Measuring accessibility to health care services for older bus passengers: A finer spatial resolution

ABSTRACT
Health care accessibility is a vital indicator for evaluating areas where there are medical shortages. However, due to the lack of population data with a satisfactory spatial resolution, efforts to accurately measure health care accessibility among older individuals have been hampered to some extent. To address this issue, we attempt to measure accessibility to health care services for older bus passengers in Nanjing, China, using a finer spatial resolution. More specifically, based on one month’s worth of bus smart card data, a framework for identifying the home stations (i.e., a passenger’s preferred station near their residence) of older passengers is developed to measure the aggregate demand at the bus stop scale. On this basis, a measurement that integrates the Gaussian two-step floating catchment area (2SFCA) and the adjusted 2SFCA methods (referred to as the adjusted Gaussian 2SFCA method) is proposed to measure accessibility to health care services for older people. The results show that: (1) almost all home stations experience inflated demand, especially those located in the suburbs; (2) despite abundant health care resources, home stations in urban districts are rarely identified as high accessibility stations, due to high demand densities among the older population; and (3) more attention should be paid to two types of home stations – those with a medical institution and those with bed shortages, respectively. The first type is predominantly distributed in the periphery of the city, in the suburbs where the travel time required to access the nearest health care service by bus is longer. The second type is mostly located in the outskirts of urban districts and in the central area of one suburb. These findings could help policy makers to implement more appropriate measures to design and reallocate health care resources.

Keywords: Older people; Bus smart card data; Health care accessibility; Demand inflation; Adjusted Gaussian 2SFCA method
1 Introduction

Demographic aging has become an increasingly pervasive and enduring societal phenomenon in many countries around the world, chiefly due to reduced levels of fertility and increased life expectancy; and China is no exception (Cheng et al., 2019a; Metz, 2003; Szeto et al., 2017). According to data released by the National Bureau of Statistics of China (2020), as of 2019, there were nearly 254 million older adults aged 60 years or over in China, accounting for 18.1% of the entire population. Various underlying physiological changes may be more likely to occur among the older population, including an increased risk of disease (World Health Organization, 2020) and reduced mobility (Ni et al., 2015; Zhang et al., 2019a). Currently, the world is facing the threat of the Coronavirus pandemic (COVID-19), caused by SARS-CoV-2, which has severely damaged the public health and disrupted the daily lives of billions of people (Liu et al., 2020). Infected older people are more susceptible to severe disease and have a higher risk of death than their younger counterparts (Guan et al., 2020; World Health Organization, 2020; Yang et al., 2020). Evidently, the anticipated rapid growth of the older population presents a significant challenge for urban planners and policy makers tasked with allocating limited health care resources. Thus, it is generally believed that greater attention needs to be paid to health issues among the older, more vulnerable, population (Luo et al., 2018; Zhang et al., 2019a).

Generally speaking, there are multiple modes of travel that senior citizens can use to access health care services in China. However, as the population ages, driving a private car is no longer the preferred mode of travel for older adults due to safety considerations (Yuan et al., 2019). Although walking accounts for a large proportion of daily trips (Cheng et al., 2019a, 2019b), as the travel distance increases, a decline in physical mobility will inhibit some older individuals from walking to health care services (Cheng et al., 2020a; Du et al., 2020). However, many Chinese cities have a well-developed bus network, relatively reliable bus services, and concessionary fare schemes, which encourages older people to travel by bus so that they can actively participate in social activities (Wong et al., 2018). For instance, in Nanjing and Fuzhou, the proportions of older people who use buses when seeking medical treatment are as high as 46% (Cheng et al., 2020a) and 60% (Du et al., 2020), respectively. It is therefore necessary to accurately measure health care accessibility for older individuals who travel by bus in order to propose relevant policies for improving their access to health care.

Health care accessibility refers to the ease with which residents (e.g., older adults) in a given area can reach health care services (Hewko et al., 2002; Wang et al., 2020). It plays an important role in identifying areas where there are medical shortages and in allocating medical facilities (Dewulf et al., 2013; Wang and Luo, 2005). Health care accessibility includes both spatial and nonspatial elements (Khan, 1992). Spatial elements are mainly concerned with spatial factors, such as the distribution of supply (e.g., physicians) and demand (e.g., patients) (Yang et al., 2006). Conversely, nonspatial elements relate to nonspatial factors, such as age, sex, socioeconomic status, and health needs (Wang and Luo, 2005). Relevant studies can also be divided into two categories according to whether or not they use actual service supply data:
revealed accessibility; and potential accessibility (Khan, 1992). The former focuses on the actual use of health care services in a given residential area, while the latter refers to the potential supply of available health care resources (Dewulf et al., 2013; Wang et al., 2020). In this study, we discuss the potential for older people to access health care systems by bus. Therefore, the focus of this study is limited to potential spatial accessibility.

1.1 Measurement of accessibility to health care services

Various methods have been proposed for measuring health care accessibility in recent years (Paez et al., 2012). Simple metrics such as provider-to-population ratios (PPR) and minimum travel time/distance are overwhelmingly favoured by health professionals and policy makers (Neutens, 2015). Although these metrics are easy to implement in practice, and to understand, they only produce a rough measurement of accessibility and hence have relatively low precision (Dewulf et al., 2013; Yang et al., 2006). Advances in geographic information systems (GIS) technology have made it possible to measure health care accessibility at a finer level of spatial granularity (Luo et al., 2018; McLafferty, 2003). A specific GIS-based method for measuring accessibility, namely, the kernel density method, was originally created by Rosenblatt (1956) and developed by Parzen (1962) and has been used by numerous other researchers since then (Yang et al., 2006). Compared to previous metrics, this method is more conceptually complete, but it does not interpret the accessibility scores in an intuitive way (Wan et al., 2012; Yang et al., 2006).

Recent health care studies on accessibility measures have mainly focused on the two-step floating catchment area (2SFCA) method (e.g., Guagliardo et al., 2004; Langford et al., 2012; Nakamura et al., 2017). The original 2SFCA method was first developed by Radke and Mu (2000) and further modified by Luo and Wang (2003). In addition to retaining the majority of the advantages of the gravity model, the 2SFCA method also provides a more straightforward way of interpreting the accessibility scores (Luo and Qi, 2009). However, it has been found that the main limitation of this method is that all population locations within a catchment area are considered to have equal access to health care services, while locations outside of the catchment area are treated as being unable to access health care services (Dai, 2011; Luo and Qi, 2009; Wang, 2018).

To remedy this issue, several advanced methods have since been developed based on the original 2SFCA method. For instance, Luo and Qi (2009) presented an enhanced 2SFCA (E2SFCA) method designed to estimate health care accessibility by assigning impedance weights to different travel time zones within a catchment area. Although the modified methods identified areas where there are medical shortage more precisely, they still assumed that all populations in each travel time zone have equal access to health care services (Dai, 2011). In reality, residents are more likely to seek services in closest proximity to where they live rather than having a uniform probability of using services at any distance within the catchment area (Chen and Jia, 2019). To continuously account for the distance decay effect within a catchment area, Dai (2010) combined a Gaussian function with the 2SFCA (known as Gaussian 2SFCA)
in order to measure health care accessibility.

Although the aforementioned approaches take distance-decay functions into consideration on the basis of the original 2SFCA method, they may still have the drawback of overestimating demand and supply. In recent years, several advanced solutions such as the three-step floating catchment area (3SFCA) method (Wan et al., 2012), the modified 2SFCA method (Delamater, 2012), and a method that integrates the Huff Model and the floating catchment area method (Luo, 2014), have been proposed in an attempt to tackle the demand overestimation issue. Furthermore, Paez et al. (2019) developed an intuitive metric by allocating demand and supply proportionally in 2SFCA measurements (known as the adjusted 2SFCA method). The major advantage of this approach is that demand and supply within the system can be maintained by adjusting the weighted values of the impedance matrix accordingly.

Given that senior citizens are more likely to seek health care services near to their homes, particularly if they have limited mobility, this can be represented mathematically by a Gaussian function that takes the distance decay effect into account (Luo et al., 2018). This study, therefore, integrates the Gaussian 2SFCA method with the adjusted 2SFCA method to overcome the overestimation issue, while accounting for the distance decay effect.

### 1.2 Accuracy of 2SFCA method parameters

Generally, the 2SFCA method uses three key parameters for measuring accessibility to health care, namely: the population demand, service supply, and spatial impedance (Cheng et al., 2016; Luo, 2004; Wan et al., 2012). The population is represented by the number of people potentially seeking health care. The service supply refers to health care resources, such as the number of medical beds or physicians available. Spatial impedance is usually measured by travel time or distance and reflects the degree to which the “distance” between the location of the service supply and population demand affects accessibility. Generally speaking, a smaller population demand will correspond to a larger service supply, while a lower level of spatial impedance signifies better accessibility to health care (Wan et al., 2012). With the enrichment of spatial databases, some recent studies have shown that the accuracy of these parameters plays an essential role in accessibility measurements (Apparicio et al., 2008; Chen., 2016; Luo and Wang, 2003).

To try to identify the population demand more accurately, a variety of spatial units from small areas (e.g., census block groups) to large areas (e.g., census tracts), have been studied by researchers. The primary areal unit of spatial analysis used to measure accessibility is the census tract (e.g., Chen and Jia, 2019; Langford and Higgs, 2006; Luo and Wang, 2003; Wan et al., 2012) because it is the easiest residential area unit from which to obtain population data. In addition, other areal units, such as census block groups (e.g., Mao and Nekorchuk, 2013; Yang et al., 2006), ZIP codes (e.g., Dai, 2010; Delamater et al., 2012), and subdistricts (e.g., Cheng et al., 2016; Tao et al., 2018; Tao and Cheng, 2019), have been used to identify the potential population within that area. However, regardless of the size of the above spatial units,
the population within each residential area is treated as if it is completely concentrated on a single point (Delamater et al., 2012; Mao and Nekorchuk, 2013; Tao et al., 2018). In reality, the population is seldom distributed homogeneously within a given residential area. Therefore, assigning it to a single point may result in large aggregation errors (Hewko et al., 2002; Spencer and Angeles, 2007; Zhang et al., 2019a), which becomes more obvious as the spatial resolution decreases (Apparicio et al., 2008).

Health care service data can usually be derived from the registration information collected by health authorities or the official website of each hospital (Luo et al., 2018). The fields of service data generally include the following fields: name, address, level, numbers of physicians, and numbers of medical beds (Tao and Cheng, 2019). Using GIS technology, the street addresses of health care locations have been successfully geocoded into the coordinates of a point on the map (e.g., Luo and Wang, 2003; Paez et al., 2010; Yang et al., 2006). For most map applications, the accuracy of the geocoded locations is satisfactory (Bell et al., 2012). Hence, compared with population data, the spatial resolution of service data is generally finer.

Furthermore, various measures have been adopted to estimate the spatial impedance between population demand and service supply locations, such as the Euclidean distance (i.e., a straight line) (Bell et al., 2012; McLafferty and Grady, 2004; Spencer and Angeles, 2007), the network distance (Chen and Jia, 2019; Dewulf et al., 2013), and the travel time (Dai, 2010; Langford and Higgs, 2006; Luo et al., 2018; Luo and Qi, 2009; Luo and Wang, 2003). The network distance and travel time respectively represent the shortest and most direct paths between two points, which better reflect the actual spatial barrier and provide more accurate results than the Euclidean distance (Apparicio et al., 2008). Both of these can be calculated using GIS software with high-resolution geometric road network files; however, the network distance does not take into account mobility and available transportation modes (Tao and Cheng, 2019). Most studies have estimated the travel time by assigning an average travel speed to each road segment and then calculating the minimum travel time between any pair of points (Haynes et al., 2003; Luo et al., 2018; McGrail and Humphreys, 2009). Because the travel speed dataset is often a rough estimate due to the technical limitations of different types of road (Langford and Higgs, 2006; Luo et al., 2018; McGrail and Humphreys, 2009), it is difficult for this approach to reflect the spatial impedance with a high degree of precision.

In summary, the accuracy of the aforementioned parameters still need to be further improved. Low accuracy in terms of population demand and spatial impedance may result in large measurement errors (Apparicio et al., 2008; Luo and Wang, 2003; Tao and Cheng, 2019). In order to guarantee the accuracy of health care accessibility for the older population in this study, it is necessary to use finer-resolution spatial units and more advanced measurement applications to set these parameters.

1.3 **Objective and contributions**

With reference to the context described in the preceding sections, the primary objective of this
study is to measure accessibility to health care services for older bus passengers (hereafter called older passengers for the sake of simplicity) using a finer spatial resolution. In particular, as urban datasets become more abundant (e.g., bus smart card data (SCD) and point of interest (POI) data) and residence identification technology improves, it has become possible to identify a passenger’s preferred station near to his/her residence (i.e., home stations) (Huang et al., 2018). Furthermore, a more reliable estimation of the travel time can be obtained using the dynamic traffic information and routing rules maintained by map developers (García-Albertos et al., 2020; Wang and Xu, 2011).

The contributions of this study to the existing literature are threefold: (1) a framework for identifying the home stations of older passengers is developed by combining bus SCD, citizen information and bus platform information. In this way, the demand for bus services among the older population can be assessed using a finer spatial resolution (i.e., bus stop scale); (2) an online map Application Programming Interface (API) is utilised to estimate the travel time by bus between any pair of home stations and health care locations more accurately; and (3) an adjusted Gaussian 2SFCA method is proposed to assess accessibility to health care services for older passengers at the bus stop scale, accounting for both distance decay and demand inflation effects.

The remainder of this paper is organised as follows. Section 2 introduces the study area and describes the data. Section 3 explains the methods adopted in this study. Section 4 illustrates the results of the parameter setting and accessibility measurements. Section 5 discusses relevant measures that can be implemented to improve health care accessibility for older individuals. Finally, our main conclusions are drawn in Section 6.

## 2 Study area and data

### 2.1 Study area

Nanjing is the political, economic, cultural, and educational centre of Jiangsu Province and the second largest city in the Eastern region of China. It consists of 11 administrative districts, comprising six urban districts and five suburban districts (Fig. 1). Nanjing is now facing an increasing demographic aging problem. By the end of 2018, the total registered population of Nanjing had reached nearly 6.97 million, of whom more than 1.46 million were aged 60 years or over (older people), accounting for 21.07% of the total; and these figures are expected to continue to increase at a faster rate for the foreseeable future (Cheng et al., 2020b; Jiangsu Commission of Health, 2020). It should be noted that Gaochun district is not included in the Nanjing study area (Fig. 1) because the spatial scope of the data (i.e., bus SCD) described in the next subsection is limited to the remaining 10 administrative districts.
2.2 Data and pre-processing

The bus smart card data (SCD), citizen information, and bus platform information are combined in this study to identify the demand for bus services among the older population in Nanjing, aggregated at the bus stop scale. The bus SCD used in this study was provided by Nanjing Citizen Card Co., Ltd. and covered a one-month period from 1st to 31st March, 2019. The SCD contain several key fields (see Fig. 2(b)). The raw data comprise 47,078,969 records of bus trips made by 3,803,147 passengers (including 2,394,163 anonymous passengers). The citizen information was also obtained from Nanjing Citizen Card Co., Ltd., and includes personal information fields such as gender and year of birth (see Fig. 2(a)). By matching the passenger identification field in the citizen information data (i.e., the card ID field in Fig. 2(a) and 2(b)), the records of bus trips made by older passengers can be extracted from the bus SCD. The older population in this study is defined as those people aged 60 years or over.

It is worth noting that the SCD on bus trip records in Nanjing does not include the boarding station of each passenger, but the transaction location is included (Fig. 2(b)). It is therefore necessary to further infer the boarding station of each trip record by combining the transaction location (Fig. 2(b)) and the bus platform information (Fig. 2(c)). For the purpose of the study, a pair of bus platforms on opposite sides of a road is regarded as a single bus boarding station (Long and Thill, 2015). Information on 768 bus routes and 7,712 bus stops (see Fig. 3(a)) was provided by Nanjing Public Transportation (Group) Co., Ltd. In this study, the Spatial Join tool, ArcGIS, was used to infer the nearest boarding platform to the transaction location of each older passenger based on the route ID field, and the platforms were then aggregated at the bus stop scale. Fig. 2(d) shows the pre-processed SCD results. The bus trip records of older passengers include the boarding station and the transaction location of each trip record.

---

1 A boarding station is a fixed bus stop where passengers wait for buses. The transaction location refers to the location of a bus at which passengers scan their cards after boarding the bus. Some passengers may not swipe their cards immediately after boarding, which results in some differences in the timestamps (Ma et al., 2012).
passengers, whose boarding stations were inferred, were used as input data for the subsequent identification of demand among the older population.

<table>
<thead>
<tr>
<th>(a) Citizen information</th>
<th>(b) Bus smart card data (SCD)</th>
<th>(c) Bus platform information</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>card_id</strong></td>
<td><strong>sex</strong></td>
<td><strong>year</strong></td>
</tr>
<tr>
<td>001</td>
<td>W</td>
<td>1938</td>
</tr>
<tr>
<td>002</td>
<td>W</td>
<td>1966</td>
</tr>
<tr>
<td>003</td>
<td>M</td>
<td>1948</td>
</tr>
<tr>
<td>004</td>
<td>M</td>
<td>1941</td>
</tr>
<tr>
<td>005</td>
<td>M</td>
<td>1956</td>
</tr>
</tbody>
</table>

Fig. 2. Flowchart of bus SCD pre-processing, showing structure of (a) citizen information, (b) bus smart card data (SCD), (c) bus platform information and (d) results of preprocessed SCD.

Note: Due to personal privacy considerations, the card ID field has been obscured.

In addition, point of interest (POI) data also constitute a critical element of our framework for identifying demand among the older population. The dataset was obtained from the Application Programming Interface (API) service provided by Baidu Map at the end of 2019, and the dataset includes several useful fields such as name, type, and address. The 476,706 POIs within the entire study area include 28,594 residential-type POIs (see Fig. 3(b)). Where it was not possible to identify the home station of older passengers by the station frequency, the percentage of residential POIs in the neighbourhood where the bus stop is located was calculated to indicate the likelihood of a bus stop being regarded as a home station (see Section 3.1 for details). In general, the larger the proportion of residential type POIs, the greater the “residential potential” of the bus stop.

In 2019, a total of 557 health care services were inventoried by the Nanjing Municipal Commission of Health and Family Planning and each district government. Since the focus of this study is on older people, children’s hospitals, cosmetic surgery hospitals, and obstetrics and gynecology hospitals were excluded. The supply data (e.g., bed numbers and street addresses) were taken from the official website of each service at the end of 2019. The average number of beds per service was 102. Each health care service was successfully geocoded to the
xy-coordinates of a point in ArcGIS based on the street address (see Fig. 3(e)).

Fig. 3. Spatial distribution of (a) bus stops, (b) residential POIs and (c) health care services in the Nanjing study area.

3 Methods

As stated earlier, the three essential parameters for measuring spatial accessibility to health care services are the population demand, service supply and spatial impedance. We utilised the demand volume, supply volume and travel time, respectively, to quantify these parameters. First, based on the pre-processed SCD, the demand volume was calculated using a framework designed to identify older passengers’ home stations (as described in Section 3.1). Second, the supply volume is a measure of the number of beds available, and was obtained from the official website of each health care service. Next, a matrix of travel times between any pair of home stations and health care locations was estimated using the online map API, which is introduced in Section 3.2. Last, the spatial accessibility to health care services for older passengers was measured using the adjusted Gaussian 2SFCA methods described in Section 3.3. A flowchart illustrating how accessibility to health care services for older passengers was measured, is shown in Fig. 4:
3.1 Identifying the demand volume

To determine the demand volume among older passengers at a finer spatial resolution, we first needed to identify the home station of each passenger. In this study, each passenger’s preferred boarding station near to his/her residence is regarded as his/her home station. This enabled the demand among the older population (i.e., the demand volume), aggregated at the bus stop scale, to be calculated. As stated earlier, the ‘older passenger’ referred to here denotes an individual with a potential need for medical treatment, not an actual patient.

We used one-month’s pre-processed SCD to identify the home station of each older passenger. Generally, the boarding station used for the first trip of the day was assumed to be the initial home station of a cardholder (Ma et al., 2017). However, in reality, the initial home stations identified for older passengers may vary from day to day. Therefore, we proposed several rules to further identify a single final home station. First, according to Long and Thill (2015), it is more reliable to infer a final home station for passengers who made transactions on two or more days in one week. Consequently, older passengers who made transactions on fewer than eight days in one month were excluded. Second, the frequency of all boarding stations used by each older passenger was counted to reflect the likelihood of a boarding station being their final home station. In addition, the “cluster” concept was used to represent a virtual polygon that encompasses all the bus stops within a certain distance threshold from each other. Many researchers have applied 500 m as the threshold distance between any adjacent bus stops (Long and Thill, 2015; Zhao et al., 2011), and hence that was the threshold adopted in this study. The framework for identifying the home stations of older passengers is shown in Fig. 5, which comprises the following three steps:

Step 1: Identify the number of clusters. If only one cluster is associated with a particular older passenger, the station that appears most frequently in this cluster is taken as his or her final home station. Otherwise, go to Step 2.
Step 2: Identify the number of stations in each cluster. If only one stop is identified in each cluster and each stop appears with equal frequency, then the final home station of a particular older passenger cannot be identified with any confidence. Otherwise, go to Step 3.

Step 3: Compute the cumulative frequency of all stations in each cluster. If only one cluster has the maximum cumulative frequency of all the stations it encompasses, then the station that appears most frequently in this cluster is regarded as the final home station for the older passenger in question. Otherwise, the passenger’s final home station is regarded as the station that appears most frequently in several clusters with the largest cumulative frequency.

![Flowchart](image)

**Fig. 5.** Framework for identifying the home station of each older passenger.

It should be noted that, in Step 2, any older passenger who does not have a dominant “cluster” is eliminated, because his/her boarding stations are geographically dispersed (i.e., the distance between any two boarding stations is greater than the threshold of 500 m) and appear with almost equal frequency, which means that these stations are less likely to be located near his/her...
residence (Long and Thill, 2015). In addition, there are some cases where more than one station appears with the same frequency within a cluster. In this case, based on the proportion of residential POIs around each bus stop, the term “residential potential” is used in relation to discovering the probability of a station being the passenger’s final home station. Generally, the larger the proportion of residential POIs, the greater the “residential potential” of the bus stop. The residential potential can be calculated as:

\[ P_k = \frac{R_k}{A_k} \]  

(1)

where \( P_k \) denotes the residential potential of bus stop \( k \), \( R_k \) denotes the total number of residential POIs in the neighbourhood (within a 500 m radius) where bus stop \( k \) is located, and \( A_k \) denotes the total number of POIs in the neighbourhood where bus stop \( k \) is located.

After identifying the home station of each older passenger based on the pre-processed SCD, it then became possible to further aggregate the potential demand volume at the bus stop scale.

3.2 Estimation of the travel time via Baidu Map’s API

In recent years, some studies have attempted to estimate travel times via the API services provided by online maps such as Baidu Map or Google Map (Cheng et al., 2016; Tao et al., 2018). This allows a more reliable estimation of bus travel times to be obtained using dynamically updated traffic information (e.g., bus routes, stations and schedules) provided by local authorities and the routing rules followed by map developers (Kuai and Wang, 2020; Wang and Xu, 2011). In this study, we utilised Baidu Map’s API (http://lbsyun.baidu.com/index.php) to automatically obtain the matrix of travel times between any pair of home stations and health care locations. An example is provided in Fig. 6. Due to the fact that older individuals are more likely to travel during off-peak times to avoid crowds (Wong et al., 2018), the range of departure times set using Baidu Map’s API was constrained to between 9 am and 5 pm, during which period there are only small variations in travel times.

It is worth noting that the travel time estimated in this study refers to the fastest real-time route from a home station to a health care location by bus. More precisely, the quickest route is the one with the shortest travel time out of all the alternative routes. The transport modes used in this study do not include rail transit as, in many Chinese cities, older people are more reluctant to travel by rail (Zhang et al., 2019b). According to data from the 2014 Nanjing Travel Survey, rail transit only accounts for 3.59% of daily trips for older people. A rail transit environment is associated with more inconvenient transfer experiences, longer walking distances, a greater likelihood of escalator-related injuries, and lower seat availability than travelling by bus (Szeto et al., 2017; Xing et al., 2019), which generally discourages older individuals from travelling by rail (Du et al., 2020). More importantly, unlike the bus SCD, Nanjing rail transit SCD does not include information about the age of passengers, so it is not feasible to exclusively extract trip records specific to the older population. In addition, the bus travel time calculated by Baidu
Map’s API consists of walking time to and from the bus stop\(^2\), in-vehicle time, and interchange time (if there is an interchange) (Gao et al., 2019).

![Fig. 6. An example showing how travel time was estimated via Baidu Map’s API.](image)

### 3.3 Measuring the spatial accessibility to health care services

The original 2SFCA method has commonly been used to measure accessibility to health care services, because it is intuitive to interpret and easy to compute (Cervigni et al., 2008; Neutens, 2015). However, the main limitation of this method is that it fails to consider the distance decay effect (as explained in Section 1.1) and also has the drawback of overestimating demand and supply. To address these two issues, an adjusted Gaussian 2SFCA method was used in this study.

#### 3.3.1 The original 2SFCA method

The original 2SFCA method has been widely applied as a simple and intuitive tool with which to measure the spatial accessibility of the population in a given area. There are two main steps involved in measuring the accessibility of older people to health care services, aggregated by their home stations.

*Step 1:* For each health care service \(j\), the catchment area within the threshold travel time \(t_0\) was generated and then used to search all home stations \(k\) and calculate the ratio of medical beds available to demand among the older population \(R_j\) via the following formula:

\[
R_j = \frac{\text{available medical beds}}{\text{demand among the older population}}
\]

\(^2\) As the bus travel time estimated in this study refers to the duration of travel from a home station (origin) to a health care location (destination), the time spent walking to the bus stop is not included.
\[ R_j = \frac{S_j}{\sum_{k \in [t_{kj} \leq t_0]} D_k} \]  

where \( S_j \) is the supply volume at health care location \( j \), which is represented by the number of medical beds; \( t_{kj} \) is the travel time between \( k \) and \( j \); and \( D_k \) is the demand volume at home station \( k \) that falls within the catchment area of \( j \) (\( t_{kj} \leq t_0 \)), which is represented by the demand among the older population aggregated at the bus stop scale.

**Step 2:** For each home station \( i \), the catchment area with a threshold travel time (\( t_0 \)) was generated and used to search all health care services \( j \) and sum up the ratios \( R_j \) via the following formula:

\[ A_i = \sum_{j \in [t_{ij} \leq t_0]} R_j \]  

where \( A_i \) denotes the accessibility at home station \( i \), \( R_j \) is the ratio of the number of medical beds available to the demand among the older population at health care service \( j \) that falls within the catchment area of \( i \) (\( t_{ij} \leq t_0 \)), and \( t_{ij} \) is the travel time between \( i \) and \( j \).

Fig.7 provides a simple example that illustrates the original 2SFCA method. In this example, there are three home stations and two health care providers. It is assumed that the older population at home stations 1, 2, and 3 is 100 persons in each case. It is also assumed that there are 10 medical beds at each of health care providers 1, 2, and 3. In this case, the catchment area of health care provider 1 contains 300 older individuals, and thus has a ratio \( R_{H1} \) (i.e., level of service) of 10/300. Likewise, the ratio of \( R_{H2} \) is 10/300. For each home station, two health care services can be sought within the relevant catchment area, so all home stations have equal accessibility, \( A_{H1} = A_{H2} = A_{H3} = 20/300 \) (10/300 + 10/300).

![Fig. 7. The original 2SFCA method (adapted from Paez et al. (2019)): (a) step 1 and (b) step 2.](image-url)
facility is 300, making the total demand for all health care services (600) within that catchment area twice that of the region’s older population (300). By contrast, the sum of the level of service of all the health care providers within that catchment area \( \frac{20}{300} = \frac{10}{300} + \frac{10}{300} \) is significantly lower than the sum of the spatial accessibility of all the home stations \( \frac{60}{300} = \frac{20}{300} + \frac{20}{300} + \frac{20}{300} \). This is caused by assigning the same level of service of health care to several home stations (Paez et al., 2019). Therefore, neither of the service levels of these two health care facilities \( (R_{H1} \text{ and } R_{H2}) \) can accurately reflect the local provider-to-population ratio (PPR) (Paez et al., 2019; Wan et al., 2012).

### 3.3.2 The adjusted Gaussian 2SFCA method

As stated earlier, many researchers have claimed that it seems unreasonable to assume that there is equal accessibility within a catchment area (Dai, 2011; Yang et al., 2006; Wang, 2018). To resolve this problem, Dai (2010) proposed the so-called Gaussian 2SFCA method, which applies a Gaussian function to continuously discount accessibility within a catchment area based on the original 2SFCA method. This method assumes that health care services that are closer to a home station are more accessible. Furthermore, an adjusted 2SFCA method (i.e., a method that allocates demand and supply proportionately) was proposed by Paez et al. (2019) to address the inflation issue. By introducing row- and column-standardised impedance weights, both the demand and the level of service in the system are adjusted to the initial level in the measurement of accessibility. In this study, we integrated the two methods and formulated an adjusted Gaussian 2SFCA method with which to measure spatial accessibility to health care.

The proposed adjusted Gaussian 2SFCA method can also be divided into two steps:

\[
R_j = \frac{S_j}{\sum_{k \in [t_j \leq t_0]} D_k W_{kj}} \quad (4)
\]

\[
A_j = \sum_{j \in [t_j \leq t_0]} R_j W_{ij} \quad (5)
\]

where \( R_j, S_j, D_k \) and \( A_j \) are explained in Eqs. (2) and (3); \( W_{kj}^k \) is the row-standardized Gaussian impedance weight from home station \( k \) to health care service \( j \), while \( W_{ij}^j \) is the column-standardised Gaussian impedance weight from home station \( i \) to health care service \( j \). The adjusted impedance weights for each of the aforementioned are then calculated as shown below:

\[
W_{kj}^k = \frac{W_{kj}}{\sum_{j=1}^{J} W_{kj}} \quad (6)
\]

\[
W_{ij}^j = \frac{W_{ij}}{\sum_{i=1}^{I} W_{ij}} \quad (7)
\]

where \( W_{kj} \) refers to the initial Gaussian impedance weight from home station \( k \) to health care service \( j \), \( J \) is the total number of health care services, and \( I \) is the total number of home stations. It is clear that these adjusted weights have two key attributes, as explained below:
\[ \sum_{j=1}^{J} W_{kj} = 1 \]  
\[ \sum_{i=1}^{I} W_{ij} = 1 \]

The initial Gaussian impedance weights employed in the preceding equations (taking \( W_{ij} \) as an example) are shown below:

\[ W_{ij} = \begin{cases} 
  e^{-1/2} \times \left( \frac{t_{ij}}{t_0} \right)^2 \cdot e^{-1/2} & \text{if } t_{ij} \leq t_0 \\
  0 & \text{if } t_{ij} > t_0 
\end{cases} \]

where \( t_{ij} \) is the travel time between home station \( i \) and health care service \( j \), and \( t_0 \) is the threshold travel time. In essence, the adjusted Gaussian 2SFCA method is obtained by replacing the initial impedance weights (i.e., \( W_{kj} \) and \( W_{ij} \)) with the adjusted impedance weights (i.e., \( W_{kj}^{\prime} \) and \( W_{ij}^{\prime} \)).

**4 Results**

**4.1 Identifying demand volume**

The results of identifying the home stations of older passengers are displayed in Fig. 8, which gives a general summary of the identification process. First, by matching the passenger identification field (i.e., the card ID field) in the citizen information, 1,409,311 passengers (37.1% of all 3,803,474 passengers) with age-related information were extracted from the raw SCD. The remaining anonymised passengers were eliminated. Second, based on the age-related information, those aged 60 years or over were further extracted, resulting in a total of 877,792 people, accounting for 62.3% of the total. Third, by identifying the boarding stations for each bus trip record, the pre-processed SCD were generated for subsequent home station identification. The pre-processed SCD comprised 15,390,506 bus trip records and 7,687 bus stops used by 867,296 older passengers. Fourth, for each older passenger, the boarding station from which he/she made the first trip of the day was extracted and assumed to be his/her initial home station. Therefore, 8,037,244 first trip records made from 7,015 bus stops were extracted from the pre-processed SCD. Next, the low-frequency older passengers who made transactions on fewer than eight days in one month (i.e., eight first trip records) were excluded, leaving 335,546 older passengers. Finally, the final home station of each older passenger was identified using the identification framework shown in Fig. 5. The identification results for 335,043 older passengers (99.85% of all 335,546 elderly passengers) were aggregated at the bus stop level, with a total of 5,064 home stations.
Fig. 8. Identification process of the home stations of older passengers in the Nanjing study area.

Fig. 9. Spatial distribution of the demand from older passengers aggregated at the bus stop scale in the Nanjing study area.

The spatial distribution of older adults who use buses in the Nanjing study area is shown in Fig. 9, which represents the total demand from older individuals aggregated at the bus stop scale. In previous studies (Luo et al., 2018; Ni et al., 2015; Tao and Cheng, 2019), the entire older population of a subdistrict has been concentrated on its centroid; however, a finer spatial
resolution is used in this study. The demand from the older population is divided into four categories (see Fig. 9) based on the Jenks Natural Breaks Classification, which identifies breakpoints between groups by minimizing the variation within each group. It is clear that home stations with a high demand from older people are predominantly concentrated in urban districts (e.g., the Gulou and Qinhuai districts), whereas home stations with a lower demand are evenly distributed in peripheral areas of the city, in the suburbs (e.g., the Liuhe and Lishui districts). The main reason for this may be that the level of bus service provision in the suburbs is lower than that in urban areas, resulting in a lower proportion of older people using buses in the suburbs.

It is necessary to verify the reliability of the estimated demand in this study, based on the local census data (Jin et al., 2018). The number of older passengers aggregated at the bus stop level is therefore summarised for each district to compare this figure with the entire older population captured from the Nanjing census data at the end of 2018 (Jiangsu Commission of Health, 2020). As can be seen from Table 1, on the whole, the results of the identification of demand among older people show that there is a similar proportional distribution of older people in each district compared to that of the census data. For instance, both data sources show that the number of older people in the Gulou, Qinhuai and Jiangning districts account for more than 10% of the total older population within the study area, whereas the proportion of older people in Jianye, Qixia, Yuhuatai, Pukou and Lishui is less than 10%. It should be noted that the estimated proportion of older people in the Liuhe and Lishui districts in this study is significantly lower than that obtained from the census data. This is probably due to the lack of reliable bus services which makes senior citizens in the Liuhe and Lishui districts less likely to use buses.

Importantly, bus SCD can be used to identify the demand among older people with a finer spatial resolution, thus providing a more accurate parameter (i.e., demand volume) for measuring accessibility to health care services for older bus passengers.

<table>
<thead>
<tr>
<th>Table 1. A comparison between the estimated demand and the Nanjing census data.</th>
</tr>
</thead>
<tbody>
<tr>
<td>District</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>Urban district</td>
</tr>
<tr>
<td>Gulou</td>
</tr>
<tr>
<td>Xuanwu</td>
</tr>
<tr>
<td>Jianye</td>
</tr>
<tr>
<td>Qinhuai</td>
</tr>
<tr>
<td>Qixia</td>
</tr>
<tr>
<td>Yuhuatai</td>
</tr>
</tbody>
</table>
### Suburban district

<table>
<thead>
<tr>
<th>Suburban district</th>
<th>Travel time (min)</th>
<th>Travel time percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pukou</td>
<td>3.24</td>
<td>9.67%</td>
</tr>
<tr>
<td>Jiangning</td>
<td>4.25</td>
<td>12.69%</td>
</tr>
<tr>
<td>Liuhe</td>
<td>2.06</td>
<td>6.15%</td>
</tr>
<tr>
<td>Lishui</td>
<td>0.93</td>
<td>2.78%</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>33.50</td>
<td>100%</td>
</tr>
</tbody>
</table>

#### 4.2 Travel time estimation

The results of estimating the travel time via Baidu Map’s API are shown in Fig. 10. Fig. 10(a) displays a matrix of travel times between any pair of home stations and health care locations. Taking home station 1 as an example, it was identified as the home station for 101 older passengers (i.e., H_1 (101)) in the previous section. The travel time from home station 1 to health care 1 by bus is 60.12 minutes, and the travel time from home station 1 to health care 2 by bus is 52.15 minutes. By comparison, the travel time from home station 1 to the nearest health care location by bus is 6.87 minutes, which is therefore designated as the minimum travel time.

![Travel time matrix](image)

Fig. 10. Results of the estimated travel time obtained via Baidu Map’s API, showing: (a) a matrix of travel times between any pair of home stations and health care locations, (b) distribution of the minimum travel time for all home stations and (c) different forms of the impedance function in the two 2SFCA methods.
The threshold travel time \((t_0)\) was then set to ensure that most older passengers can access at least one health care provider within the catchment area centred on the home station. The minimum travel time distribution of all home stations is shown in Fig. 10(b). As can be seen from the histogram, most of the minimum travel times are under 30 minutes. This view is further supported by the cumulative percentage curve (i.e., the green curve in Fig. 10(b)) according to which those home stations from which the nearest health care facility can be accessed in less than 30 minutes account for 93% of all home stations. In addition, the State Council of the People’s Republic of China (2017) has committed to improving primary medical and health care services to ensure that all patients can receive medical assistance within 30 minutes. Therefore, the threshold travel time \((t_0)\) used in this study was set to 30 minutes. It is worth noting that the minimum travel time for urban home stations is mostly within the threshold travel time while the minimum travel time distribution for suburban home stations is more discrete, as shown in the scatter plot in Fig. 10(b). This is because urban home stations are more evenly distributed than those in the suburbs, and health care resources are more abundant; thus, older passengers in urban districts usually spend less time seeking health care services than those in suburbs.

In addition, different forms of the impedance function used in the original 2SFCA and Gaussian 2SFCA methods are depicted in Fig. 10(c). The binary impedance function (red polyline) is a constant of one within the threshold travel time (i.e., 30 minutes) and zero outside of the threshold, and is mainly used in the original 2SFCA method. This dichotomy may include or exclude points near the boundary of the catchment area, and the statistical bias caused would have a significant impact on the accessibility results (Chen, 2017). In contrast, the Gaussian impedance function (blue curve) converges to zero when it approaches the boundary of the catchment area, thus alleviating the aforementioned bias.

4.3 Spatial accessibility measurement

Three critical parameters for measuring accessibility were set as explained in previous sections, namely, the supply volume obtained from local authorities, the demand volume identified from the bus SCD, and the travel time estimated via Baidu Map’s API. The results of using the adjusted Gaussian 2SFCA method used to measure accessibility to health care for older passengers are shown in Fig. 11. The Jenks Natural Breaks Classification was applied to separate the accessibility scores into the following five classes (see Fig. 11(a)): lowest accessibility (<0.0089), lower accessibility (0.0089-0.0233), medium accessibility (0.0233-0.0847), higher accessibility (0.0847-1.8049) and highest accessibility (>1.8049). All the health care services beyond the threshold travel time were defined as inaccessible.

Fig. 11(a) shows the spatial distribution of accessibility to health care services for older passengers which was obtained using the adjusted Gaussian 2SFCA method. First, the home stations with the lowest levels of accessibility are mainly distributed in peripheral areas of the city, in the suburbs (e.g., the Pukou, Liuhe, and Lishui districts) and the Qixia district. Second, home stations with medium accessibility are predominantly concentrated in the central areas
of Jiangning and surrounding urban districts, where bus networks and health care services are usually more developed. However, it is surprising that home stations with high accessibility scores are mainly scattered across the peripheral areas of the Pukou and Jiangning districts, rather than in the urban core. This may be because the provider-to-population ratio (PPR) within the catchment area cantered on these home stations is high. Although health care resources are relatively unabundant overall, some of the home stations in the peripheral areas of the Pukou and Jiangning districts were still identified as high accessibility stations because the demand density is low in those areas. In contrast, those stations located in the densely populated urban centres were rated as having only moderate accessibility due to their low PPR.

![Fig. 11. Spatial distribution of (a) accessibility score and (b) accessibility inflation factor in the Nanjing study area obtained using the adjusted Gaussian 2SFCA method.](image)

To demonstrate the merits of the adjusted Gaussian 2SFCA method in terms of overcoming the inflation issue, the accessibility inflation factor – that is, the ratio of the accessibility measured by the Gaussian 2SFCA and adjusted Gaussian 2SFCA methods – was calculated in this study (see Fig. 11(b)). It is clear that the issue of demand and level of service inflation is associated with nearly all the home stations, as shown by the fact that their initial accessibility scores are at least three times higher than their adjusted scores. Moreover, this accessibility inflation factor is not uniformly distributed from the city centre to the periphery. In the case of the Liuhe and
Lishui districts, most stations have an inflation factor of more than 20. One reason for this is that the 30-minute catchment areas of these marginal stations may overlap with neighbouring districts, thereby increasing the demand and level of service inflation effect (Paez et al., 2019). Another reason could be that the service level of the same health care facility is allocated to several home stations multiple times within the relevant catchment areas. These issues are likely to be particularly pronounced in the Liuhe and Lishui districts where medical resources are insufficient.

Rather than simply superimposing demand and supply, the adjusted Gaussian 2SFCA method allocates them proportionately in order to ensure that the total level of service for all health care services in the region is consistent with the total accessibility for all home stations. Thus, the adjusted Gaussian 2SFCA method produces more accurate and robust results than traditional approaches.

5 Discussion

More health care resources could be allocated to those home stations with low levels of health care accessibility or those which are deemed inaccessible (designated as ‘shortage stations’) because they are most in need of improvement. According to the accessibility measurement results, shortage stations can be classified into two types: firstly, home stations where there is a shortage of medical institutions, which are those from which it takes more than 30 minutes (threshold travel time) to reach the nearest health care facility by bus. Secondly, home stations with shortages of medical beds, which are those with the lowest accessibility scores (i.e., less than 0.0089). Furthermore, those stations with a higher demand density are also regarded as stations that need to be prioritised for improvement. Out of all the home stations with the lowest levels of accessibility, these ‘shortage stations’ account for 19% of the total number of stations, and 80% of the older population’s demand for health care is concentrated in these locations.

Fig. 12 shows the spatial distribution of these two types of under-served stations. The former is mostly distributed in the peripheral areas of the city, in the suburbs (e.g., the Liuhe and Lishui districts in Fig. 12(a)), where bus services and the associated infrastructure are less developed. Therefore, new medical institutions (e.g., hospitals or surgeries) should be built close to these shortage stations to reduce the bus travel time required for older passengers to access them. However, the planning and construction of these institutions often depends on the long-term decisions taken by the relevant local authorities, and it is difficult to implement them in a short period of time. Therefore, improving the level of bus services in the suburbs would be the most obvious measure that local agencies could take to improve health care accessibility. Greater attention should also be paid to promoting an inclusive design within the travel environment, such as installing barrier-free facilities for boarding and alighting from buses, and adding dedicated seats for older passengers to use while waiting for buses. In addition, reducing the bus travel time may also promote bus travel among older individuals who may have limited mobility. Specific measures could include: (a) enhancing the accessibility from densely populated home stations to health care services by increasing the number of direct bus routes;
(b) introducing dedicated bus lanes to guarantee that public transport is given priority; (c) increasing the service frequency of some bus routes primarily used by older people, so they do not have to spend as long waiting at bus stops.

Fig. 12. Spatial distribution of home stations with (a) shortage of medical institutions and (b) shortage of medical beds.

In contrast, the second type of shortage stations are mainly located in the outskirts of the city’s urban districts and the central areas of some suburbs (e.g., the Pukou and Lishui districts), as shown in Fig. 12(b). In line with our findings, the Health Care Facilities Construction Plan published by the Qinhuai district government proposed gradually reallocating health care resources from the central area to the eastern area (see the white dashed circle in Fig. 12(b)) from 2021 to 2023 (Nanjing Qinhuai District People's Government, 2017). In addition, the Pukou district local government has committed to increasing the number of medical beds available for health care services (see the black dashed circle in Fig. 12(b)) because the number of medical beds per thousand people is currently below the provincial target value (Nanjing Pukou District People's Government, 2020). These suggestions further support our view that more attention should be paid to these shortage stations so as to improve the spatial accessibility to health care services for older people in Nanjing. As stated earlier, these shortage stations are identified based on bus smart card data, which only reflects the distribution of health care accessibility among older bus passengers. Rail transit also exists in some areas with low
accessibility (e.g., Pukou) and older people use them—though infrequently—for accessing healthcare services. However, due to the lack of the usage data of rail transit, it is not feasible in this research to measure the health care accessibility for rail transit riders. By and large, the findings in relation to these two types of shortage stations could still help policy makers to take more appropriate measures when designing and reallocating health care resources in the future.

6 Conclusions

Faced with the pervasive societal phenomenon of demographic aging, many researchers and practitioners worldwide have focused on the need to assess spatial accessibility to health care services. However, efforts to accurately measure the health care accessibility of older individuals have been hampered to some extent by the lack of a satisfactory spatial resolution within current population data. Therefore, this study aimed to measure accessibility to health care services for older bus passengers in Nanjing at a finer spatial resolution.

First, a framework for identifying the home stations of older passengers was developed by combining bus SCD, citizen information, and bus platform information. The bus trip records of 335,043 older passengers were then aggregated at the bus stop scale with a total of 5,064 home stations. Unlike previous studies that have used a way of using a finer spatial unit such as census tracts, ZIP codes, and subdistricts, this study offered a way of using a finer spatial resolution for setting the demand volume parameters. Next, a matrix of travel times by bus between any pair of home stations and health care locations was estimated via Baidu Map’s API. The estimation results showed that those home stations that are less than 30 minutes from the nearest health care service accounted for 93% of all home stations. Unlike most prior methods that have relied on road travel speeds, the method used in this study of estimating travel time via the online map API was more reliable and easier to calculate.

A method that integrated the Gaussian 2SFCA and adjusted 2SFCA methods (known as the adjusted Gaussian 2SFCA method) was then developed to assess the spatial accessibility to health care services for older passengers at the bus stop scale. The results showed that the majority of home stations with the lowest levels of accessibility were distributed in peripheral areas of the city, in the suburbs (e.g., the Pukou, Liuhe, and Lishui districts) and the Qixia district, whereas the home stations with high levels of accessibility were mainly scattered across the peripheral areas of the Pukou and Jiangning districts, rather than in the urban core. In addition, the results of the accessibility inflation factor calculations showed that nearly all home stations were affected by the issue of demand and level of service inflation, especially those in the suburbs. Overall, the adjusted Gaussian 2SFCA method produced more accurate and robust results than traditional approaches, because it can be used to assess spatial accessibility at a finer spatial resolution.

Several policy implications can be drawn from this study. First, given that the travel time by bus from many home stations in the suburbs (e.g., the Liuhe and Lishui districts) to the nearest
health care service exceeds 30 minutes (travel time threshold), the level of bus service provision could be improved, through measures such as installing barrier-free facilities and creating bus-only lanes, in order to improve health care accessibility for older passengers. Second, more health care resources could be allocated to the outer areas of urban districts and the central areas of some suburbs (e.g., the Pukou and Lishui districts). These areas contain a number of home stations with shortages of medical beds, and account for 19% of the lowest accessibility stations, but 80% of the demand from the older population is concentrated there. Therefore, the findings of this study could help policy makers to take more appropriate measures when designing and reallocating health care resources, so that they can allocate limited resources to the places that need them most.

This study has some limitations that could be addressed in future studies. First, it only considers older people who travel by bus, and does not take into account members of the older population who use other modes of transport. If relevant data become available in the future, we will conduct further analysis of older people’s accessibility to healthcare services using other modes of transport and draw comparisons with bus passengers. Second, although the distance decay effect was considered in this study, the catchment area of health care providers was still regarded as unvarying (i.e., 30 minutes). However, in reality, the catchment size may differ accordingly to the capacity of health care providers, i.e. hospitals with a higher capacity have a larger catchment area (Luo and Whippo, 2012). Moreover, the focus of this study lies in evaluating the potential health care accessibility of the older population. The revealed accessibility to health care services for older people who travel by bus is also an interesting topic worthy of further research.
References


Tao, Z., Yao, Z., Kong, H., Duan, F., Li, G., 2018. Spatial accessibility to healthcare services in Shenzhen, China: Improving the multi-modal two-step floating catchment area method by estimating travel time via online map APIs. BMC Health Serv. Res. 18, 1–10.


