

# Estimating spatial availability/mismatch using singly constrained accessibility measures

Author One<sup>a</sup>, Author Two<sup>a</sup>

<sup>a</sup>Address

## Abstract

This is the abstract.

It consists of two paragraphs.

---

\*Corresponding Author

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

## Introduction

The concept of accessibility (*opportunity access*) 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 *opportunity access* 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.

*Opportunity access* analysis 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 access measures sensitive to the number of opportunities and the characteristics of the transportation network, raw accessibility 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 *opportunity access*. 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 *opportunity access* 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 *opportunity access* 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 *opportunity access*. 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 *opportunity access* measures (excluding utility-based measures) are derived from the gravity model, and are known as *gravity-based* accessibility. Briefly, consider the following widely used accessibility measure  $A_i$  :

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

where:

- $i$  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 such as a negative exponential, lognormal, or gamma .

The accessibility value  $A_i$  is the weighted sum of opportunities that can be reached from location  $i$ , given the cost of travel  $c_{ij}$  and an impedance function. 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. Depending on the impedance function, the measure could be cumulative opportunities (if  $f(\cdot)$  is a binary or indicator function) or a more traditional gravity measure, for instance with a Gaussian impedance function or an inverse cost impedance function .

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.

### Numerical Example

In this section we present a simple numerical example. The setup for the example is a system with three employment centers and nine population centers, as seen in Table 1.

The accessibility to employment of each of the population centers can be calculated using the expression above for  $A_i$ . As noted, this yields the number

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

of jobs (opportunities) that are accessible (i.e., can be reached) from each population center, given the cost. In this example we use the straight line distance between the population and jobs for  $c_{ij}$ , and a negative exponential function with  $\beta = 0.0015$ .

Figure 1 shows the 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. At a glance:

- Population centers (triangles) in the middle of the map 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  $A_i = 17.12$ );

#### *What are the Issues?*

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, however, despite the accessibility value for P1 being low it is still better than zero. This evaluation leaves decision makers unclear on how to interpret accessibility values, particularly extreme values and particularly when comparing across scenarios.

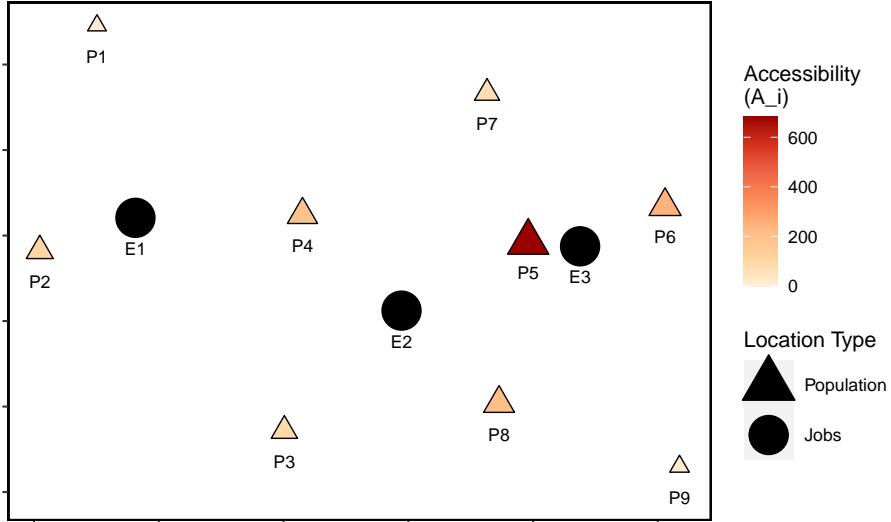


Figure 1: Accessibility results using simple numerical example

or regions . The obscurity in interpretation of accessibility values arises from the following two ways.

First, origins with accessibility to a high density of opportunities are susceptible to having disproportionately higher accessibility values than origins with a lower density of opportunities. This occurs since total accessibility ( $\sum_{i=1}^I A_i$ ) depends on the number of origins; every additional origin in the analysis causes at least one and possibly more opportunities at destinations to be multiple-counted. As opportunities are multiple-counted, origins which are in proximity to high density of opportunities have accessibility values which are relatively conglomerated compared to origins in proximity to a lower density of opportunities. This issue means total accessibility is vulnerable to the modifiable areal unit problem (MAUP) as opportunities and origins which are spatially aggregated are multiple-counted inconsistently.

Second, the calculated accessibilities do not take competition into account. For example, an individual at P5 has accessibility to 680.6373657 jobs. But since this is also a large population center, there is potentially large competition for those accessible jobs. In other words, the value of  $A_i$  is not sensitive to the size of population at the origin seeking the opportunity (in this case jobs), let alone the population at other locations. This unfortunately limits the interpretability of the measure. Floating catchment areas purport to account for competition/congestion effects, but as discussed by Paez, Higgins, and Vivona (2019), they are vulnerable to conglomeration, which makes them prone to bias unless corrected.

To address these two shortcomings of the accessibility measure, we propose a singly-constrained gravity measure that corresponds to the concept of *spatial availability*.

### Spatial Availability: Proportional Allocation of Opportunities Based on Demand

#### *Analytical Framework*

As recent research on accessibility shows (see for instance Paez, Higgins, Vivona (2019) and Allen and Farber (2019)), accounting for competitive access in a meaningful way requires the proportional allocations of quantities in the accessibility calculations, or their normalization. At issue is the fact that multiple-counting is commonplace when calculating conventional accessibility  $A_i$  for  $i = 1, \dots, n$  as every opportunity enters the weighted sum once for every origin  $i$  that can reach it. This has the unfortunate effect of obscuring the interpretability of  $A_i$  and fails to answer for a individual at a specific population center the question: “*many jobs are accessible, but the same jobs are also accessible to my (possibly) numerous neighbors... what does high accessibility actually mean to me?*”

In the spatial availability framework proposed, and in line with the gravity tradition, we distinguish between opportunities at a destination and demand for opportunities at the origin. To explain the analytical framework, the example of access to employment is illustrative, with “population” in the role of demand (i.e., the number of individuals in the labour market who “demand” employment) and “jobs” in the role of opportunities.

As an overview, spatial availability ( $V_{ij}$ ) is the proportional allocation of opportunities ( $O$ ) and allocation factor for population ( $f^p$ ) and cost of travel ( $f^c$ ). Since spatial availability ( $V_{ij}$ ) consists of these two allocation factors, this example first details how the population allocation factor  $f_{ij}^p$  produces  $V_{ij}^p$ , next details the role of travel cost allocation factor  $f_{ij}^c$  in producing  $V_{ij}^c$ , and finally combines both allocation factors in the final general form of spatial availability  $V_{ij}$  as follows:

$$V_{ij} = O_j \frac{f_{ij}^p \cdot f_{ij}^c}{\sum_{k=1}^K f_{kj}^p \cdot f_{kj}^c}$$

#### *Population Allocation Factor*

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:

$$f_{ij}^p = \frac{P_{i \in r}^\alpha}{\sum_{k=1}^K P_{k \in r}^\alpha}$$

On the right hand side of the equation,  $P_{i \in r}$  is the population at location  $i$  that is eligible for jobs of type  $r$  (maybe those with a certain level of training, or in a

designated age group). The summation in the bottom is over  $k = 1, \dots, K$ , the number of population centers in the region. The resulting factor  $f_{ij}^p$  corresponds to the proportion of the population in zone  $i$  relative to the rest of the region's population centres  $K$ . The factors  $f_{ij}^p$  satisfy the property that  $\sum_i^I f_{ij}^p = 1$ . We can also add an empirical parameter  $\alpha$  that can be used to modulate the effect of size in the calculations (i.e.,  $\alpha < 1$  places greater weight on smaller centres relative to larger ones while  $\alpha > 1$  achieves the opposite effect).

This factor ( $f_{ij}^p$ ) can now be used to proportionally allocate a share of the jobs at the employment centre  $j$  to population center  $i$  and population center  $k$ . The share of jobs at  $j$  allocated to (i.e., available to) each population center is:

$$V_{ij}^p = O_j f_{ij}^p$$

and since  $\sum_i^I f_{ij}^p = 1$  it follows that:

$$\sum_{i=1}^I V_{ij}^p = O_j$$

In other words, the number of jobs is preserved. The result is a proportional allocation of available jobs to population centres based on demand. As an example, consider an employment center  $j$  in a region with two population centers (say  $i$  and  $k$ ). For simplicity, assume that the all the population in the region is eligible for these jobs, that is, that the entirety of the population is included in the set  $r$ . The allocation factors for the jobs at  $j$  would be:

$$f_{ij}^p = \frac{P_i^\alpha}{P_i^\alpha + P_k^\alpha}$$

$$f_{kj}^p = \frac{P_k^\alpha}{P_i^\alpha + P_k^\alpha}$$

Suppose that there are three hundred jobs in the employment center ( $W_j = 300$ ), and that the populations are  $P_i = 240$  and  $P_k = 120$ . The jobs are allocated as follows (assuming that  $\alpha = 1$ ):

$$V_{ij}^p = O_j f_{ij}^p = O_j \frac{P_i^\alpha}{P_i^\alpha + P_k^\alpha} = 300 \frac{240}{240+120} = 300 \frac{240}{360} = 200$$

$$V_{kj}^p = O_j f_{kj}^p = O_j \frac{P_k^\alpha}{P_i^\alpha + P_k^\alpha} = 300 \frac{120}{240+120} = 300 \frac{120}{360} = 100$$

It can be seen that proportionally more jobs are allocated to the bigger center and also that the total number of jobs is preserved. However, the factors above account for the total number of opportunities at the destination (i.e., the number of jobs at the employment center), but they do not account for their location relative to the population centers. The proportional allocation procedure above is insensitive to how far population centers  $i$  or  $k$  are from employment center  $j$ . To account for this effect we define a second set of allocation factors based on distance to the employment centers.

### *Travel Cost Allocation Factor*

These are defined as:

$$f_{ij}^c = \frac{f(c_{ij})}{\sum_{k=1}^K f(c_{kj})}$$

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, this allocation factor ( $f_{ij}^c$ ) serves to proportionally allocates more jobs to closer locations through an impedance function. To illustrate, assume that the impedance function is a negative exponential function as follows, and assume that  $\beta$  (which modulates the steepness of the impedance effect and is an empirical parameter) is the value of 1:

$$f(c_{ij}) = \exp(-\beta c_{ij})$$

Continuing the example, suppose that the distance from population center  $i$  to employment center  $j$  is 0.6 km, and the distance from population center  $k$  to employment center  $j$  is 0.3 km. Being closer, we would expect more jobs to be allocated to the population of  $k$ . The jobs would be sorted as follows:

$$\begin{aligned} f_{ij}^c &= \frac{\exp(-\beta D_{ij})}{\exp(-\beta D_{ij}) + \exp(-\beta D_{kj})} \\ f_{kj}^c &= \frac{\exp(-\beta D_{kj})}{\exp(-\beta D_{ij}) + \exp(-\beta D_{kj})} \end{aligned}$$

This step normalizes the impedance between  $i$  and  $j$  and  $k$  and  $j$  by the total impedance in the study area. Numerically, the jobs allocation is:

$$\begin{aligned} V_{ij}^c &= O_j f_{ij}^c = O_j \frac{\exp(-D_{ij})}{\exp(-D_{ij}) + \exp(-D_{kj})} = 300 \frac{\exp(-0.6)}{\exp(-0.6) + \exp(-0.3)} = 3 \times 0.426 = 127.8 \\ V_{kj}^c &= O_j f_{kj}^c = O_j \frac{\exp(-D_{kj})}{\exp(-D_{ij}) + \exp(-D_{kj})} = 300 \frac{\exp(-0.3)}{\exp(-0.6) + \exp(-0.3)} = 3 \times 0.574 = 172.2 \end{aligned}$$

A larger share of jobs (172.2 jobs) is allocated to the population center ( $k$ ) that is closest as assumed. As before, the sum of jobs allocated to the population centers matches the total number of jobs available. Nevertheless, in isolation, this step does not account for the allocation of jobs based on demand.

### *Putting Spatial Availability Together*

We can combine the proportional allocation factors by population and distance and calculated spatial availability ( $V_{ij}$ ) as follows:

$$V_{ij} = O_j \frac{f_{ij}^p \cdot f_{ij}^c}{\sum_{k=1}^K f_{kj}^p \cdot f_{kj}^c}$$

When applied to the example of population center  $i$  and  $k$  (i.e., the demand) traveling to employment center  $j$  (i.e., the opportunity  $O$ ):

$$V_{ij} = O_j \cdot \frac{f_{ij}^p \cdot f_{ij}^c}{f_{ij}^p \cdot f_{ij}^c + f_{kj}^p \cdot f_{kj}^c} = 300 \frac{\left(\frac{2}{3}\right)\left(0.426\right)}{\left(\frac{2}{3}\right)\left(0.426\right) + \left(\frac{1}{3}\right)\left(0.574\right)} = (300)\left(\frac{0.284}{0.475}\right) = 179.4$$

$$V_{kj} = O_j \cdot \frac{f_{kj}^p \cdot f_{kj}^c}{f_{ij}^p \cdot f_{ij}^c + f_{ik}^p \cdot f_{ik}^c} = 300 \frac{\left(\frac{1}{3}\right)\left(0.574\right)}{\left(\frac{2}{3}\right)\left(0.426\right) + \left(\frac{1}{3}\right)\left(0.574\right)} = (300)\left(\frac{0.191}{0.475}\right) = 120.6$$

Notice how fewer jobs are allocated to population center  $i$  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.,  $k$ ), but since it is smaller, it also gets a smaller share of the jobs overall. Again, the sum of jobs at employment center  $j$  that are allocated to population centers  $i$  and  $k$  simultaneously based on *population-* and *distance-* based allocation is preserved (i.e.,  $W_{ij} + W_{kj} = W_j$ ).

Availability is simply the sum of the above by origin:

$$V_i = \sum_{j=1}^J V_{ij}$$

This quantity represents opportunities (e.g., jobs) that can be reached from  $i$  (i.e., they are accessible), 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 (call it lower-case  $v_i$ ) as follows:

$$v_i = \frac{V_i}{P_i}$$

In the example above:

$$v_{ij} = \frac{V_{ij}}{P_i} = \frac{179.4}{240}$$

$$v_{kj} = \frac{V_{kj}}{P_k} = \frac{120.6}{120}$$

Less competition ( $P_k$  is the smallest population center in the region) and being closer to the jobs clearly works in favor of individuals at  $k$ . 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.

In the following sections we use the same numerical example presented above to illustrate how availability  $V_i$  is calculated. We conclude by contrasting the two measures in the final section.

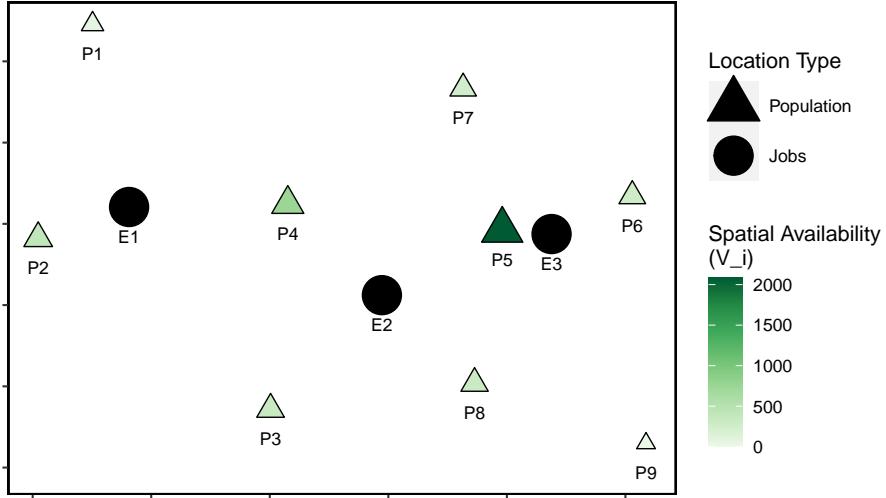


Figure 2: Spatial availability of jobs

### 1st Use Case: Available Jobs for The Working Population

In the following examples we use the same impedance function that was used to illustrate the accessibility calculations in the numerical example. 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 2 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 3).

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.

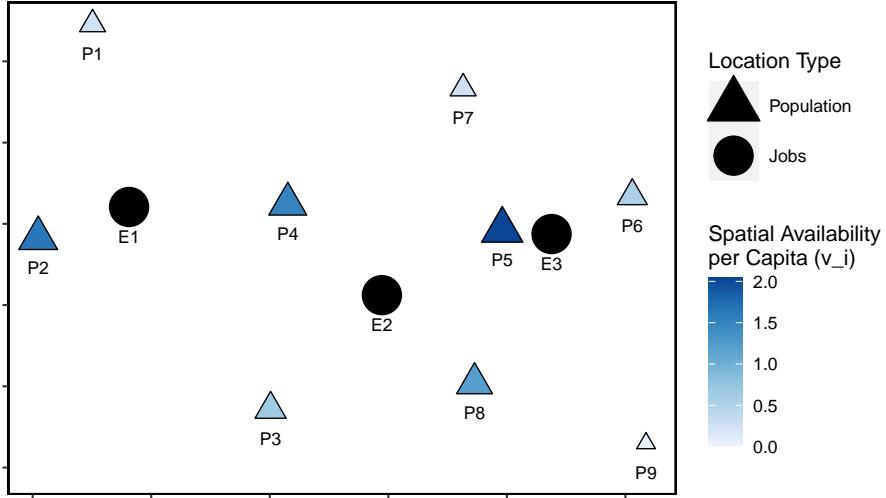


Figure 3: Spatial availability of jobs per capita

#### 2nd Use Case: Available Jobs for Specialized Working Populations

In this 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  $V_{ijr}$ ):

Plot the availability estimates:

Available jobs per person with catchment/eligibility conditions:

The plot in Figure 5 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.

#### 3rd Use Case: 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.

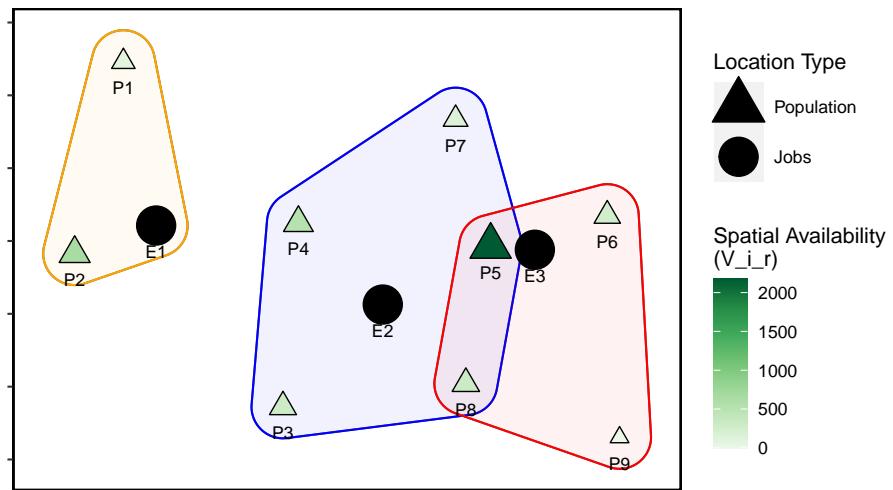


Figure 4: Spatial availability of jobs with catchment restrictions

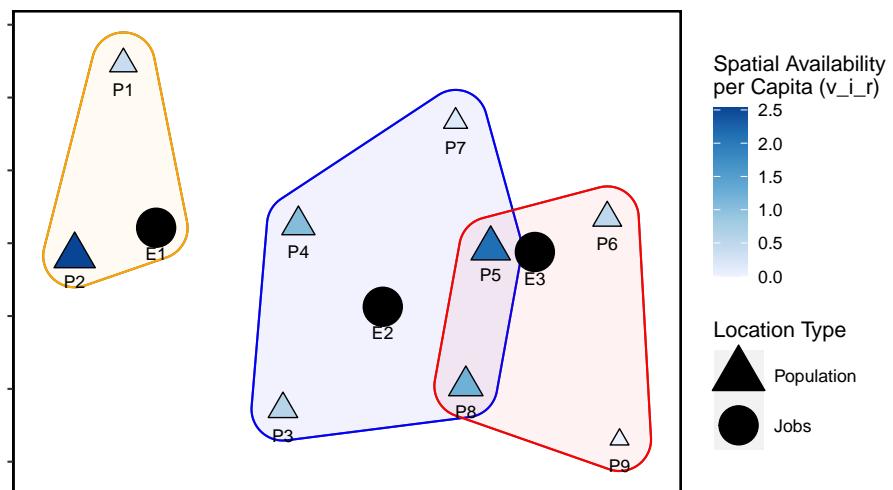


Figure 5: Spatial availability of jobs per capita with and without catchment restrictions

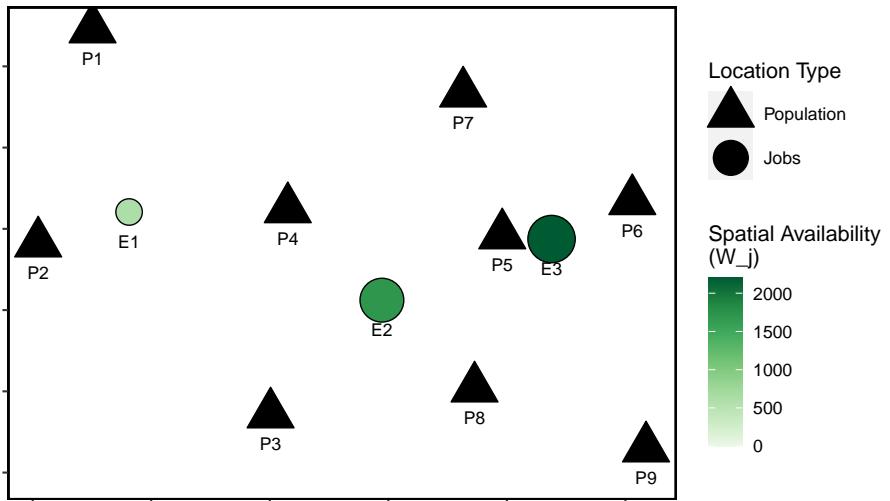


Figure 6: Spatial availability of workers

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:

### **Empirical Example: Spatial Availability and Accessibility of Jobs in Toronto**

For the reasons demonstrated in the hypothetical toy example; the spatial availability measure produces different and more interpretable results than accessibility. In this section we will demonstrate both measures and a comparison using empirical data for home-based work trips to places of employment in Toronto, Ontario.

#### *Data*

The 2016 Transportation Tomorrow Survey (TTS) data for 20 municipalities contained within the the Greater Golden Horseshoe (GGH) area in the province of Ontario, Canada ( $43.6^{\circ}\text{N}$   $79.73^{\circ}\text{W}$ ) is analysed (Figure 8). This data set includes home-based origins and employment destinations defined by centroids of Traffic Analysis Zones (TAZ) of varying areas ( $n=3764$ ), 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 calculated travel times by car (calculated

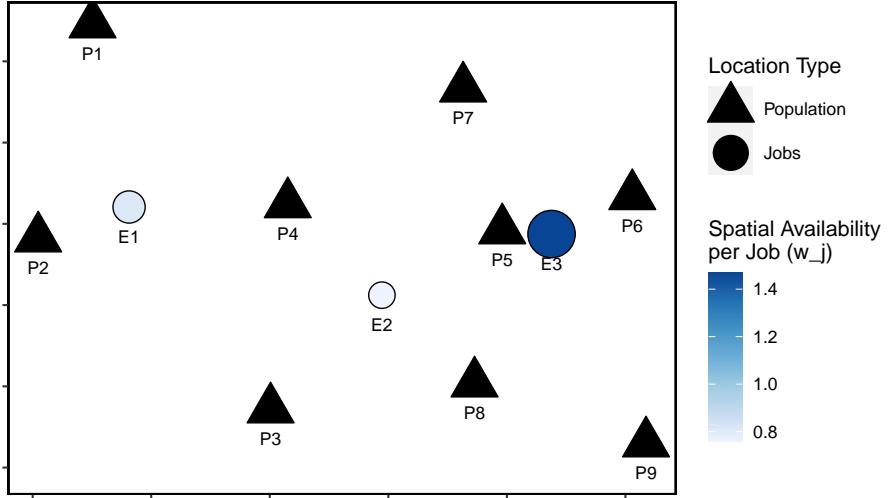


Figure 7: Spatial availability of workers per job

via `r5r`) and a derived impedance function values corresponding to the cost of travel based on the trip length distribution (TLD). The descriptive statistics are presented in Table 2 and a data-package is available to explore the data in greater detail.

#### *Calibrating an Impedance Function*

In the hypothetical toy data set, an arbitrary negative exponential function describing the increase in travel cost as distance increases was used as the impedance function to derive both accessibility and availability. In this data set, an impedance function can be derived from the empirical trip length distribution (TLD) as the number of trips and their travel cost (in the case, travel

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



Figure 8: The TTS 2016 study area within the Greater Golden Horseshoe in Ontario, Canada.

time in minutes) are known (see black data points in Figure 10).

The TLD density plot 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 using the maximum likelihood estimation method and Nelder-Mead method for direct optimization available within the `fitdistrplus` package in R (Delignette-Muller and Dutang, 2015; **R\_2021?**) . Based on goodness-of-fit criteria, the gamma distribution was selected (see red line in Figure 10). The gamma distribution function is given in the following general form:

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

where the estimated ‘shape’ is  $\alpha$ , the estimated ‘rate’ is  $\beta$ , and  $\Gamma(\alpha)$  is defined as:

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

The calibrated impedance function has a shape parameter of 2.019 and a rate parameter of 0.094. The function and diagnostics are plotted in Figure 9 and the predicted values are plotted alongside the empirical TLD in Figure 10. As a point of summary, values calculated by this calibrated impedance function captures our theoretical cost of travel (i.e. probability of travel based on the trip length distribution of empirical trips and associated travel time) within the following accessibility and spatial availability calculations.

#### *Accessibility and Spatial Availability of Jobs in Toronto*

Using the impedance function defined, the accessibility to employment destinations in Toronto from home-based origins anywhere in the GGH is calculated. The higher the accessibility value, the more accessible places of employment are to home-based origins as seen in Figure 11. It can be briefly summarized that the 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 using the same impedance function is calculated for the same survey area; the higher the spatial availability value the more available employment destinations in Toronto are to any home-based origins in the GGH (Figure 12). If a TAZ has a spatial availability value which is higher or lower than 637.813 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). 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 13, 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

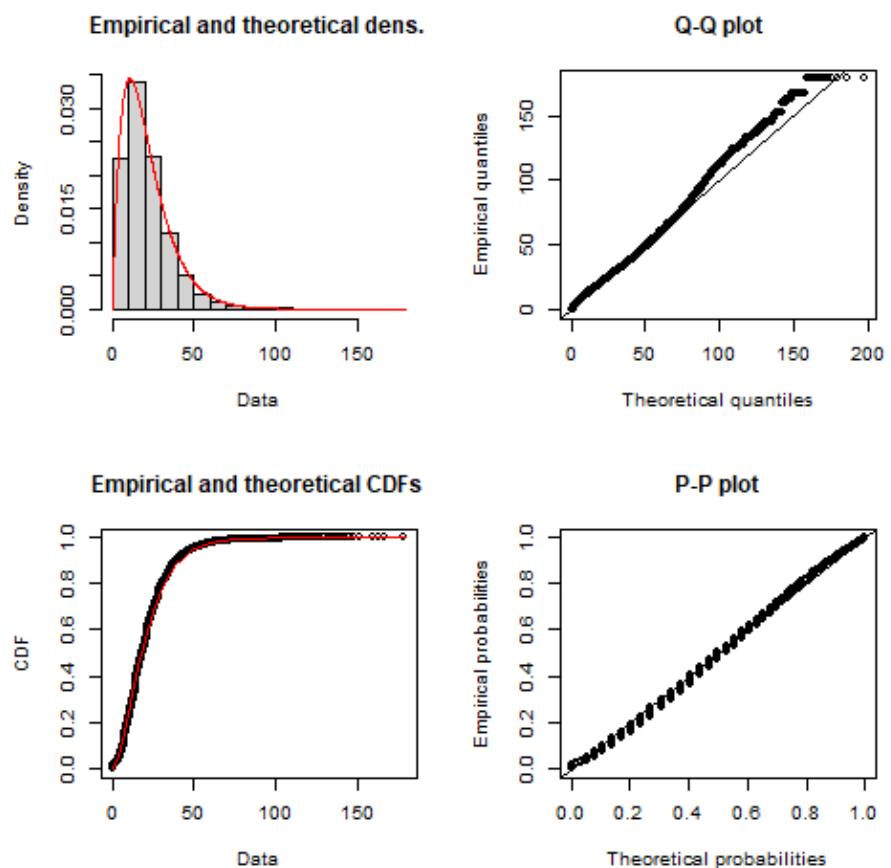


Figure 9: Impedance function and diagnostics.

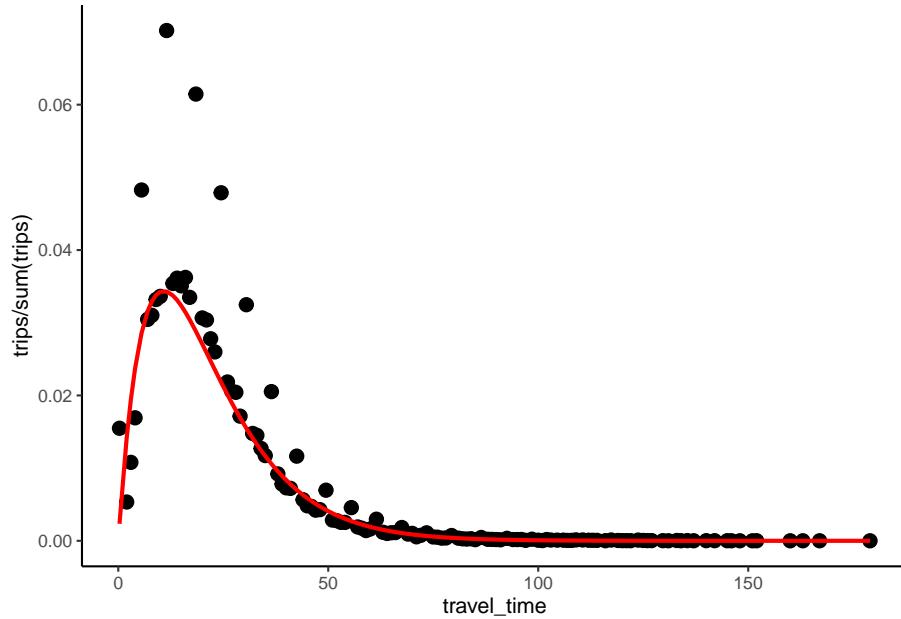


Figure 10: Empirical trip length distribution (black) and calibrated gamma distribution impedance function (red)

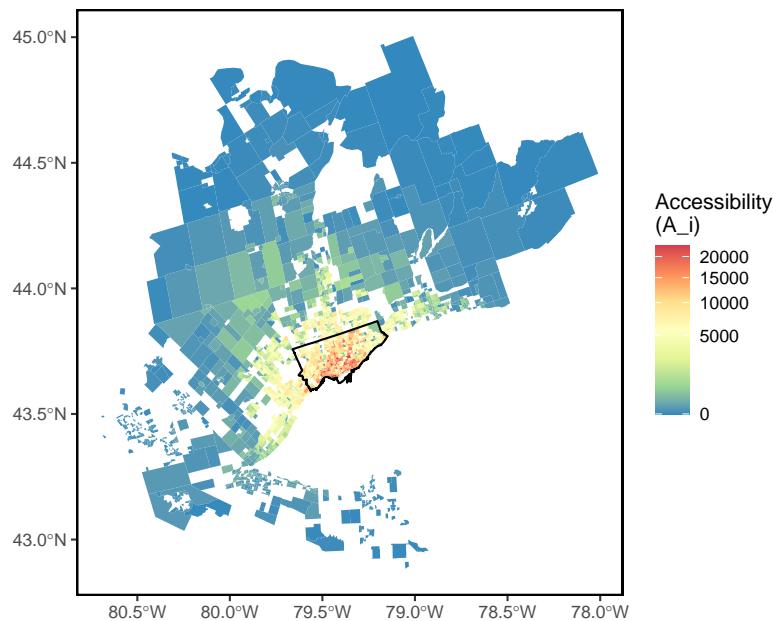


Figure 11: Calculated accessibility of employment from origins in the GGH to destinations in the City of Toronto

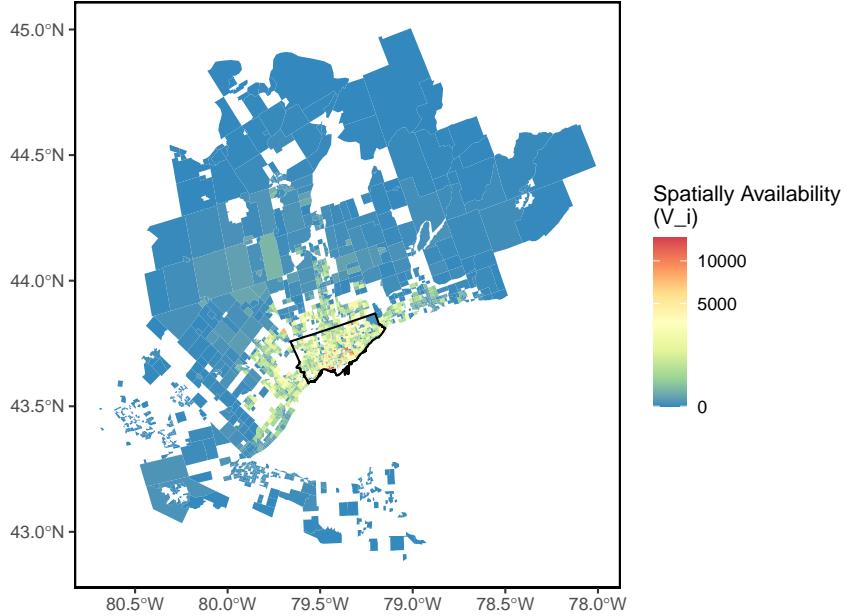


Figure 12: Calculated spatial availability of employment from origins in the GGH to destinations in the City of Toronto

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.

## Discussion

### *Comparing Job Access Measures*

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 14. The quantile ranks range from 1 to 5 where 1 represents the highest percentage difference between spatial availability and accessibility. In other words, the

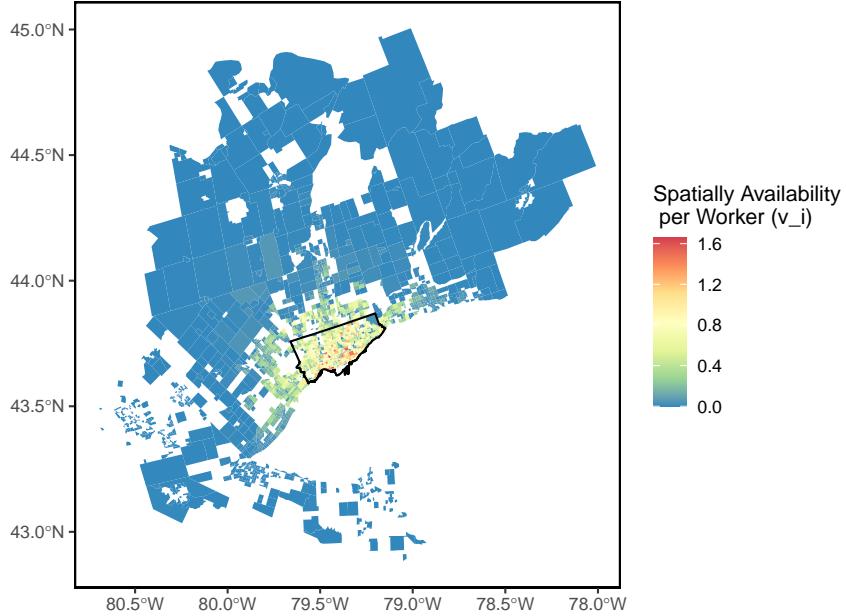


Figure 13: Calculated spatial availability of employment, per worker, from origins in the GGH to destinations in the City of Toronto.

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 rescaled measures are shown in grey.

We can see from Figure 14 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

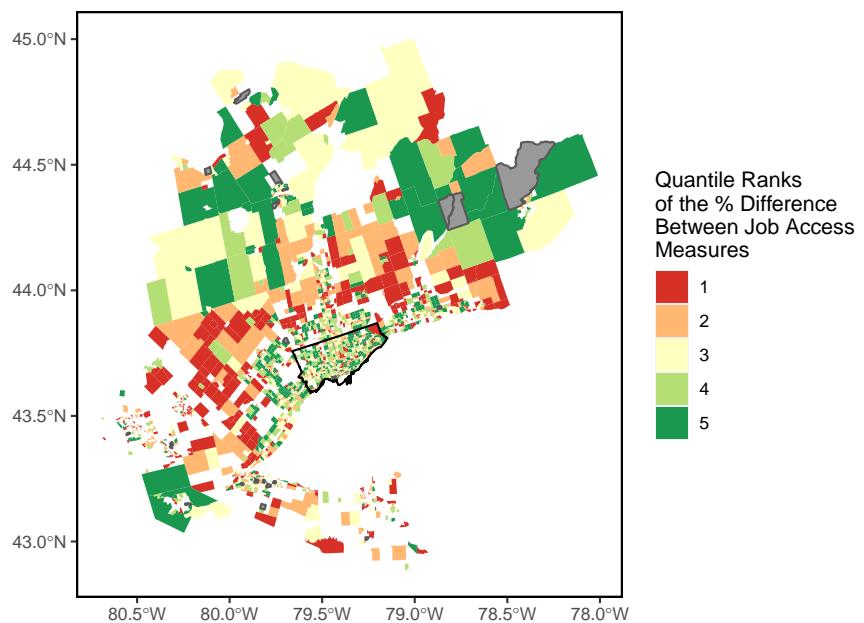


Figure 14: Difference in rescaled accessibility and spatial availability values in the context of employment from origins in the GGH to destinations in the City of Toronto. TAZs with rescaled spatial accessibility values that are lower than rescaled accessibility values are presented in five quantiles from 1 to 5 with 1 representing the highest negetative change between spatial availability and accessibility. TAZs with values which are higher than rescaled accessibility are shown in grey.

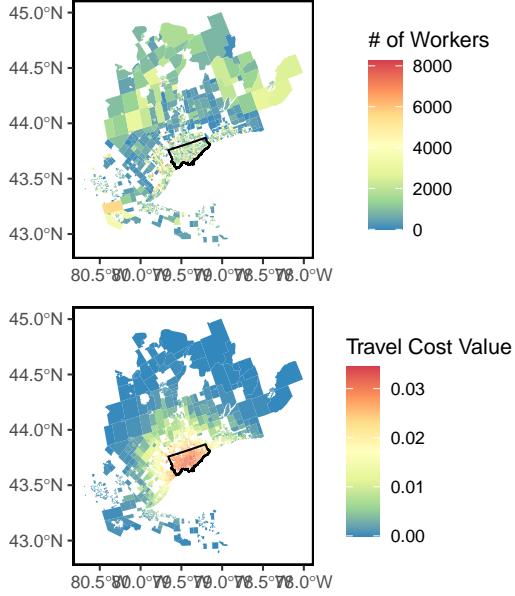


Figure 15: Number of workers (top) and average impedance function value (bottom) for each origin TAZ in the GGH.

for all TAZs are presented in Figure 15.

Accessibility when used to measure *job access* obscures the *availability* of jobs. Overall, this can be seen when comparing Figure 11 and Figure 12 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 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.

#### *Contextualizing Spatial Availability Use Cases*

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

sustible 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 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.

#### *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 for equity analysis.

### **Conclusion**

Words go here.

### **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
- 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