Modal shift potential of trips to bicycle modal
Case study of São Paulo, Brazil

Pedro Gigeck Freire
Supervisors: Fabio Kon, Higor Amario de Souza

June, 2021

Related Work

Research on modal shift to cycling has grown considerably in the past decade. Much of this work focuses on the effects of bike-sharing systems (BSS) implementation (Fishman et al., 2014; Shaheen et al., 2013; Ma et al., 2020). These systems have a significant impact on changing the modal share of the region they are implemented at, arriving at 20% of trips that were made by car migrating to bicycle (Fishman et al., 2014).

Another stream of research, pointed by Larsen et al. (2013), aimed to understand what are the policies with better response on cycling increase, such as segregated lanes for bicycles (bike lanes), marketing campaigns, cyclist education, and driver training. Besides, the great relevance of cycling infrastructure to modal shift influenced the development of literature focusing on technical aspects of these structures, for example, what are the best material, signaling, lanes width, etc (AASHTO, 2012).

While literature established homogeneous conclusions on *what* are the policies necessary to modal shift and *how* to implement them, there is limited research on *where* to build them. (Larsen et al., 2013; Silva et al., 2019). Recently, some methods have been proposed to indicate which regions in a city (districts, neighborhoods) have greater *cycling potential*, i.e, which are the places where these policies have better efficiency, and what are the attributes influencing this potential.

The studied methods for calculating cycling potential, also called cycling potential indexes, were developed by both academic papers and governmental organizations, and vary according to the available data of the studied region. Lovelace et al. (2017) categorize the measures considered on each index into three general classes: individual-based measures (demographic data), area-based measures (environmental data, such as infrastructure), and route-based measures (trips data, obtained from origin-destination surveys). In this section, we list some of these methodologies, presenting what kind of information is considered and how this data is aggregated in each one.

1. Copenhagenize Index (Copenhagenize Design Company, 2011)
One of the most established tools to assess the cycling environment of a city is the Copenhagenize Index (Copenhagenize Design Company, 2011; Silva et al., 2019). Copenhagenize aims to rate the most bicycle-friendly cities in the world,

updating the rank annually. It considers variables such as the existence and quality of cycling infrastructure, traffic calming, growing on bicycle share in the past 10 years, and others.

Although this method process important data about a region bicycle conditions, it is only accessible on a macro scale, considering whole metropolitan areas and not local neighborhoods. In addition, the Copenhagenize Index is not scalable, since each city needs deep data extraction and technical adaptation.

2. Prioritization Index (Larsen et al., 2013) The second approach to estimate where the investments on cycling infrastructure should be prioritized was proposed by Larsen et al. (2013). This method consists of a grid-cell model on the studied region, the grid-cell size of 300 meters was chosen after empirical tests.

For each cell, a value from 0 to 1, called the prioritization index, is calculated. The higher the prioritization index, the higher the priority a grid cell is ascribed in terms of the addition of cycling infrastructure. The authors applied that method in Montreal Island, Canada, as a study case.

Five attributes are considered in this index: the number of observed cycling trips; the number of potential cycling trips; the number of collisions with cars; the connectivity of cycling infrastructure; and the opinion of cyclists about where the investments should be prioritized.

The observed cycling trips are those that already occur in the city, while the *potential* cycling trips are those currently made by car with distance smaller than 75% of the observed cycling trips, which corresponded to trips shorter than 2km in the Montreal study case. The authors cite that this criteria for the potential trips could be improved to consider other parameters besides the distance, such as sociodemographic variables. The connectivity of the cycling infrastructure is calculated by counting the "dangling nodes", i.e., the places where the bicycle lanes end abruptly. In the studied region, the last attribute was obtained from an online survey with 3000 cyclists, where they were asked to tell which were the roads that should receive cycling infrastructure.

The methodology to aggregate these data for each grid cell i consists in summing the quantity of the attributes percentage in that cell. For example, if a cell contains 5% of the observed cycling trips and 8% of the collisions (and 0 in all other attributes) then it would receive a value of $x_i = 0.05 + 0.08 = 0.13$. Finally, this value is divided by the sum of values of all cells $\sum_j x_j$, getting a "priority percentage" of this cell concerning to the whole region.

3. Willingness Index (Zhang et al., 2014) Zhang et al. (2014) proposed another method to define where the investments in cycling infrastructure should be prioritized. The method consists in identifying the regions of the studied city where people are most willing to use a bicycle. The studied case was the city of Belo Horizonte (BH), Brazil, in the context of cycling infrastructure expansion between 2010 and 2014.

The considered variables were obtained from a survey with 2008 full answers from citizens of BH. The interviewees were asked what were their willingness to use bicycle if there were appropriate bicycle paths in their daily trips routes. In addition, the survey collected the following "influential factors": educational level; commuting time; monthly income; if the respondent lives in a house or not (in an apartment for example); house ownership; car ownership; current mode of transport; if there are kids in the household; and cycling network density in the respondent's daily trips.

After collecting the data, an ordered outcome discrete model was applied to quantify the effects of each influential factor on the interest variable (willingness to cycle). With this model, a coefficient was calculated for each influential attribute, the higher this coefficient is, the most relevant this attribute is to influence people to use bicycle. Inversely, the smaller this coefficient, the stronger this attribute is in influencing people to not use bicycle.

The most strong correlations were found in the attributes of commuting time (higher commuting time has negative influence), monthly income (the group with medium household income has a great willingness to cycle), and current mode (walk mode has the strong positive influence). These conclusions, although, represent this city specific context. It is known that the relations between bicycle use and sociodemographic factors are country-specific and unlikely causal (Parkin et al., 2007). So it is necessary to collect local data in order to apply this method in other cities (conducting a similar survey).

The coefficients obtained from the model were then used to calculate the willingness index on each administrative region, multiplying the value of each attribute in the region by the respective coefficient, using census and OD survey data.

4. Steer Cycling Potential Index (Steer, 2015) Steer is a consultancy that develops tools for urban complex problems. The Cycling Potential Index (Steer, 2015) was developed to rank regions in England and Wales in terms of cycling investments effectiveness.

This index considers three dimensions: hilliness, sociodemographics, and trip length. The hilliness is calculated as the standard deviation of the heights in the region, using a regular grid of 90m resolution. The sociodemographic dimension is conceived by classifying the predominant lifestyle of the region, considering 23 sociodemographic attributes. Each lifestyle class has a precalculated cycling potential value. Finally, the trip length dimension aims to consider trips that could be migrated to cycling mode. The data were obtained from local OD surveys and the limit length established was 8km, although the volume of actual bicycle trips longer than 5km was small.

For each dimension, the regions are ranked, and the overall rank is the average of the three ranks, though with the trip-length dimension having half of the weight of the other two, because the considered trips were trips for work (commutation) but the index aimed to reach other purpose journeys, that could have greater length.

5. Analysis of Cycling Potential (Transport for London, 2017) Analysis of Cycling Potential (ACP) is a tool developed by Transport for London to identify cyclable trips in London. The trips data comes from the London Travel Demand Survey from 2012 to 2020.

In ACP, a trip is considered cyclable if it satisfies the following four criteria: the person making the trip is carrying no encumbrance; the trip length is no longer than 3km, if the person's age is greater than 80, 5km if the person's age is between 65 and 79, or 10km if the person's age is between 5 and 64; the current mode is cyclable (is not boat or plane, for example); and the trip is not part of a wider chain of trips that can not be cycled in its entirety.

With these filters, Transport for London calculated the number of potential trips in each London sub-region and analyzed these trips' main characteristics (location, purpose, individual age and income, etc).

6. Propensity to Cycle Tool (Lovelace et al., 2017) The Propensity to Cycle Tool (PCT) is an open source planning support system that calculates how many trips could be cycled in a given region. The study case was administrative regions from England and Wales and the data came from the UK 2011 Census.

Each administrative region was divided into smaller zones with the same amount of residents (around 7500 individuals each). Then, the census data provided the proportion of bicycle trips between each pair of these smaller zones. Then, a logistic regression model was used to discover trip distance and slope influence in the proportion of cycling.

The logistic model considered the distance, distance squared, and the distance square root to capture non-linear influences. The slope measure was calculated as the average slope along the three shortest routes between each zone.

The final "propensity to cycle" P_i of a region i is given by the formula $c_i + t_i * p_i$, where c_i is the current number of cyclists, t_i is the total number of trips and $p_i \in [0, 1]$ is the proportion of bicycle use calculated with the logistic model coefficients. P_i measures the number of cyclists that the region would have in a future based on the geographical characteristics of its trips. PCT also generates other future scenarios with some independent assumptions, such as the assumption of women and men cycle in the same proportion, or the assumption that British people cycle as much as Dutch people.

7. Cycling Potential Tool (Phillips and Range, 2017) Aiming to identify Scotland areas where cycling investments would bring higher benefits, the governmental organization Cycling Scotland developed the Cycling Potential Tool (CPT).

CPT is used on a local scale, dividing each studied city into a grid of 1 meter sized cells. For each cell, 8 attributes are considered: the existence of physical barriers; population density; slope; distance to cycling infrastructure; average road speed; average distance to work or school; cyclists proportion; and access to services (hospitals, schools, etc). The data was provided by local authorities and geographic information systems (GIS).

For each attribute, the range of values is reclassified to a common scale from 1 to 10. Finally, the cycling potential of a cell is the average between all 8 attributes, where the final value 1 is considered to have low cycling potential and 10 representing the greatest potential.

8. Potential for Cycling Assessment Method (Silva et al., 2019) Focusing on starter cycling cities, with lack of cycling tradition and data about cyclists profile, Silva et al. (2019) reviewed the exiting work on what variables affect bicycle usage to develop the Potential for Cycling Assessment Method.

This method does not consider trips data, but a list of more than 20 variables to evaluate 3 dimensions of a given point in the city: target population, target area, and political commitment. The first dimension considers factors such as average age, population density, educational level, and car ownership. The target area dimension aims to identify geographic attributes that benefit cycling, including streets inclination, distance to education facilities, etc. The last dimension evaluates the existence of local policies to stimulate cycling, such as bicycle parking, integration with public transport, and cycling infrastructure coverage.

Each variable receives a weight based on the relevance of that factor in cycling choice. These weights were defined after an extensive literature review on relevant attributes for cycling. Then, for each attribute, the region is classified from 1 to 5 (most cyclable to least cyclable). This classification is tabulated and generally considers a function of the city average value.

The Assessment Method was applied in 8 Portuguese cities. The generated maps and cycling potential values are being used by authorities to direct investments on cycling.

9. Increasing Cycling in Canada (Verlinden et al., 2019) Verlinden et al. (2019) developed a guide reporting different strategies effectiveness to encourage cycling in large cities. Cycling Potential analysis to identify regions with bigger bicycle demand is one of the suggested policies. The guide does not provide a precise method to calculate a cycling potential, but suggests important attributes to be considered in this analysis.

The elements cited as relevant in cycling potential include: short trips proportion (smaller than 8km); average number of trips per person; average number of vehicles per person; cycling infrastructure density; current cycling mode share in adjacent neighborhoods; weather (number of snow days); and hilliness (maximum change of elevation). All these elements are not aggregated together, for instance, in a study case of Toronto, the specialists mapped the short trips and the cycling mode share in the city and identified separately higher potential regions.

10. Data Science Framework (Olmos et al., 2020) Recently, Bogota received massive investments to support cycling (Olmos et al., 2020). Using data from bike apps, Olmos et al. developed a framework to identify trips that could be supported by bicycle paths.

The methodology is based on percolation theory and considers two attributes: the trips distance and the bicycle infrastructure connectivity. The distances are calculated with the route generated by the Open Street Routing Machine, using a shortest path algorithm, prioritizing bicycle lanes. The route inclination was not considered since Bogota is generally flat. Then, the bike apps distances distribution was used as weights in a percolation model to consider bicycle infrastructure connectivity. Thus, a trip is assigned with greater potential if it is short or if its route connects isolated clusters of bike lanes.

11. Nodal Approach (Hitge and Joubert, 2021) Hitge and Joubert (2021) investigate the Cape Town cycling potential by filtering the population that corresponds the most with the cyclists profile around the world. This cyclists' profile is defined after a literature review on the individual factors that influence choosing bicycle as vehicle.

In this filter, a person is a potential cyclist if they are currently studying or if they belong to low-middle income economic classes. Also, a potential cyclist does not live in an informal household and does not have access to cars. The authors ponder that this is a conservative approach to consider only the most probable people that would migrate to cycle.

After the initial filter, the person receives a weight considering the gender, age, and the number of family members. This weight is used as a probability. For example, if the individual gets a 0.7 weight, it means that he/she is gonna be considered a potential cyclist with a chance of 70%.

This individual approach shows where the nodes (regions) are concentrated and would better benefit from investments in cycling mobility.

Methods Relation

Most of the studied methods to calculate cycling potential consider some common attributes, even in different detail levels, such as trips distance and inclination, for trip-based approaches, or age and car ownership, for individual-based indexes, or cycling infrastructure, for geographical proposals. Table 1 sum up all variables somehow used by the different methodologies.

Table 1: Attributes considered

attribute \ method	01	02	03	04	05	06	07	08	09	10	11
Age					x			x			x
Income			x								x
Gender	x										x
Car ownership			x					x	x		x
Education			x					x			x
House style			x								x
Family members			x								x
Number of trips									x		
Life Style				x							
Trip Distance/Time		x	x	x	x	x	x	x	x	x	
Encumbrance					x						
Vehicle modal			x		x						
Average speed	x						x	x			
Inclination				x		x	x	x			
Physical barriers							x				
Cyclists proportion	x	x				x	x		x		
Cycling infrastructure	x	x	x				x	x	x	x	
Cyclists opinion	x										
Traffic accidents	x	x						x			
Demographic density							x	x			
Occupational diversity								x			
Public transport								x			
Interest places (e.g. schools, hospitals)							x	x			

References

AASHTO (2012), ‘Guide for the development of bicycle facilities’.

Copenhagenize Design Company (2011), ‘Our methodology’.

Fishman, E., Washington, S. and Haworth, N. (2014), ‘Bike share’s impact on car use: Evidence from the united states, great britain, and australia’, *Transportation Research Part D: Transport and Environment* **31**.

Hitge, G. and Joubert, J. W. (2021), ‘A nodal approach for estimating potential cycling demand’, *Journal of Transport Geography* **90**.

Larsen, J., Patterson, Z. and El-Geneidy, A. (2013), ‘Build it. but where? the use of geographic information systems in identifying locations for new cycling infrastructure’, *International Journal of Sustainable Transportation* **7**.

Lovelace, R., Goodman, A., Aldred, R., Berkoff, N., Abbas, A. and Woodcock, J. (2017), ‘The propensity to cycle tool: An open source online system for sustainable transport planning’, *Journal of Transport and Land Use* **10**.

Ma, X., Yuan, Y., Oort, N. V. and Hoogendoorn, S. (2020), ‘Bike-sharing systems’ impact on modal shift: A case study in delft, the netherlands’, *Journal of Cleaner Production* **259**.

Olmos, L. E., Tadeo, M. S., Vlachogiannis, D., Alhasoun, F., Alegre, X. E., Ochoa, C., Targa, F. and González, M. C. (2020), ‘A data science framework for planning the growth of bicycle infrastructures’, *Transportation Research Part C: Emerging Technologies* **115**.

Parkin, J., Wardman, M. and Page, M. (2007), ‘Estimation of the determinants of bicycle mode share for the journey to work using census data’, *Transportation* **35**.

Phillips, L. and Range, A. (2017), Development of a gis based toolbox for mapping cycling potential across scotland.

Shaheen, S., Martin, E. and Cohen, A. (2013), ‘Public bikesharing and modal shift behavior: A comparative study of early bikesharing systems in north america’, *International Journal of Transportation* **1**.

Silva, C., Teixeira, J. and Proença, A. (2019), ‘Revealing the cycling potential of starter cycling cities’, *Transportation Research Procedia* **41**.

Steer (2015), ‘Cycling potential index’.
URL: *https://www.yumpu.com/en/document/view/41572043/1-cycling-potential-index-steer-davies-gleave*

Transport for London (2017), ‘Analysis of cycling potential 2016’.

Verlinden, Y., Manaugh, K., Savan, B., Lea, N. S., Tomalty, R. and Winters, M. (2019), ‘Increasing cycling in canada: A guide to what works’.

Zhang, D., Magalhães, D. J. A. V. and Wang, X. C. (2014), ‘Prioritizing bicycle paths in belo horizonte city, brazil: Analysis based on user preferences and willingness considering individual heterogeneity’, *Transportation Research Part A: Policy and Practice* **67**.