

Estimating spatial availability/mismatch using singly constrained accessibility measures

Author One^a, Author Two^a

^aAddress

Abstract

Accessibility measures are widely used in transportation, urban, and health care planning, among other applications. By giving a weighted sum of the opportunities that can be reached given the cost of movement and are interpreted as the potential for spatial interaction. While these measures are useful to understand spatial structure there are issues in interpretability and spatial bias which have been partially addressed in recently introduced measures such as the balanced floating catchment areas (BFCA) and competitive measures of accessibility. In following this spirit, we propose a new measure of *spatial availability* which is calculated by imposing a single constraint on the conventional accessibility measure. Similar to the gravity model from which it is derived, a single constraint ensures that the marginals at the origin and destination are met and thus the number of opportunities are preserved. In this paper, we detail the formulation of the proposed spatial availability measure, use-cases of the measure using a simple numerical example, and contrast how the access to jobs changes between spatial availability and conventional accessibility measure using empirical 2016 travel survey data in the City of Toronto, Canada. We show that measure of spatial availability should be used for opportunities which are indivisible.... the values are more interpretable, less biased than conventional accessibility measure. All data presented and original manuscript are openly available.

*Corresponding Author

Email addresses: author.1@example.com (Author One), author.2@example.com (Author Two)

Introduction

The concept of accessibility is a relatively simple one whose appeal derives from combining the spatial distribution of opportunities and the cost of reaching them (Hansen, 1959). Numerous methods for calculating accessibility have been proposed that can be broadly organized into infrastructure-, place-, person-, and utility-based measures (Geurs and van Wee, 2004). Of these, the place-based family of measures is arguably the most common, capturing the number of opportunities reachable from an origin using the transportation network. This type of measure is also referred to as a gravity-based measure of accessibility that captures the potential for interaction.

What accessibility measures is sometimes referred to as *opportunity access* and the analysis of opportunity access is widely employed in transportation, geography, public health, and many other areas, and there is increasing emphasis on a shift from mobility-oriented to access-oriented planning (Deboosere et al., 2018; Handy, 2020; Proffitt et al., 2017; Yan, 2021). However, while these types of opportunity access measures are excellent indicators of the intersection between urban structure and transportation infrastructure, they have been criticized in the past for not being highly interpretable. Previous research has highlighted how the weighting of opportunities using an impedance function can make gravity measures more difficult for planners and policymakers to interpret compared to simpler cumulative opportunity measures (Geurs and van Wee, 2004; Miller, 2018). Moreover, because place-based measures are sensitive to the number of opportunities and the characteristics of the transportation network, raw values cannot be easily compared across study areas (Allen and Farber, 2019).

Intra- and inter-regional comparisons are challenging because gravity-based accessibility indicators are spatially smoothed estimates of the total number of opportunities, however, the meaning of their magnitudes is unclear. This is evident when we consider the “total accessibility” in the region, a quantity that is not particularly meaningful since it is not constrained to resemble, let alone match the number of opportunities available. Furthermore, while accessibility depends on the supply of destination opportunities weighted by the travel costs associated with reaching them, the calculated accessibilities are not sensitive to the demand for those opportunities at the origins. Put another way, traditional measures of place-based accessibility do not capture the competition for opportunities. This theoretical shortcoming (Geurs and van Wee, 2004) is particularly problematic when those opportunities are “non-divisible” in the sense that, once they have been taken by someone, are no longer available to other members of the population. Examples of indivisible opportunities include jobs (when a person takes up a job, the same job cannot be taken by someone else) and placements at schools (once a student takes a seat at a school, that particular opportunity is no longer there for another student). From a different perspective, employers may see workers as opportunities, so when a worker takes a job, this particular individual is no longer in the available pool of candidates for hiring.

To remedy these issues, researchers have proposed several different approaches for calculating competitive accessibility values. On the one hand, this includes several approaches that first normalize the number of opportunities available at a destination by the demand for them from the origin zones and, second, sum the demand-corrected opportunities which can be reached from the origins (e.g. Joseph and Bantock, 1984; Shen, 1998). These advances were popularized in the family of two-step floating catchment area methods (Luo and Wang, 2003) that have found widespread adoption for calculating competitive accessibility to healthcare and other uses. In principle floating catchment areas purport to account for competition/congestion effects, although in practice several researchers (e.g., Delamater, 2013; Wan et al., 2012) have found that they tend to over-estimate the level of demand and/or service. The underlying issue, as demonstrated by Paez et al. (2019), is the multiple counting of both population and level of service, which can lead to biased estimates if not corrected.

A second approach is to impose constraints on the gravity model to ensure flows between zones are equal to the observed totals. Based on Wilson's (1971) entropy-derived gravity model, researchers can incorporate constraints to ensure that the modeled flows match some known quantities in the data inputs. In this way, models can be singly-constrained to match the row- or column-marginals (the trips produced or attracted, respectively), whereas a doubly-constrained model is designed to match both marginals. Allen and Farber (2019) recently incorporated a version of the doubly-constrained gravity model within the floating catchment area approach to calculate competitive accessibility to employment using transit across eight cities in Canada. But while such a model can account for competition, the mutual dependence of the balancing factors in a doubly-constrained model means they must be iteratively calculated which makes them more computationally-intensive. Furthermore, the double constraint means that the sum of opportunity-seekers and the sum of opportunities must match, which is not necessarily true in every case (e.g., there might be more people searching for work than jobs exist in a region).

In this paper we propose an alternative approach to measuring competitive accessibility. We call it a measure of **spatial availability** (SA), and it aims to capture the number of indivisible opportunities that are not only *accessible* but also *available* to the opportunity-seeking population, in the sense that they have not been claimed by a competing seeker of the opportunity. As we will show, spatial availability is a singly-constrained measure of accessibility. By allocating opportunities in a proportional way based on demand and distance, this method avoids the issues of conglomeration that result from multiple counting of opportunities in traditional accessibility measures. The method returns meaningful accessibilities that correspond to the rate of available opportunities per person. Moreover, the method also returns a benchmark value for the region under study against which results for individual origins can be compared.

In the following sections we will describe and illustrate this new measure using simple numerical examples. First, we will describe the measure. Second, we will calculate the SA using a simple hypothetical population and employment centers data set for three use-cases: one of jobs from the perspective of the

population, another considering catchment restrictions, and another of workers from the perspective of employers. Thirdly, we calculate the SA using real world data for the Transportation Tomorrow Survey (TTS) home-to-work commute in 2016 for the Greater Golden Horseshoe (GGH) area in Ontario, Canada. Finally, we discuss the differences between accessibility estimates to the proposed measure of SA and the potential range of uses of the SA measure.

Background

Most accessibility measures (excluding utility-based measures) are derived from the gravity model and follow the widely used formulation in (1). A thorough explanation of this conventional accessibility measure through a numerical example and its limitations are detailed in this section. The limitations of the conventional accessibility measure, namely issues in interpretation and basis, are the motivation for the *spatial availability* measure which we propose and describe in the following sections.

$$A_i = \sum_{j=1}^J O_j f(c_{ij}) \quad (1)$$

where:

- A is accessibility.
- i is a set of origin locations.
- j is a set of destination locations.
- O_j is the number of opportunities at location j . These are opportunities for activity and add some sort of *supply* to the area;
- c_{ij} is a measure of the cost of moving between i and j
- $f(\cdot)$ is an impedance function of c_{ij} ; it can take the form of any monotonically decreasing function (e.g., negative exponential distribution) .

The accessibility value A_i is the weighted sum of opportunities that can be reached from location i , given the cost of travel c_{ij} determined by the impedance function $f(\cdot)$. Summing the opportunities in the neighborhood of i (the neighborhood is defined by the impedance function) estimates of the total number of opportunities that can be reached from i at a certain cost. The type of accessibility value A_i can be changed depending on the impedance function, the measure could be cumulative opportunities (if $f(\cdot)$ is a binary or indicator function e.g., [XX]) or a more traditional gravity measure (e.g., a Gaussian impedance function [XX], inverse cost impedance function [XX], ...).

We use a simple numerical example to introduce the key concepts, and we will use the usual accessibility measure for comparison. In this way, we aim to show the differences between accessibility and spatial availability, which helps to explain how spatial availability can improve interpretability in the analysis of spatially dispersed opportunities.

Table 1: Numerical example

id	number	type
E1	750	jobs
E2	2250	jobs
E3	1500	jobs
P1	260	population
P2	255	population
P3	510	population
P4	495	population
P5	1020	population
P6	490	population
P7	980	population
P8	260	population
P9	255	population

Numerical Accessibility Example

The setup for the simple numerical example is a system with three employment centers and nine population centers, as summarized in Table 1. The access to jobs for each population center is calculated using the conventional accessibility measure A_i (1). In this numerical example we use the straight line distance between the population and jobs for c_{ij} and a negative exponential function with $\beta = 0.0015$. As noted, A_i represents the number of jobs (i.e., opportunities) that can be reached from each population center given the estimated cost as depicted in Figure 1.

Figure 1 depicts three employment centers locations (black circles), where the size of the symbol is in proportion to the number of jobs at each location. We also see nine population centers (triangles), where the size of the symbol is proportional to the accessibility (A_i) to jobs. The accessibility values illustrates the following:

- Population centers (triangles) in the middle of the plot are relatively close to all three employment centers and thus have the highest levels of job accessibility. Population center P5 is relatively central and close to all employment centers, and it is the closest population to the second largest employment center in the region. Unsurprisingly, this population center has the highest accessibility (680.64);
- Population centers (triangles) near the left edge of the map (only in proximity to the small employment center) have the lowest levels of job accessibility. Population center P1 is quite peripheral and the closest employment center is also the smallest one. Consequently, it has the lowest accessibility with $A_i = 17.12$);

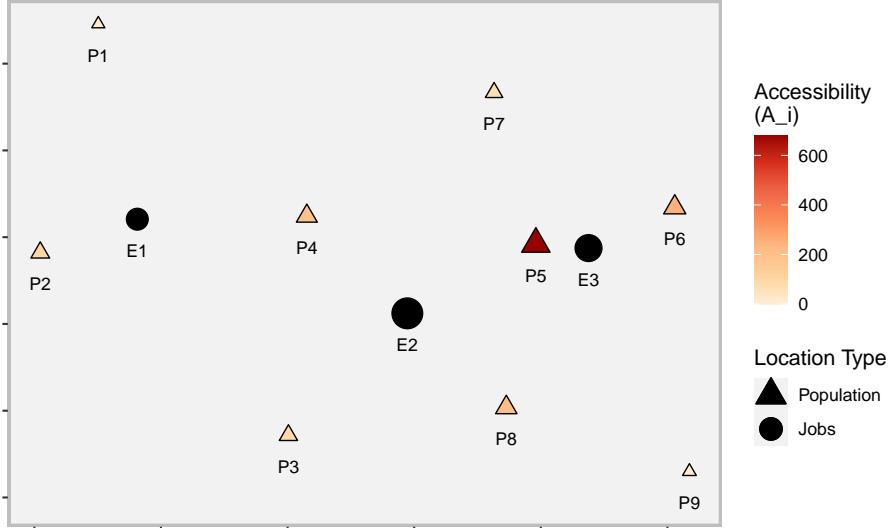


Figure 1: Accessibility of jobs from population centers for the simple numerical example

Issues With the Conventional Accessibility Measure

Accessibility measures are excellent indicators of the intersection between urban structure and transportation infrastructure . However, beyond enabling comparisons of relative values they are not highly interpretable on their own. For instance, from Figure 1, P1 has lower accessibility than P5 but despite the accessibility value for P1 being relatively low it is still better than *zero*. To address this interpretability issue, previous research index and normalize values on a per-capita basis . However, as recent critical research on accessibility discusses (see for instance Paez, Higgins, Vivona (2019) and Allen and Farber (2019)), these steps do not address the bias introduced through the uneven multiple-counting of opportunities which are unconstrained by demand-side competition. We call this issue the ‘conglomeration effect’ and it arises as a result of the underlying mathematical assumption that for conventional accessibility A_i all opportunities for all origins $i = 1, \dots, n$ are divisible and non-competitive. This results in every opportunity entering the weighted sum once for every origin i that can reach it. Put another way, if a densely populated population center pops up next to P5 this center too will have a high accessibility score since A_i does not consider competition of opportunities from neighbouring demand centers. This neglect to constraint opportunity counts (i.e., single-constraint) obscures the interpretability of accessibility which may manifests in the following two ways of biased estimation:

- 1) Demand centers in the less dense outer limits of the urban core may be

assigned disproportionately *high* accessibility values. These periphery areas are traditionally located in proximity to more dense urban demand centers and large urban opportunity centers and thus may have low travel cost to these large opportunity centers. Accessibility A_i does not consider opportunity-constraints and as such these periphery demand centers benefit from the high accessibility to opportunities without competition considerations from their more dense and more centrally located neighbours.

- 2) remote/isolated areas which are still within the region of the analysis and have regionally-relative low demand centers, opportunity centers, and travel cost to opportunity centers, are assigned disproportionately *low* accessibility values. These remote/isolated areas may be sufficiently supplied with opportunities proportionate to their demand but is obscured by the artificially high accessibility awarded to periphery/other areas in which conglomeration disproportionately occurs.

The spatially uneven multiple-counting of opportunities (i.e., the conglomeration effect) leave decision makers unclear on how to interpret resulting access values and all recent accessibility measures which seek to improve interpretability are either vulnerable to this impact or require potentially unrealistic assumptions . For instance, the floating catchment areas (FCA) method increases interpretability by purporting to account for competition, however, as discussed by Paez, Higgins, and Vivona (2019), FCA methods are vulnerable to conglomeration effect. On this same note, the doubly-constrained gravity model proposed by Allen and Farber (2019) which is based on the FCA method accounts for competition but requires that the magnitude of demand matches the opportunities. As already mentioned, this assumption is not always realistic for many opportunity types such as in the case of job seekers and jobs.

To address the conglomeration effect and introduce a more realistic assumption for opportunities, we propose a singly-constrained gravity measure called **spatial availability**. This measure fundamentally seeks to answer for an individual at a specific population center the following questions: “*many jobs are accessible, but the same jobs are also accessible to my (possibly) numerous neighbors... what does a high accessibility actually mean to me?*” and “*few jobs are accessible but I am located in a remote area with proportionally few neighbours... what does low accessibility mean to me?*”.

The Analytical Framework of Spatial Availability

Spatial availability V_{ij} is defined by opportunities O which are proportionally allocated based on the relative population allocation factor F_{ij}^p and cost of travel allocation factor F_{ij}^c for all origins i to all destinations j as detailed in (2). In line with the gravity tradition, the proposed framework distinguishes between opportunities at a destination and demand for opportunities at the origin.

$$V_{ij} = O_j \frac{F_{ij}^p \cdot F_{ij}^c}{\sum_{i=1}^K F_{ij}^p \cdot F_{ij}^c} \quad (2)$$

where:

- V_{ij} is spatial availability.
- i is a set of origin locations in the region K .
- j is a set of destination locations in the region K .
- O_j is the number of opportunities at location j in the region K .
- F_{ij}^p is the proportional allocation factor of the population in i relative to the population in region K .
- F_{ij}^c is the proportion allocation factor of travel cost for i relative to the travel cost in region K ; it is a product of a monotonically decreasing (i.e., impedance) function associated with the cost of travel between i and j .

To explain the analytical framework, the calculation of *job access* is illustrated with a simple step-by-step example for two population centers (P_2 and P_3) in the role of demand (i.e., the number of individuals in the labour market who “demand” employment) and one employment center (O_1) in the role of opportunities.

Additionally, since spatial availability V_{ij} consists of these two allocation factors, we detail first how the role of population allocation factor F_{ij}^p in producing V_{ij}^p , next the role of the travel cost allocation factor F_{ij}^c in producing V_{ij}^c , and finally how both allocation factors in the final general form of spatial availability V_{ij} are combined.

Population and Travel Cost Allocation Factors

We begin with allocation based on demand; consider an employment center j with O_j^r jobs of type r . In the general case where there are K population centers in the region, the following factor can be defined (3).

$$F_{ij}^p = \frac{P_{i \in r}^\alpha}{\sum_{i=1}^K P_{i \in r}^\alpha} \quad (3)$$

The population allocation factor F_{ij}^p corresponds to the proportion of the population in origin i relative to the rest of the region's population centres K . On the right hand side of the equation, the numerator $P_{i \in r}$ is the population at origin i that is eligible and ‘demand’ jobs of type r (maybe those with a certain level of training or in a designated age group). The summation in the denominator is over $i = 1, \dots, K$, the number of population at all origins i in the region K . To modulate the effect of the size in from this factor we also add an empirical parameter α (i.e., $\alpha < 1$ places greater weight on smaller centres relative to larger ones while $\alpha > 1$ achieves the opposite effect). This population allocation factor F_{ij}^p can now be used to proportionally allocate a share of the jobs at a destination j to all origin pairs.

More broadly, since the factor F_{ij}^p is a proportion, when it is summed over $i = 1, \dots, K$ it always equals to 1 (i.e., $\sum_i^K F_{ij}^p = 1$). This is notable since the share of jobs (the spatial availability based on population V_{ij}^p) at each destination j allocated to (i.e., available to) each origin is equal to $V_{ij}^p = O_j \cdot F_{ij}^p$ and since

the sum of \hat{F}^p_{ij} is equal to 1 it follows that $\sum_{i=1}^I V_{ij} = O_j$. In other words, the number of jobs across the region is preserved. The result is a proportional allocation of jobs (opportunities) to origins based on population demand; this factor does not consider travel cost, that is defined in the travel cost allocation factor F_{ij}^c which is introduced shortly.

To illustrate the population allocation factor, consider an employment center has 300 jobs ($O_1 = 300$) in a region with two population centers which have 240 and 120 people, respectively, ($P_2 = 240$ and $P_3 = 120$). For simplicity, assume that all the population in the region is eligible for these jobs, that is, that the entirety of the population is included in the set r . Also assume that $\alpha = 1$ meaning the only impact on value is the population size for each center. The population allocation factors F_{ij}^p for the jobs at O_1 for each population center P_2 and P_3 would be defined as (4).

$$\begin{aligned} F_{2,1}^p &= \frac{P_2^\alpha}{P_2^\alpha + P_3^\alpha} = \frac{240}{240+120} = \frac{240}{360} \\ F_{3,1}^p &= \frac{P_3^\alpha}{P_2^\alpha + P_3^\alpha} = \frac{120}{240+120} = \frac{120}{360} \end{aligned} \quad (4)$$

The F_{ij}^p values can be used to find a *partial* spatial availability of jobs which is only defined by the relative population demanding jobs; this partial spatial availability V_{ij}^p for each population center would be calculated as follows in (5).

$$\begin{aligned} V_{2,1}^p &= O_1 \cdot F_{2,1}^p = 300 \cdot \frac{240}{360} = 200 \\ V_{3,1}^p &= O_1 \cdot F_{3,1}^p = 300 \cdot \frac{120}{360} = 100 \end{aligned} \quad (5)$$

It can be seen that when using only the proportional allocation factor F_{ij}^p to calculate spatial availability (differentiated here by being defined as V_{ij}^p instead of V_{ij}), proportionally more jobs are allocated to the bigger population center (i.e., 2 times more jobs as it is 2 times larger in population). We can also see that the sum of spatial availability for all population centers is equal to the sum of jobs, i.e., total opportunities are preserved. However, as mentioned, using only the proportional allocation factor F_{ij}^p to calculate spatial availability does not account for how far population centers P_2 or P_3 are from employment center O_1 . To account for this effect we introduce a second allocation factor F_{ij}^c based on distance to the employment centers defined in (6).

$$F_{ij}^c = \frac{f(c_{ij})}{\sum_{i=1}^K f(c_{ij})} \quad (6)$$

Where c_{ij} is the cost (e.g., the distance, travel time, etc.) from population center i to employment center j , and $f(\cdot)$ is an impedance function that is a monotonically decreasing function of cost (c_{ij}); in other words, the travel cost allocation factor F_{ij}^c serves to proportionally allocates more jobs to closer locations through an impedance function. To continue illustrating, assume that the impedance function is a negative exponential function and assume that β (which modulates the steepness of the impedance effect and is an empirical parameter) is the value of 1 for simplicity. Also suppose that the distance from population center P_2 to employment center O_1 is 0.6 km, and the distance from

population center P_3 to employment center O_1 is 0.3 km. The proportional allocation factor F_{ij}^p for the jobs at O_1 for both population centers P_2 and P_3 is defined as follows (7).

$$\begin{aligned} F_{2,1}^c &= \frac{\exp(-\beta \cdot D_{2,1})}{\exp(-\beta \cdot D_{2,1}) + \exp(-\beta \cdot D_{3,1})} = \frac{\exp(-0.6)}{\exp(-0.6) + \exp(-0.3)} = 0.426 \\ F_{3,1}^c &= \frac{\exp(-\beta \cdot D_{3,1})}{\exp(-\beta \cdot D_{2,1}) + \exp(-\beta \cdot D_{3,1})} = \frac{\exp(-0.3)}{\exp(-0.6) + \exp(-0.3)} = 0.574 \end{aligned} \quad (7)$$

We can see that the proportional allocation factor for P_3 is larger than P_2 since the distance to O_1 is shorter. Using the travel cost proportional allocation factors F_{ij}^c as defined in (7), we can now calculate the spatial availability of jobs for each population center based only on F_{ij}^c and the jobs available O_1 to these two competing population centers (note: V_{ij}^c not the complete V_{ij}) as follows in (8).

$$\begin{aligned} V_{2,1}^c &= O_1 \cdot F_{2,1}^c = 300 \times 0.426 = 127.8 \\ V_{3,1}^c &= O_1 \cdot F_{3,1}^c = 300 \times 0.574 = 172.2 \end{aligned} \quad (8)$$

As shown, the spatial availability defined by F_{ij}^c (i.e., V_{ij}^c) allocates P_3 a larger share of jobs since the population center is closer to O_1 . However, as previously discussed, P_3 has a smaller population than P_2 , so P_2 receives a larger share of jobs when spatial availability when it is defined by F_{ij}^p (i.e., V_{ij}^p). It is necessary to combine both population and travel cost factors to accurately reflect demand; these two components are in line with how demand is conventionally modelled in accessibility calculations. Fortunately, since both F_{ij}^c and F_{ij}^p preserve the total number of opportunities (jobs) as they independently sum to 1, they can be combined multiplicatively to calculate the proposed spatial availability (V_{ij}) which considers demand to be based on both population and travel cost.

Putting Spatial Availability Together

We can combine the proportional allocation factors by population F_{ij}^p and travel cost F_{ij}^c and calculate spatial availability V_{ij} as introduced in (2) and repeated below:

$$V_{ij} = O_j \frac{F_{ij}^p \cdot F_{ij}^c}{\sum_{i=1}^K F_{ij}^p \cdot F_{ij}^c}$$

To complete the illustrative example of employment center O_1 and population centers P_2 and P_3 , the resulting spatial availability V_{ij} is calculated for both population centers is calculated in (9).

$$\begin{aligned} V_{2,1} &= O_1 \cdot \frac{F_{2,1}^p \cdot F_{2,1}^c}{F_{2,1}^p \cdot F_{2,1}^c + F_{3,1}^p \cdot F_{3,1}^c} = 300 \frac{\left(\frac{2}{3}\right)(0.426)}{\left(\frac{2}{3}\right)(0.426) + \left(\frac{1}{3}\right)(0.574)} = 179.4 \\ V_{3,1} &= O_1 \cdot \frac{F_{3,1}^p \cdot F_{3,1}^c}{F_{2,1}^p \cdot F_{2,1}^c + F_{3,1}^p \cdot F_{3,1}^c} = 300 \frac{\left(\frac{1}{3}\right)(0.574)}{\left(\frac{2}{3}\right)(0.426) + \left(\frac{1}{3}\right)(0.574)} = 120.6 \end{aligned} \quad (9)$$

As can be seen, fewer number of jobs are allocated to population center P_2 compared to the allocation by population only, to account for the higher cost of reaching the employment center. On the other hand, distance alone allocated more jobs to the closest population center (i.e., P_3), but since it is smaller, it also gets a smaller share of the jobs overall. To reiterate, the sum of jobs at employment center O_1 that are allocated to population centers P_2 and P_3 simultaneously based on *population-* and *travel cost* allocation factors are preserved (i.e., $V_{2,1} + V_{3,1} = O_1$).

In the common case that population centers have multiple destination opportunities j , availability is simply the sum of (2) for all opportunities J (i.e., $V_i = \sum_{j=1}^J V_{ij}$). This quantity represents opportunities (e.g., jobs) that can be accessed from i and that are *not* allocated to a competitor: therefore the weighted sum of available opportunities. Compare V_i to the singly-constrained gravity model (see Wilson (1971)). In essence, V_i is the result of constraining A_i to match one of the marginals in the origin-destination table, the known total of opportunities.

Since the sum of opportunities is preserved in the procedures above, it is possible to calculate a highly interpretable measure of spatial availability per capita (lower-case v_i) as follows in (10).

$$v_i = \frac{V_i}{P_i} \quad (10)$$

To complete the illustrative example, the per capita spatial availability of jobs would be calculated as follows in (11).

$$\begin{aligned} v_{2,1} &= \frac{V_{2,1}}{P_2} = \frac{179.4}{240} = 0.8 \\ v_{3,1} &= \frac{V_{3,1}}{P_3} = \frac{120.6}{120} = 1.0 \end{aligned} \quad (11)$$

We can see that since P_3 is closer to O_1 and has less competition (as it has a smaller population than P_2), P_3 benefits with a higher spatial availability of jobs per capita.

Where the overall ratio of jobs to population in the region is $300/(240 + 120) = 0.83$, the spatially available jobs per capita at k is closer to unity.

Empirical Example: Spatial Availability and Accessibility of Jobs in the GGH

In this section, we use two empirical examples to demonstrate how the spatially uneven multiple-counting of opportunities (i.e., conglomeration effect) inherent to the accessibility measure (1) introduces spatial bias and obscures interpretability. This conglomeration effect is highlighted by calculating spatial availability (2) and comparing the relative difference between the two measures. The first example demonstrates how accessibility broadly overestimates *job access* (relative to spatial availability) and particularly overestimates job access for areas in the less dense outer limits of the urban core. The second example demonstrates how accessibility underestimates job access for areas in the

Table 2: Descriptive statistics of the TTS 2016 dataset for the Greater Golden Horseshoe Area

	Trips	Travel_Time
Min. : 1.00	Min. : 0.00	
1st Qu.: 14.00	1st Qu.: 13.00	
Median : 22.00	Median : 20.00	
Mean : 33.44	Mean : 23.39	
3rd Qu.: 38.00	3rd Qu.: 30.00	
Max. :1129.00	Max. :179.00	
NA	NA's :3507	

urban periphery. Both examples are based on the same empirical data set for home-based work trips in the Greater Golden Horseshoe (GGH) area. We first introduce the data used, then calibrate the impedance function, and finally illustrate the two examples.

Data

The 2016 Transportation Tomorrow Survey (TTS) data for 20 municipalities contained within the GGH area in the province of Ontario, Canada (43.6°N 79.73°W) is used within this section (Figure 2). This data set includes home-based origins and employment destinations defined by centroids of Traffic Analysis Zones (TAZ) ($n=3764$), the number of jobs ($n=3081900$) and workers ($n=3446957$) at each origin and destination, and the trips from origin to destination for the morning home-to-work commute ($n=3446957$).

Also included are travel times and cost of travel from origin to destination by car; travel times are calculated using the R package **r5r** (Rafael H. M. Pereira et al., 2021) and an impedance function based on these travel times is derived. It is important to note that for simplicity, all trips within this data set are assumed to be taken by car, and the travel time is calculated from an origin TAZ centroid to a destination TAZ centroid. The centroid is snapped to the nearest street line by **r5r** and the travel time is calculated for all trips assuming a car travel mode and a departure of 7:00am on Wednesday October 2021, an arbitrary date selected by the authors. Additionally, only travel times less than or equal to 180 mins (3 hrs); this threshold represents 99% of trip's travel times which are summarized in the descriptive statistics in Table 2. All data and data preparation steps can be freely explored in companion R data-package **AccessPack**.

Calibrating an Impedance Function

In the toy data set introduced in the Numerical Accessibility Example section, an arbitrary negative exponential function describing travel cost (as a function of distance) was used as the impedance function to derive both accessibility and spatial availability. In this empirical data set, since the travel time is calculated based on the existing street network in the GGH, an empirical trip



Figure 2: The TTS 2016 study area within the Greater Golden Horseshoe in Ontario, Canada.

length distribution (TLD) is defined and used to derive an impedance function. For background, a TLD is the representation of the likelihood that a proportion of trips are taken at a specific travel cost and are widely used to derive impedance functions in accessibility research .

The empirical TLD for this data set is represented by black data points in Figure 3). As can be observed, this empirical distribution appears to follow a gamma distribution, as such, this theoretical distribution along with other common TLD distributions such as log-normal and exponential distributions were fitted to the empirical TLD. The maximum likelihood estimation method and Nelder-Mead method for direct optimization available within the `fitdistrplus` package (Delignette-Muller and Dutang, 2015) was used . Based on goodness-of-fit criteria and diagnostics presented in Appendix Figure 18 and 17, the gamma distribution was selected and is represented as a red line in Figure 3.

The resulting calibrated gamma distribution, which serves as the impedance function for the empirical data set, is given in the following general form where the estimated ‘shape’ is α , the estimated ‘rate’ is β , and $\Gamma(\alpha)$ is defined in (12). The calculated shape and rate parameter is 2.019 and 0.094 respectively. We would like to reiterate that this impedance function, though it is derived using empirical data, assumes all home-to-work trips in the 2016 TTS data set are all taken by car. In reality there is a modal split but for illustrative purposes, the same impedance function is used to calculate accessibility and spatial availability in the following examples.

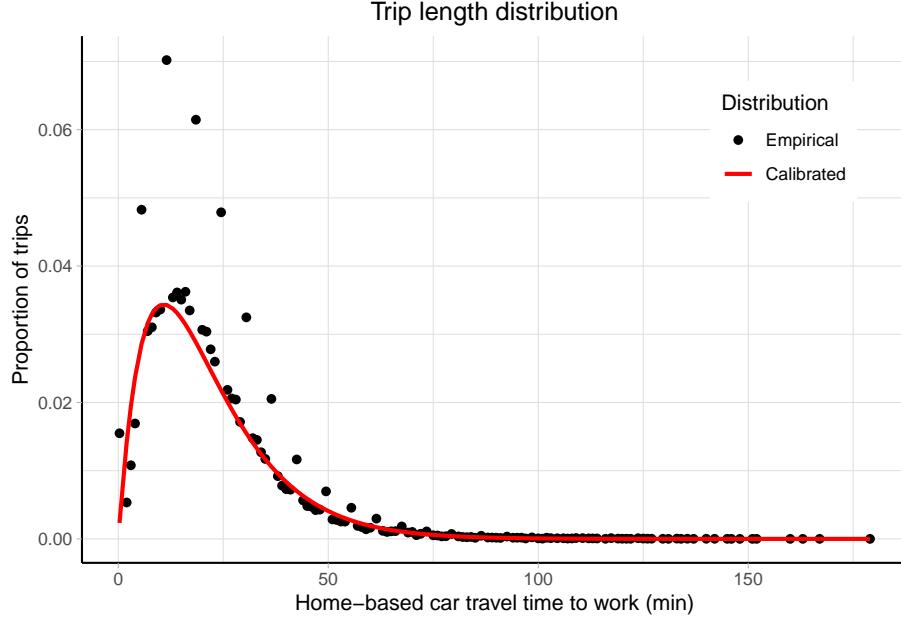


Figure 3: Empirical TTS 2016 home-based car trip length distribution (black) and calibrated gamma distribution impedance function (red)

$$f(x, \alpha, \beta) = \frac{x^{\alpha-1} e^{-\frac{x}{\beta}}}{\beta^\alpha \Gamma(\alpha)} \quad \text{for } 0 \leq x \leq \infty \quad (12)$$

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$$

Access to jobs in Toronto

In this section we calculate *job access* using both accessibility and spatial availability measures for jobs in city of Toronto and all origins in the GGH using the impedance function defined in the previous section. Toronto is the largest city in the GGH and represents a significant subset of workers and jobs in the GGH; 50% of workers in the GGH travel to jobs in Toronto and 72% of jobs are located within Toronto. As it will be later discussed, this example demonstrates the first issue in accessibility: neglecting to include the single-constraint overestimates *job access* for TAZ with a low proportion of workers that are next to TAZs with a high proportion of jobs and workers.

Accessibility is first calculated and presented in Figure 4. The higher the accessibility value, the more accessible places of employment are to home-based origins. It can be briefly summarized that the accessibility values follow a radial trend where the majority of TAZs in Toronto have high accessibility values and values decrease in TAZs which are further from the city boundary.

Next, the spatial availability is calculated and also presented in Figure 4. It can be observed that the higher the spatial availability value the more available employment destinations in Toronto are to any home-based origins in the GGH.

If a TAZ has a spatial availability value which is higher or lower than NA it has above or below average spatial availability of jobs relative to all job opportunities within the City of Toronto (for all origins in the GGH).

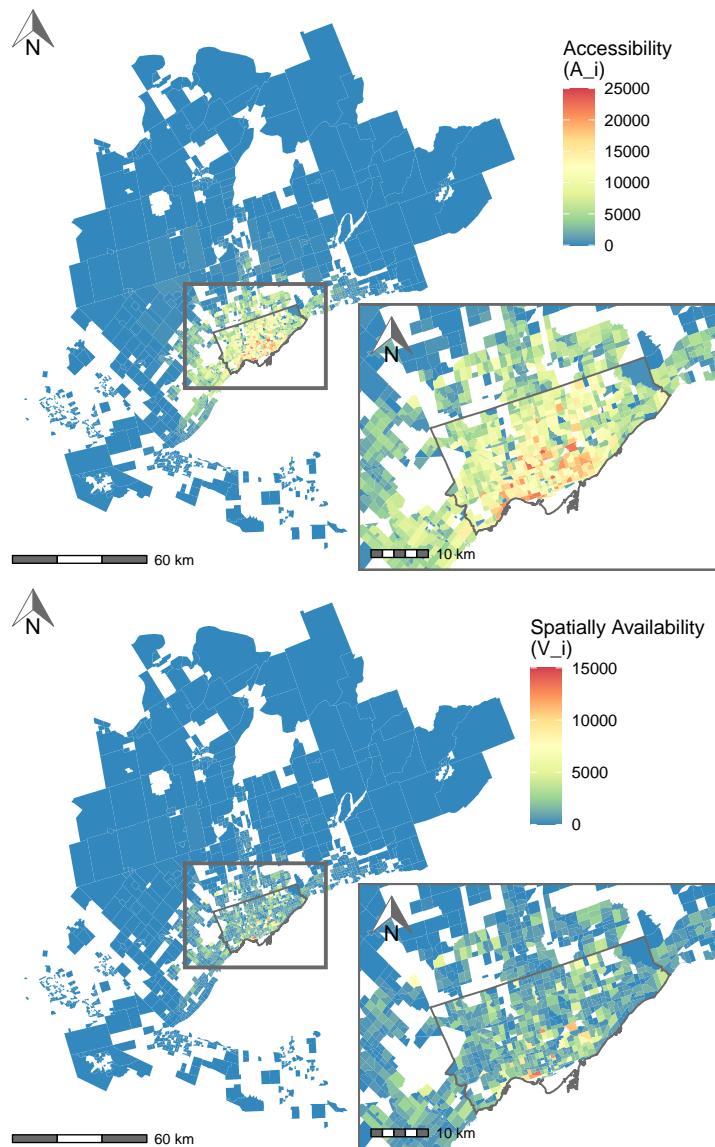


Figure 4: Calculated accessibility (top) and spatial availability (bottom) of employment from origins in the GGH to destinations in the City of Toronto.

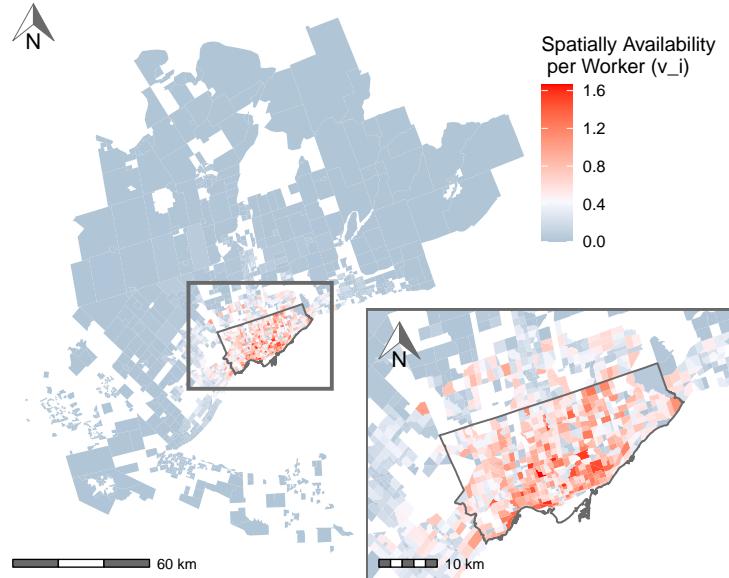


Figure 5: Calculated spatial availability of employment, per worker, from origins in the GGH to destinations in the City of Toronto.

Spatial availability can also be normalized per workers within each TAZ to provide meaningful insight on how many jobs are *available* on average for each TAZ. This normalization, shown in Figure 5, demonstrates which TAZ have above and below the average (0.422) available jobs per worker in the GGH (to jobs located within the city of Toronto). It can be briefly summarized that the spatial availability measure has a smaller and less pronounced radial impact than the accessibility measure; TAZs within and immediately outside the boundary of Toronto have above average spatial availability of jobs but at a lower relative intensity than compared to the availability measure. Additionally, more TAZs outside of Toronto have low spatial availability values.

Access to jobs in the GGH

In this section we calculate job access for all jobs in the GGH for all origins in the GGH using both accessibility and spatial availability measures. As will be later elaborated, this example demonstrates the second issue in the accessibility measure: neglecting to include the single-constraint underestimates job access for TAZ which are on the periphery of the GGH and have relatively-low proportion of workers. These TAZs are relatively isolated (by travel cost) from the city of Toronto which has the highest density of opportunities (jobs). The first issue, namely the over estimation of the TAZs which are on the outer limits outside the Toronto border is also observed.

Accessibility and spatial availability are both calculated and presented in Figure 6. Interestingly, despite the majority of jobs being located within Toronto - general trends are similar for both accessibility plots but not for both spatial availability plots. For instance, the accessibility values follow a radial trend where the majority of TAZs in Toronto have high accessibility values and values decrease in TAZs which are further from the city boundary. However, it can be seen that the radial effect is less pronounced and TAZs outside of Toronto appear to have a higher accessibility score. Conversely, spatial availability does not appear to follow a radial trend as depicted in the previous plot in Figure 4. Job access, as measured by spatial availability, appears much more even throughout the GGH (in comparison to job access as measured by accessibility). This can be noted in higher values around the north east and south west periphery TAZs and more moderate values in and around Toronto.

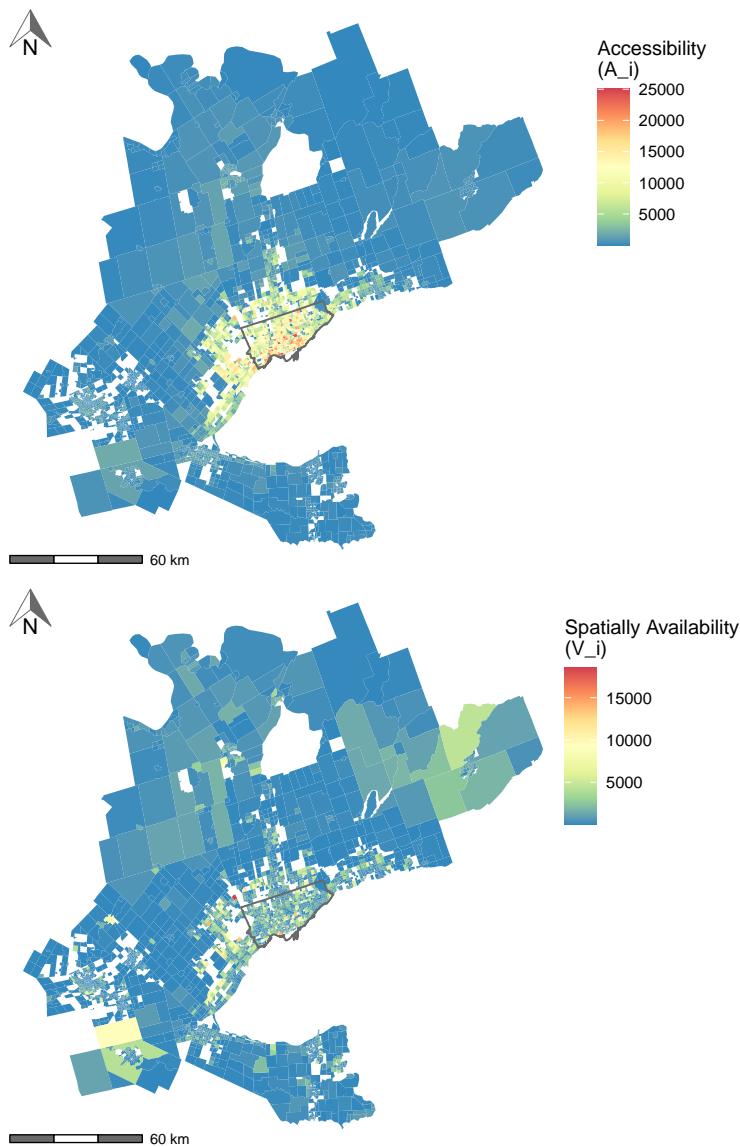


Figure 6: Calculated accessibility (top) and spatial availability (bottom) of employment from origins in the GGH to destinations in the GGH.

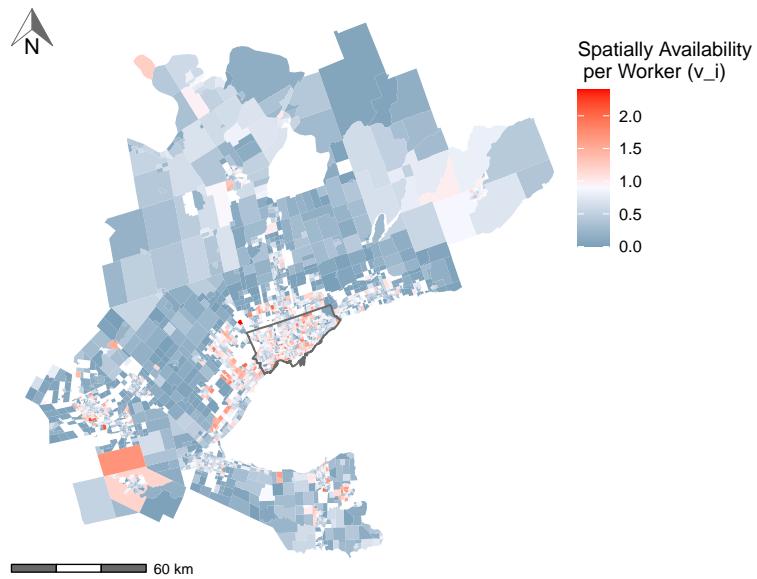


Figure 7: Calculated spatial availability of employment, per worker, from origins in the GGH to destinations in the City of Toronto.

Spatial availability within this example can also be represented on a per-worker basis and a similar difference in trend between the previous plot in Figure 5 can be seen in the plot for all GGH jobs in Figure 7.

Discussion

Comparing Opportunity Access Measures - Toronto

Accessibility and spatial availability in our example both measure *access to jobs*. However, these two measures cannot be meaningfully compared through normalization on a per worker basis since the accessibility methodology inherently multiple-counts opportunities. We thus take a different approach to compare the relative magnitudes (within each measure) by re-scaling both measures from 0 to 100 where each value of the measure is divided by the maximum value. The percentage change between re-scaled spatial availability and accessibility for each TAZ is then calculated; 98.9% of TAZs have lower re-scaled spatial availability values than accessibility values and they are on average -80.64% different and range between -2.95% to -99.9% different than re-scaled accessibility. To simplify the presentation of percentage change in difference between measures, five quantiles are calculated and coloured to represent each TAZ in Figure 8. The quantile ranks range from 1 to 5 where 1 represents the highest percentage difference between spatial availability and accessibility. In other words, the higher the quantile rank, the more pronounced the overestimation of *job access* is in that TAZ when referring to the accessibility measure relative to the spatial availability measure. It should be noted that the 1.1% of TAZs which experience a positive change between the re-scaled measures are shown in grey.

We can see from Figure 8 that when we re-scale the accessibility values, they consistently overestimate *job access* compared to the re-scaled spatial availability measure. Overestimation ranges from -3% to -100% and is on average -81%. However, while accessibility consistently overestimates *job access* relative to spatial availability, it does so in a variable way. *Job access* is more overestimated in TAZs that have a low number of workers and is less overestimated in TAZs with a high number of workers; the correlation between the number of workers and the percentage difference between measures is 0.877. As such, the TAZs that have a relatively low number of workers are located between the border of the GGH and the City of Toronto and are also mostly coloured by red and orange indicating Quantiles 1 and 2 (highest percentage of overestimation) and vice versa. Both *job access* measures are more similar when the number of workers correlates with travel cost and the number of jobs at the destination. However, when the calibrated travel cost outpaces the number of workers, as is the case for the low-worker TAZs within this empirical example, then these measures increasingly diverge. This difference emphasizes the importance in using an accurate and well-informed impedance function and origin-destination data. For reference to these trends, the plots for workers and impedance function values for all TAZs are presented in Figure 9.

Accessibility when used to measure *job access* obscures the *availability* of jobs. Overall, this can be seen when comparing Figure ?? and Figure ?? where the intra-measure magnitude of accessibility is much higher within and closer to the boundary of the City of Toronto while spatial availability is more level as it is proportioned by the number of workers within each origin TAZs. Since spatial availability is proportioned by the number of workers, the resulting *job*

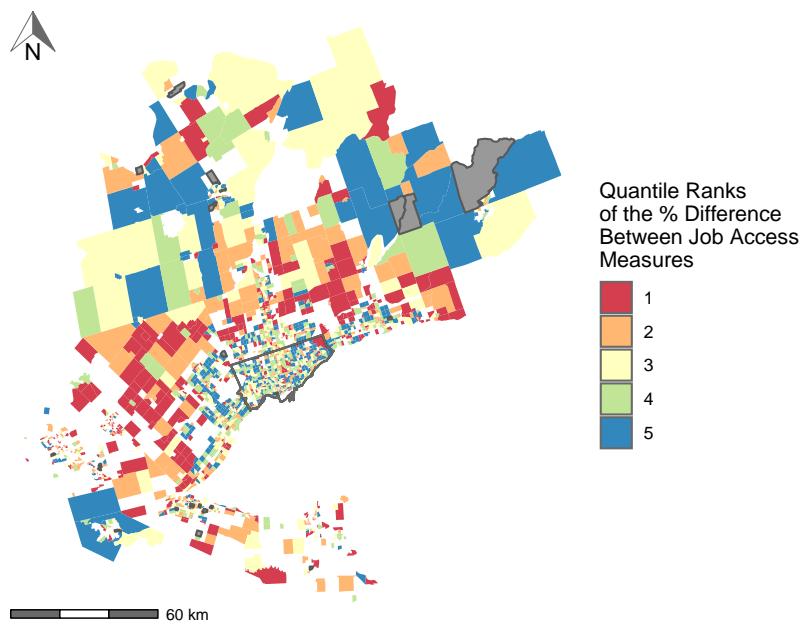


Figure 8: Difference in rescaled accessibility and spatial availability values in the context of employment from origins in the GGH to destinations in the City of Toronto. TAZs with rescaled spatial accessibility values that are lower than rescaled accessibility values are presented in five quantiles from 1 to 5 with 1 representing the highest negative change between spatial availability and accessibility. TAZs with values which are higher than re-scaled accessibility are shown in grey.

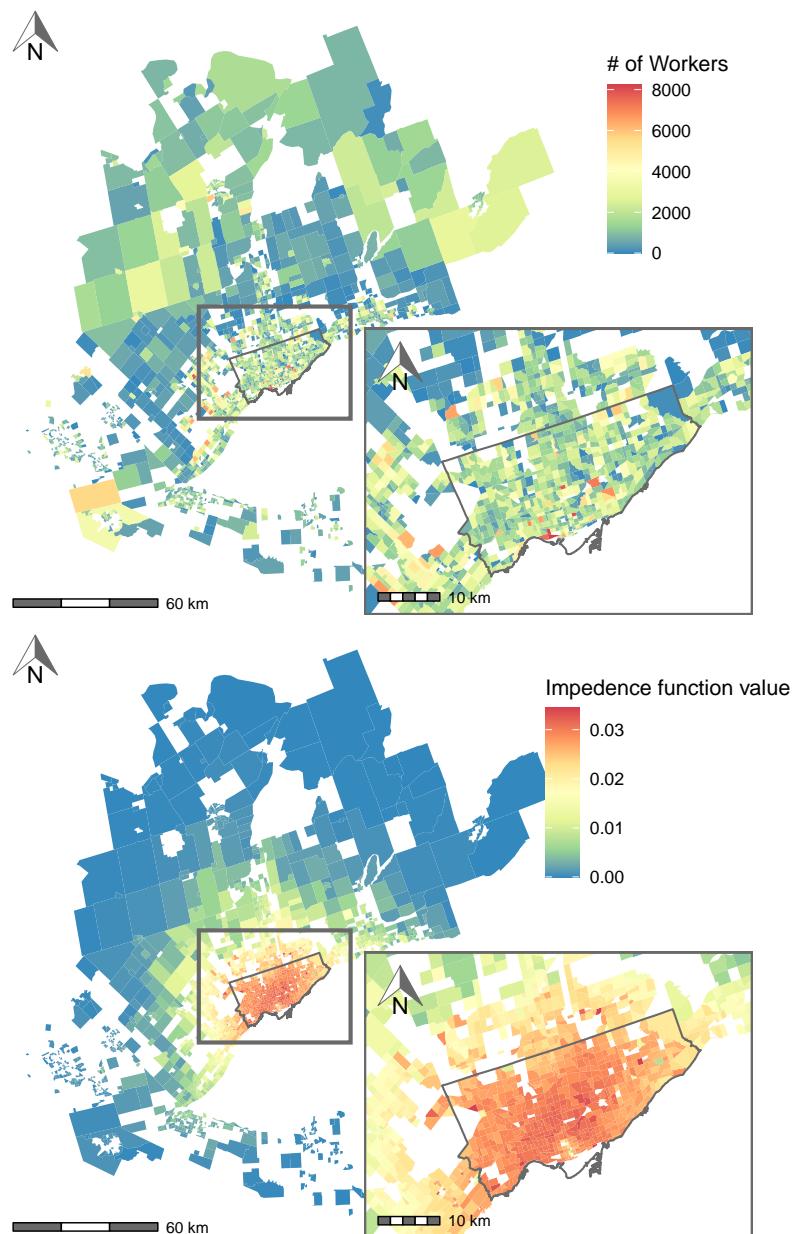


Figure 9: Number of workers (top) and average impedance function value (bottom) for each origin TAZ in the GGH.

access value is numerically meaningful to determine what *job access* is available per worker and how it compares to the average *job access* for the region. In this way, spatial availability is more interpretable than accessibility as it can be stated that the average Toronto jobs per worker is XX for the GGH and XX for Toronto; interestingly, all TAZs in Toronto have a higher *job access* than the GGH value but still range between XX and XX.

Comparing Opportunities Access Measures - All GGH

Contextualizing Spatial Availability Use Cases

Note: A_i is vulnerable to the modifiable areal unit problem (MAUP)

With the toy and empirical example presented, we propose that spatial availability be seen as a type of *spatial mismatch* and an evolution of the *balanced floating catchment approach* (BFCA). Historically, literature has iteratively improved measures assessing access to opportunities. In the context of employment and healthcare opportunities, simple container counting solutions such as the population-to-provider ratio (PPR) for healthcare services and jobs to housing ratio were implemented. This approach, while straight forward, is highly susceptible to the modifiable unit area problem (MUAP). Recognizing this and harnessing the computation power and access to finer resolution data as it became available, scholars proposed the next evolution, namely the accessibility measure which is widely used today and we implemented in our examples for comparison. It partially addresses the MUAP by considering opportunities outside of the conventional ‘containers’ which represented opportunities in a census areas/neighbourhoods by counting opportunities informed by an impedance function based on travel cost. Our measure, spatial availability, iterates on the accessibility measure by proportionally allocating travel cost and population (workers) of origins to opportunities at destinations. This single-constrained approach ensures that the population is mutually exclusive in essence replicating the properties of the self-contained unit which is *not* limited to a zoning system proposed; the pros of accessibility solution in using the impedance function and pros of the container solutions.

Similar to the accessibility measure, it should be noted that spatial availability is only as robust to the MUAP as the input data allows. For instance, in the empirical example present, the measure of *job access* still only considers population, opportunities, and travel times from the centroids of TAZs. However, unlike the accessibility measure, spatial availability can be meaningful calculated on a per population basis at higher resolutions because of the proportionality property.

Further, the measure of spatial availability can be a useful way to distinguish between low accessibility/low population centers, which may enjoy higher availability than the accessibility value may suggest, and contrariwise, high accessibility/high population centers (which potentially can result in lower availability due to competition). For instance, as presented in the numerical example, more remote, smaller population centers can have sufficient spatial availability by being in close proximity to the smaller employment centers; however this

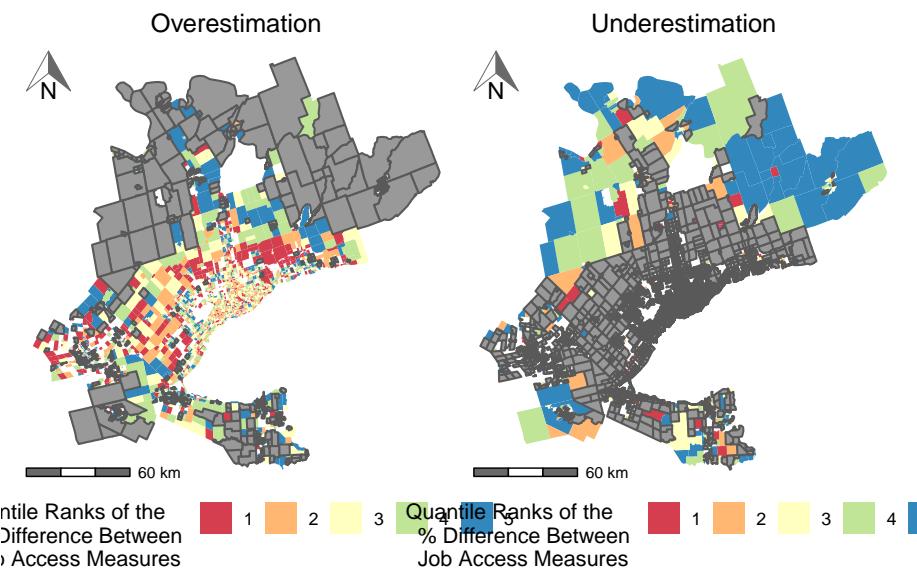


Figure 10: Difference in rescaled accessibility and spatial availability values for access to employment in the GGH. TAZs with rescaled spatial accessibility values that are lower (top) and higher (bottom) than rescaled accessibility values are presented in five quantiles from 1 to 5 with 1 representing the highest magnitude of change between spatial availability and accessibility.

sufficiency is obscured by accessibility measure by conglomerating accessibility of population centers which are more central to more (and larger) employment centers. Conversely, referring to the empirical GGH example, the TAZs that are relatively close to the Toronto job destinations but have a relatively low number of workers receive sufficiently high values of accessibility and significantly lower spatial availability values; this trend is opposite to what is experienced in the numerical example and is a result of higher job competition (i.e. job opportunities are more highly allocated to TAZs with lower travel costs and higher worker populations). Measuring access to opportunities is a multi-scale problem; since the number of opportunities are preserved, different scales, populations, and time-windows can be incorporated within the measure without introducing additional spatial basis.

Fundamentally, accessibility measure's methodology results in the overestimation of *jobs access* since it simply sums the count of destination opportunities based on travel cost to opportunities; the measure does not include factors to bound the summation so origins, hypothetically, can have infinite *opportunity access*. This property can be practical for calculating *opportunity access* for opportunities which have large capacities and are *non-competitive* such as large natural parks and beaches. These non-competitive opportunities can be considered infinitely divisible as they offer more 'spots' at any given time than the population. However, often times opportunities are not divisible but are in fact indivisible and competitive meaning, such as the numerical and GGH empirical example of *jobs access*, where only 1 worker can access 1 job at any given time. In these cases, spatial availability measure can be used to calculate the *opportunity access* in which the report values are numerically meaningful, the MUAP is potentially addressed, and *opportunity access* values across regions, neighbourhoods, and spatial scales can be compared.

Spatial Availability Use Cases

With the analytical framework for spatial availability outlined, in the following sub-sections, we return to the numerical example which was introduced at the start of this section and illustrate how spatial availability V_i can be used for two other use cases.

Available Jobs for The Working Population

In this use case, we calculate the spatial availability of jobs using the same impedance function that was used to calculate job accessibility in Figure 1. The spatial availability calculations are implemented in the function `sp_avail`. The inputs are an Origin-Destination table with labels for the origins (`o_id`), labels for the destinations (`d_id`), the population (`pop`) and number of opportunities (`opp`), an indicator for catchments or other eligibility constraints (`r`), and a pre-calculated impedance function (`f`). For this example, we assume that there are no catchment restrictions by setting `r` to 1.

The value of the function (its output) is a vector with V_{ij} given the inputs, that is, the opportunities available to i from j :

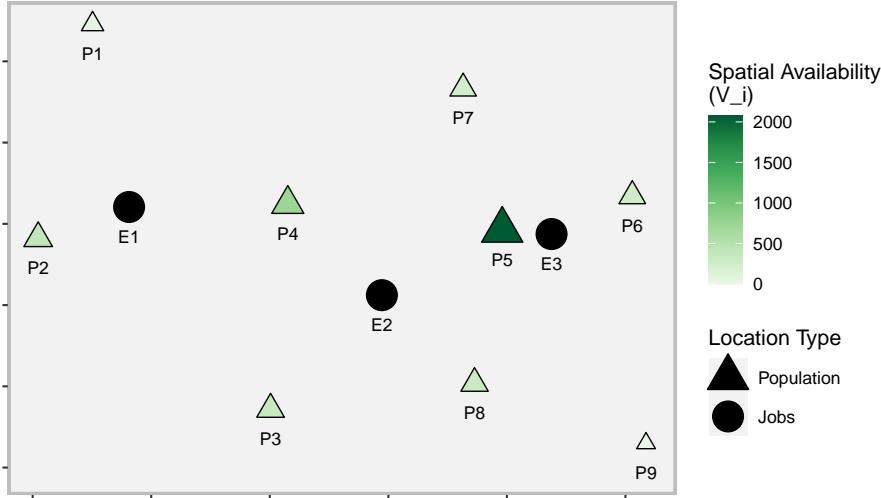


Figure 11: Spatial availability of jobs from population centers for the simple numerical example

Figure 11 shows the estimates of spatial availability:

We see that population center 5 has the highest level of spatial availability, due to being a large population center that is moreover relatively close to jobs. To improve the interpretability of this measure, we first note that the regional measure of jobs per capita is 0.994. We then calculate the spatially available jobs per person by dividing each spatial availability value by the population at each population center (Figure 12).

Some population centers have almost two jobs available per person (compared to the overall regional value of approximately one job per person), while others have less than one job available per person. This does not mean that people are not taking some of the jobs. It means that controlling for the cost of reaching jobs, they are worse off than those with more jobs spatially available.

Alternative Use Case 2: Available Jobs for Specialized Working Populations

In this sub-section we introduce catchment/eligibility constraints. Due to differences in educational achievement among the population, the jobs in Employment Center 1 can only be taken by individuals in population centers 1 and 2. Jobs in Employment Center 2 can be taken by individuals in population centers 3, 4, 5, 7, and 8. Lastly, jobs in Employment Center 3 require qualifications available only among individuals in population centers 5, 6, 8, and 9.

Calculate the spatial availability by proportionally allocating *specialized* workers to jobs (we refer to spatial availability in this case as V_{ijr}):

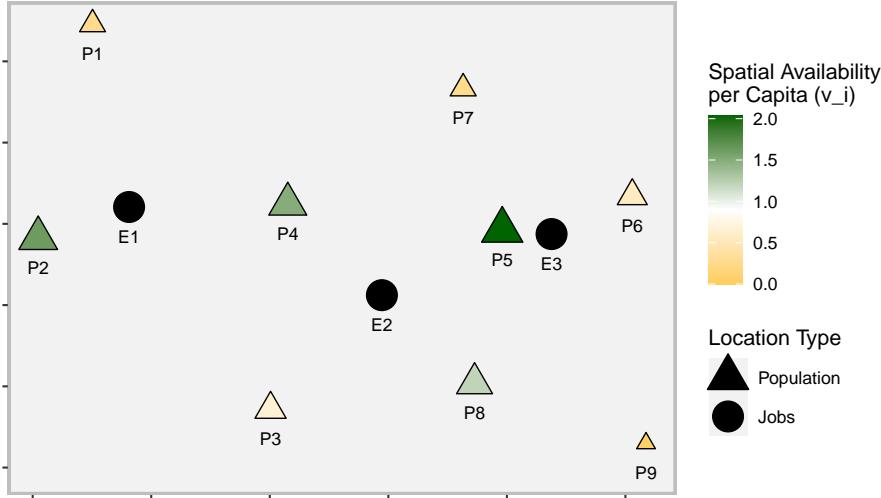


Figure 12: Spatial availability of jobs per capita from population centers for the simple numerical example

Plot the availability estimates:

Available jobs per person with catchment/eligibility conditions:

The plot in Figure 14 shows the availability per person without and with catchment restrictions.

We can see that when there are catchment restrictions population center 2, despite being relatively peripheral, has higher spatial availability due to spatial specialization. With catchments, the spatial availability of jobs declines from the perspective of population center 4: the population here has skills required for jobs at a small employment center, where they face substantial competition from other population centers.

Alternative Use Case 3: Available Workers for Employment Centers

We can also examine the pool of workers available to each employment center by considering the workers as the opportunities and the jobs as the population.

Calculate the spatial availability by proportionally allocating jobs to workers (we refer to spatial availability in this case as W_{ji}) as follows:

Plot the availability estimates:

Plot the spatial availability of workers per job:

Limitations

still does not tell us how many accessible opportunities will lead to participation/positive outcomes; it does come closer... this is an important question

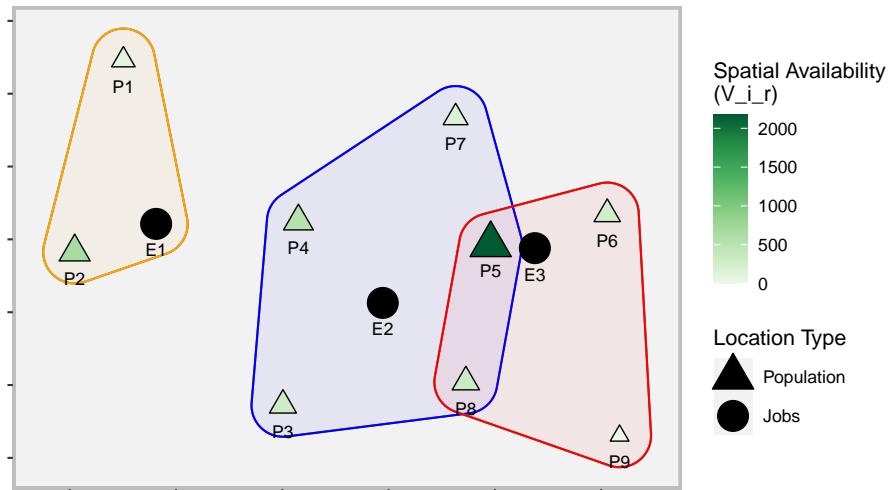


Figure 13: Spatial availability of jobs from population centers assuming catchment restrictions for the simple numerical example

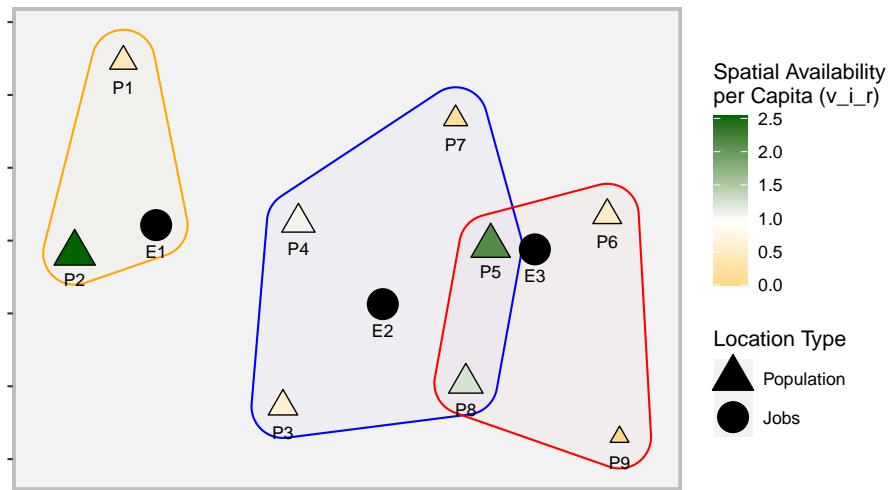


Figure 14: Spatial availability of jobs per capita of jobs from population centers with and without catchment restrictions for the simple numerical example

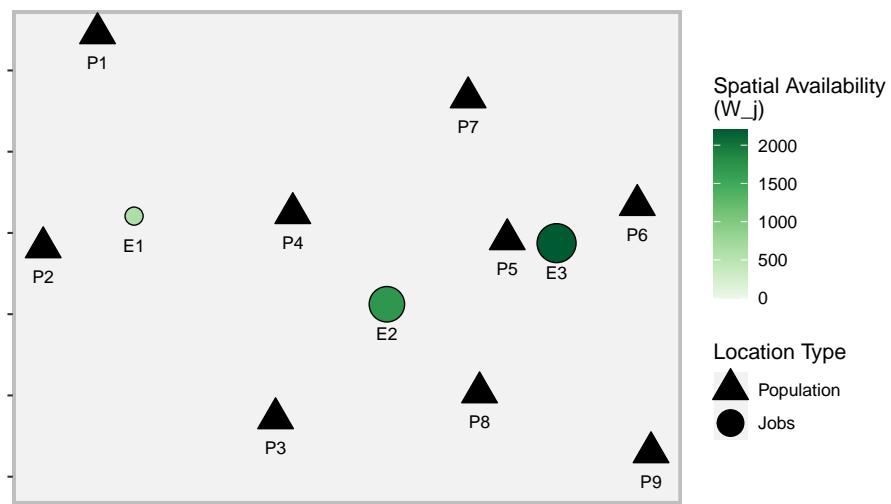


Figure 15: Spatial availability of workers

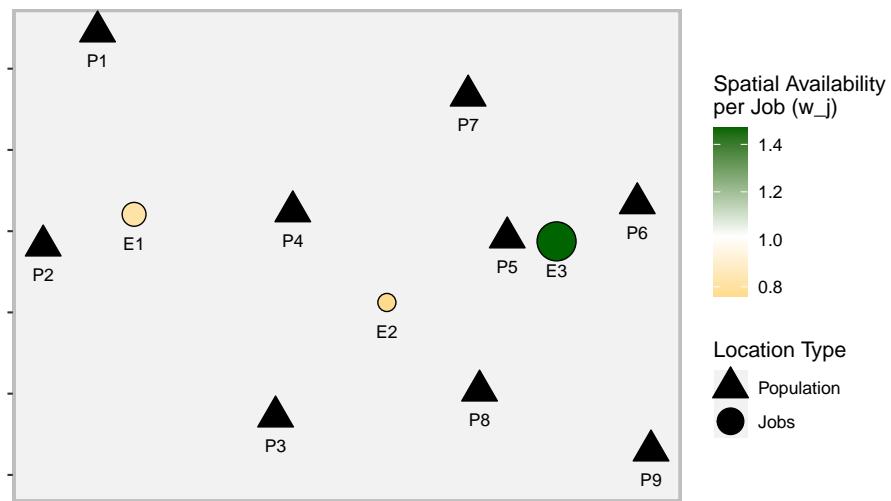


Figure 16: Spatial availability of workers per job

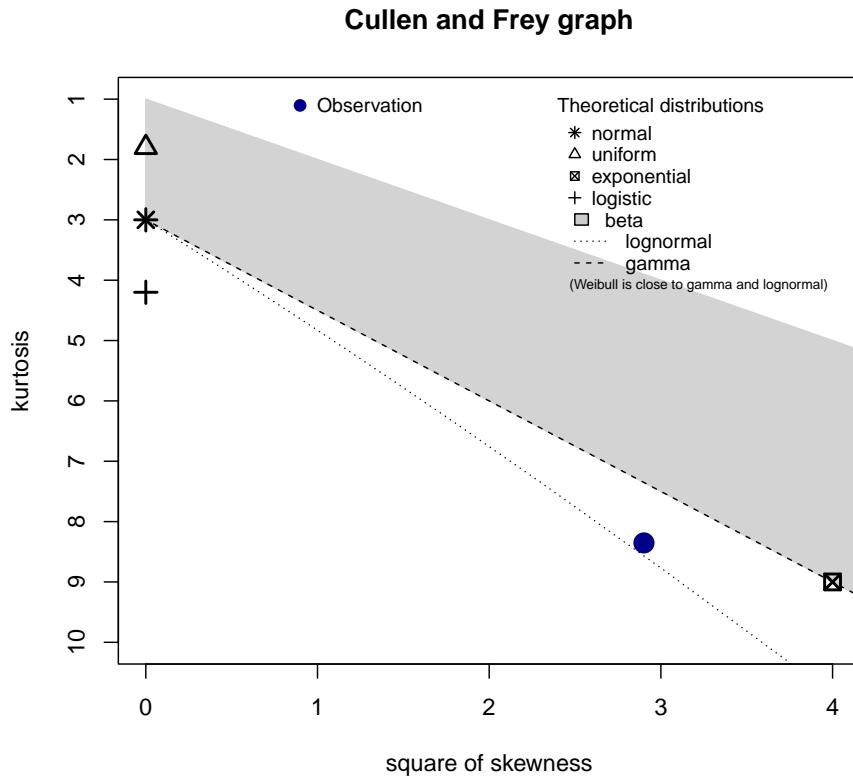


Figure 17: Cullen and frey graph for the complete empirical 2016 TTS travel time data set.

for equity analysis.

Conclusion

Words go here.

Appendix

```
summary statistics
-----
min: 0.1   max: 179
median: 18
mean: 21.4344
estimated sd: 14.61254
estimated skewness: 1.703326
estimated kurtosis: 8.353363
```

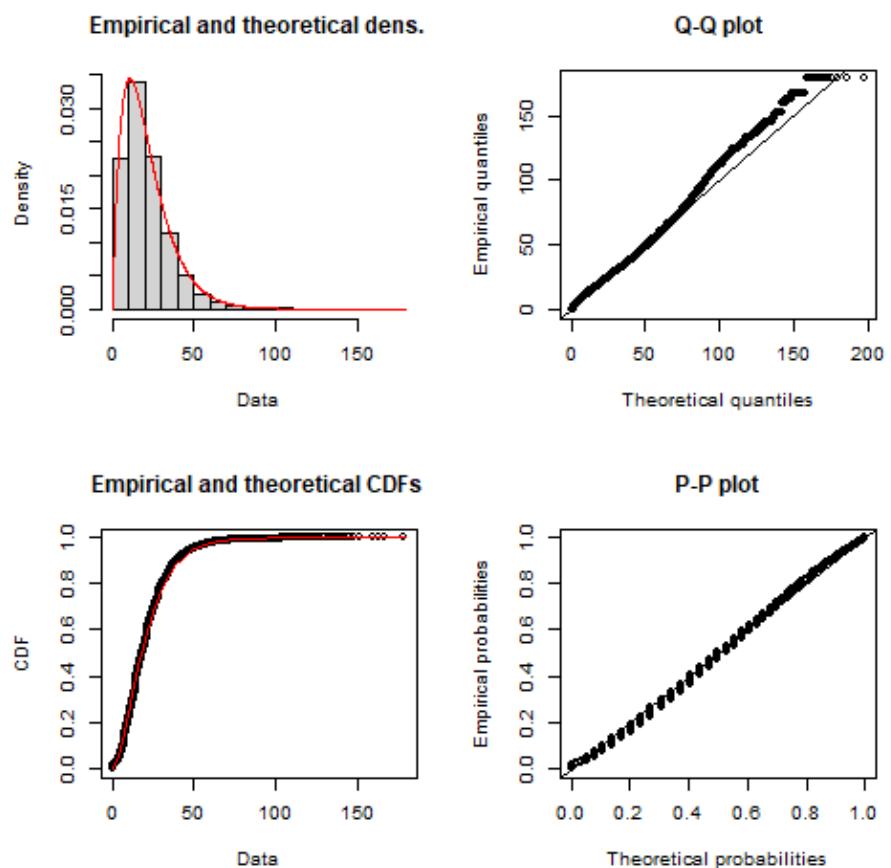


Figure 18: Diagnostics associated with the fitted gamma distribution.

References

- Allen, J., Farber, S., 2019. A Measure of Competitive Access to Destinations for Comparing Across Multiple Study Regions. *Geographical Analysis* 52, 69–86. doi:10.1111/gean.12188
- Deboosere, R., El-Geneidy, A.M., Levinson, D., 2018. Accessibility-oriented development. *Journal of Transport Geography* 70, 11–20. doi:10.1016/j.jtrangeo.2018.05.015
- Delamater, P.L., 2013. Spatial accessibility in suboptimally configured health care systems: A modified two-step floating catchment area (M2SFCA) metric. *Health & Place* 24, 30–43. doi:10.1016/j.healthplace.2013.07.012
- Delignette-Muller, M.L., Dutang, C., 2015. fitdistrplus: An R package for fitting distributions. *Journal of Statistical Software* 64, 1–34.
- Geurs, K.T., van Wee, B., 2004. Accessibility evaluation of land-use and transport strategies: review and research directions. *Journal of Transport Geography* 12, 127–140. doi:10.1016/j.jtrangeo.2003.10.005
- Handy, S., 2020. Is accessibility an idea whose time has finally come? *Transportation Research Part D: Transport and Environment* 83, 102319. doi:10.1016/j.trd.2020.102319
- Hansen, W.G., 1959. How Accessibility Shapes Land Use. *Journal of the American Institute of Planners* 25, 73–76. doi:10.1080/01944365908978307
- Joseph, A.E., Bantock, P.R., 1984. Rural Accessibility of General Practitioners: the Case of Bruce and Grey Counties, ONTARIO, 1901–1981. *The Canadian Geographer/Le Géographe canadien* 28, 226–239. doi:10.1111/j.1541-0064.1984.tb00788.x
- Luo, W., Wang, F., 2003. Measures of Spatial Accessibility to Health Care in a GIS Environment: Synthesis and a Case Study in the Chicago Region. *Environment and Planning B: Planning and Design* 30, 865–884. doi:10.1068/b29120
- Miller, E.J., 2018. Accessibility: measurement and application in transportation planning. *Transport Reviews* 38, 551–555. doi:10.1080/01441647.2018.1492778
- Paez, A., Higgins, C.D., Vivona, S.F., 2019. Demand and level of service inflation in Floating Catchment Area (FCA) methods. *PLOS ONE* 14, e0218773. doi:10.1371/journal.pone.0218773
- Proffitt, D.G., Bartholomew, K., Ewing, R., Miller, H.J., 2017. Accessibility planning in American metropolitan areas: Are we there yet? *Urban Studies* 56, 167–192. doi:10.1177/0042098017710122
- Rafael H. M. Pereira, Marcus Saraiva, Daniel Herszenhut, Carlos Kaeue Vieira Braga, Matthew Wigginton Conway, 2021. r5r: Rapid realistic routing on multimodal transport networks with R5 in r. *Findings*. doi:10.32866/001c.21262
- Shen, Q., 1998. Location characteristics of inner-city neighborhoods and employment accessibility of low-wage workers. *Environment and Planning B: Planning and Design* 25, 345–365. doi:10.1068/b250345
- Wan, N., Zou, B., Sternberg, T., 2012. A three-step floating catchment area method for analyzing spatial access to health services. *International Journal of Geographical Information Science* 26, 1073–1089. doi:10.1080/13658816.2011.624987
- Wilson, A.G., 1971. A Family of Spatial Interaction Models, and Associated Developments. *Environment and Planning A: Economy and Space* 3, 1–32.

doi:10.1068/a030001

Yan, X., 2021. Toward Accessibility-Based Planning. *Journal of the American Planning Association* 87, 409–423. doi:10.1080/01944363.2020.1850321