



A novel big data approach to measure and visualize urban accessibility



Yuqin Jiang* , Diansheng Guo, Zhenlong Li and Michael E. Hodgson

Abstract

Accessibility is a topic of interest to multiple disciplines for a long time. In the last decade, the increasing availability of data may have exceeded the development of accessibility modeling approaches, resulting in a modeling gap. In part, this modeling gap may have resulted from the differences needed for single versus multimodal opportunities for access to services. With a focus on large volumes of transportation data, a new measurement approach, called Urban Accessibility Relative Index (UARI), was developed for the integration of multi-mode transportation big data, including taxi, bus, and subway, to quantify, visualize and understand the spatiotemporal patterns of accessibility in urban areas. Using New York City (NYC) as the case study, this paper applies the UARI to the NYC data at a 500-m spatial resolution and an hourly temporal resolution. These high spatiotemporal resolution UARI maps enable us to measure, visualize, and compare the variability of transportation service accessibility in NYC across space and time. Results demonstrate that subways have a higher impact on public transit accessibility than bus services. Also, the UARI is greatly affected by diurnal variability of public transit service.

Keywords: Geospatial big data, Accessibility, Geo-visualization, Spatiotemporal analysis, Transportation

1 Introduction

Accessibility is a term commonly used in geographical research, transportation, and urban planning. A definition for accessibility varies in different research contexts, including “interactions between human and lands” (Hansen, 1959), “the ease or difficulty for people to reach their opportunities or services” (Wachs & Kumagai, 1973), and “the benefits provided by a transportation/land-use system” (Ben-Akiva & Lerman, 1979). Metrics used in accessibility studies include those based on distance, such as straight-line, Manhattan,¹ or network distance, and distance derivatives, such as times or “cost” (e.g. in dollars/euros). In studies evaluating accessibility to a facility, the service areas or catchment areas are derived as polygons based on a threshold distance or cost (Delamater,

Messina, Shortridge, & Grady, 2012; Vadrevu & Kanjilal, 2016).

While many accessibility studies are based on the shortest path/cost metric, a more realistic measure is based on observational data. Observational data of one’s travel experience was typically only from volunteers. However, such data (e.g. static travel scenario or self-reported travel diary) not only lack accuracy but also involve subjective bias due to limited samples. Recently, many public transportation companies (e.g., taxi, bus, and subway) have released their transportation data, in much finer spatiotemporal resolutions, such as the individual taxi trips with pickup-dropoff location and time, which offer unprecedented opportunities for data-driven accessibility measurement.

Public transit usually refers to buses and subway systems in urban areas. Accessibility to public transit has drawn intensive attention in transportation research from different perspectives. As an environment-friendly commuting and travel mode, public transportation can help reduce greenhouse gas emissions, traffic congestion,

¹Technically, the Manhattan distance is a form of network distance although it is often treated as something unique.

* Correspondence: yuqin@email.sc.edu
Department of Geography, University of South Carolina, Columbia, USA

car accidents, and oil price vulnerability (Litman, 2003). Research also found that public transportation users have a lower obesity rate and present better physical and mental health (Sallis, Frank, Saelens, & Kraft, 2004). Moreover, accessibility of public transportation can be used as an indicator to measure social equality. For example, measurements of access to public transportation helped researchers to identify socially disadvantaged groups in gender (Kwan, 1999; Kwan, Murray, O'Kelly, & Tiefelsdorf, 2003), age (Hess, 2009), socio-economic status (Niedzielski & Boschmann, 2014), races (Tribby & Zandbergen, 2012), and disability (Church & Marston, 2003). Traditionally, public transit accessibility was measured by service frequency estimation or travel situation simulation. Recently, the increasing availability of public transit usage data opens the possibility to measure public transit accessibility more realistically and dynamically with a big data approach.

Taxis, unlike buses or subways, provide private and convenient location-to-location transportation services. In the past few years, taxi companies worldwide have installed Global Navigation Satellite System (GNSS) equipment in taxicabs. In general, two types of travel data can be collected from an in-car GNSS. The first type of data has the entire route recorded with GNSS positions regularly sampled at a specified time interval (Herring, Hofleitner, Abbeel, & Bayen, 2010). The second type of data only contains information on the origin (pick-up location) and destination (drop-off location) of each taxi trip, and the travel distance, duration, and cost, without recording the actual route of each trip (Guo & Zhu, 2014; Guo, Zhu, Jin, Gao, & Andris, 2012).

Both public transit accessibility and taxi trip data have been individually studied to reveal different urban characteristics. Public transit data have been used to analyze job opportunities (Farber & Fu, 2017; Lei, Chen, & Goulias, 2012), food deserts (Burns & Inglis, 2007; Paez, Gertes Mercado, Farber, Morency, & Roorda, 2010), and activity-based research (Mavoa, Witten, McCreanor, & O'Sullivan, 2012). Continuous taxi data have been used for travel condition monitoring and road network analysis (Veloso, Phithakkitnukoon, Bento, Fonseca, & Olivier, 2011). Trip-based taxi data is useful for urban land use and human mobility analysis (Peng, Jin, Wong, Shi, & Liò, 2012). However, few studies have integrated more than one type of transportation mode in accessibility measurements.

This paper proposes a novel approach to measure and visualize urban accessibility, with big data of taxi trips and public transit uses (including both bus and subway), using New York City (NYC) as a study case. Specifically, a new Urban Accessibility Relative Index (UARI) was developed by integrating multiple transportation modes and big data of daily mobility, and subsequent analyses were then carried out to visualize and understand the

spatiotemporal distribution of accessibility patterns in NYC.

2 Literature review

2.1 Accessibility

In the simplest case, two places (or points) are connected, which means accessibility exists between these two places. As noted earlier, access based on distance can be straight-line, Manhattan, or network distance. Thus, using a transportation road network, accessibility between two places can be measured as the length of road, or the travel time/cost that connects the two places. When measuring accessibility from a social or economic perspective, an "attraction" variable is often added to a distance decay function. Hansen (1959) introduced a gravity model in accessibility and land-use. For example, as the distance between home and a shopping center increases, the possibility for one to go to that shopping center decreases. Different functional forms are applied to calculate the distance decay between two locations, including power, exponential, and Gaussian (Scott & Horner, 2008).

Accessibility in transportation research generally measures how easy, or how difficult, for people to get to their opportunities or services (Wachs & Kumagai, 1973). Based on different applications, Geurs and van Wee (2004) grouped accessibility into four groups: 1) infrastructure-based accessibility, 2) location-based accessibility, 3) person-based accessibility, and 4) utility-based accessibility. Infrastructure-based accessibility measures the performance of a road network, such as travel speed and congestion conditions. Location-based accessibility measures the number of places of interest that can be reached from an origin. Person-based accessibility comes from space-time geography, which measures places that can be reached given individual's time and space constraints (Kwan, 1999; Miller, 1991). Utility-based accessibility measures the usage of a certain transportation mode or the market share of a transportation mode (Ben-Akiva & Lerman, 1979).

Based on different data types used in accessibility measurements, Páez, Scott, and Morency (2012) grouped accessibility measurements into two categories: normative measurement and positive measurement. *Normative measurements* do not use observational data (i.e., people's behavior) and consider only the performance of transportation. Larsen and Gilliland (2008) measured food deserts in urban London, ON, Canada, based on minimum walking and public transit accessibility. Farber, Morang, and Widener (2014) used public transit data to study the temporal variability of public transportation. *Positive measurements* use people's travel behavior data, which come from surveys or behavioral models. For example, Minocha, Sriraj, Metaxatos, and Thakuriah

(2008) used local trip data to estimate demand factors. Pasch, Hearst, Nelson, Forsyth, and Lytle (2009) surveyed teenagers to study the association between teenagers' alcohol use and alcohol outlet locations. Scott and Horner (2008) conducted a travel diary survey for urban opportunity accessibility.

Accessibility in urban areas has been studied in many applications using different types of data. In this article, three major areas of urban accessibility are reviewed. The first area includes studies related to *public transit accessibility* (Section 2.2). Each public transit trip can be divided into three segments: from one's origin to transit network, travel inside transit network, and from transit network to destination. Because of this division, public transit accessibility has two sub-areas: to/from transit network and in a transit network. The second major area of research involves *taxi travels* (Section 2.3). Taxi travel data is a direct indicator of accessibility in terms of monetary or time cost, and more complicated and meaningful information can be retrieved from taxi travel data. The third major area focuses on *relative accessibility* (Section 2.4), which includes comparison among more than one type of accessibility measurements. Such comparison can be conducted between accessibility using different transportation modes, during different time periods, or for different groups of public transit users.

2.2 Accessibility in public transit

Accessibility measurements in public transit can be further grouped into two categories: to/from transit network and in-transit network. The to/from transit network accessibility measures the combined travel from their original location to the transit stop and travel from the last transit stop to their final destination. The in-transit-network accessibility considers the cumulative travel time/cost on buses or subways, transfer time among different lines, etc. Malekzadeh and Chung (2020) provides a review of transit accessibility models.

A to/from transit network accessibility measures the physical access to a transportation network, defined as how far one person needs to travel to a transit stop, either a bus stop or a subway station (Currie, 2010). If the accessibility is based on a threshold distance the resulting polygonal area is a service area/catchment area of a transit stop. There are two traditional methods to measure a service area: circular buffer analysis and road network analysis. However, the buffer analysis is rarely used for human-related service areas as the result is often an overestimation of service areas and served population. Biba, Curtin, and Manca (2010) and Guti and Garc (2008) compared service areas and population inside service areas using circular buffer analysis and road network analysis. Their results showed that buffer zone overestimates one-third to half of the service areas

and served population. This overestimation is due to the difference between distances constrained by travel along a road network and the straight-line distance used for the buffer analysis. Studies also showed that individuals' walking speed and maximum tolerant walking distance affect accessibility of public transit accessibility (Hess, 2009; Mavoia et al., 2012). In addition, other factors, such as safety, weather, infrastructure (Walton & Sunseri, 2010) and even an individual's social identity (Murtagh, Gatersleben, & Uzzell, 2012) also affect accessibility to/from the public transit network.

Derived from the general definition of accessibility, the in-transit accessibility measures the ease, or difficulty, often expressed simply as "cost", for passengers to get to their destinations using public transit systems. The measurement of travel cost varies in different applications and different research preferences. The most straightforward measurement is the travel distance, or the length of a public transit route between the origin and the destination, which can be calculated with road network data (Liu & Zhu, 2004). These measurements, however, often assume that the travel speed in the network is constant, which obviously varies by roadway type, speed limits, congestion, and such variability diurnally and episodically.

Compared to route length, travel time is a more commonly used metric in measurements for travel cost. While an accessibility polygon based on a maximum travel cost is often used as a service area, the algorithm for calculating costs can be done for an infinite number of origins or a systematic sample, as with a raster. Thus, a cost surface can be created from the a large set of origins. For example, O'Sullivan, Morrison, and Shearer (2000) developed a tool to draw isochrones lines over a cost surface to represent public transit accessibility. Given an origin, their tool can draw accessible areas that can be reached within a given time threshold using public transit. Their travel time estimation is based on average travel speed along road network. Later studies implemented public transit timetables and potential adjustments into their travel time calculation models to increase accuracy (Cheng & Agrawal, 2010; Lei & Church, 2010; Zhang, Dong, Zeng, & Li, 2018).

Another important development in accessibility measurement is to measure accessibility for different time periods of a day. Polzin, Pendyala, and Navari (2002) included supply and demand during different time periods in 1 day. Based on the predicted ridership and available public transit service, their model calculates the availability of transit services as the daily trips per capita. Chen et al. (2011) applied a temporal component in job-based accessibility. Liao, Gil, Pereira, Yeh, and Verendel (2020) analyzed travel time between cars and transit in four different cities around world and found car traveling time is shorter in majority parts of the all four cities.

2.3 Accessibility with taxi trip data

Taxis in urban areas provide convenient and private origin-to-destination transportation services based on customers' requests. In the last decade, with the advancement of GNSS technology, many taxi companies have installed in-car GNSS trackers in taxicabs. This tracking data not only help monitor taxicabs' movements for better navigation and dispatching but also guarantees a safer environment for taxi drivers. Taxi data have been used for different research objectives (e.g., measuring accessibility) in cities all over the world, including San Francisco, USA (Herring et al., 2010; Hoque, Hong, & Dixon, 2012; Hunter, Herring, Abbeel, & Bayen, 2009), Lisbon, Portugal (Veloso et al., 2011), Shanghai, China (Peng et al., 2012), Stockholm, Sweden (Jenelius & Koutsopoulos, 2013), and Delft, Netherlands (Zheng & Van Zuylen, 2013).

Taxi data can be grouped into two categories: tracking data and origin-destination (OD) data. Tracking data consists of taxi trajectories obtained from GNSS receivers that also record taxi locational information at a specified time interval, usually at a 30 s or 60 s interval. They are useful to monitor road network conditions and are commonly used to measure infrastructure-based accessibility. Since GNSS locations are less accurate in urban settings a pre-processing step is typically used to project any 'off-road' points onto road networks. Based on the changes in location, driving speed can be inferred and thus travel time can be predicted. Hunter et al. (2009) used about 60,000 observations from 50 taxicabs in San Francisco to calculate travel time. Herring et al. (2010) used a probabilistic model to estimate arterial traffic condition from 500 taxicabs in San Francisco. Veloso et al. (2011) used a Gamma distribution to model taxicabs' distribution over Lisbon, Portugal area. Jenelius and Koutsopoulos (2013) used data from 1500 vehicles in Stockholm, Sweden to estimate travel time between any two points on the road network. Zheng and Van Zuylen (2013), based on probe cars, created a three-layer neural network to simulate travel conditions in Delft, Netherlands.

2.4 Relative accessibility

Relative accessibility compares different accessibility measures for population groups. These comparisons can be between poverty and non-poverty (Niedzielski & Eric Boschmann, 2014), personal identities (Murtagh et al., 2012), genders (Kwan, 1999), and age groups (Hess, 2009). The most commonly compared population groups in transportation and planning research are between public transportation and private car driving. Implemented by O'Sullivan et al. (2000), mapping service areas from a given location of public transit and driving provides the most direct comparison between accessibility measurements.

Besides visual comparison, travel time ratios and location-based accessibility measurements are commonly used to quantify differences in accessibility.

Mapping the ratio of travel time between public transit and private vehicle is one of the common methods. Hess (2005) focused on low-wage job accessibility for low-income adults in the Buffalo-Niagara region. Using the centroids of each neighborhood, they calculated a 30-min travel buffer for automobile driving and public transit riding, respectively. Their job accessibility measurements were the summations of jobs within a 30-min travel time. They used the ratio of automobile and public transit job accessibility in each neighborhood. Their results showed that automobile drivers have 2 to 3 times more job accessibility than public transit users. Salonen and Toivonen (2013) used the travel time ratio of public transit to private vehicle to map accessibility to public libraries in Greater Helsinki area, Netherland. They used three different models for public transit and driving: (1) a simple model that ignores congestion and parking, (2) an intermediate model that includes congestion but ignores parking, and (3) an advanced model considering both congestion and parking. Not surprisingly, in all three models, the average travel time for public transit is longer than average travel time for private vehicle. In addition, public transit travel time to the closest destination is also longer than that for private vehicles. However, this ratio is calculated at the infrastructure level, by measuring how public transit and taxi are performing locally. Only one origin can be used at each time. When the origin is determined, it measures accessibility as a property to all possible destinations. Temporally, Farber and Fu (2017) measured public transit accessibility using origin-destination travel time cubes to model fluctuations during a day.

There are other studies on location-based accessibility measurements, which often employ an opportunity index, to compare accessibility among different locations. Shen (2001) studied job opportunities differences between public transit and private vehicles. They defined an accessibility score as the ratio of the total number of opportunities to the total of opportunity seekers for each zone. They generated 775 transportation analysis zones in the Boston Metropolitan Area and calculated a job opportunity index for each zone using public transit and private cars. Their study showed that not many job opportunities exist if one only uses public transit. For these opportunity-based calculation methods, researchers need to subjectively assign a weight or score to different opportunity or activity types, or subjectively classify service or importance levels for different opportunities. Compared to ratio measurements, opportunity-based measurements consider accessibility as a property of the origin, but these weights or classifications are usually based on survey results or arbitrary assignments.

In deference to the previous studies we have proposed and demonstrate the use of a new index – urban accessibility relative index (UARI) – that combines a location-based focus and connection-based focus using observational data. The UARI is ideal for comparing accessibility between travel modes, such public transit and alternative transit measurements (e.g. taxi transit).

3 Study area and data

3.1 Study area: New York City

This research focuses on New York City (NYC), consisting of five boroughs: Brooklyn, Queens, Manhattan, the Bronx, and Staten Island (Fig. 1). Each of the five boroughs is a separate county in the state of New York. NYC had an estimated population of 8,491,079 in year 2014 and an area of about 800 km² (U.S. Census Bureau). Since Staten Island has a separate subway system that is not connected to the main subway system and has very limited bus routes connecting to the main areas of NYC, Staten Island was excluded in this study.

NYC has the highest population density of all major cities in the United States, which makes NYC an ideal study area for this research for two main reasons. First, the public transportation ridership is very high in NYC. Due to limited spaces and high land price, NYC has the lowest car ownership in the United States, with 66% households not owning a private car (Salon, 2009). Therefore, residents’ daily commuting and traveling rely heavily on public transportation. According to the data from Metropolitan Transportation Authority (MTA)’s ridership report in 2013, the annual ridership was more

than 1.7 billion for subway and 0.67 billion for transit buses. On an average weekday, the ridership was about 5 million for subway and 2 million for transit buses. According to the data obtained from the New York City Taxi & Limousine Commission, there were more than 14 million taxi trips for each month in 2013 (i.e., roughly half a million taxi trips per day). Second, NYC has a complex, dense and effective public transportation network and a large fleet of taxi cars. In the Manhattan area, bus stops or subway stations are within walking distance. In 2013, there were 13,437 taxicabs operating in NYC. In addition, there are 21 subway routes with 494 subway stations in NYC. Transit bus network consists of 237 local routes and 65 express routes (MTA, 2013). This high density of the public transportation network and high diversity in transportation modes makes NYC an ideal test bed with abundant data for the studying, understanding, and use of accessibility.

3.2 Data

3.2.1 GIS-based data

Basic geographical information data, including city and borough boundaries, are available from the New York State GIS Clearinghouse (<https://gis.ny.gov/>). Subway and transit bus routes are digitized and maintained by the City University of New York Mapping Service at the Center of Urban Research (<http://www.gc.cuny.edu/CUR>). Hospital data used in the application parts can be found from New York State Open Data – Health Data (<https://health.data.ny.gov/Health/Health-Facility-Map/875v-tpc8>). This dataset includes all types of healthcare



Fig. 1 Study area of New York City

facilities. For demonstration purposes, only hospital data were used in the application presented in Section 5.2.

3.2.2 Taxi trip data

The New York City Taxi & Limousine Commission is one of the major taxi companies operating in NYC. Trip data are available for taxicabs holding license from the New York City Taxi & Limousine Commission. This research used the taxi trip data for the entire year of 2013, which has 13,437 registered taxicabs and 173,179,759 taxi trips in 2013 (i.e. about half a million taxi rides each day), with a total of 1.99 billion dollars for taxi fare (tips were not included). Information associated with each trip includes: pick-up data and time, drop-off date and time, passenger count, trip time in second, trip distance, pick-up location (latitude and longitude), drop-off location (latitude and longitude), payment type, fare amount, surcharge, MTA tax, toll amount, and total amount. The average taxi trip time was 799 s (about 13 min), the average trip distance was 4.65 km and the average fare for a taxi trip was 11.49 dollars.

In this research, it was assumed that the actual driving route of each recorded taxi trip (which was not available in the source data) is the shortest path (in terms of network travel time) from the origin to the destination.

3.2.3 Public transit data

Public transit data for NYC subway and transit buses are published and maintained by the MTA, which is the company that operates NYC subways and major transit bus routes. The public transit data are in General Transit Feed Specification (GTFS) format, containing public transportation schedules and associated geographical information. The structure of GTFS data includes agency, routes, trips, stops, stop times, and calendar. Detailed explanation of the GTFS data format can be found at <https://developers.google.com/transit/gtfs/reference>.

4 Methodology

This research proposes a UARI to compare the relative accessibility between public transit and taxi. As reviewed in Section 2.4, there are two main approaches to measure relative accessibility: travel time ratio measurement and opportunity-based measurement. Travel time ratio is used to measure the relative performance between taxi and public transit and considers accessibility as a property of the connection between origin and destination (rather than a property of the location, such as either the origin or the destination). Opportunity-based measurement, on the other hand, views accessibility as a property of the origin (rather than the connection), which involves arbitrary decisions on different types of destinations.

The proposed UARI measurement is derived with a regression approach. The UARI for a given location is defined as the slope of the regression line, with public transit travel time on the y-axis and taxi travel time on the x-axis, for all (or a selected group of) destinations from the given location. If the regression slope is 1.0 then no difference in travel time exists between travel modes. If the slope is greater than 1.0 the public transit time is greater than a taxi mode of travel. This new index enables both location-based measurement (by mapping ratio-based relative accessibility of the given location) and connection-based measurement (by comparing the relative accessibility of different locations), both of which can vary across space and time.

4.1 Computation of public transit accessibility

Given an origin (or a destination) and a departure time (or arrival time), the total travel time can be estimated based on the complete public transit schedule, with the arrival and departure time for each bus or subway train and the estimate of walk time for transfer connections within the network and to/from origin/destination. The Dijkstra shortest path algorithm (Dijkstra, 1959) was used to find the expected travel time using public transportation between the origin and the destination, including walking time to/from stations, waiting time, riding subways and/or buses, walking for transfers, and waiting time during the transfer. A threshold of 500 m was used to define a maximum “walkable” distance from an origin location to a public transit and from the public transit to a destination, calculated using network distance (perhaps more appropriately called Manhattan distance in NYC).

4.2 Computation of taxi accessibility

Given an origin (or a destination) and a departure time (or arrival time), related taxi travel records will be retrieved and processed to estimate the travel time, which can be the average time of all taxi trips that started from the neighborhood of the origin around the given departure time and ended near the destination. Taxi trips within a 500-m distance from the selected location are eliminated from the calculation, as we assume 500-m is the walking distance threshold. Since this method uses actual historical taxi trip records (instead of modeled travel over road networks) to derive the actual travel time between a given origin and a destination for a specific departure time, it implicitly considers traffic conditions and other unknown variable factors in calculating the driving time.

4.3 Urban accessibility relative index (UARI)

4.3.1 Computing UARI for one location

UARI can be calculated for either an origin or destination and a selected time period. To calculate the UARI

for a given origin, the first step is to build a 500-m buffer zone, as the origin zone, around a hypothesized origin. All taxi trips leaving from this origin zone are selected and all the public transit stations, including both bus stops and subway stations, are also selected. Public transit accessibility and taxi accessibility are calculated using methods described in Sections 4.1 and 4.2, respectively. For each pair of origin-destinations (O-D) with existing historical travel records, the UARI for this O-D pair can be calculated with Eq. 1:

$$UARI = \frac{accessibility_{public\ transit}}{accessibility_{taxi}} \tag{1}$$

To calculate UARI for a given destination, trips arriving at the selected destination zone (500-m around a destination) at the given time are selected. Public transit accessibility and taxi accessibility are then calculated. Similar to the UARI calculation process for origins, the UARI of each pair of destination-origins (D-O) with existing travel records are calculated. Of note is the UARI derived in Eq. 1 implicitly does not include walking time or waiting time for either public transit or taxi stops.

When calculating the UARI for an origin, all the O-D pairs have the same origin. Therefore, the UARI can be viewed as a property associated with the destinations corresponding with the trips leaving from the given origin. For the same reason, the UARI for a given destination is a property associated with the origins where trips arriving at this given destination start from.

The UARI for Pennsylvania Station (Penn Station) as the origin was used to demonstrate the calculation process in this study. Penn Station is an important transit hub connecting commuting trains from New Jersey and public transit network in NYC. The UARI for Penn Station was calculated for two different times: 3:00 am and 3:00 pm, to compare the dynamics of UARI during different times of a day. A 10-min time window was applied to both starting time so that trips starting between 2:55 am and 3:05 am are included for 3:00 am to allow more flexibility.

4.3.2 Computing UARI for multiple locations

To calculate UARI for multiple origins, the first step is to conduct the calculation for each of the available origins. Then for each of the given origins, its public transit time and taxi time to all potential destinations are plotted, with the taxi travel time on the x-axis and the public transit time on the y-axis. A linear regression analysis is used to derive the UARI between the public transit travel time and taxi travel time. Specifically, the total least squares regression method is used:

$$y = \alpha + \beta x \tag{2}$$

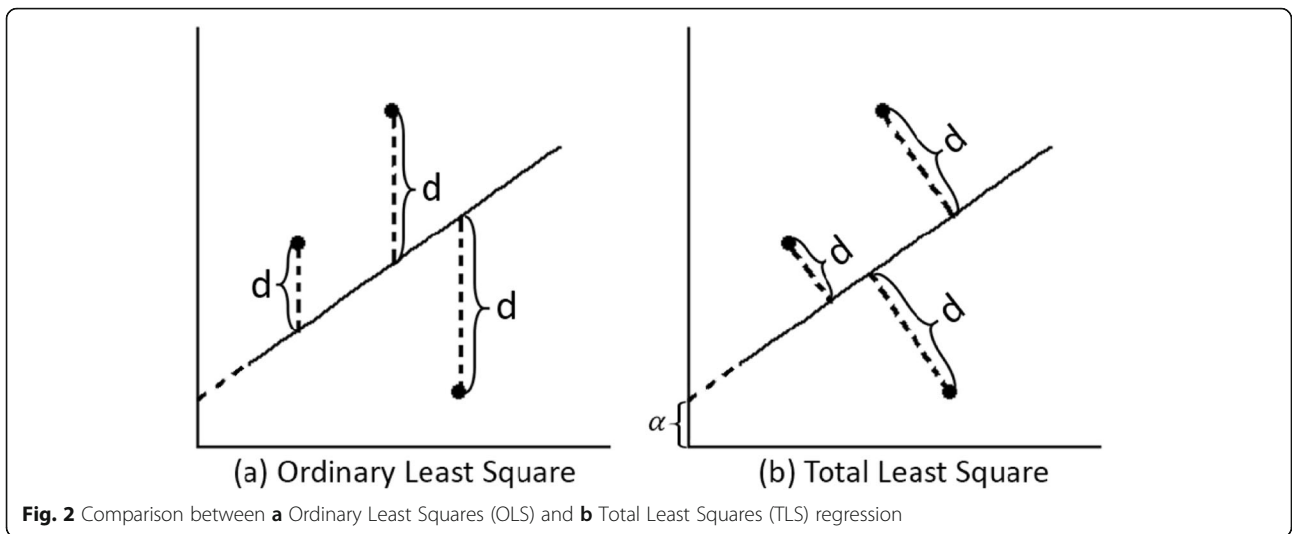
where the slope of this regression line, β , is an overall measurement of how efficient the public transit travel time is compared to taxi travel time. We use β as the UARI. Compared to the ordinary least square regression (Fig. 2a), the total least square regression (Fig. 2b) calculates residuals for both x and y, which allows us to treat α and β symmetrically (Golub & Van Loan, 1980). In this research, errors exist in both taxi and public transit measurements and thus the total least square regression is more suitable for this case.

In Eq. 2, α (intercept with y-axis) can be interpreted as the sum of the walking time to the public transit origin and the waiting time for next bus or subway. Because taxi waiting time is not available, α is not reported or discussed in this study.

For a specific location the UARI is the expected change in public transit travel time given a 1.0 unit change in travel time by taxi. For example, if one location has a UARI of 7.5, it means, for each minute in taxi travel time, public transit riders should expect 7.5 min of public transit – or 7.5 times the taxi travel time. A UARI of 1.0 means taxi and public transit travel time have essentially similar performances. Therefore, a higher UARI means the location has a lower relative accessibility (i.e. not convenient for people to use public transit system compared to using taxi) and a lower UARI means a higher relative accessibility, that public transit has similar performance with taxi (assuming no walking and waiting time).

Figure 3 is an example of trips starting at Penn Station at 3:00 pm. Each point on this scatter plot represents one Penn Station-to-destination pair. For this Penn Station-to-destination pair, the origin is within a 500-m buffer zone around Penn Station and the destination is outside this 500-m buffer zone. Location of the point in this coordinate system is determined by the travel time using taxi and public transit. The slope of the red line (2.8) is the UARI for the cell containing Penn Station during 3:00 pm (Fig. 3).

For each O-D pair, both public transit travel time and taxi travel time (if existing) are retrieved. Because all subway stations and bus stops are connected in the public transit system network, cells in which subway stations or bus stops are located, as well as cells reached by walking, have values for transit travel time. However, not all O-D pairs have taxi trips. For O-D pairs having more than one taxi trip, the average time for all taxi trips is used for taxi travel time of that O-D pair. UARI values are only calculated for origins with 10 or more destinations. A minimum frequency of 10 was subjectively chosen in this case study. Similarly, when calculating UARI for a destination, that destination must have no



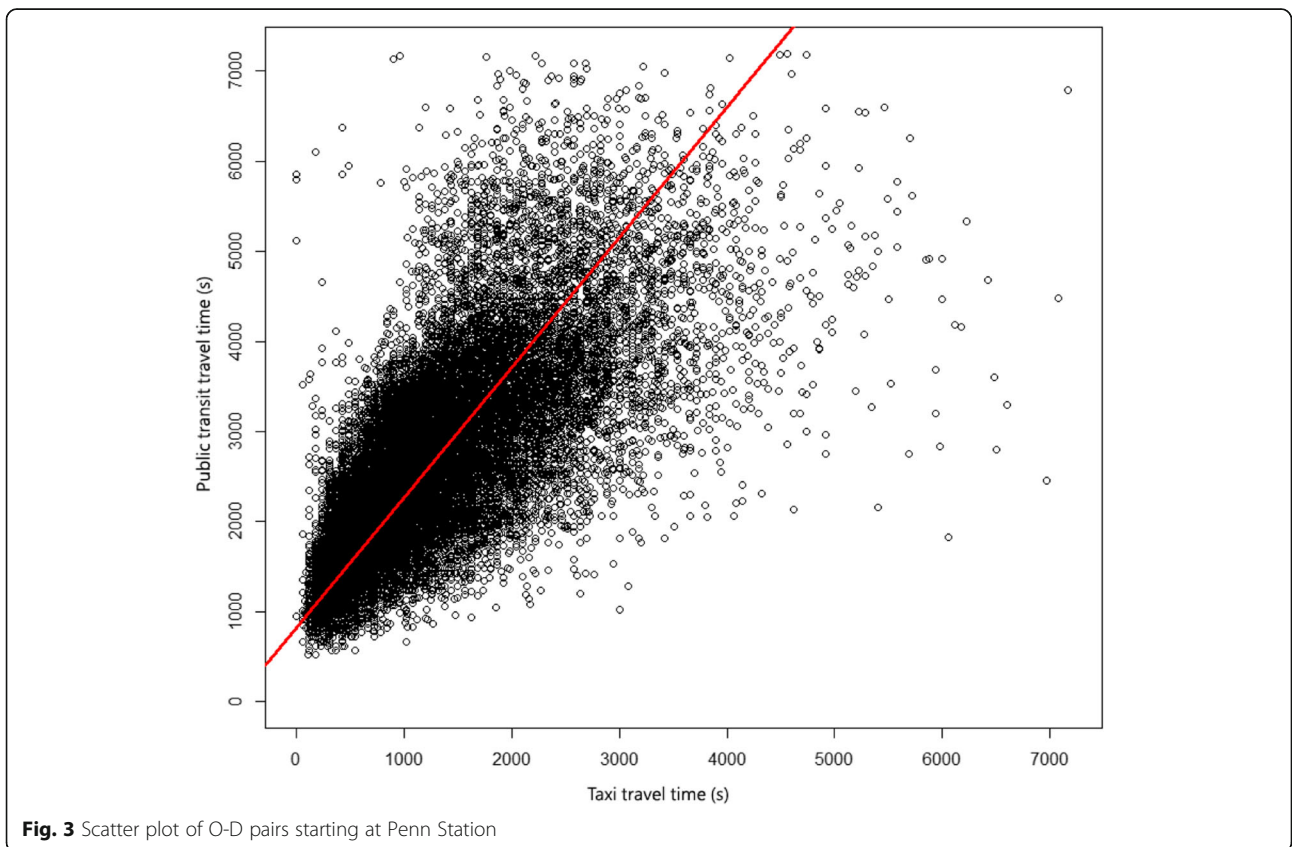
less than 10 origins. In other words, to run the regression, no less than 10 points must exist on the scatter plot.

To calculate UARI for multiple destinations, trips arriving at each destination are selected first. The UARI is calculated using the same procedure as calculating UARI for origins.

5 Results and discussion

5.1 UARI for one origin

Relative urban accessibility values can be geographically displayed by mapping the UARI values for all locations. The UARI for Penn Station as an origin is shown at 3:00 am (Fig. 4) and 3:00 pm (Fig. 5). The origin (Penn Station) is at the center of the ‘black hole’ on Staten



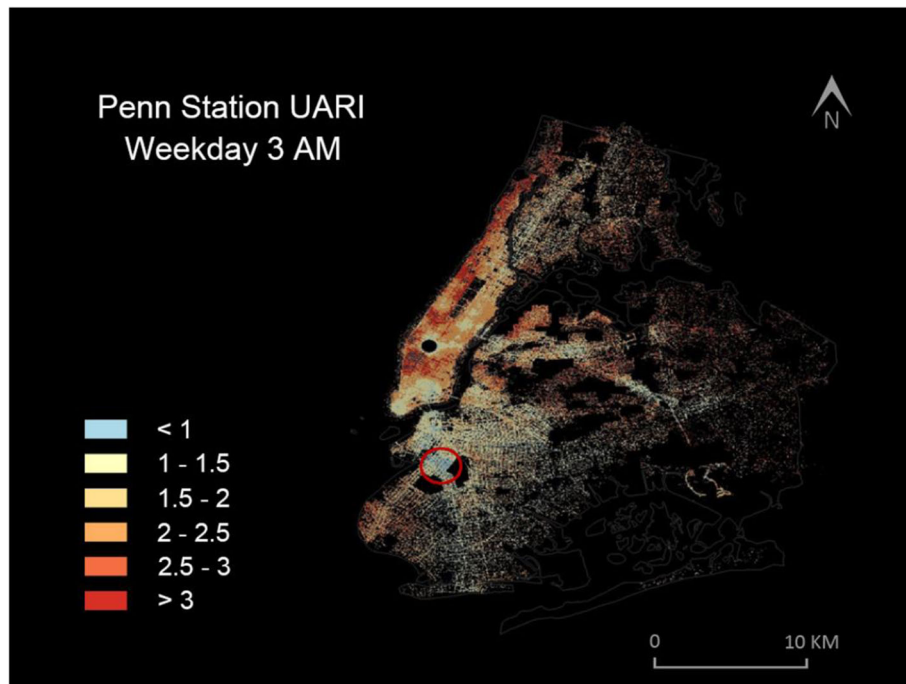


Fig. 4 UARI for 3:00 am using Penn Station as the origin

Island as short trips (trip distance less than 500 m) were excluded in this analysis. In these two maps, blue colors indicate that the public transit time is shorter than taxi time, while yellow to red colors indicate that a taxi takes less time than public transit to travel from Penn Station to the destination. For some destinations (e.g. upper Staten Island), the public transit time can be more than three times longer than taxi travel time. Noticably, the Park Slope neighborhood (red circle in Fig. 4) in Brooklyn, the public transit time is less than taxi travel time. This neighborhood is located near Barclays Center, the home to the NBA Brooklyn Nets basketball team. A total of nine subway lines are going around this neighborhood, which reduces public-transit travel time.

The frequency of travel by subway, expressed as the times between arriving subway train, has a major impact on the UARI values in this example. The early morning time (i.e. 3:00 am) was chosen as an example when very few public transit services are available. At this time of day, the Manhattan area shows 2 to 3 times longer travel time using public transit than using taxi (Fig. 4). Even though Manhattan is generally considered to have both the most road congestion and the most convenient public transit system, during the night hours, public transit accessibility is markedly reduced by limited public transit services. The afternoon travel time of 3:00 pm illustrates a notable difference in relative travel mode times on Manhattan Island (Fig. 5). Curiously, public travel time north of Penn Station is less than taxi travel time

while travel time to the south is more efficient by taxi. One reasonable explanation is that subway network is more accessible, compared to taxi travel, south of Penn Station.

5.2 UARI for multiple origins

Figures 6 and 7 illustrate the UARI values for all origins in NYC during two different time periods in a day. Figure 6 indicates UARI for 3:00 am, representing transportation conditions during the early morning hours and Fig. 7 shows UARI for 3:00 pm. In both figures, the red color indicates steep slopes for the regression line (i.e. the UARI value), which means longer travel time with public transit compared to taxi travel, while the blue color indicates shorter travel time using public transit. These hypsometric classification scheme for Figs. 6 and 7 was based on quantile divisions of all possible slopes. In other words, the array of slope values (combined 3:00 am and 3:00 pm) was divided into 9 classes with equal numbers of observations in each class.

For each origin, we first plotted travel time for all the destinations onto a scatter plot, similar to Fig. 3, and then we calculated UARI using the method described in Section 4.3.2. In other words, we generated a scatter plot for each origin cell, and the UARI for the origin was calculated as we solved the β in the regression equation (Eq. 2). Due to a limited number of taxi trips and reduced public transit services at 3:00 am, many cells have

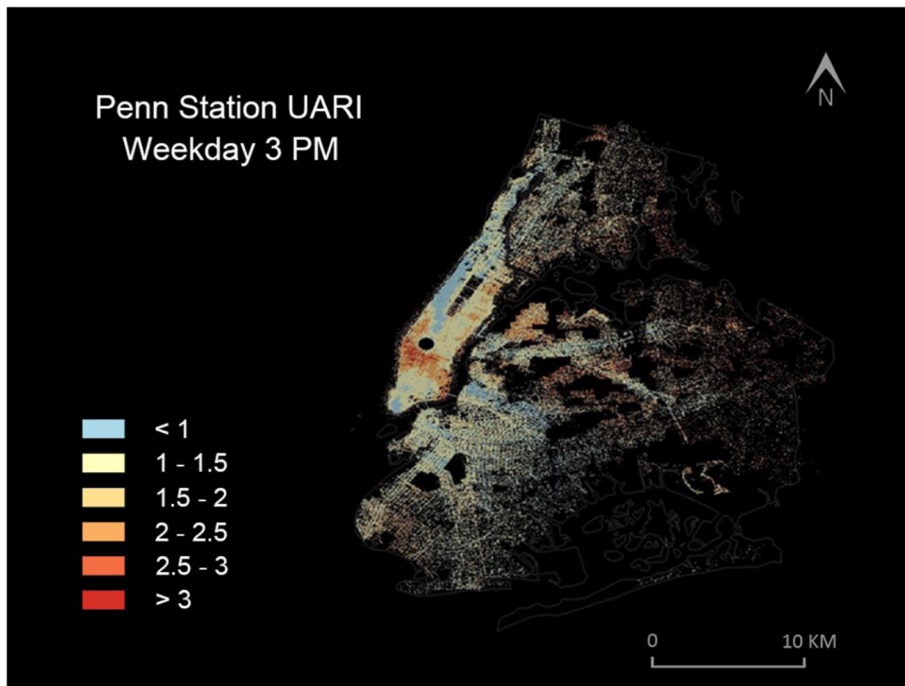


Fig. 5 UARI for 3:00 pm using Penn Station as origin

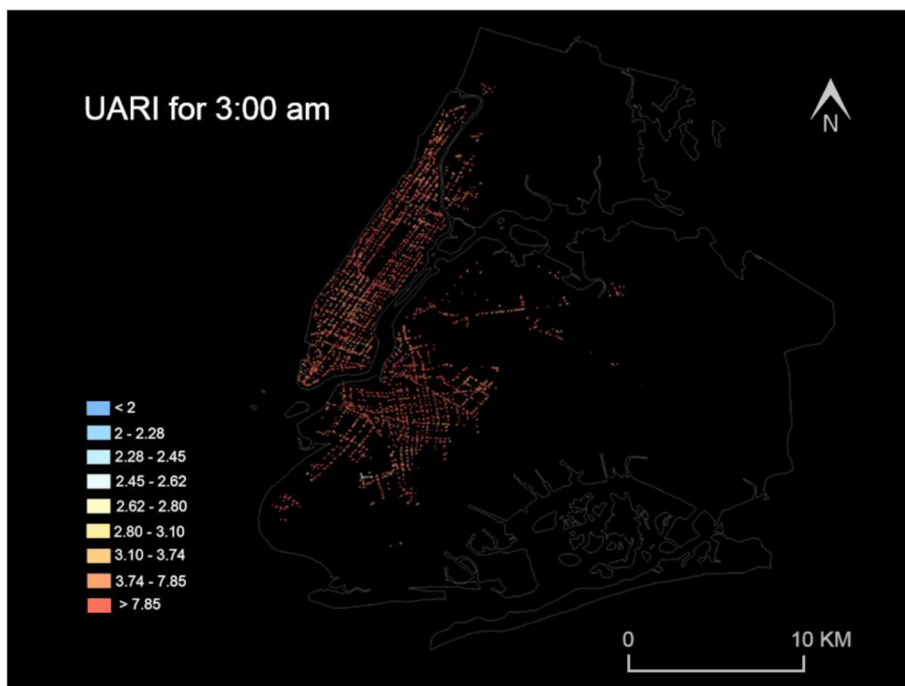


Fig. 6 UARI for all origins at 3:00 am

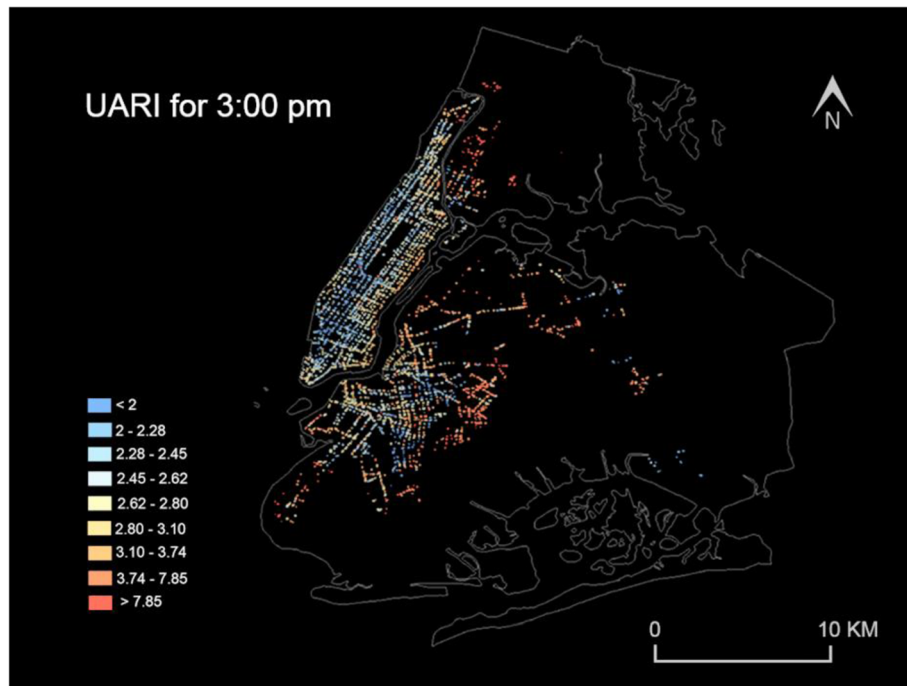


Fig. 7 UARI for all origins at 3:00 pm

no value, meaning no UARI at 3:00 am. Thus, a missing UARI value means either the cell was inaccessible using public transit service or no taxi trips were available in records at the given time. All cells with no values are shown in black color. Therefore, Figs. 6 and 7 are not comparable to Figs. 4 and 5 as the UARIs were calculated for multiple origins.

In Figs. 6 and 7, the majority of Manhattan areas and some parts of Brooklyn have very similar values. In Queens, only areas along major subway lines and around some stations have values.

In the map for 3:00 am (Fig. 6), most areas are shown in red color, which means public transit riders should expect at least 4 times the taxi travel time needed. In some areas (shown in dark red), public transit riders should expect almost 8 times or more public transit time than taxi time. The map for 3:00 pm (Fig. 7) is quite different from the one for 3:00 am, where the majority of Manhattan is covered by UARI values of only 2 to 3. This means the majority of Manhattan areas have low UARI, indicating high public transit accessibility. During daytime hours, for every minute traveling in taxi, public transit riders should expect about 2 min travel time using public transit system.

In Figs. 6 and 7, not all locations had an ample frequency (i.e. 10 or more) trip records to be considered as a valid location in UARI measurement. Visual examination of Figs. 6 and 7 provided evidence of travel demands for taxis. Since public transit network time

covers all of NYC, whether a location is valid or not was actually determined by the number of taxi trips starting from that location. In Figs. 6 and 7, most areas of Manhattan have enough taxi trips to be considered as valid. Outside of Manhattan, most valid locations are only along subway lines. This distribution pattern indicated taxi travel demand. From this visual examination, NYC provided a good public transit service to meet travel demands.

At 3:00 am, UARI results indicated that for the majority of NYC, public transit takes three to four times comparing to taxi travel times, even in the Manhattan area, where people would expect the most convenient public transit services. At 3:00 pm, the Manhattan area shows results as expected with public transit times only slightly longer than taxi times. This difference between 3:00 am and 3:00 pm is indicative of the frequency of subway services during day hours and night hours. With a reduced number of subways during night hours and consequently longer waiting times, accessibility for public transit during night is much lower than accessibility during daytime.

5.3 UARI for multiple destinations

This section presents results from UARIs calculated for selected destinations. Nine major hospitals were selected as destinations. These results provide a practical scenario for comparing the accessibility of hospitals at 3:00 am and 3:00 pm. Similarly as before, for each hospital as

a destination, the original cell must have more than 10 trips to be considered as a valid origin for that O-D pair. Table 1 and Fig. 8 provide the UARIs for these nine hospitals as destinations at 3:00 am and 3:00 pm.

Seven of these nine major hospitals are located in lower Manhattan and two are located in Brooklyn. Not surprisingly, UARIs at 3:00 am are much higher for all hospitals than UARIs at 3:00 pm, indicating lower accessibility of public transit system during night hours. The Woodhull Medical & Mental Health Center hospital has the lowest UARI at 3:00 pm (1.8). This hospital also has the second lowest UARI at 3:00 am. Compared to other hospitals. NYU Hospitals Center and Bellevue Hospital Center are located farther away from subway routes. These two hospitals have the highest UARI during both time periods, which indicates low public transit accessibility to reach these two hospitals.

6 Limitations

The methods and data sources used in this research have some limitations and need further improvement. First, the taxi data at early morning hours for many locations have very low frequencies. For example, in the map of UARI for 3:00 am (See Fig. 6), not many places have historical taxi trips. Therefore, only a limited number of places have UARIs. In future research, additional efforts may be needed in the early stages of data to either remove outliers in taxi trips or to include a confidence measurement based on the frequency/variance of taxi times for a location. A second limitation of this research is with defining a valid location and the 500-m walking time. Since the actual origin of the commuter is not known we subjectively used a 500-m walking distance threshold to derive O-D pairs to an origin or a destination. Also, to ensure a large enough frequency (e.g. 10 in this study) we used the 500-m buffer area. Other possible sources of information (voluntary surveys, social media, etc.) could be exploited to derive more refined starting or ending locations emanating from the public transit or taxi stops. Many areas, especially areas other

than Manhattan, did not contain enough observations. In future research, a smoothing algorithm or a scalable filter could be applied to increase the number of valid locations. Also, a large temporal period of observations, either additional years of public transit data or taxi data, or larger hourly periods of origins (e.g. 1:00 am to 3:00 am) would increase the frequencies. Another limitation with public transit data was of possible unexpected delays. The UARI could be applied to city development or travel planning to include real-time information and improve estimates of actual time using public transit. In addition, the walking distance of people may vary in different places. More details about local people’s travel behaviors require further analysis.

7 Conclusion

This paper introduces and demonstrates a new measurement of accessibility (UARI), aiming to bridge current methodologies with the increasing availability of multimodal transportation data. The UARI developed in this study has three main innovations. First, it considers accessibility as a property of a location (either origin or destination) but derives the measure as the collective property of all connections that involve the location. The new measure is empirically derived and calculated, generally, with a large number of actual travel records so that there are less arbitrary decisions or biases involved compared to traditional methods. Second, the method uses historical taxi travel records and public transit timetables to accurately model travel time, rather than using road network and properties such as speed limits on road segments. Third, the new method and data can enable accessibility measurements at a fine spatiotemporal resolution, such as for different time periods of a day and different days of a week, based on big data of taxi trips and transit schedules (which, for example, differ significantly for weekdays and weekend days). As such, the new method can enable the analysis and understanding of dynamic accessibility patterns with time-varying and multimodal accessibility measurements.

Table 1 UARI for nine major hospitals

ID	Hospital Name	3:00 am	3:00 pm
①	Wyckoff Heights Medical Center	4.1	2.5
②	New York-Presbyterian/Lower Manhattan Hospital	5.5	2.9
③	Bellevue Hospital Center	7.5	2.7
④	Mount Sinai Beth Israel	6.3	2.7
⑤	NYU Hospital for Joint Diseases	6.1	2.7
⑥	New York Eye and Ear Infirmary of Mount Sinai	6.0	2.5
⑦	NYU Hospitals Center	7.8	2.6
⑧	Woodhull Medical & Mental Health Center	4.3	1.8
⑨	Lenox Health Greenwich Village	5.7	2.6

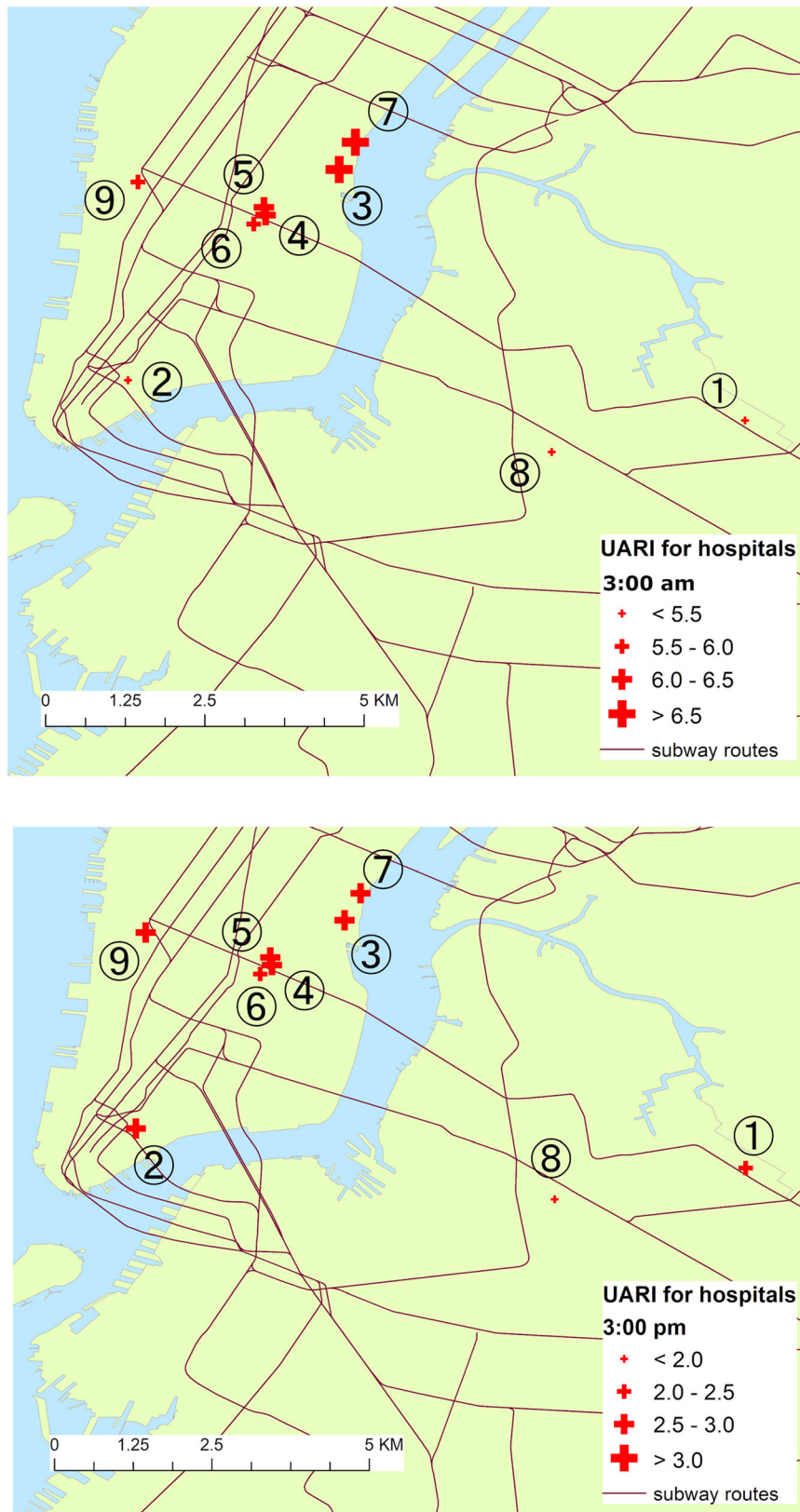


Fig. 8 UARI for the nine hospitals in NYC (refer to Table 1 for the hospital names)

Using NYC as a study case, UARI for different time periods showed the temporal changes in accessibility patterns in NYC. Comparing UARI of the same location during different times of a day indicates the temporal and spatial variation of accessibility. Examples in this study demonstrate that UARI can be used for both origins and destinations, and that the number of origins or destinations can be varied according to different applications. Potential applications of this method include, but not limited to, measuring the accessibility of hospitals, grocery stores, voting stations, and other public facilities for transportation planning. This method can further be integrated with more transportation modes, such as bikes and e-scooters sharing to encourage more environment-friendly transportation in city. In addition, transportation authorities can adjust current public transit routes or schedules according to this accessibility measurement. Also, given a standardized data format for taxi trip records and public transit timetable, this research can be applied other cities when relevant data are available.

Code availability

Not applicable.

Authors' contributions

Conceptualization: Y.J., D.G.; Methodology: Y.J., D.G., Z.L., M.H.; Analysis: Y.J., D.G., Z.L., M.H.; Writing-original draft: Y.J.; writing-review and editing: D.G., Z.L., M.H. The author(s) read and approved the final manuscript.

Funding

Not applicable.

Availability of data and materials

Data used in study are open data. Data sources are described in data section in the manuscript.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors have no conflicts of interest.

Received: 16 April 2021 Accepted: 27 May 2021

Published online: 15 June 2021

References

- Ben-Akiva, M., & Lerman, S. R. (1979). Disaggregate travel and mobility choice models and measures of accessibility. In *Behavioural travel modelling* (pp. 654–679).
- Biba, S., Curtin, K. M., & Manca, G. (2010). A new method for determining the population with walking access to transit. *International Journal of Geographical Information Science*, 24(3), 347–364. <https://doi.org/10.1080/13658810802646679>.
- Burns, C. M., & Inglis, A. D. (2007). Measuring food access in Melbourne: Access to healthy and fast foods by car, bus and foot in an urban municipality in Melbourne. *Health & Place*, 13(4), 877–885. <https://doi.org/10.1016/j.healthpla.2007.02.005>.
- Chen, Y., Ravulaparthi, S., Deutsch, K., Dalal, P., Yoon, S. Y., Lei, T., Goulias, K., Pendyala, R., Bhat, C., & Hu, H.-H. (2011). Development of opportunity-based accessibility indicators. *Transportation Research Record*, 2255, 19.
- Cheng, C.-L., & Agrawal, A. W. (2010). TTSAT- A new approach to mapping transit accessibility. *Journal of Public Transportation*, 13(1), 18.
- Church, R. L., & Marston, J. R. (2003). Measuring accessibility for people with a disability. *Geographical Analysis*, 35(1), 83–96. <https://doi.org/10.1111/j.1538-4632.2003.tb01102.x>.
- Currie, G. (2010). Quantifying spatial gaps in public transport supply based on social needs. *Journal of Transport Geography*, 18(1), 31–41. <https://doi.org/10.1016/j.jtrangeo.2008.12.002>.
- Delamater, P. L., Messina, J. P., Shortridge, A. M., & Grady, S. C. (2012). Measuring geographic access to health care: Raster and network-based methods. *International Journal of Health Geographics*, 11(15), 15. <https://doi.org/10.1186/1476-072X-11-15>.
- Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, 1(1), 269–271. <https://doi.org/10.1007/BF01386390>.
- Farber, S., & Fu, L. (2017). Dynamic public transit accessibility using travel time cubes: Comparing the effects of infrastructure (dis) investments over time. *Computers, Environment and Urban Systems*, 62, 30–40. <https://doi.org/10.1016/j.compenvurb.2016.10.005>.
- Farber, S., Morang, M. Z., & Widener, M. J. (2014). Temporal variability in transit-based accessibility to supermarkets. *Applied Geography*, 53, 149–159. <https://doi.org/10.1016/j.apgeog.2014.06.012>.
- Geurs, K. T., & van Wee, B. (2004). Accessibility evaluation of land-use and transport strategies: Review and research directions. *Journal of Transport Geography*, 12(2), 127–140. <https://doi.org/10.1016/j.jtrangeo.2003.10.005>.
- Golub, G. H., & Van Loan, C. F. (1980). An analysis of the total least squares problem. *SIAM Journal on Numerical Analysis*, 17(6), 883–893. <https://doi.org/10.1137/0717073>.
- Guo, D., & Zhu, X. (2014). Origin-destination flow data smoothing and mapping. *IEEE Transactions on Visualization and Computer Graphics*, 20(12), 2043–2052. <https://doi.org/10.1109/TVCG.2014.2346271>.
- Guo, D., Zhu, X., Jin, H., Gao, P., & Andris, C. (2012). Discovering spatial patterns in origin-destination mobility data. *Transactions in GIS*, 16(3), 411–429. <https://doi.org/10.1111/j.1467-9671.2012.01344.x>.
- Guti, J., & Garc, J. C. (2008). Distance-measure impacts on the calculation of transport service areas using GIS. *Environment and Planning B: Planning and Design*, 35(3), 480–503.
- Hansen, W. G. (1959). How accessibility shapes land use. *Journal of the American Institute of Planners*, 25(2), 73–76. <https://doi.org/10.1080/01944365908978307>.
- Herring, R., Hofleitner, A., Abbeel, P., & Bayen, A. (2010). Estimating arterial traffic conditions using sparse probe data. In *Paper presented at the Intelligent Transportation Systems (ITSC), 2010 13th International IEEE Conference on*.
- Hess, D. B. (2005). Access to employment for adults in poverty in the Buffalo-Niagara region. *Urban Studies*, 42(7), 1177–1200. <https://doi.org/10.1080/00420980500121384>.
- Hess, D. B. (2009). Access to public transit and its influence on ridership for older adults in two US cities. *Journal of Transport and Land Use*, 2(1). <https://doi.org/10.5198/jtlu.v2i1.11>.
- Hoque, M. A., Hong, X., & Dixon, B. (2012). Analysis of mobility patterns for urban taxi cabs. In *Paper presented at the Computing, Networking and Communications (ICNC), 2012 International Conference on*.
- Hunter, T., Herring, R., Abbeel, P., & Bayen, A. (2009). Path and travel time inference from GPS probe vehicle data. *NIPS Analyzing Networks and Learning with Graphs*, 12, 1.
- Jenelius, E., & Koutsopoulos, H. N. (2013). Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B: Methodological*, 53, 64–81. <https://doi.org/10.1016/j.trb.2013.03.008>.
- Kwan, M. P. (1999). Gender and individual access to urban opportunities: A study using space-time measures. *The Professional Geographer*, 51(2), 210–227.
- Kwan, M.-P., Murray, A. T., O'Kelly, M. E., & Tiefelsdorf, M. (2003). Recent advances in accessibility research: Representation, methodology and applications. *Journal of Geographical Systems*, 5(1), 129–138. <https://doi.org/10.1007/s101090300107>.
- Larsen, K., & Gilliland, J. (2008). Mapping the evolution of food deserts' in a Canadian city: Supermarket accessibility in London, Ontario, 1961–2005. *International Journal of Health Geographics*, 7(1), 1.
- Lei, T., Chen, Y., & Goulias, K. (2012). Opportunity-based dynamic transit accessibility in Southern California. *Transportation Research Record: Journal of the Transportation Research Board*, 2276(1), 26–37. <https://doi.org/10.3141/2276-04>.

- Lei, T. L., & Church, R. L. (2010). Mapping transit-based access: Integrating GIS, routes and schedules. *International Journal of Geographical Information Science*, 24(2), 283–304. <https://doi.org/10.1080/13658810902835404>.
- Liao, Y., Gil, J., Pereira, R. H. M., Yeh, S., & Verendel, V. (2020). Disparities in travel times between car and transit: Spatiotemporal patterns in cities. *Scientific Reports*, 10(1), 4056. <https://doi.org/10.1038/s41598-020-61077-0>.
- Litman, T. (2003). Integrating public health objectives in transportation decision-making. *American Journal of Health Promotion*, 18(1), 103–108. <https://doi.org/10.4278/0890-1171-18.1.103>.
- Liu, S., & Zhu, X. (2004). Accessibility analyst: An integrated GIS tool for accessibility analysis in urban transportation planning. *Environment and Planning B: Planning and Design*, 31(1), 105–124. <https://doi.org/10.1068/b305>.
- Malekzadeh, & Chung. (2020). A review of transit accessibility models: Challenges in developing transit accessibility models. *International Journal of Sustainable Transportation*, 14(10), 733–748. <https://doi.org/10.1080/15568318.2019.1625087>.
- Mavoa, S., Witten, K., McCreanor, T., & O'Sullivan, D. (2012). GIS based destination accessibility via public transit and walking in Auckland, New Zealand. *Journal of Transport Geography*, 20(1), 15–22. <https://doi.org/10.1016/j.jtrangeo.2011.10.001>.
- Miller, H. J. (1991). Modelling accessibility using space-time prism concepts within geographical information systems. *International Journal of Geographical Information System*, 5(3), 287–301. <https://doi.org/10.1080/02693799108927856>.
- Minocha, I., Sriraj, P., Metaxatos, P., & Thakuria, P. (2008). Analysis of transit quality of service and employment accessibility for the greater Chicago, Illinois, region. *Transportation Research Record: Journal of the Transportation Research Board*, 2042(1), 20–29. <https://doi.org/10.3141/2042-03>.
- MTA. (2013). New York City Transit. <https://new.mta.info/agency/new-york-city-transit>.
- Murtagh, N., Gatersleben, B., & Uzzell, D. (2012). Multiple identities and travel mode choice for regular journeys. *Transportation Research Part F: Traffic Psychology and Behaviour*, 15(5), 514–524. <https://doi.org/10.1016/j.trf.2012.05.002>.
- Niedzielski, M. A., & Boschmann, E. (2014). Travel time and distance as relative accessibility in the journey to work. *Annals of the Association of American Geographers*, 104(6), 1156–1182. <https://doi.org/10.1080/00045608.2014.958398>.
- O'Sullivan, D., Morrison, A., & Shearer, J. (2000). Using desktop GIS for the investigation of accessibility by public transport: An isochrone approach. *International Journal of Geographical Information Science*, 14(1), 85–104. <https://doi.org/10.1080/136588100240976>.
- Paez, A., Gertes Mercado, R., Farber, S., Morency, C., & Roorda, M. (2010). Relative accessibility deprivation indicators for urban settings: Definitions and application to food deserts in Montreal. *Urban Studies*, 47(7), 1415–1438. <https://doi.org/10.1177/0042098009353626>.
- Páez, A., Scott, D. M., & Morency, C. (2012). Measuring accessibility: Positive and normative implementations of various accessibility indicators. *Journal of Transport Geography*, 25, 141–153. <https://doi.org/10.1016/j.jtrangeo.2012.03.016>.
- Pasch, K. E., Hearst, M. O., Nelson, M. C., Forsyth, A., & Lytle, L. A. (2009). Alcohol outlets and youth alcohol use: Exposure in suburban areas. *Health & Place*, 15(2), 642–646. <https://doi.org/10.1016/j.healthplace.2008.10.002>.
- Peng, C., Jin, X., Wong, K.-C., Shi, M., & Liò, P. (2012). Collective human mobility pattern from taxi trips in urban area. *PLoS One*, 7(4), e34487. <https://doi.org/10.1371/journal.pone.0034487>.
- Polzin, S., Pendyala, R., & Navari, S. (2002). Development of time-of-day-based transit accessibility analysis tool. *Transportation Research Record: Journal of the Transportation Research Board*, 1799, 7.
- Sallis, J. F., Frank, L. D., Saelens, B. E., & Kraft, M. K. (2004). Active transportation and physical activity: Opportunities for collaboration on transportation and public health research. *Transportation Research Part A: Policy and Practice*, 38(4), 249–268.
- Salon, D. (2009). Neighborhoods, cars, and commuting in New York City: A discrete choice approach. *Transportation Research Part A: Policy and Practice*, 43(2), 180–196. <https://doi.org/10.1016/j.tra.2008.10.002>.
- Salonen, M., & Toivonen, T. (2013). Modelling travel time in urban networks: Comparable measures for private car and public transport. *Journal of Transport Geography*, 31, 143–153. <https://doi.org/10.1016/j.jtrangeo.2013.06.011>.
- Scott, D., & Horner, M. (2008). Examining the role of urban form in shaping people's accessibility to opportunities: An exploratory spatial data analysis. *Journal of Transport and Land Use*, 1(2). <https://doi.org/10.5198/jtlu.v1i2.25>.
- Shen, Q. (2001). A spatial analysis of job openings and access in a US metropolitan area. *Journal of the American Planning Association*, 67(1), 53–68. <https://doi.org/10.1080/01944360108976355>.
- Tribby, C. P., & Zandbergen, P. A. (2012). High-resolution spatio-temporal modeling of public transit accessibility. *Applied Geography*, 34, 345–355. <https://doi.org/10.1016/j.apgeog.2011.12.008>.
- Vadrevu, L., & Kanjilal, B. (2016). Measuring spatial equity and access to maternal health services using enhanced two step floating catchment area method (E2SFCA) – A case study of the Indian Sundarbans. *International Journal for Equity in Health*, 15(1), 87. <https://doi.org/10.1186/s12939-016-0376-y>.
- Veloso, M., Phithakitnukoon, S., Bento, C., Fonseca, N., & Olivier, P. (2011). Exploratory study of urban flow using taxi traces. In *Paper presented at the First Workshop on Pervasive Urban Applications (PURBA) in conjunction with Pervasive Computing, San Francisco, California, USA*.
- Wachs, M., & Kumagai, T. G. (1973). Physical accessibility as a social indicator. *Socio-Economic Planning Sciences*, 7(5), 437–456. [https://doi.org/10.1016/0038-0121\(73\)90041-4](https://doi.org/10.1016/0038-0121(73)90041-4).
- Walton, D., & Sunseri, S. (2010). Factors influencing the decision to drive or walk short distances to public transport facilities. *International Journal of Sustainable Transportation*, 4(4), 212–226. <https://doi.org/10.1080/15568310902927040>.
- Zhang, T., Dong, S., Zeng, Z., & Li, J. (2018). Quantifying multi-modal public transit accessibility for large metropolitan areas: A time-dependent reliability modeling approach. *International Journal of Geographical Information Science*, 32(8), 1649–1676. <https://doi.org/10.1080/13658816.2018.1459113>.
- Zheng, F., & Van Zuylen, H. (2013). Urban link travel time estimation based on sparse probe vehicle data. *Transportation Research Part C: Emerging Technologies*, 31, 145–157. <https://doi.org/10.1016/j.trc.2012.04.007>.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.