Planning with accessibility measures II: unconstrained and constrained accessibility

Anastasia Soukhov

*This blog post is the second part of a multi-part series. This series aims to walk readers through the potential uses of accessibility measures in transportation equity planning. This post explores different accessibility measures and how the selection of the measure and the impedance function impacts outputs. See the first* [*post*](**LINK%20TO%20POST1**) *for an introduction to accessibility measures and how assumptions about travel behavior enters accessibility measures through the impedance function.*

*Accessibility* (or *potential access*) has many definitions. Within the context of transportation planning, accessibility can be defined as a measure of the amount of interaction a “population” at an origin potentially has with “opportunities” at destinations in a given region. It is a product of the land-use and the population’s means of transportation.

In this post, we distinguish between two types of accessibility measures: *unconstrained* and *constrained* measures. The distinction is important to help accessibility analysts to more clearly interpret outputs.

First, let’s detail **unconstrained** accessibility. The general form of the unconstrained measure, let’s call it , is the measure proposed by (Hansen 1959). Many accessibility measures are derived and continue to be derived from this proposed formulation. is an accessibility value that is calculated for each spatial unit, and is appropriately termed *place-based accessibility*. This value is the summation of all the available (i.e., reachable) at each spatial unit according to some impedance function . is defined in [Equation 1](#eq-hansen-access):

Where:

* is a measure of the cost of moving between and .
* is an impedance function of ; it can take the form of any monotonically decreasing function chosen based on positive or normative criteria (Paez, Scott, and Morency 2012).
* is a set of origin locations in the region ().
* is a set of destination locations in the region().
* is the number of opportunities at location ; is the total supply of opportunities in the study region.

Many variations of have been proposed - but largely they focus on tweaks to the type of used. This measure counts (after being weighted by ) for each . This means that the score assigned to each is the summation of all the opportunities that can *potentially* be interacted with.

In lay terms, is a measure of the number of opportunities that someone at can potentially interact with, given their travel behaviour. However, counting all opportunities of potential interaction may not suit certain opportunities. Consider the following hypothetical example.

One can live in a part of the city where they have relatively high accessibility to jobs as a result of land-use (e.g., close to a commercial business district) and transportation options (e.g., great roads, excellent transit). Say is a value of “10,000 potential job opportunities” for a neighbourhood . Now, imagine if their adjacent neighbourhoods have 15,000 people who can also reach those same 10,000 job opportunities. Though they can potentially interact with a *relatively* high value of 10,000 opportunities, they may have less *available* opportunities as a result of relatively high neighbouring demand for opportunities. When compared to other areas of the city with relatively lower values but with a similar level of job opportunities and population demand, those other areas in the city may have more potential *spatial availability* than the neighbourhood where our hypothetical person lives.

This concept is formally known as *competition*, and has been applied within the influential accessibility works of Shen (1998) and Weibull (1976) as well as the widely used floating catchment areas methods (e.g., the two step floating catchment area (2SFCA) approach of Luo and Wang (2003)). We can think of these works as adjustments to (unconstrained accessibility) that account for the population’s demand for opportunities in the region of interest.

In a recently published journal article, an alternative derivation of competitive accessibility that *constrains* the results to match known quantities in the system is proposed (Soukhov et al. 2023). These known quantities can be the total number of opportunities or the total population in the region under analysis. Since we constrain one of those two (opportunities or population), we think of this measure as a *singly-constrained* competitive accessibility measure ().

In , the total number of opportunities in the study region are preserved. So if a urban region has 100,000 opportunities, at the end of the analysis, the sum of *all* values in the region is 100,000. How is this achieved? As a result of the proportional allocation feature: opportunities are allocated proportionally to each spatial units in the region based on the relative travel impedance and relative population density. We call this measure *spatial availability* which we denote as to distinguish it from unconstrained accessibility . Spatial availability’s mathematical formulation is defined in [Equation 2](#eq-spatial-avail) as follows:

where contains the opportunities, just as in the *unconstrained* accessibility measure , but the allocation depends on balancing factor . The sum of of in the region adds up to 1, which is how the sum of the spatial availability is equal to the sum of . Revising our hypothetical example of a urban region with 100,000 opportunities: it can be understood that both measures ( and ) are weighted sums of opportunities, but in the sum of all opportunities may be more or less than the 100,000. In contrast, thanks to the proportional allocation balancing factors in , the sum of values across the region is constrained to equal 100,00 opportunities.

For additional context, within :

* is a balancing factor defined the population balancing factor () and travel impedance balancing factor ()

The balancing factor corresponds to the proportion of the population in origin relative to the population in the region. On the right hand side of the equation, the numerator is the population in neighbourhood . The summation in the denominator is over , and adds up to the total population of the region under analysis. What does this mean practically? It means, neighbourhoods with a higher density of people get allocated more opportunities (i.e., a larger value), and less population dense neighborhoods get allocated smaller amounts. This measure is sensitive to demand: more people who are seeking opportunities get allocated more opportunities.

The second balancing factor, is the travel impedance balancing factor. It uses the impedance function (i.e., probability of travel given travel costs) to proportionally allocate more opportunities to neighbourhoods that are closer to (or contain) a higher density of opportunities. That is, this balancing factor assumes that populations within neighbourhoods that have lower travel impedance (less costly travel) to opportunities are *more willing* to take these opportunities, resulting in a higher value of for the neighbourhood. Indeed, the travel cost balancing factor can be thought of as the proportion of the population at an ‘origin’ neighbourhood willing to travel to a ‘destination’ neighbourhood , conditional on the travel behavior as described by the impedance function. What does this mean practically? It means, the higher the for a neighbourhood, the more opportunities get allocated to this neighbourhood than other neighbourhoods.

Overall, and can be complex to understand - but the outputs may clarify their intuition!

# Differences in unconstrained and constrained accessibility

For this demonstration, **unconstrained** and **constrained** accessibility are calculated using data taken from the R data package {TTS0216R}. This package contains a subset of home to work (full-time) flows from the 2016 Transportation Tomorrow Survey (TTS) as well as travel time by car calculated using {r5r}. {TTS2016R} is detailed in this publication (Soukhov and Páez 2023) and is freely available to be explored [here](https://soukhova.github.io/TTS2016R/). The focus of this demonstration is the City of Hamilton, Canada, a city approximately 70 km south-west of Toronto, and within the TTS survey area.

A calibrated gamma distribution probability density function serves as the impedance function for the analysis and is shown in [Figure 1](#fig-gamma) ( (shape) is 3 and (rate) is 0.2). The data set and parameters were fit using the empirical data and discussed in the first blog [post](**LINK%20TO%20POST1**). As a refresher, a gamma distribution form was selected as it best fits the sample of home to full-time work trips beginning and ending in the City of Hamilton. The values along the y-axis can be interpreted as the probability density of a trip at a certain travel time of occurring i.e., trips of length 9 minutes are the most likely to occur and hence are assigned the highest relative value.

|  |
| --- |
| Figure 1: The fitted theoretical gamma distribution travel impedance function of home-to-work trips (in estimated minutes by car) for the City of Hamilton. |

The {accessibility} package is used to conveniently calculate unconstrained accessibility ([Equation 1](#eq-hansen-access)) and singly-constrained competitive accessibility ([Equation 2](#eq-spatial-avail)). The resulting (Purples) and (Greens) are shown in [Figure 2](#fig-raw-con-and-unconstrained-access) below:

|  |
| --- |
| Figure 2: The calculate unconstrained (Si) and constrained (Vi) accessibility (a.k.a. spatial availability) scores for the City of Hamilton. Calculated using an empirically calibrated gamma distribution travel impedance function. |

In [Figure 2](#fig-raw-con-and-unconstrained-access), the raw and scores are presented. Both measures reflect *potential interaction* with jobs based on empirical home-to-work travel behaviour in Hamilton: people who reside in each spatial unit make certain trips to other spatial units (in Hamilton) and these trips have a certain travel time (travel cost ) with associated gamma impedance value. These observed trip patterns inform the calculated and values for each spatial unit. What’s notable is the difference in the **magnitude** and the **interpretation** of unconstrained () and constrained () values, let’s discuss.

Looking at the left plot in [Figure 2](#fig-raw-con-and-unconstrained-access), the values reflect the sum of jobs that can be potentially interacted with by the population at each multiplied by their probability of being reached as informed by the calculated travel impedance value. The maximum value is the darkest purple, and that value means that people who reside in those spatial units have the *lowest* travel impedance and *highest* concentration of potential job interaction for the region. The value itself does not have a specific meaning as it is just the sum of ‘weighted’ jobs: it can be interpreted as a relative score of potential interaction based on the observed trip patterns of people who reside in the City. For instance, areas within the centre of Hamilton have the highest values, this is both where jobs are largely clustered as well as major roadways (highways are pictured) and denser street networks.

Now looking at the right plot (constrained) in [Figure 2](#fig-raw-con-and-unconstrained-access), the values reflect the sum of the proportionally allocated (based on travel impedance and population) *potential job interactions*. In other words, each value of is the number of jobs that the spatial unit can interact with based on the observed trips made from that spatial unit’s impedance values relative to how others in the region can interact with the jobs. Unlike , the raw values of do have a meaning in addition to being a relative score of *competitive* potential job interactions. This score reflects the potential *availability* of jobs: potential job interactions are less likely to occur if the concentration of jobs is low and the density of people interacting with those jobs are high. So as we can observe in the plot, considers population demand, and as the centre of Hamilton is the most densely populated area in the city, the centre does not share the same intensity of trend as .

The value in itself is also meaningful. values communicates the number of potentially *available* job interactions per each spatial unit out of all the jobs in the region. If all values are added together, the sum equals 108,526 - the total number of jobs taken by people who live and work (full-time) within the City from the data set. So in the most green spatial unit, 3,160 potentially available job interactions can occur out of the total 108,526 jobs. Again, is produced through proportional allocation, the total number of jobs is divided up and assigned to each spatial unit based on the impedance to reach jobs and the population who also interact with these potential opportunities.

To more equally compare and and make sense of the ‘highs’ and ‘lows’, it may be useful to standardise the values onto a similar scale. In [Figure 3](#fig-perc-con-and-unconstrained-access), the values as presented as a percentage of the regional sum (i.e., a % of the sum of all values and the % of the sum of all values) are visualized for the centre of Hamilton:

|  |
| --- |
| Figure 3: The scaled unconstrained and constrained accessibility score for Hamilton Centre. Values are presented as a percentage of the total sum of scores within the region. Major highways are shown in purple for spatial reference. Values are calculated using an empirically calibrated gamma distribution travel impedance function. |

Let’s examine (left plot) in [Figure 3](#fig-perc-con-and-unconstrained-access): unconstrained accessibility. Neighbourhoods with ‘high’ accessibility (e.g., greens, that start at 1.3% relative values or higher) can potentially interact with 1417 jobs or more as informed by observed travel behaviour. These raw values are difficult to interpret, so seeing a neighbourhood as being an area of relative ‘high’, ‘medium’ (yellows) or ‘low’ (reds) accessibility value simplifies the interpretation of ‘potential interaction’ with jobs; as long as we ignore *competition*.

(the plot on the right side in [Figure 3](#fig-perc-con-and-unconstrained-access)) visuals singly-constrained competitive accessibility (spatial availability). This measure does not ignore competition for potential job interaction. The general trend between both plots are similar, but a handful of spatial units that are more intensely green or red/orange can be seen. These differences are a result of competition. Within this region, spatial units that are more densely populated as well as having below average travel impedance have higher standardized values than values. Conversely, below average population and above average travel impedance yields spatial availability values that are lower than .

In essence, reflects travel impedance like does, but it also considers competition. Spatial units with orange/green that have red are a cause for concern: they likely have low travel impedance but high competition that makes their relatively low. Conversely, spatial units with orange/red that have green have high travel impedance but low competition for their opportunities so their spatial availability of jobs may in fact be alright. Spatial availability adds an additional layer of consideration into the accessibility measure, and as such, reveals more about the region (under the travel behaviour and opportunity accessed assumptions).

Across both and in [Figure 3](#fig-perc-con-and-unconstrained-access), we can see some common low values (red) located in the north end of the city. From unconstrained accessibility, we know these TAZ have high relative travel impedance - this may be because people who work in the north end do not live relatively close to these opportunities so have high relative travel times. Interestingly though, we can confirm that there is a high relative number of jobs within these TAZ (see [Figure 4](#fig-worker-job-plot) below), however, even the number of jobs does not balance the impedance value and higher demand for those jobs. Hence, the constrained accessibility measure is also low.

|  |
| --- |
| Figure 4: The number of workers and jobs in Hamilton Center. Note: only workers who reside and work within Hamilton Center are considered in the accessibility calculations for this demonstration. |

## Concluding remarks

Accessibility is a unique measure that characterises the relationship between land-use (where populations reside and the opportunities they can interact with) in addition to transportation travel impedance. How the relationship is conceptualization (i.e., if there’s competition or not) and how the travel impedance is calibrated (i.e., what function describes travel behaviour) are critical in determining what the final values are and how to interpret them.

In this post, we outlined two branches of accessibility measures: unconstrained ( in this case) and constrained (spatial availability ). Unconstrained accessibility only considers opportunities that could be interacted with while constrained considers *both* those opportunities and the demand for those opportunities in addition to having that property of proportional allocation. This property allows the raw values of to be interpreted without any sort of transformation or standardizing; is simply the number of opportunities that can be potentially interacted with out of *all* opportunities in the region.

provides insights that does not. Firstly, it considers considers competition from demand. Secondly, it does not need to be standardized to be understood. These two insights are important:

* **Considering competition**: places of employment are a *non-divisible* type of opportunity, they only allow one person to take one job. Unless there’s a reason to *not* consider competition, measuring access to opportunities that have some capacity using an unconstrained measure does not make much theoretical sense from an opportunity-planning perspective; this will be explored in subsequent post(s).
* **Interpretation**: a spatial unit has a certain number of , opportunities that are spatially available for interaction. We can tangibly interpret if that’s high or low (out of the total number of opportunities). Further, we can also divide by population at that origin to obtain opportunity per capita values. This value can be used as a benchmark to compare opportunity per capita or levels of service across areas of the region, between regions, and/or across time; again, this will be explored in subsequent post(s).

Ultimately, unconstrained accessibility tells you how many opportunities can be potentially reached. Spatial availability tells you how many opportunities are *available* based on how many can be potentially reached and demanded.

Accessibility analysis sheds light on regions of inequitable potential access. Assumptions on what region to analyse, what population and opportunities are the subject of analysis, the travel cost unit and calculation, the impedance function and the measure used all impact the final results. But ultimately, the output represents the number of opportunities that could *potentially* be reached from each origin. It is critical that the assumptions embedded within each step of analysis are understood so that the final value can be interpreted and inequities be identified.

Once these spatial inequities have been identified - what do we do about it? That is the subject for a future post(s).

Openness is legitimacy: this blog post was written in a R environment and can be fully reproduced from the materials available at this GitHub ([repository](https://github.com/soukhova/MJ-Accessibility-Blogs)). If interested, see the open access PDF of the full article (which includes the mathematical formulation for the spatial availability function) in the references (Soukhov et al. 2023).

*The data used in this post is a subset of data from {TTS2016R}, the plots are created using {tmap}, and spatial objects are manipulated using {sf}, along with base {R} functions.*

## References

Hansen, Walter G. 1959. “How Accessibility Shapes Land Use.” *Journal of the American Institute of Planners* 25 (2): 73–76. <https://doi.org/10.1080/01944365908978307>.

Luo, Wei, and Fahui Wang. 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 (6): 865–84. <https://doi.org/10.1068/b29120>.

Paez, A., D. M. Scott, and C. Morency. 2012. “Measuring Accessibility: Positive and Normative Implementations of Various Accessibility Indicators.” Journal Article. *Journal of Transport Geography* 25: 141–53. <https://doi.org/10.1016/j.jtrangeo.2012.03.016>.

Shen, Q. 1998. “Location Characteristics of Inner-City Neighborhoods and Employment Accessibility of Low-Wage Workers.” *Environment and Planning B: Planning and Design* 25 (3): 345–65. <https://doi.org/10.1068/b250345>.

Soukhov, Anastasia, Antonio Paez, Christopher D. Higgins, and Moataz Mohamed. 2023. “Introducing Spatial Availability, a Singly-Constrained Measure of Competitive Accessibility | PLOS ONE.” *PLOS ONE*, 1–30. https://doi.org/https:// doi.org/10.1371/journal.pone.0278468.

Soukhov, Anastasia, and Antonio Páez. 2023. “TTS2016R: A Data Set to Study Population and Employment Patterns from the 2016 Transportation Tomorrow Survey in the Greater Golden Horseshoe Area, Ontario, Canada.” *Environment and Planning B: Urban Analytics and City Science*, January, 23998083221146781. <https://doi.org/10.1177/23998083221146781>.

Weibull, Jörgen W. 1976. “An Axiomatic Approach to the Measurement of Accessibility.” *Regional Science and Urban Economics* 6 (4): 357–79. <https://doi.org/10.1016/0166-0462(76)90031-4>.