independent of the movement of the user. Therefore, it is reasonable to assume that the user moved in-between two events. The preprocessing is done as follows:

- (1) From midnight until the first event of the day, assume the user to be at the antenna of the first event.
- (2) Estimate the user position using all pairs of consecutive events (see below). If there are multiple events during a minute, only the antenna that appears most often is kept.
- (3) From the last event to the end of the day, assume the user to be at the antenna of the last event.

Given two consecutive events in time $e_1 = (t_1, c_1)$ and $e_2 = (t_2, c_2)$, where t_i is the time and c_i is the antenna where the event occurred, and we define the user position as

$$p(t) = \begin{cases} c_1, & \text{for } t_1 \leq t < \max\left\{t_1 + \frac{1}{2}(t_2 - t_1), t_2 - T_{\Delta \max}\right\}, \\ c_2, & \text{for } \max\left\{t_1 + \frac{1}{2}(t_2 - t_1), t_2 - T_{\Delta \max}\right\} \leq t < t_2, \end{cases} \quad (1)$$

where $T_{\Delta \max}$ is a parameter that limits how much back in time the switch between antennas can be moved at most.

3.2. Stop-Based Trip Extraction. A simple approach to extract trips is by detecting the stops that a user makes. This approach has been used, for example, by Calabrese et al. [12]. Using the preprocessed user positions, the user's stops are

TABLE 2: Parameters used in the trip extraction methods.

Parameter	Value
Preprocessing	
$T_{\Delta \max}$	Maximum move of antenna switch
STOP	
D_{\max}	Maximum distance during stop
T_{\min}	Minimum dwell time
MOVEMENT	
α	Speed weight
β	Efficiency weight
T_V	Speed window
T_E	Efficiency window
v_{\max}	Maximum speed
M_L	Movement threshold, low
M_H	Movement threshold, high
D_{\min}	Minimum distance

detected. Finally, the user's trips are extracted: A trip starts when a stop ends and ends when the next stop starts.

In the stop detection (see Algorithm 1), a stop continues as long as the next position is considered to be close to the previous positions included in the stop. A distance threshold D_{\max} is used as the maximum distance between the positions of an ongoing stop. Once the position of the user moves outside of the current stop area, the stop has ended and is saved if it fulfilled the minimum dwell time T_{\min} . If the stop was shorter, it is disregarded and is considered to be part of the ongoing trip. In this paper, the distance threshold $D_{\max} = 1$ km and a minimum dwell time of $T_{\min} = 40$ min has been used.

3.3. Movement-Based Trip Extraction. In contrast to the `STOP` algorithm described in the previous section, the `MOVEMENT` algorithm extracts trips by detecting movements instead of detecting stops first. A more detailed description has also been given in Breyer et al. [14]. Using the user positions resulting from preprocessing as described in Section 3.1, the *movement indicator* $M(t)$ is calculated for each minute t . $M(t)$ is calculated as the weighted average of the *speed indicator* $0 \leq V(t) \leq 1$ and the *efficiency indicator* $0 \leq E(t) \leq 1$ with α and β being the weights:

$$M(t) = \alpha V(t) + \beta E(t). \quad (2)$$

The speed indicator $V(t)$ is defined as

$$V(t) = \min\left[\frac{v(t)}{v_{\max}}, 1.0\right]. \quad (3)$$

Here $v(t) = (d_v(t)/2T_V)$ is the speed estimated using $d_v(t)$, the distance between antenna positions $p(t - T_V)$ and $p(t + T_V)$ (T_V is the window size parameter). To prevent very high speeds from gaining too much influence, the speed indicator is limited by the v_{\max} parameter.

The efficiency indicator $E(t)$ is defined as

$$E(t) = \frac{d_s(t)}{d_a(t)}, \quad (4)$$

where

```

Data: user position  $p(t)$  for every minute  $0 \leq t \leq T$ 
Result: list of stops  $S$ 
 $L := [p(0)]; t_s := 0; t_e := 0$  /* Initialize with stop */
for each  $t \in 1, \dots, T$  do
  if  $\forall p' \in L: d(p(t), p') \leq D_{\max}$  then /* Within distance? */
     $L := L \cup [p(t)]; t_e := t;$  /* Stop continues */
  else
    if  $t_e - t_s \geq T_{\min}$  then /* Stop long enough? */
       $S := S \cup [(t_s, t_e)];$  /* Save valid stop */
    end
     $L := [p(t)]; t_s := t; t_e := t;$  /* Start new stop */
  end
end

```

ALGORITHM 1: Stop detection as used by the STOP trip extraction algorithm. A stop ends once the next position is not in distance of all positions belonging to the current stop (L).

$$d_s(t) = d(p(t - T_E), p(t + T_E)),$$

$$d_a(t) = \sum_{t - T_E \leq t' < t + T_E} d(p(t'), p(t' + 1)). \quad (5)$$

This indicator serves as a filter against switching between neighbour antennas that is often caused by other reasons than physical movements of the mobile phone (network balancing and variations in the signal strength). It compares the straight-line distance within a time window defined by the parameter T_E to the travelled distance (including all antennas used along the way). Note that the $E(t)$ does not exceed 1 as $d_s(t) \leq d_a(t)$ at all times.

Given that $M(t)$ has been calculated for every t , a trip is defined using a low and a high threshold $M_L < M_H$. Any timespan (t_s, t_e) for which all of the following criteria is met generates a trip:

- (i) $M(t) > M_H$ is fulfilled for some t within the trip
- (ii) $M(t) > M_L$ is fulfilled for all t within the trip
- (iii) The trip distance (including via antennas) exceeds D_{\min}

The minimum distance requirement is an additional measure against short hops between antennas being considered as a trip. The start and end time of the trip is when the first and last, respectively, change of antennas takes place within this timespan according to the estimated user position (see Section 3.1).

3.4. Trip Extraction Example. The trip extraction using both algorithms is illustrated with an example of a morning commuting trip for a given user (see Figure 1). Figure 2 shows that STOP and MOVEMENT detect the start of the trip both at the same time but differ in the detected end time of the trip. While the trip detected by STOP ends earlier, the trip detected by MOVEMENT also includes a switch back and forth between two antennas at the destination, which most likely is not related to a physical movement. The underlying indicators used in the MOVEMENT algorithm for the same morning are shown in the bottom timeline in Figure 2.

After the actual trip is finished, there is a small spike in the indicators, which is an example of a situation where the minimum distance threshold is not fulfilled, and therefore, no trip is extracted by STOP as well as MOVEMENT.

3.5. Parameter Setup. All parameter values used in this paper for the trip extraction algorithms above are presented in Table 2. These values have been set to fit the data characteristics of both datasets used in this paper (see Section 5). We have tested different values for most parameters using a small GPS validation dataset (see Section 5.1) and inspected the results manually. A systematic calibration has not been done, given the nonrepresentativeness of the dataset and a high risk of overfitting given the small size of the dataset.

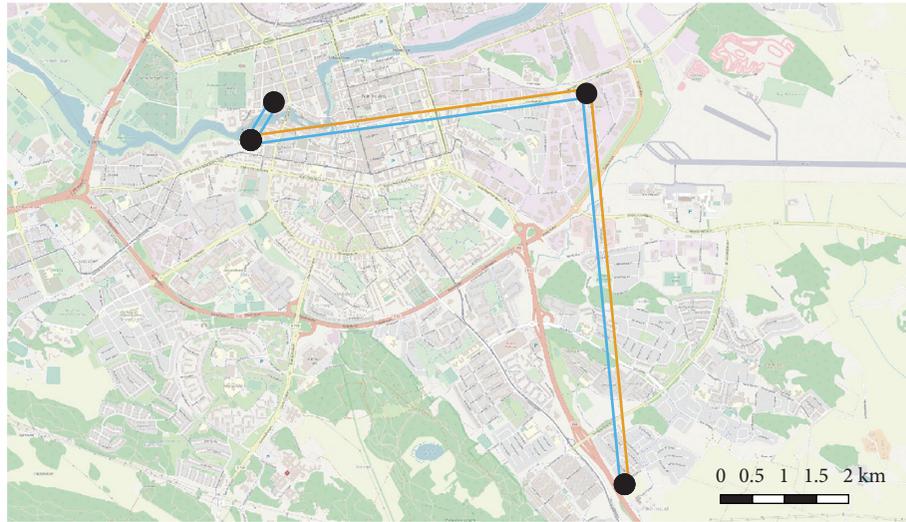
The datasets used in this paper contain periodic updates every 30 minutes. If only periodic updates are available, it is reasonable to assume that the movement happened in the middle between the periodic updates, which is enabled by setting $T_{\Delta \max} = 15$ minutes in the preprocessing.

For the STOP algorithm, the trade-off for T_{\min} is between detecting trips even between short activities (low T_{\min}) and detecting transit trips with transfers in one trip (high T_{\min}). $T_{\min} = 40$ min allows to detect trips between activities shorter than an hour but still keeps most transit trips in one trip as we assume that transfers in the datasets are usually taking less than 40 min.

The parameter D_{\max} in STOP and D_{\min} in MOVEMENT, respectively, is a trade-off between detecting even short trips (low value) and including noise in the data, that is, trips caused from switches between antennas without any physical movement. The parameters were set to 1 km which is appropriate for a city scenario with a rather dense network of antennas.

4. Travel Demand Inference

The next step is to aggregate the trips extracted from the cellular network data in order to infer the travel demand (see Figure 3). For the first step of the process, a remote access server setup, as described by de Montjoye et al. [26], is used to ensure privacy. After aggregation and scaling (see Section



● Cellular network event
 — Stop
 — Movement
 Background: OpenStreetMap

FIGURE 1: Example of trips extracted for one user starting southeast and ending northwest. The black dots are the positions of the antennas the user is connected to during the trip.

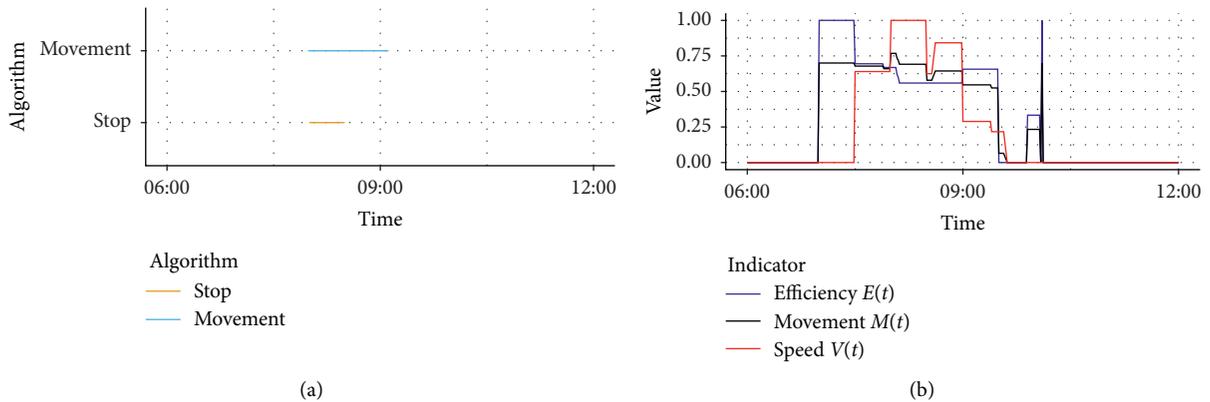


FIGURE 2: Example of trips extracted for one user between 6am and 12am. The top figure shows when the two cellular network-based methods have detected a trip. The bottom figure shows the movement indicators used in the MOVEMENT algorithm for the same time interval.

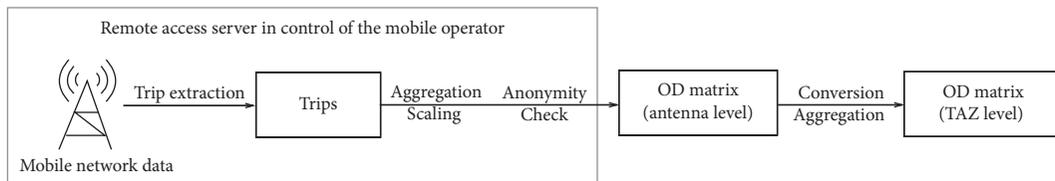


FIGURE 3: Overview of the travel demand inference process from cellular network data. The trip extraction is executed on a dedicated environment controlled by the mobile operator. Only the aggregated OD matrix is transferred for further analysis.

4.1) and checking anonymity, the resulting OD matrix is converted (see Section 4.2) and aggregated further (see Section 4.3) to make it usable in a traffic analysis context.

4.1. Aggregation of Trips and Scaling. To describe the travel demand, the previously extracted trips are aggregated into an OD matrix sliced by weekday and hour. The origins and

destinations in this initial OD matrix are the start and end antennas of the trips. The travel demand in this matrix is calculated as the number of trips in each pair of antennas at a certain hour. To ensure that the resulting OD matrix does not reveal any information on individual travellers, the travel demand for an OD pair is only saved if there are trips made by multiple users at the given weekday and hour.

The OD matrix calculated by summing the number of trips detected in the cellular network dataset describes the mobility patterns of the mobiles used by customers of one operator. For traffic planning purposes, however, the mobility patterns of the whole population are of interest. This establishes the need for some scaling method which scales the travel demand to the whole population. Even under the assumption that the dataset is representative for the whole population, this is not trivial.

In this paper, the focus is the structure of the OD matrix, consisting both of how the travel demand is distributed between different areas and how the travel demand varies over time rather than to estimate the total demand. We leave the scaling problem for future research. However, to be able to compare the inferred travel demand from cellular network data with an existing travel demand model, we multiply the demand inferred from cellular network data in each OD pair with a scale factor. The scale factor is global; that is, every hour, day, and OD pair is scaled with the same factor. The scale factor is set such that the total travel demand for an average weekday (Monday–Thursday) is equal to total demand in the existing travel demand model used for comparison.

4.2. Conversion to Traffic Analysis Zones. To be able to use the OD matrix generated from cellular network data in a traffic analysis context, the OD matrix on the antenna level needs to be converted to an OD matrix using traffic analysis zones (TAZs) instead. The conversion to a TAZ level OD matrix allows also to compare the converted OD matrix to an existing modelled OD matrix which is available for the same traffic analysis zones (TAZs).

We have implemented a conversion process which starts by defining an estimated coverage area for the origin and destination antenna in which a user is assumed to possibly have started/ended a trip in the given OD pair. If available, a polygon describing the coverage area of the antenna can be used. In absence of detailed descriptions of the antenna coverage, however, in this paper, we use the Voronoi cell of the antenna plus an additional buffer of 1 km as the estimated coverage area. For a given antenna's coverage area \bar{C} , a weight is assigned to every TAZ Z :

$$w(Z, \bar{C}) = \frac{\text{area}(Z \cap \bar{C})}{\text{area}(\bar{C})} \cdot \text{population}(Z). \quad (6)$$

Each TAZ outside of the area of effect will be assigned the weight zero by definition. Using the population (number of people living in the area according to the census) in the weighting is to assign travel demand to populated areas rather than areas without any buildings. In Figure 4, for example, without including population data, the highest

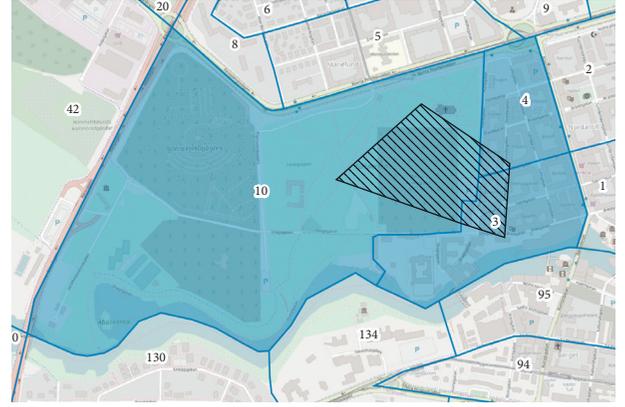


FIGURE 4: Example of the conversion to traffic analysis zones (TAZs). The travel demand connected to the black-shaded Voronoi cell is distributed among the blue traffic analysis zones (TAZs) which overlap with the Voronoi cell.

weight for the black-shaded Voronoi cell would be assigned to TAZ 10 which is a park and graveyard, while the residential areas 3 and 4 would get a lower weight. Including the population statistics, most weight will be given to the residential areas, which seems more realistic.

For a TAZ-OD pair O, D , the flow (travel demand) is calculated as the sum of flows assigned to the OD pair from all antenna-OD pairs \bar{O}, \bar{D} :

$$\text{flow}(O, D) = \sum_{\bar{O}, \bar{D} \in \text{CELL}} \frac{w(O, \bar{O}) \cdot w(D, \bar{D})}{\sum_{O', D' \in \text{TAZ}} w(O', \bar{O}) \cdot w(D', \bar{D})} \cdot \text{flow}(\bar{O}, \bar{D}). \quad (7)$$

Note that, after the conversion to TAZ, the total flow is maintained:

$$\sum_{O, D \in \text{TAZ}} \text{flow}(O, D) = \sum_{\bar{O}, \bar{D} \in \text{CELL}} \text{flow}(\bar{O}, \bar{D}). \quad (8)$$

A problem occurs in areas close to the border of the area of investigation for which data are available. Using the same conversion process as above for these areas would assign all external travel demand (travellers leaving or entering the area of investigation) to traffic analysis zones (TAZs) close to the border. To separate the external travel demand, antennas close to the border are marked as external. A number of traffic analysis zones (TAZs) are defined that correspond to different external areas around the area of investigation. The flows to and from external antennas are then mapped to the closest external TAZ.

4.3. Aggregation in Time and Space. The time-sliced OD matrix by weekday and hour, converted to TAZ, is the most detailed level we find useful to analyse travel patterns. We use further aggregation in both time and space wherever appropriate. Aggregation in time refers to merging certain timespans of the OD matrix time-sliced by weekday and hour. A typical aggregation used in this paper is the

aggregation to 24-hour typical weekday flows. Here, the flow in each OD pair is calculated as the average for the days from Monday to Thursday. This aggregation allows to compare the inferred travel demand from cellular network data to an existing model that is not time-sliced. To be able to understand the travel demand on a more spatially aggregated level, we define a new aggregated zoning for some comparisons by grouping several traffic analysis zones (TAZs). The new spatially aggregated matrix then contains the sum of the flows for each group of traffic analysis zone (TAZ).

5. Datasets

Two separate datasets of cellular network data have been used to evaluate different aspects of the presented method. A first dataset, described in Section 5.1, is used for a small-scale verification of the trip extraction algorithms. A second dataset, described in Section 5.2, is used to infer travel demand on a city level. Following the classification of different types of cellular network data in Gundlegård [27], both datasets contain billing data and location updates extracted from the core network. The billing data include data records for calls, SMS, and data service requests, sometimes collectively referred to as x-detail record (xDR) data [26]. The location updates include periodic, location area (LA), routing area (RA), tracking area (TA), and cell updates. No events related to handover or measurement reports are included in the datasets.

5.1. Cellular Network and GPS Validation Dataset. The first dataset of cellular network signalling data has been extracted from the network by the operator for 20 dedicated SIM cards (the data does not include any other user's data apart from the participants of the study, which have explicitly given their consent to the collection of their location updates). This small dataset is not representative of the whole population. The signalling data that are extracted from the cellular network in this dataset include location area updates, periodic updates, as well as call detail records (CDRs) generated by telephone calls and SMS. Periodic updates are performed every 30 minutes when connected to a 4G network (LTE). The average time between events in the dataset is approximately 25 minutes.

During the test period, location data for the same devices have also been collected using the Google location history service. Google location history data are collected locally on the device based on a combination of global positioning system (GPS), WiFi, and cellular positioning supported by local sensors like accelerometer and gyroscope to detect mobility. Since Google location history uses local sensors for positioning and movement detection, the spatiotemporal accuracy is higher compared to the cellular network signalling data. Google location history also classifies the travel mode ("activity") and splits trips when the travel mode changes. For the comparison of trips extracted from a cellular network, Google trips with less than 30 minutes in-between have been merged into one trip.

Several issues have been identified in the raw data. We use a number of filters to clean the data from these issues and remove whole day of the affected user (dataday) if a filter has found a problem. For the cellular network data, days with 31 minutes or more without an event ("missing data") and days with more than 50 kilometres between two consecutive events ("large hop") are removed. Days are also removed when a problem with the data from Google location history is detected. This includes activities classified as "moving" by Google location history. By manual inspection, we found that most of these activities were not connected to any actual travel and should therefore not be included in the comparison. Also, days are removed when trips with an unrealistic speed for the activity have been recorded. For the driving activity, for example, trips slower than 5 km/h or faster than 130 km/h are considered to be unrealistic. Finally, days with more than 1 km between the end of the last and start of the next Google location history trip or more than 24 hours without a trip ("missing data") are removed. More than half of the Google trips and cellular events have been removed in the process (see Table 3). The filters used for Google location history removed a large part of the data, indicating that there are frequent problems with the data collection using Google location history with the used devices. Note that the total can be less than the sum of the individual filters, as the same day might be marked as spurious by more than one filter.

5.2. City-Level Cellular Network Dataset. The city-level cellular network dataset is based on 37 million cellular network events (billing data and location updates) from three weeks during 2017/2018. The weeks were selected to not include major holidays. The user id in the dataset has been rehashed every day. On average, there are about 47000 users in the dataset per day. The average time between events for a mobile device in the dataset is about 14 minutes. However, there are large differences depending on the circumstances, e.g., which cellular network is used, if the device is moving etc., such that the interevent time can vary between seconds to many hours. As described in Figure 3, the first steps have been controlled by the cellular network operator using the algorithms we provided to extract trips.

To represent the antenna coverage areas, Voronoi tessellation is used [28]. By the nature of the cellular network, one base station typically hosts three antennas at the same position with antenna each covering different angles. Bachir et al. [18] have proposed a method to improve the Voronoi tessellation when having three sectors per base station. For this dataset, we use a simple approach to improve the representation of sectors. Instead of applying the Voronoi tessellation to the initial antenna positions, we move each antenna using its azimuth a few meters into the direction of its coverage. The result is a better representation of the different sectors by the resulting Voronoi cells.

5.3. Gravity Model. We compare the inferred travel demand from mobile phone data to the travel demand model used by Norrköping municipality. The model is a classical four-step

TABLE 3: Result of the data filtering of the validation dataset. If a day for a certain user (dataday) contains spurious data, both the Google trips and cellular events on that day are removed from the data.

Filter	% spurious		% removed	
	Datadays	Google trips	Events	
Google: missing data	28.63	45.16	38.46	
Google: moving activity	13.26	9.50	10.79	
Google: unrealistic speed	17.20	11.82	14.65	
Mobile data: large hop	2.98	1.94	3.24	
Mobile data: missing data	10.57	10.85	8.66	
Total (all filters)	48.32	58.53	55.15	

model consisting of trip generation, trip distribution, mode choice, and route choice steps [6].

The municipality’s traffic model is based on census data from 2014 as the night and day population (workplaces). For the trip generation also, two travel surveys have been used (one survey from 2010 and one from 2014) as input. In total, 4880 reported trips have been used from the surveys. The trip distribution is modelled per activity using a gravity model [7]. The activity types modelled are work, school, bringing kids to school, shopping, free-time activity, and others. The attraction of destinations is modelled using the number of facilities relevant for the activity (for example, the number of workplaces). As the cost in the gravity model, the distance of each pair of zones is used. Each activity generates a trip from home to the activity and a symmetric return trip. For the mode choice step, a logit model [29] is used with travel times and distance for each mode as input. The travel modes modelled are public transport, car, heavy goods vehicle (HGV), bicycle, and walking. Finally, in the route choice step, a network loading is repeated until the user equilibrium state is reached. The link flows calculated by the model have been validated using traffic counts for each travel mode.

For the comparison with OD matrices from cellular network data, we use the total travel demand generated by all these modes except for walking for each OD pair. The model is not time-sliced and models the average traffic Monday–Thursday. However, efforts have been made by Lindström and Persson [30] to ex-post time-slice the car traffic OD matrix using traffic counts.

The model uses a zoning with 189 traffic analysis zones (TAZs) which is referred to as TAZ-189. Among these, 167 zones are within the municipality (see Figure 5) and the remaining 22 zones represent external traffic to neighbouring municipalities. To allow a comparison on a more aggregated level, we define an additional alternative zoning (TAZ-24) where the original zones are grouped into only 24 zones (19 internal and 5 external zones) represented by different colours in Figure 5.

6. Results

To validate the trip extraction methods described in Section 3, we use the dataset described in 5.1 and compare the individual trips to trips detected by Google location history on the same device. The results of the validation including an analysis of the limitations of the trip extraction from cellular network data are presented in Section 6.1. The results in

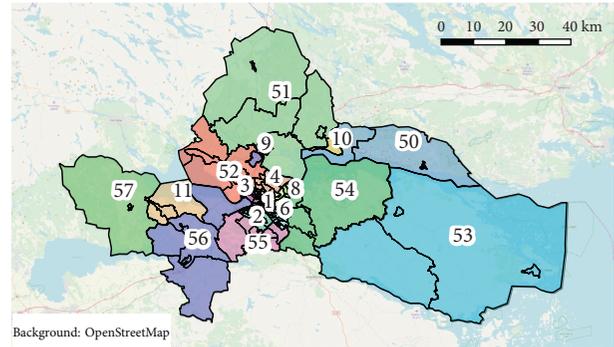


FIGURE 5: Map of Norrköping municipality. The coloured 24 zones labelled with their ID are an aggregation of the 189 zones used by the municipality’s traffic model, which are shown with gray outlines. External zones are not displayed.

Section 6.2 are based on the large-scale dataset described in Section 5.2 with focus on the travel demand inferred, which is compared to the municipality’s classic travel demand model.

6.1. Trip Validation. Running the trip extraction algorithms STOP and MOVEMENT (see Section 3) on the cleaned validation dataset (see Section 5.1) yields 393 and 450 trips, respectively. This compares to 548 trips (see Table 4) that have been detected by Google location history. When only including trips of a certain minimum length in kilometres in the comparison, the recall increases.

Figure 6 shows how the trips distribute among the hours of a day according to their start time. The time distribution of the raw events shows that slightly more events are generated when movements take place; based on this way, the data were collected. In general, the difference between the trip extraction algorithms is marginal when it comes to the time distribution. Notably, however, the morning peak is earlier according to cellular network data compared to Google location history. One reason can be the way the user position is estimated in the preprocessing used for both algorithms (see Section 3.1). It seems also that Google location history sometimes takes a while to be triggered by movement and thereby misses the actual start time of a trip on the devices we have used. This occurs frequently when the phone has not been active overnight. Another visible difference is the smaller peak around lunch hours in Google location history, which is not captured by the trip extraction from cellular network data. Most of these trips are short walking trips to a lunch restaurant near work, and these short trips are particularly difficult to detect from cellular network data.

To further judge the performance of the trip extraction from cellular network data in comparison to the global positioning system- (GPS-) based Google location history, we use two values:

- (i) Recall: the share of Google location history trips for which there exists a matching trip extracted from cellular network data

TABLE 4: Recall and precision achieved by the different trip extraction methods. The counts show the total number of Google trips and trips from cellular network data.

	Algorithm	# trips		Performance	
		Google	Cellular	Recall	Precision
All trips	Stop	548	450	0.69	0.83
	Movement	548	393	0.53	0.74
≥ 2 km	Stop	414	370	0.76	0.84
	Movement	414	346	0.65	0.75
≥ 4 km	Stop	332	306	0.80	0.83
	Movement	332	295	0.70	0.76

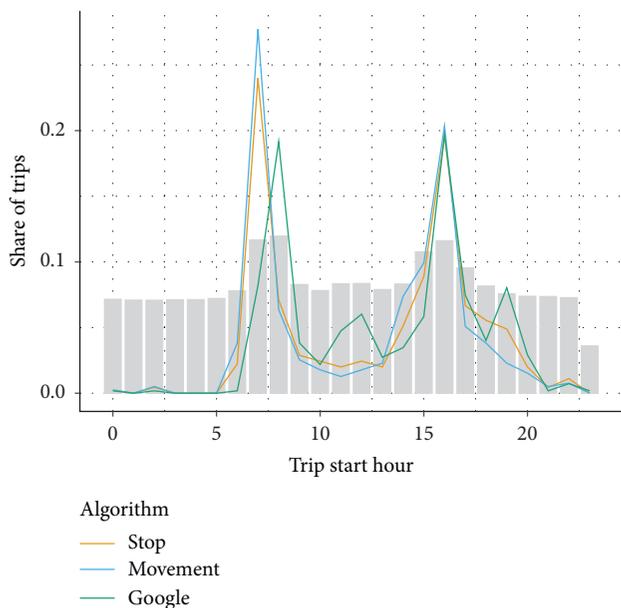


FIGURE 6: Time distribution of trips by hour detected using the different algorithms and Google location history. The gray bars show the time distribution of the cellular network events.

- (ii) Precision: the share of trips extracted from cellular network data for which there exists a matching Google location history trip

The recall measures how successful the algorithms are in reidentifying the trips which Google location history has identified, while the precision makes sure that this comes not to the expense of detecting many false positives. For both values, a higher value is better. When nothing else is stated, a trip detected from cellular network data and a trip from Google location history are considered as *matching* if all of the following are fulfilled:

- (1) There are at most 45 minutes between the start times of the trips. The same holds for the end times.
- (2) The distance between the start positions of the trips is at most 2 kilometres. The same holds for the end positions.

The overall recall achieved is 0.69 for STOP and 0.53 for MOVEMENT (see Table 4). However, these values depend largely on the composition of this validation dataset, which is not representative for the whole population. The recall is increasing significantly when only trips of a certain minimum length are compared. Shorter trips, especially those shorter than 2 kilometres, are particularly difficult to detect from cellular network data. Trips shorter than 1 kilometer cannot be detected due to the minimum distance thresholds used in the STOP and MOVEMENT algorithms. Even an optimal algorithm cannot realistically reach a recall or precision of 1 as the data collected from cellular network as well as from Google location history is not perfect (see Sections 7.1 and 7.4).

Since the goal is to use trips extracted from cellular network data to infer travel demand, it is important to understand the limitations which may lead to certain types of trips being over- or underrepresented. An important aspect is the possibility to detect trips of different travel modes. Table 5 shows that the recall largely differs for different modes: for walking trips, STOP detects only about half of the trips and MOVEMENT even less with only a about a quarter of the Google location history trips. In contrast, for most public transit modes, the recall is very high, with the caveat that there are only a limited number of trips in the dataset.

Besides the travel mode, it is relevant to understand how well trips of different length can be captured. Figure 7 shows the recall, and Figure 8 shows the precision depending on the trip distance. While the recall first increases and reaches its peak for trips around 8 kilometres, there is a slight decrease in the recall after that point. This can be explained by the fact that the longer the trips become, the higher the probability that they have been split up differently from cellular network data and in Google location history (for example, when a transfer was made). MOVEMENT performs significantly worse than STOP especially for very short trips.

To understand the spatial accuracy of the trips extracted from cellular network data, we define the spatial error of the start or end position of a cellular network trip as the distance to the start or end position, respectively, of the Google location history trip which matches best in time. For both STOP and MOVEMENT, we observe that half of the start/end positions of the extracted trips are within 500 meters of their counterparts according to Google location history (see Figure 9). More than 90% of the positions are within 2 kilometres of the respective Google location history position.

6.2. City-Level Travel Demand. This section focusses on the aggregated travel demand inferred from large-scale cellular network data for the municipality of Norrköping. The matrices discussed here are based on the dataset described in Section 5.2 and have been computed using the trip extraction methods in Section 3 and the demand inference procedure described in Section 4. We analyse both the structure of the OD matrices by comparing to the municipality's travel demand model and the time profiles of the matrices.

TABLE 5: Recall achieved for different activities and trip extraction methods, that is, the share of Google location history trips with a matching cellular network trip. Count shows the total number of Google trips of the activity in the dataset.

Activity	# trips Google	Recall	
		Stop	Movement
Driving	306	0.74	0.63
Walking	136	0.53	0.26
On a train	40	0.80	0.68
Cycling	30	0.53	0.37
On a bus	29	0.83	0.72
On the subway	6	1.00	1.00
On a tram	1	1.00	0.00

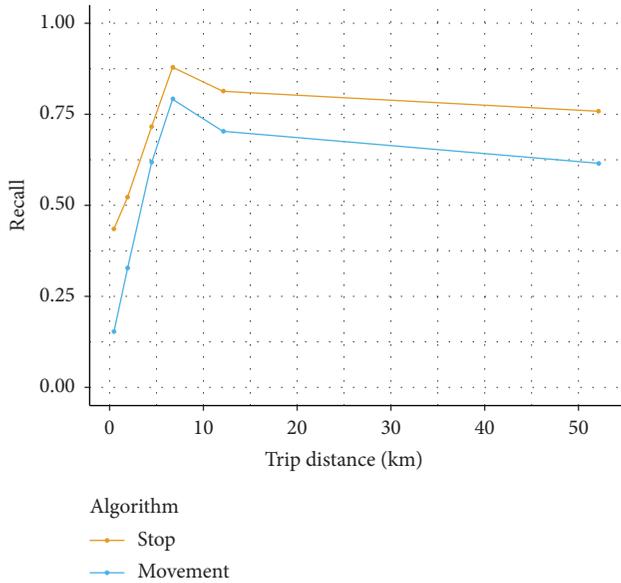


FIGURE 7: Recall depending on the trip distance. Each datapoint corresponds to 1/6 of the Google trips and shows the share for which of these trips a matching cellular trip was detected by the different trip extraction algorithms.

6.2.1. *Structure of the OD Matrix.* To analyse the structure of the OD matrices inferred from cellular network data, the travel demand after scaling to match the total flow is compared to the municipality’s gravity model. This allows to understand if there are structural differences in how the travel demand distributes among different OD pairs in the matrices inferred from cellular network data and the gravity model. Figure 10 gives a side-by-side overview of the travel demand in the densest areas as calculated using cellular network data and in the urban travel demand model. While there are some differences, the travel demand estimated from cellular network data resembles the gravity model quite well in general.

Based on the construction and thresholds used by the STOP algorithm, it detects more short-distance trips than the MOVEMENT algorithm. This is visible in Figure 7 where trips shorter than a few kilometres show a considerably higher recall using the STOP algorithm in the trip validation dataset. It also explains the higher recall for travel modes used

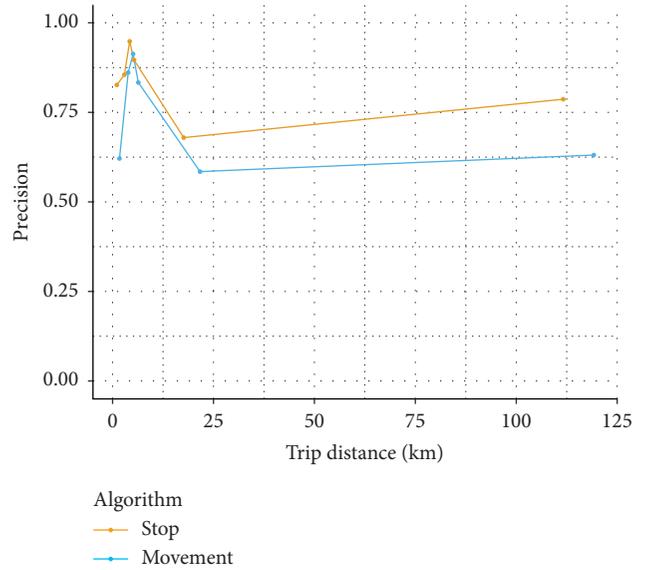


FIGURE 8: Precision depending on the trip distance. Each datapoint corresponds to 1/6 of the cellular trips and shows the share for which of these trips a matching Google trip was present.

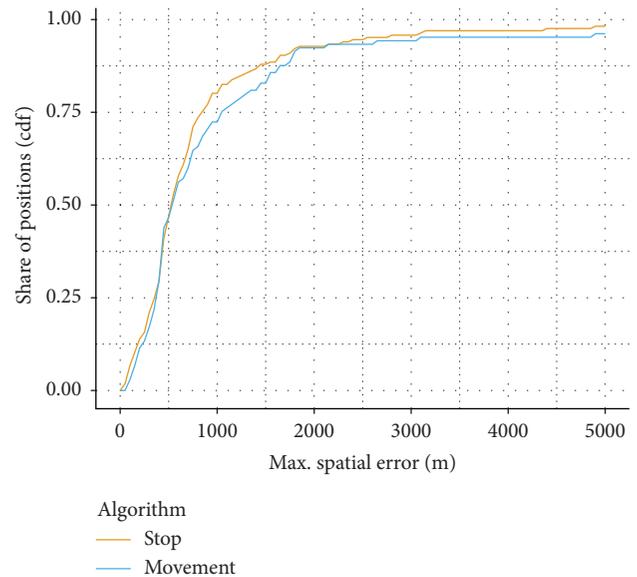


FIGURE 9: Spatial error at the start and end of trips. For each Google trip, the distance at the start and end positions is calculated to the cellular trip that matches best in time. Google trips without a cellular trip that matches in time (start and end within ± 45 min) are not included.

frequently for shorter trips, such as walking and cycling (see Table 5). A consequence of this is visible in the OD matrix where relatively higher flows are assigned to OD pairs with short distance (see Figure 11). MOVEMENT, in contrast, produces a distance distribution which is very much in line with the gravity model. Analysing the zoneflow shows the structural effect on the OD matrix of the different trip extraction methods. The term zoneflow here refers to the total of the OD flows starting and ending in a zone. A significant

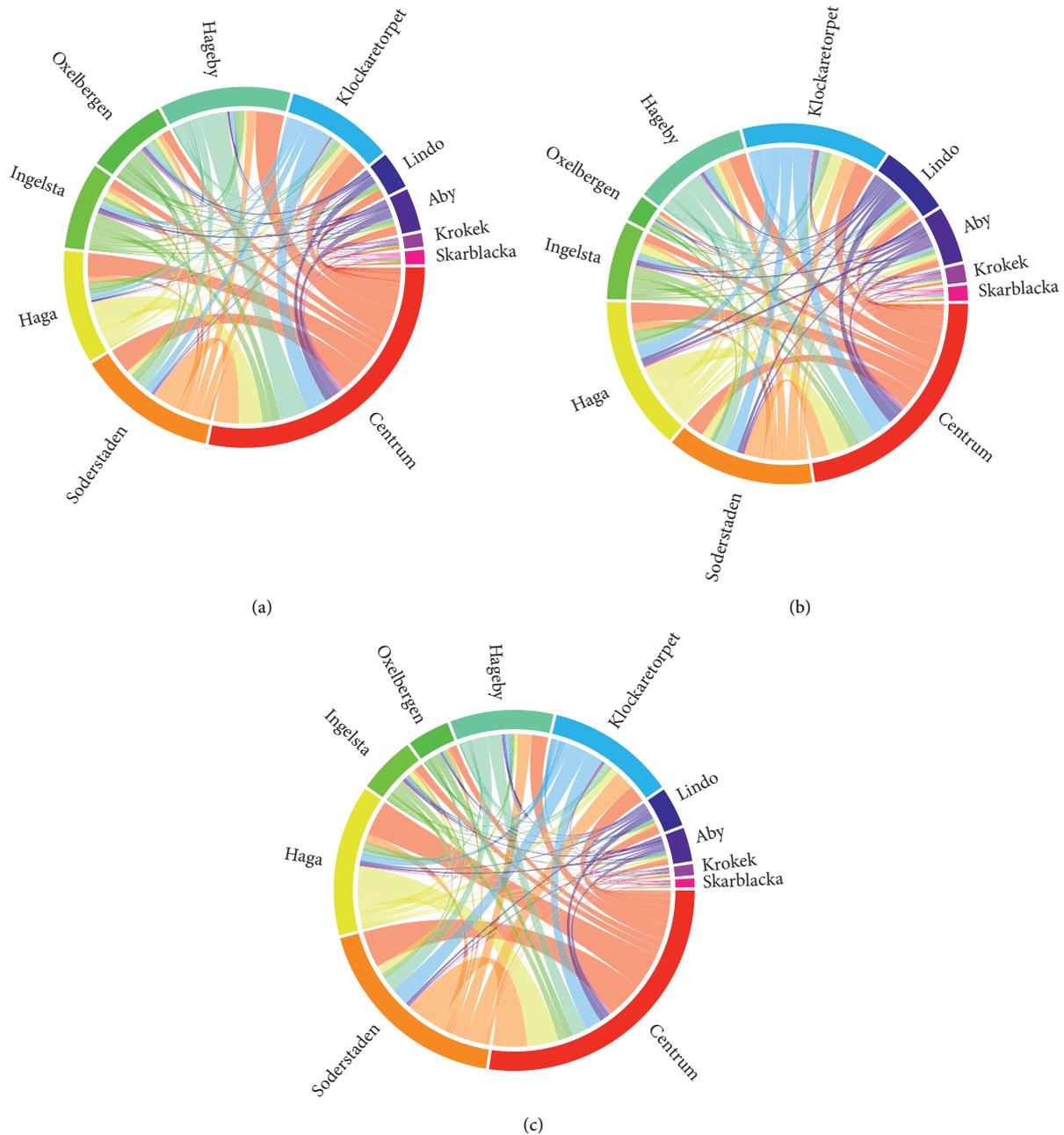


FIGURE 10: The OD flows inferred from cellular network data using the STOP (top left), MOVEMENT (top right) algorithms, and the municipality's travel demand model (bottom). For readability, only 11 zones (with all densely populated areas) out of the 24 are displayed. The width of the connectors corresponds to the flow in each OD pair.

difference can be observed in the densely populated areas where MOVEMENT produces much lower zoneflow than STOP. For the rural areas of the municipality, the opposite can be observed. An explanation can be that short trips might be more common in the central city and that those trips are better captured by the STOP algorithm.

To systematically compare the structure of the OD matrices to the existing gravity model, we investigate the correlation between the OD-matrices (see Figure 12). A large difference is found between the two levels of aggregation (see Table 6). While the correlation for the original

TAZ-189 zoning is weak, the MOVEMENT- and STOP-based matrices reach an R^2 value of 0.82 and 0.81, respectively for the more aggregated TAZ-24 zoning. This indicates that we are not capable, with the used method and data, to infer the travel demand for the detailed TAZ-189 zoning (where a zone often only contains a few housing blocks), while the travel demand inference works well for the TAZ-24 zoning. Similar results have been found by Batran et al. [22]. While the two methods show a similar correlation to the model, this does not imply that there is no difference between the two cellular network data-based matrices as the R^2 value of

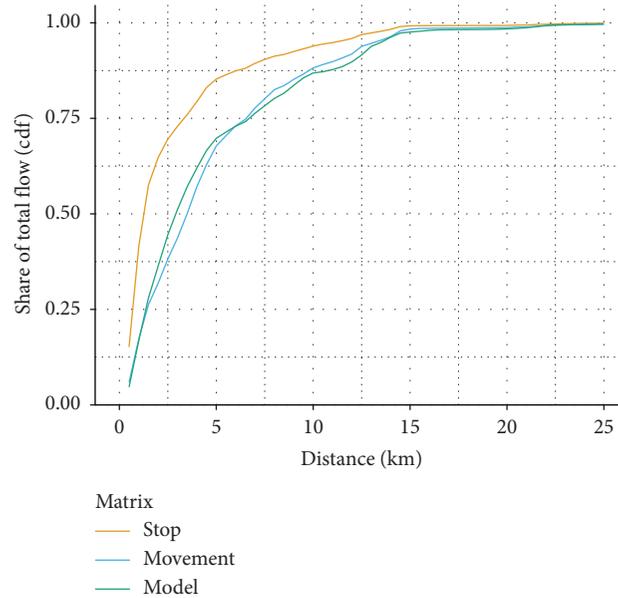


FIGURE 11: Distribution of the average travel demand Monday–Thursday by distance of the OD pair for the OD matrices estimated from cellular network data as well as the model (TAZ-189). The distance is based on the centroid of the TAZ.

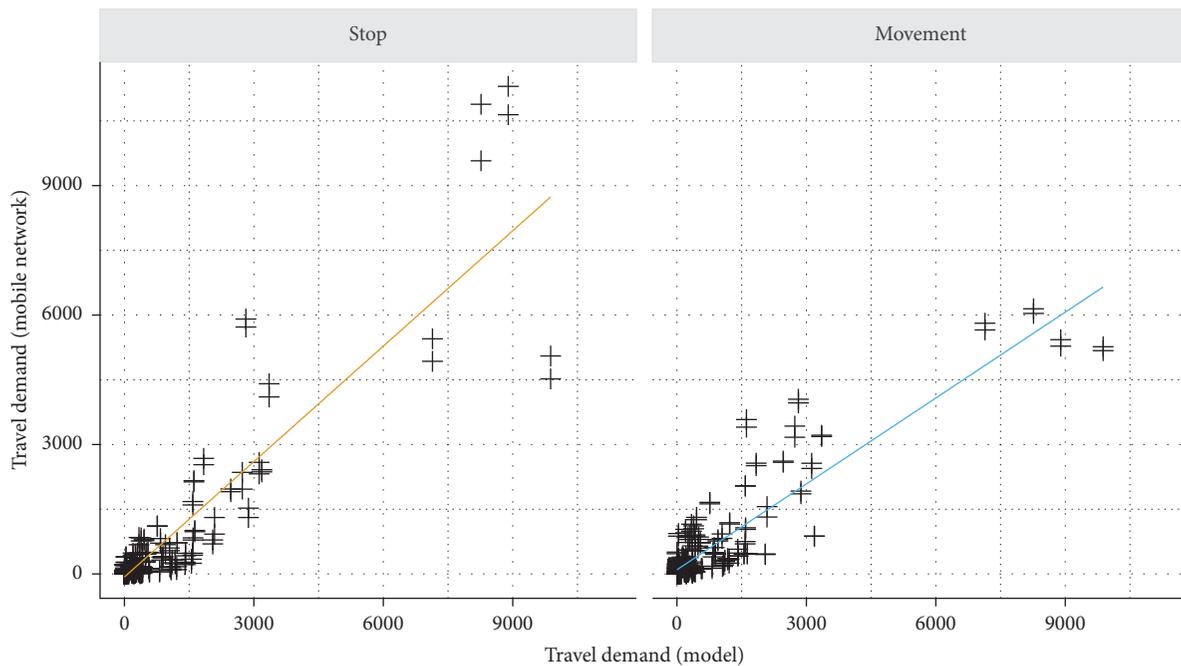


FIGURE 12: Correlation of the OD matrices inferred from cellular network data in comparison with the model. Each data point corresponds to an OD pair, and its average travel demands for Monday–Thursday (TAZ-24 zoning). The line shows a fitted linear model.

0.84 between the MOVEMENT- and STOP-based OD matrices shows.

The correlation of the zoneflow (see Figure 13) indicates how similar the travel activity is distributed among the zones; thus, if using cellular network data, the same zones are identified as the zone which generate the most trips, without considering between which zones the trips are made. The R^2 values of 0.85 and 0.90 for MOVEMENT and STOP (TAZ-24)

show that, using cellular network data, the zone activity is well correlated to the gravity model (see Table 6).

6.2.2. *Time Profile of the Travel Demand.* Using the time-sliced OD-matrix from cellular network data, we can obtain time profiles for specific traffic analysis zones (TAZs) or OD pairs. Figure 14 shows an example for the city center of

TABLE 6: Correlation between the travel demand inferred from cellular network data by different algorithms and the municipality’s demand model for both TAZ-189 and TAZ-24 zoning regarding the flow in each OD pair and the zoneflow (total inflow + outflow) for each zone. Zone internal flows and OD pairs with origin/destination outside of the municipality are not included.

Algorithm	OD flows (R^2)		Zoneflows (R^2)	
	TAZ-189	TAZ-24	TAZ-189	TAZ-24
Stop vs. model	0.16	0.81	0.31	0.90
Movement vs. model	0.11	0.82	0.21	0.85
Stop vs. movement	0.52	0.84	0.79	0.94

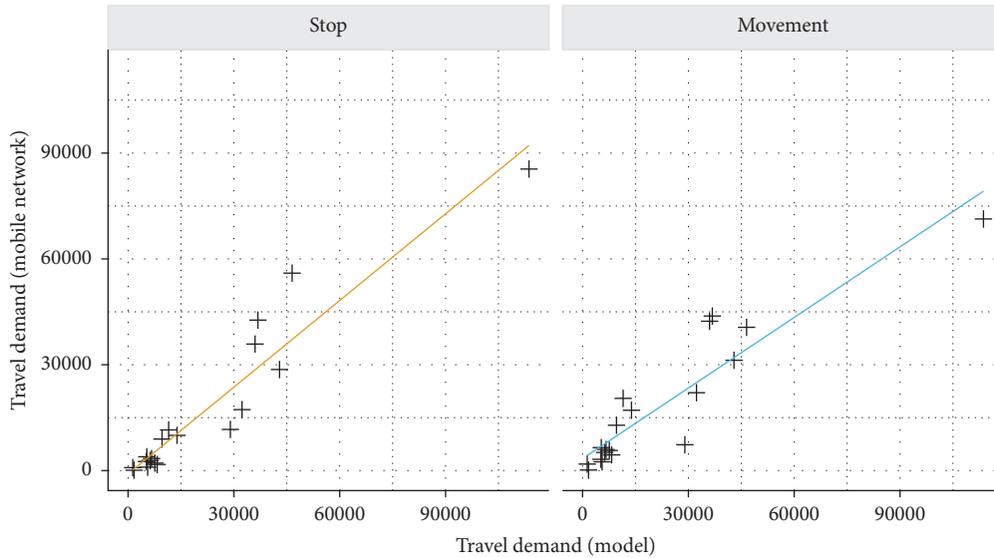


FIGURE 13: Correlation of the zoneflow (total inflow + outflow) inferred from cellular network data in comparison with the model. Each data point corresponds to an average zoneflow for a TAZ Monday–Thursday (TAZ-24 zoning). The line shows a fitted linear model.

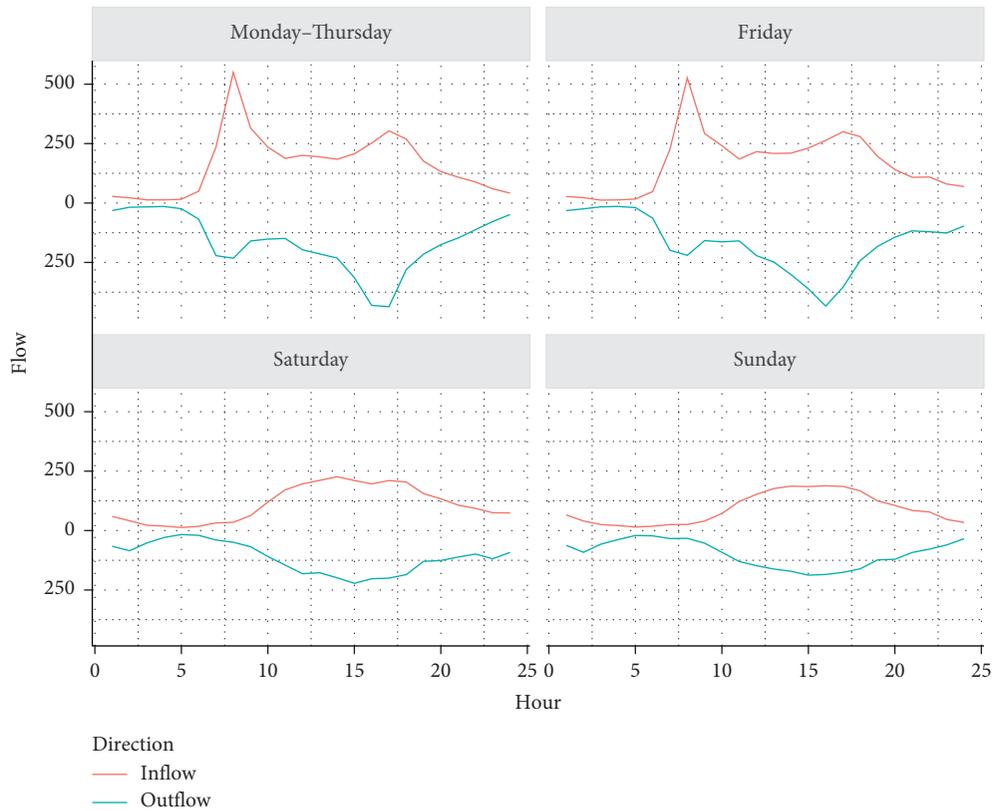


FIGURE 14: Time distribution of the inflow and outflow of the zone containing the city center of Norrköping using the MOVEMENT algorithm. For Monday–Thursday, an average is presented as these days show very similar patterns.

Norrköping obtained using the *MOVEMENT* algorithm. We find a typical commuting pattern during workdays with an arrival peak around 8:00 in the morning and a departure peak around 16:00-17:00 in the afternoon. Given the sample size, it is typically not possible to get similar time profiles for specific traffic analysis zones (TAZs) or OD pairs from travel surveys. Sensors providing traffic counts can provide time profiles for certain traffic modes but only measure the total traffic on a link within the traffic network, which is difficult to decompose into OD pairs.

Figure 15 shows the time profiles for the total demand inferred by both *MOVEMENT* and *STOP*. The time profiles from cellular network data show a clear morning/afternoon peak pattern for weekdays and a different pattern on weekends. The municipality's travel demand model is a static model of a typical weekday. Therefore, the model does not contain any time profile to compare to. However, Lindström and Persson [30] have made attempts to create a time-sliced version of the original model containing a time profile for the typical weekday (labelled "Model (car, 24)" in Figure 15). This time profile has been estimated using data from traffic sensors.

We find that the time profile inferred from cellular network data resembles this time-sliced demand model based on sensor data very well (see Figure 15). The total flow of this model is lower because it is only based on the travel demand for cars, while the total of the cellular network OD matrices sums up to the total of car, public transport, heavy goods vehicle (HGV), and cycling. To complement the time-sliced car matrix, we also add the number of tap-ins made in the municipality's public transit system on average (Monday–Thursday). Combining the time-sliced car matrix and the number of tap-ins ("Model (car, 24) + Tap-in") fits very well with the time profile only inferred from cellular network data.

7. Discussion

The results above show potential as well as a number of issues when inferring travel demand from cellular network data. The comparisons of individual trips and aggregated OD flows allow us to discuss the potential challenges linked to both trip extraction and travel demand inference. By comparing two trip extraction methods, we can better understand the effect of the trip extraction method on the resulting travel demand. The challenges identified are linked to data collection, the methods used for trip extraction, and travel demand inference as well as the comparison methods.

7.1. Data Collection. Investigating the recorded events in the cellular network dataset described in Section 5.1, we found situations where the raw data contain errors. This includes, for example, large hops of more than 50 km from one event to another, where normally a connection to other antennas should have been made in-between. In rare cases, there appear to be interlaced periodic updates for some users and days. This means that there are not only periodic updates in antenna A every 30 minutes for a user but also periodic updates in antenna B every 30 minutes in-between. As this

can last over many hours, it may lead to trips continuing infinitely, while this situation lasts. We suspect that this could be related to the different cellular network types (Global System for Mobile communications (GSM), Universal Mobile Telecommunications System (UMTS), and Long-Term Evolution (LTE)), creating data simultaneously.

A large portion of the trips not detected from cellular network data can be explained by limited resolution in space which is determined by the density of antennas. Trips which are too short to trigger a switch of antennas can never be detected from cellular network data which only consist of billing data or location updates. However, information about the actual coverage area of each antenna could replace the static minimum distance thresholds in the trip extraction algorithms, allowing to detect even shorter trips where the antenna density allows and could be used to improve the spatial accuracy of the start and end of trips. The accuracy of the start and end time could be improved by using data with a lower interevent time. Data spanning over longer periods in time for individual users would allow making use of recurring patterns.

7.2. Trip Extraction. We chose two trip extraction algorithms which work rather differently in order to get an idea of how much impact the choice of the algorithm has. We see differences in recall and precision between the algorithm and believe that especially the inference of the exact start and end time and position for trips in the *MOVEMENT* algorithm could be improved. A sensitivity analysis of the parameter values should be carried out to improve the detailed understanding of their effects on the results. A thorough calibration of the parameters of both algorithms, which has not been done for this paper, could lead to an improvement for both recall and precision. A recall of 100% is not achievable by any algorithm, however given that some trips are too short to trigger a switch of antennas. However, we see potential in using data on the estimated antenna coverages instead of a static distance threshold to be able to detect more short trips where the antenna density allows for it as well as to improve the precision in areas with few antennas and large coverage areas. Making use of recurring patterns, like places which are visited regularly, could improve the quality of the detected trips further but would require data for the same user to be available for more than one day.

The results for the *MOVEMENT* and *STOP* algorithms show that adding more complexity to the trip extraction method does not necessarily improve the trip extraction. In our case, the relatively simple *STOP* algorithm often gave better results than the more complex *MOVEMENT* algorithm. However, we believe that improvements to the trip extraction algorithms can be made, for example, in the estimation of the exact start and end time of trips.

7.3. Travel Demand Inference. Inferring travel demand by aggregating the extracted trips carries the risk of a trip bias. The travel demand for short OD pairs is, for example, underestimated if too few short trips are detected. From Figure 11, we can conclude that this seems not to be the case

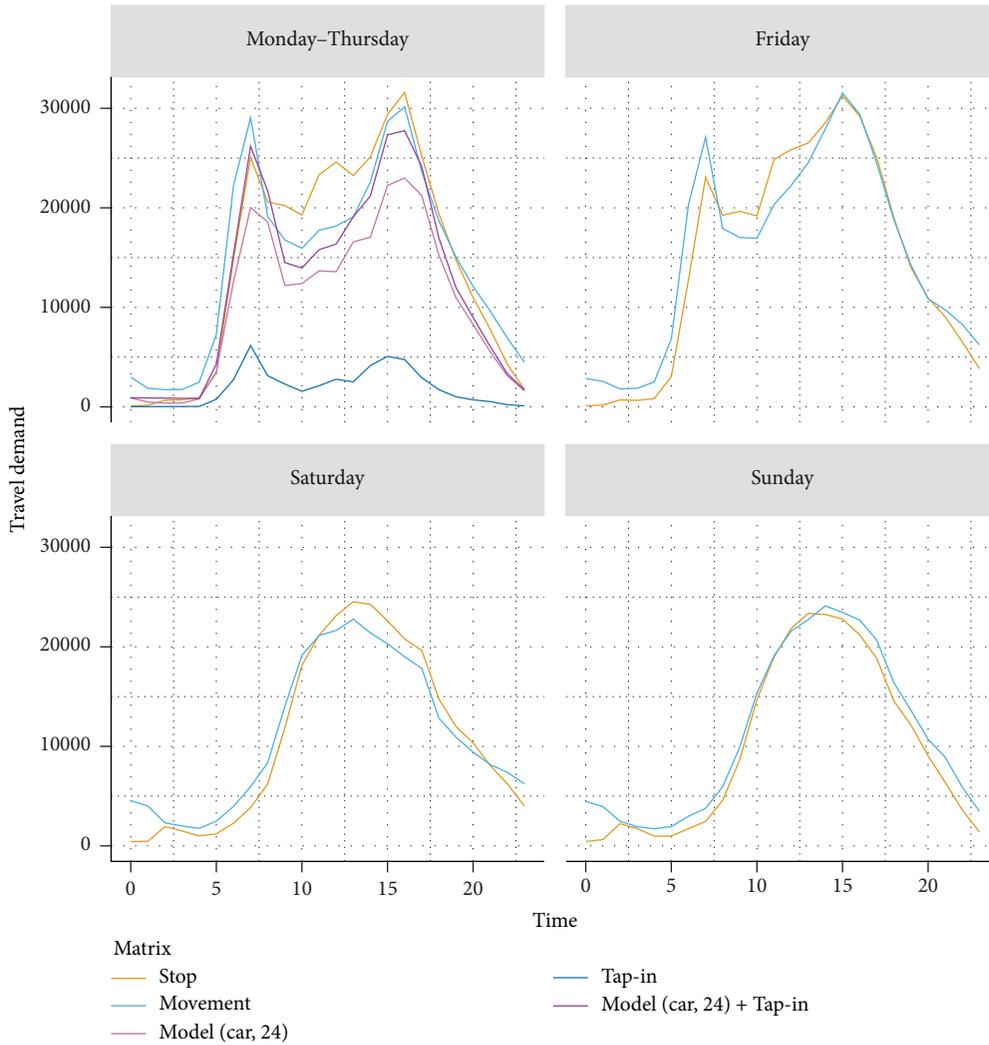


FIGURE 15: Time distribution of the scaled travel demand in the different OD matrices. For Monday–Thursday, an average is presented as these days show very similar patterns.

when comparing to the existing travel demand model. Contrary, STOP overestimates flows for short distances. But it is known that, in travel surveys, shorter trips are often being underreported as well [31]. Hence, a good match with the model does not necessarily need to be positive. Given the overall correlation with the model being very similar for both methods presented, it is an interesting result that the different trip extraction methods cause significant structural differences in the resulting OD matrices. This can be seen, for example, in differences in the zoneflow shown in Figure 16.

We are able to reproduce travel patterns on the aggregated TAZ-24 zoning, but not for the detailed TAZ-189 zoning (see Section 6.2). With the method and the cellular network data used, the spatial resolution is limited by the antenna densities. However, it is likely that even the gravity model cannot accurately model flows for the detailed TAZ-189 zoning. A limitation of the travel demand calculation used is that the demand is scaled with a constant factor to match the model demand. To estimate the total travel

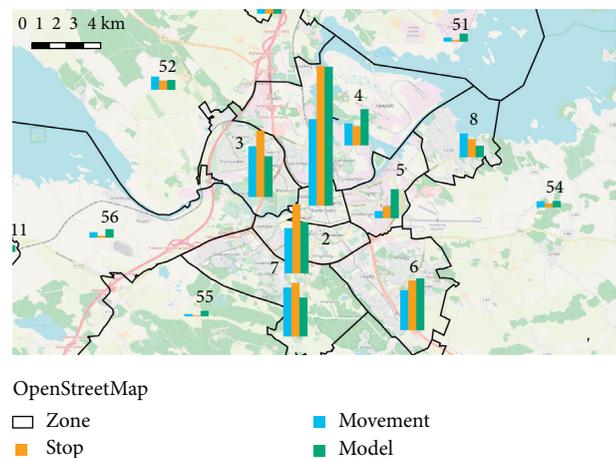


FIGURE 16: Zoneflow (total inflow + outflow) inferred from cellular network data in comparison with the model (TAZ-24 zoning). Zone internal flows and OD pairs with origin/destination outside of the municipality are not included.

demand independently of the existing travel demand model, a more advanced scaling method is necessary, which could be based on census data, market shares of the operator, and traffic counts. In the conversion of the OD matrix to TAZ, we used population statistics to distribute the demand among zones close to the antennas Voronoi cell. This explains why the travel demand in zone 8 (a residential area) seems to be overestimated, while the demand in the neighbouring zone 5 (dominated by industry and workplaces) is underestimated (see Figure 16). Alternatives are to use land use data or include data on the number of workplaces when converting to traffic analysis zones (TAZs).

7.4. Comparison Methods. Comparing trip-by-trip to Google location history allowed us to validate trips in a more detailed way than most other studies. However, we found that Google location history is not a perfect ground truth. Significant portions of the data showed problems and needed to be removed from the comparison (see Section 5.1). The resolution of Google location history is not always as good as one might expect from Global Positioning System (GPS) tracks. Possible reasons could be battery-saving techniques or problems with the specific devices used. It also needs to be considered that the dataset collected using 20 phones is not representative for the whole population. Furthermore, the matching definition counts it as an error if a trip is split in two trips in one dataset but one connected trip in the other.

While the urban travel demand model used for comparing the OD matrix inferred from cellular network data is the best available description of the actual travel demand, it is not the ground truth. There are a number of assumptions in the model which do not hold in reality. The gravity model, for example, is based on the distance between traffic analysis zones (TAZs) only. The model is also symmetric as all activities generate trips starting at home to the activity and a trip back home, while there are activity chains possible in reality. External trips from/to outside the municipality as well as heavy goods vehicle (HGV) traffic have been taken from the national demand model for Sweden, which does not provide details within the municipality.

8. Conclusions

We have presented a process to infer travel patterns from cellular network data by first extracting trips which then are aggregated and converted to an OD matrix. Comparing trip-by-trip as well as the aggregated travel demand to other available data, we can understand the performance of the method and identify potential issues.

For the trip extraction, the simple STOP algorithm performed in some ways better than the more complex MOVEMENT algorithm for our validation dataset from Google location history. The biggest difference can be observed for shorter trips (≤ 2 km), while the difference becomes small for longer trips. The recall achieved is more than 80% (STOP) for trips made with public transit, while it is poor with only 25%–50% for walking or cycling trips.

We find a reasonable correlation between the inferred travel demand from cellular network data and the existing travel demand model of Norrköping municipality after aggregating the zoning to 24 zones. The difference between the travel demand inferred using the two trip extraction methods is marginal when it comes to the correlation to the model with R^2 values of 0.82 for MOVEMENT and 0.81 for STOP. This can however not hide the fact that the inferred OD matrices exhibit significant structural differences. The R^2 value of 0.84 between the two cellular network data-based matrices is not significantly higher than the correlation between the cellular network data-based matrices and the model. The choice of the trip extraction method is crucial considering that systematic differences, as the over- or underrepresentation of short trips, have a substantial impact on the inferred travel demand.

Future research is needed regarding the separation of travel demand by a travel mode. The presented method also needs to be completed with an adequate scaling method allowing to obtain new information on the total demand. The comparison with the existing travel demand model suffers from the fact that the model is not a ground truth. Major potential of large-scale data as cellular network data lies in the fact that the much larger amount of data compared to, for example, travel surveys allows for zooming into specific OD pairs including hourly time profiles which potentially even could be obtained in real time.

Acronyms

CDR:	Call detail record
GPS:	Global Positioning System
GSM:	Global System for Mobile Communications
HGV:	Heavy goods vehicle
LTE:	Long-term evolution
RMSE:	Root mean square error
TAZ:	Traffic analysis zone
UMTS:	Universal mobile telecommunications system
xDR:	x-detail record.

Data Availability

The cellular network data used to support the findings of this study have not been made available to protect privacy.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

This work was supported by the Swedish Agency for Innovation Systems (Vinnova) (grant number 2013-03077).

References

- [1] A. Cuttone, S. Lehmann, and M. C. González, "Understanding predictability and exploration in human mobility," *EPJ Data Science*, vol. 7, no. 1, pp. 2193–1127, 2018.

- [2] H. T. Marques-Neto, F. H. Z. Xavier, W. Z. Xavier et al., "Understanding human mobility and workload dynamics due to different large-scale events using mobile phone data," *Journal of Network and Systems Management*, vol. 26, no. 4, pp. 1079–1100, 2018.
- [3] A. Janecek, D. Valerio, K. A. Hummel, F. Ricciato, and H. Hlavacs, "The cellular network as a sensor: from mobile phone data to real-time road traffic monitoring," *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 5, pp. 2551–2572, 2015.
- [4] C. Anda, A. Erath, and P. J. Fourie, "Transport modelling in the age of big data," *International Journal of Urban Sciences*, vol. 21, no. sup1, pp. 19–42, 2017.
- [5] Z. Wang, S. Y. He, and Y. Leung, "Applying mobile phone data to travel behaviour research: a literature review," *Travel Behaviour and Society*, vol. 11, pp. 141–155, 2018.
- [6] M. G. McNally, *The Four-Step Model, Chapter 3*, Emerald Group Publishing, Bingley, UK, 2007.
- [7] S. Erlander and N. F. Stewart, *The Gravity Model in Transportation Analysis: Theory and Extensions*, vol. 3, VSP, London, UK, 1990.
- [8] N. Caceres, J. P. Wideberg, and F. G. Benitez, "Deriving origin-destination data from a mobile phone network," *IET Intelligent Transport Systems*, vol. 1, no. 1, pp. 15–26, 2007.
- [9] E. Toch, B. Lerner, E. Ben-Zion, and I. Ben-Gal, "Analyzing large-scale human mobility data: a survey of machine learning methods and applications," *Knowledge and Information Systems*, vol. 58, no. 3, pp. 501–523, 2018.
- [10] M. von Mörner, "Application of call detail records—chances and obstacles," *Transportation Research Procedia*, vol. 25, pp. 2233–2241, 2017.
- [11] L. Alexander, S. Jiang, M. Murga, and M. C. González, "Origin-destination trips by purpose and time of day inferred from mobile phone data," *Transportation Research Part C: Emerging Technologies*, vol. 58, pp. 240–250, 2015.
- [12] F. Calabrese, F. C. Pereira, G. Di Lorenzo, L. Liu, and C. Ratti, "The geography of taste: analyzing cell-phone mobility and social events," in *Proceedings of the 8th International Conference on Pervasive Computing, Pervasive'10*, pp. 22–37, Springer-Verlag, Helsinki, Finland, May 2010.
- [13] E. Graells-Garrido, D. Caro, and D. Parra, "Inferring modes of transportation using mobile phone data," *EPJ Data Science*, vol. 7, no. 1, p. 49, 2018.
- [14] N. Breyer, D. Gundlegård, C. Rydbergren, and J. Bäckman, "Trip extraction for traffic analysis using cellular network data," in *Proceedings of the 2017 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS)*, pp. 321–326, IEEE, Naples, Italy, June 2017.
- [15] P. Widhalm, Y. Yang, M. Ulm, S. Athavale, and M. C. González, "Discovering urban activity patterns in cell phone data," *Transportation*, vol. 42, no. 4, pp. 597–623, 2015.
- [16] D. Gundlegård, C. Rydbergren, N. Breyer, and B. Rajna, "Travel demand estimation and network assignment based on cellular network data," *Computer Communications*, vol. 95, pp. 29–42, 2016.
- [17] A. Bwambale, C. F. Choudhury, and S. Hess, "Modelling trip generation using mobile phone data: a latent demographics approach," *Journal of Transport Geography*, vol. 76, pp. 276–286, 2017.
- [18] D. Bachir, G. Khodabandelou, V. Gauthier, M. El Yacoubi, and J. Puchinger, "Inferring dynamic origin-destination flows by transport mode using mobile phone data," *Transportation Research Part C: Emerging Technologies*, vol. 101, pp. 254–275, 2019.
- [19] F. Calabrese, M. Diao, G. Di Lorenzo, J. Ferreira, and C. Ratti, "Understanding individual mobility patterns from urban sensing data: a mobile phone trace example," *Transportation Research Part C: Emerging Technologies*, vol. 26, pp. 301–313, 2013.
- [20] M. Fillekes, "Reconstructing trajectories from sparse call detail records," Master's thesis, University of Tartu, Tartu, Estonia, 2014.
- [21] F. Liu, D. Janssens, J. Cui, Y. Wang, G. Wets, and M. Cools, "Building a validation measure for activity-based transportation models based on mobile phone data," *Expert Systems with Applications*, vol. 41, no. 14, pp. 6174–6189, 2014.
- [22] M. Batran, M. G. Mejia, Y. Sekimoto, and R. Shibasaki, "Inference of human spatiotemporal mobility in greater maputo by mobile phone big data mining," in *Proceedings of the 10th International Workshop on Agents in Traffic and Transportation*, Stockholm, Sweden, 2018.
- [23] L. J. J. Wismans, K. Friso, J. Rijdsdijk, S. W. de Graaf, and J. Keij, "Improving a priori demand estimates transport models using mobile phone data: a Rotterdam-region case," *Journal of Urban Technology*, vol. 25, no. 2, pp. 63–83, 2018.
- [24] T. Pollard, N. Taylor, T. van Vuren, and M. MacDonald, "Comparing the quality of OD matrices in time and between data sources," in *Proceedings of the European Transport Conference 2013*, Frankfurt, Germany, September 2013.
- [25] M. S. Iqbal, C. F. Choudhury, P. Wang, and M. C. González, "Development of origin-destination matrices using mobile phone call data," *Transportation Research Part C: Emerging Technologies*, vol. 40, pp. 63–74, 2014.
- [26] Y.-A. de Montjoye, S. Gams, V. Blondel et al., "On the privacy-conscious use of mobile phone data," *Scientific Data*, vol. 5, no. 1, Article ID 180286, 2018.
- [27] D. Gundlegård, *Transport analytics based on cellular network signalling data*, Ph.D. thesis, Linköping University, Communications and Transport Systems, Faculty of Science & Engineering, Linköping, Sweden, 2018.
- [28] A. Aggarwal, L. J. Guibas, J. Saxe, and P. W. Shor, "A linear-time algorithm for computing the voronoi diagram of a convex polygon," *Discrete & Computational Geometry*, vol. 4, no. 1, pp. 591–604, 1989.
- [29] C.-H. Wen and F. S. Koppelman, "The generalized nested logit model," *Transportation Research Part B: Methodological*, vol. 35, no. 7, pp. 627–641, 2001.
- [30] A. Lindström and F. Persson, "Estimation of hourly origin destination trip matrices for a model of Norrköping," Master's thesis, Linköping University, Communications and Transport Systems, Faculty of Science & Engineering, Linköping, Sweden, 2018.
- [31] P. Stopher, C. FitzGerald, and M. Xu, "Assessing the accuracy of the Sydney household travel survey with GPS," *Transportation*, vol. 34, no. 6, pp. 723–741, 2007.