Transportation Research Part A 179 (2024) 103916

Contents lists available at ScienceDirect

Transportation Research Part A

journal homepage: www.elsevier.com/locate/tra

A novel route-based accessibility measure and its association with transit ridership

João Lucas Albuquerque-Oliveira [a,*], Francisco Moraes Oliveira-Neto [a], Rafael H. M. Pereira [b]

[a] Postgraduate Program in Transportation Engineering, Department of Transportation Engineering, Federal University of Ceará, UFC, Campus do Pici, Bloco 703, Fortaleza, Ceara Zip 60440-900, Brazil
[b] Institute for Applied Economic Research - Ipea, Brazil

ARTICLE INFO

Keywords:
Transit accessibility
Route-level
Transit ridership
Determinants of ridership
Directed acyclic graph
Transit network planning

ABSTRACT

Transit systems play a key role in improving access to job opportunities and basic services such as health and education. Most studies in the literature calculate transit accessibility using traditional place-based indicators that measure accessibility at a given location. However, because transit routes are the main unit of analysis in most approaches for planning and operation of transit systems, these accessibility indicators provide limited information to inform transport planning at the route-level. In addition, previous studies have demonstrated the methodological limitations of traditional place-based accessibility metrics to study the association between accessibility and transit ridership. In this paper, we propose a novel route-based accessibility measure to fill the mentioned gaps. The indicator measures the average level of access to opportunities provided by a given transit route to the population in its extended catchment area. This indicator is flexible enough that it can be calculated using different travel cost functions and can be applied to measure access to different activity types for the whole population or for specific groups. To illustrate the applicability of the proposed indicator, we calculated the employment accessibility provided by all routes of the transit system of Fortaleza, Brazil. We also show that the proposed indicator has greater predictive power of transit ridership than other route-level accessibility measures found in the literature. This paper provides a methodological contribution that could help transport planners incorporate accessibility analysis into transit system redesign projects, and help practitioners anticipate what accessibility impacts and subsequent changes in transit ridership could be expected from route-level service changes, and to examine the influence of accessibility on transit ridership.

1. Introduction

There is a growing consensus in the transportation planning literature that one of the main benefits of public transit systems is the promotion of accessibility, i.e. the ease with which people can reach activities or destinations (Bertolini et al., 2005; Geurs and van Wee, 2004). As a result, an increasing number of researchers, transportation and funding agencies have been using accessibility metrics

* Corresponding author.
E-mail addresses: joao@det.ufc.br (J. Lucas Albuquerque-Oliveira), moraes@det.ufc.br (F. Moraes Oliveira-Neto), rafael.pereira@ipea.gov.br (R.H.M. Pereira).

https://doi.org/10.1016/j.tra.2023.103916
Received 26 April 2023; Received in revised form 20 September 2023; Accepted 23 November 2023
0965-8564/

to evaluate the performance of transportation systems and their integration with land use policies (Boisjoly and El-Geneidy, 2017; Manaugh et al., 2015; Silva et al., 2016). There is also growing evidence on the association between accessibility and levels of mobility and/or the participation in out-of-home activities (Diab et al., 2020b; Fransen et al., 2018; Martens, 2019). However, practitioners still face persistent challenges in incorporating accessibility analyses to transit operations planning (Silva et al., 2016; Silva and Larsson, 2018). El-Geneidy and Levinson (2022) argue that transport planners could more easily adopt accessibility analyses if such analyses would speak more directly to the goals of transit systems. Recent studies have investigated the association between accessibility and transit ridership (Bree et al., 2020; Cui et al., 2022; Merlin et al., 2021), which is a key indicator of public transit performance (Cui et al., 2022). However, these studies present some important methodological limitations (see Section 2). One of these limitations is that these studies use place-based accessibility indicators in ways that cannot distinguish the accessibility benefits of each transit route. Ultimately, such limitations preclude one's ability to understand how accessibility levels and subsequent changes in transit ridership could result from the creation of new route alternatives or changes to the transit network.

In this work we propose a new route-based accessibility measure and examine how this metric can help us understand the association between transit accessibility and ridership. The proposed metric captures the average accessibility that a given route provides to the population in its area of influence, taking into account both land use patterns and the spatial–temporal connectivity of that route with the rest of the transit network. This route-based accessibility measure can be calculated using different types of cost functions (e.g. cumulative or gravity models) and can be applied to measure access to different types of activities (e.g. jobs, schools, etc.) for the population as a whole or for specific groups (e.g. income or ethnic groups). To demonstrate the new route-based accessibility measure, we calculate the level of access to jobs conveyed by each of the bus routes in the transit system of Fortaleza (Brazil) using both cumulative and gravity functions. We also compared the proposed metric against other route-based accessibility measures found in the literature, and showed how our proposed indicator has a stronger association with transit ridership. This comparison was carried out using correlation and multivariate regression analyses. The selection of control variables of the regression model was informed by a Directed Acyclic Graph (DAG) that we propose to map the determinants of transit ridership based on our literature review. This strategy allows us to estimate the effect of accessibility on ridership without eventual biases that could be caused by confounding variables.

To the best of our knowledge, only two studies have previously proposed route-based accessibility measures. Karner (2018) proposed to calculate a transit route accessibility by first calculating a gravity-based accessibility from all the stops of the transit system, and then calculating the level of accessibility for each route by taking the average accessibility of its stops. Similarly, Diab et al. (2020a) proposed to calculate the accessibility of each route by taking the average accessibility of the census tracts that are intersected by each route. Meanwhile, a few studies have examined the marginal accessibility gains from particular transit routes by comparing how accessibility levels change with and without the route of interest being considered in the transit network (Farber and Fu, 2017; Liu and Miller, 2023; Niehaus et al., 2016; Pereira et al., 2019). As it will be discussed in Section 2, these approaches underestimate the contribution of sub-optimal routes and do not necessarily capture how the accessibility of a route would be affected by changes in its frequency or operational speed, for example.

This paper advances the literature by proposing a new route-based accessibility measure that simultaneously takes into account transit supply factors (e.g. speed, frequency and connectivity), land use factors (how activities and people are distributed in space), and time constraints of individuals' travel time budgets. The proposed indicator can capture the average accessibility that a transit route provides to the population located at the immediate walking catchment area of the route, as well as to the population that could indirectly benefit from it given the spatial–temporal connectivity of the route with the rest of the transit system. This work also advances the literature by showing how the proposed route-based accessibility metric can advance the study of the association between accessibility and transit ridership, and thus inform the operational planning of transit systems. Finally, this work also presents the first DAG mapping the causal paths of the key variables that influence transit ridership at the route level.

The remainder of the paper is organized as follows: Section 2 presents a literature review of studies that analyze the association between accessibility and transit ridership, and two other route-based accessibility metrics found in the literature; Section 3 introduces the new route-based accessibility measure proposed in this paper; Section 4 introduces the case study; Section 5 discusses the materials and methods used; Section 6 presents the results and discussion from a case study and; Section 7 presents the conclusions.

2. Literature review

Accessibility measures have increasingly been used by researchers and funding agencies to evaluate the performance of the transportation system and land use (Boisjoly and El-Geneidy, 2017; Manaugh et al., 2015; Silva et al., 2016). However, their application in day-do-day practice is still incipient, especially among managers and operators of transit agencies. According to Silva et al. (2016), there are various constraints, such as data availability and the separation of urban planning and transportation agencies. Conforming to El-Geneidy and Levinson (2022), one way to improve the adoption of accessibility analysis in planning practice would be to devise accessibility methods that could more directly help decision makers achieve their desired objectives.

In the case of public transit, a key goal of policy makers is to increase ridership. Increasing transit ridership is associated with more sustainable mobility patterns and lower environmental externalities (Adler and van Ommeren, 2016; Beaudoin et al., 2015; Diab et al., 2020b). Transit ridership can also be understood as an important metric that reflects how much people value a transit system or the direct utility people derive from it, in other words how successful it is (Merlin et al., 2021).

Recent studies suggest that accessibility estimates can improve the performance of models that attempt to predict transit ridership (Boisjoly and El-Geneidy, 2016; Bree et al., 2020; Cui et al., 2022; Merlin et al., 2021). According to these studies, it is to be expected that there is a positive association between accessibility and ridership. This is based on the idea that one of the main benefits of a transportation system is to improve peoples' access to opportunities, and so one should expect at least to a certain point that more passengers would use the transit system if it provided access to a larger number of opportunities that could be of interest to a greater number of people.

The works of Bree et al. (2020) and Cui et al. (2022) measure the association between accessibility and ridership at the stop level, while Merlin et al. (2021) measure at the city level. Cui et al. (2022) argues that measuring this association at an aggregate level such as census tracts or city level are not easily translatable into service changes. However, doing this analysis at the stop level is also problematic because the accessibility from a transit stop is influenced by all of the transit routes that pass through it.

Measuring accessibility at the route level can be important for different reasons. Different routes provide varying levels of accessibility based on their geographical path, frequency, speed, and connectivity. Assessing the accessibility conferred by each route allows for a more complete examination of the benefit created by each component of the transit network considering its integration to the rest of the network and the numerous alternatives available to users (Van Wee, 2016). In addition, because most transit planning decisions are made at the route level, and so it becomes particularly important to understand how design and operational changes to specific transit routes could impact the accessibility of its users. In that sense, a route-based accessibility metric can help decision makers adopt accessibility analysis in practice. This type of metric can more directly capture, for example, how changes in a transit route (e.g. in terms of service frequency or stop sequence, etc.) would impact its accessibility, and thus help inform the allocation of resources to specific routes that could generate larger accessibility benefits to underserved and disadvantaged communities.

To date, there are only two papers in the literature that have attempted to measure accessibility at the transit route level. The first paper was that of Karner (2018), who focuses on equity issues in transit services in Phoenix (USA). His indicator is calculated in two steps. In the first step, he estimates stop-level jobs accessibility using a gravity-based indicator with a mixed decay function. The function considers a cumulative character up to a certain point (39.9 min, which is the average travel time found in the case study), after which it follows a non-linear form fitted to the distribution of travel times observed in a travel survey. In the second step, the author normalizes this accessibility and weights it by the number of people (of a given income level) who can access that stop on foot. Finally, Karner (2018) calculates the accessibility of a transit route as the average of the accessibility of the stops that are part of the route.

The second work that seeks to measure accessibility at the route level is the study by Diab et al. (2020a). The study examined the determinants of transit demand in Montreal (Canada) between the years 2012 and 2017. The authors used a regression model with random coefficients covering supply control variables (such as headway, route speed, fare), socioeconomic and demographic variables (unemployment, population density, average income), and a variable of the job accessibility conferred by each route. To measure the accessibility of each route, the authors first used a cumulative opportunity measure to calculate the number of jobs accessible within 45 min from each census tract. Then, the authors calculated the accessibility of each transit route as the average accessibility of census tracts intersecting with the route weighted by the size of the intersection area between the census tract and a 400-meter buffer zone around the route.

A common limitation of the indicators by Karner (2018) and Diab et al. (2020a) is that they only consider walking catchment areas, overlooking how the accessibility benefits of a given transit route also accrue to passengers transferring from other routes upstream in the transit network (e.g. feeder routes). More importantly, neither of them restrict the route taken by individuals in the routing analysis to calculate travel times and accessibility estimates. Both studies rely on shortest-path algorithms that only consider the quickest trip for each origin–destination pair. This results in accessibility estimates that always consider optimal network solutions, but which cannot capture the accessibility provided by specific routes, especially suboptimal ones. In other words, the methods adopted by Karner (2018) and Diab et al. (2020a) may capture the average accessibility at the places where a given route passes by, but they do not necessarily reflect what accessibility a person could have by using that route. This limitation can produce different problems. For example, these indicators would not necessarily capture, for instance, how the accessibility of a route would be affected by an increase in service frequency, speed or connectivity if the route remained as a suboptimal choice for transit users despite those changes. Additionally, these methods could produce biased results by wrongly estimating high accessibility values for poorly connected routes with low frequency, but which share stops with other frequent and highly connected routes. In an extreme example, these indicators would suggest that two bus routes with overlapping trajectories and stops provide the same accessibility level, even if these routes have drastically different operational characteristics.

Finally, other studies with a focus on project impact assessment measure the accessibility impact of particular routes using a common method. In these studies, the impact of a route is calculated based on how accessibility levels change when computed considering the full transit network versus a synthetic network that includes all routes except the route of interest (Liu and Miller, 2023; Niehaus et al., 2016; Pereira et al., 2019; Pereira et al., 2019). This method has been recently referred to as the "accessibility derivative" measure by Liu and Miller (2023), but perhaps it would be more appropriately named as the "marginal accessibility change" method. Conceptually, this approach is rather different from a route-based accessibility metric. A route-based metric, as we detail in the next section, captures the average accessibility benefit of a transit route (e.g. the average number of opportunities that can be reached by using that route). By contrast, the accessibility derivative measure only captures the marginal change of accessibility brought about by adding (or removing) a transit route.

Moreover, like the works of Karner (2018) and Diab et al. (2020), the method of marginal accessibility change does not restrict the route taken by individuals in the routing analysis, and it relies on shortest-path algorithms. A direct implication here is that this method tends to underestimate the importance of sub-optimal routes. Consider the following hypothetical scenario: a person at bus stop A can utilize transit route 1 to reach transit stop B in 30 min, thereby accessing 5000 job opportunities. However, an alternative route, route 2, exists, which is both faster and more frequent, allowing access to those same 5000 jobs at stop B and an additional 3000 jobs at a subsequent stop C, all within the same 30-minute timeframe. Employing the accessibility derivative method, transit route 2 would yield a marginal accessibility gain of 3000 jobs. Nonetheless, the same method would suggest that the contribution of route 1 to the transit system is seemingly zero. This result emerges due to route 2's superior efficiency, leading to route 1 being overlooked in the routing analysis. This might be particularly problematic if one wants to examine the relationship between transit ridership and accessibility at the route level. This is because the marginal accessibility change method might suggest that certain routes could have relatively low accessibility levels, even though they still attract passenger ridership. Yet, even transit routes that are sub-optimal in terms of travel times can still attract passengers due to reasons such as service reliability, to avoid crowding, because they are cheaper or require less walking (Arriagada et al., 2022; Raveau et al., 2014).

Next, we discuss some features that an accessibility indicator should have to overcome these limitations found in the literature.

3. Proposed route-based accessibility measure

The accessibility promoted by a transit route refers to its potential to facilitate people's access to opportunities. From a theoretical point of view, a route-based accessibility measure should consider a few criteria. It should account for the population potentially benefited by the route either directly (i.e. the population living in its immediate catchment area within walking distance to the routes' stops) and indirectly (i.e. those that could benefit from the route through connectivity with the rest of the transportation system). A route-based accessibility metric also needs to account for land use patterns (i.e. the spatial distribution of population and opportunities) and time constraints considering passengers' travel time budgets. Finally, route-based accessibility needs also to consider some key characteristics of transit routes both in terms of operational performance such as frequency, speed, in terms of monetary costs and its network connectivity.

Given the network nature of public transit systems, the accessibility benefit provided by a route needs to account for its upstream and downstream connectivity with other routes in the transportation network. For example, consider a feeder route that goes from a neighborhood to a terminal. This route provides people with access to opportunities that go beyond the neighborhood-terminal path. This is because this feeder route is connected to other routes in the transit system, allowing passengers to make transfers and reach further opportunities down the network. For example, if a person's travel time budget is say 60 min, a feeder route could significantly expand the opportunities accessible to a person if she could in 15 min using that feeder route transfer to other transit routes and continue the rest of her journey for 45 more minutes.

The route-based accessibility measure proposed in this paper is calculated in three steps. First, the accessibility levels from all geographic zones are calculated using each transit route. The second step is to estimate the population potentially benefited by each route. Finally, the third step is to calculate the route-level accessibility by aggregating the accessibility of geographic zones (step 1) weighted by the population benefited by that route (step 2). These steps are described in detail below. For the sake of simplicity, the route-based metric calculated in this paper covers all of the criteria listed above except for the consideration of monetary costs, which could be incorporated in future developments.

In step 1, we calculate the number of opportunities accessible from each geographic unit using each route in the transit system. This step is similar to a traditional place-based accessibility measure and can be calculated considering different types of impedance functions (Eq. (1)). The difference, however, is the restriction that accessibility only considers trips that necessarily use route r in at least some part of the journey. The result of this step is a series of accessibility estimates for each point of origin and transit route. All the geographic units from which it is possible to access at least one opportunity using a given route comprise what we refer to as the route's "area of influence".

$$A_{ir} = \sum\nolimits_{j=1}^{n} O_j \times f(C_{ijr}) \tag{1}$$

where A_{ir} = Accessibility from a origin i using a route r, and n is a set of destinations j; O_j = Number of opportunities in a destination j; $f(C_{ijr})$ = Impedance function that considers the cost, C_{ijr}, of traveling from an origin i to a destination j using the route r in at least some portion of the trip.

The second step is to calculate the size of the population potentially benefited by route r in a given time period, i.e. the population in its area of influence. This includes all the population for whom this route could be potentially useful to access at least one opportunity. This step can be calculated considering either the population as a whole, or it can be calculated separately for different population groups (such as income or racial groups). While step 1 allows us to determine the zones from which the route r could help improve access, in step two we count the number of people in those zones as the population who could potentially benefit from the route.

Finally, in step 3, the accessibility provided by a route r is calculated as the average accessibility from each geographic unit as estimated by Eq. (1), weighted by the benefited population of each unit obtained in step 2. The calculation of the route-level accessibility is represented by Eq. (2):

$$A_{wr} = \frac{\sum_{i=1}^{n} A_{ir} \times P_{wir}}{\sum_{i=1}^{n} P_{wir}} \tag{2}$$

where A_{wr} = Route-level accessibility of route r for a group of people w; A_{ir} = Accessibility from a origin i using route r; P_{wir} = Population of group w located in the spatial unit of origin i benefited from route r.

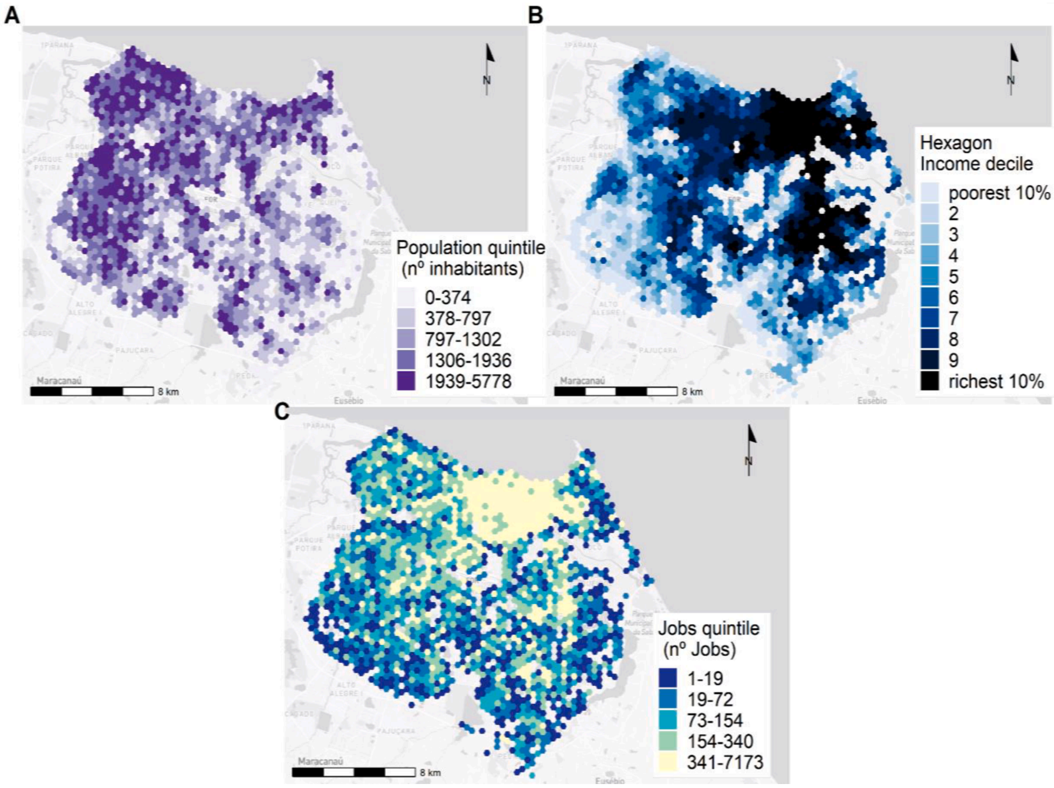

Fig. 1. Spatial distribution population (A), income per capita (B) and number of jobs in Fortaleza (C).

4. Case study: Fortaleza (Brazil)

Fortaleza is the fifth largest city in Brazil, with approximately 2.7 million people (IBGE, 2020) and an area of around 313 km^2, making it the ninth most densely inhabited city in the country (IBGE, 2020). The distribution of population, income and job opportunities in Fortaleza is shown in Fig. 1. Employment opportunities are more concentrated in the northernmost region of the city, which coincides with the region with the highest per capita income. Meanwhile, low-income communities are more concentrated in the outskirts of the city.

According to the most recent household travel survey (Pasfor, 2019), the population of Fortaleza commutes predominantly by private modes (car or motorcycle) (37.7%), followed by walking (30.8%), bus transit (23.2%), bicycle (6.2%), subway (0.4%), and other modes (1.7%). Bus routes account for the vast majority of public transit journeys. In 2018, the bus system of Fortaleza included approximately 300 bus routes in service and 9 terminals. Seven of the nine terminals have access control, while the other two serve primarily as integration points in the city core (Fig. 2).

The bus system also includes a smart card system, which recorded an average of almost 1 million passengers per day in 2018. Fig. 2 illustrates the public transit system, where the links between the stops of the bus network are colored to reflect the service frequency in the morning peak (between 6am and 8am). The rail system includes the south and west subway lines, and a light rail line that was in test phase in 2018. The bus transit system has a high spatial coverage and is organized as a hybrid network largely built on a few trunk routes connecting terminals and the city center, and multiple feeder routes that take passengers to terminals. This helps explain higher concentration of routes and service frequencies near the city center. It is also worth noting that there is no fare integration between the bus and subway systems, and only poor physical integration between the two modes in very few stations. This means that the bus and subway systems operate largely independently in practice, which helps explain at least in part why the subway responds for only a small share of the transit demand in Fortaleza.

5. Materials and methods

To demonstrate the route-based accessibility metric proposed in this paper, we use the indicator to calculate the average access to jobs conveyed by each bus route of the transport system of Fortaleza, Brazil. Additionally, we compare the performance of our proposed indicator against similar metrics found in the literature (Karner 2018; Diab et al., 2021) and the traditional method of accessibility impact assessment (Liu and Miller, 2023) by examining the extent to which these measures are associated with the ridership of transit routes. The assumption here is that, if accessibility is indeed one of the main benefits promoted by public transit systems

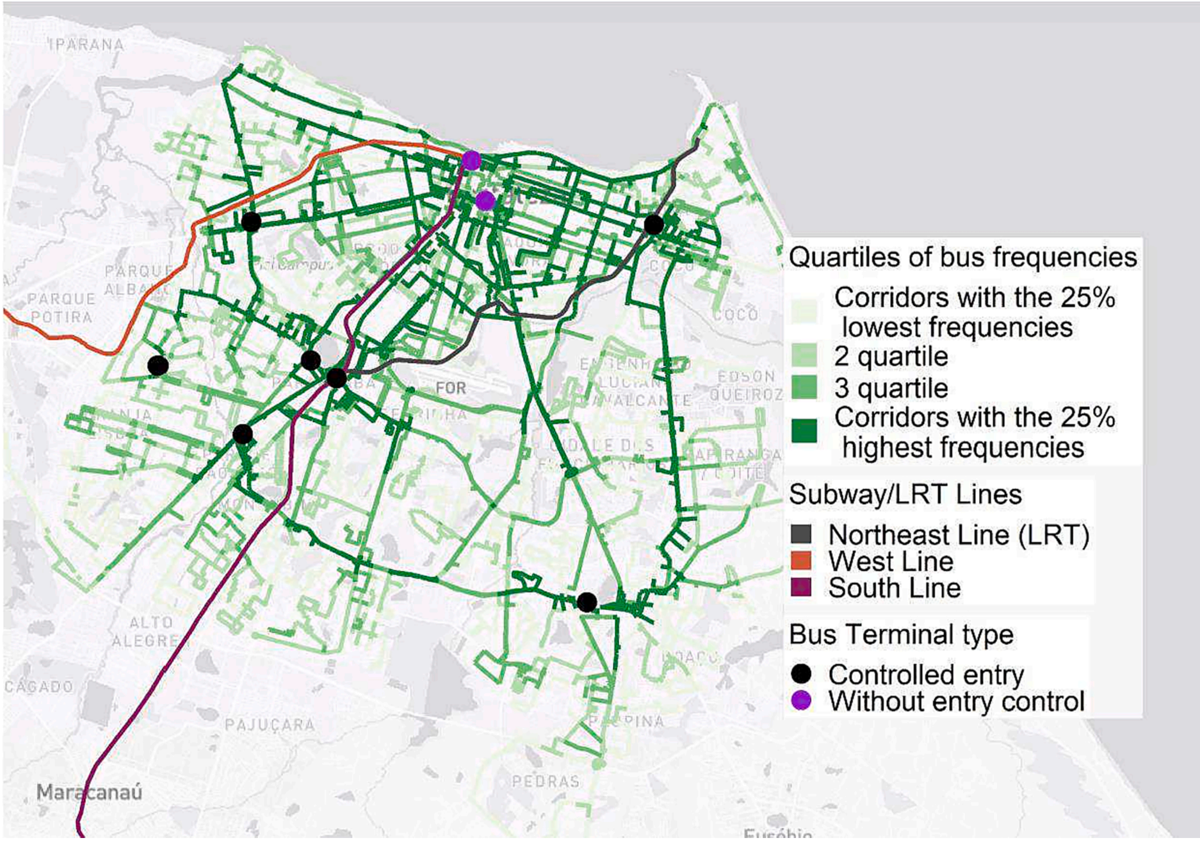

Fig. 2. Public transit system of Fortaleza, 2018.

Table 1
Information used in the study, according to year and source.

Data	Information	Year	Source	Indicators calculated
Population characteristics	Spatial distribution of the population and income	2010	Demographic census, IBGE	– Population near transit routes – Accessibility indicators
Economic Activity	Spatial distribution of jobs	2018	Annual Social Information Report (RAIS), Ministry of Economy	– Jobs near transit routes – Accessibility indicators
Road Network Characteristic	Road network	2018	Open Street Map	– Accessibility indicators
General Transit Feed Specification (GTFS)	Itineraries, stops, schedule, among others	2018	Etufor, 2018	– Accessibility indicators – Route frequency
Smart card	Ridership of each route	2018	Sindionibus, 2018	– Transit ridership
Origin-DestinationSurvey	Number of boardings per route at transit terminals	2019	Pasfor, 2019	– Impedance functions of accessibility metrics – Transit ridership

(Bertolini et al., 2005; Geurs and van Wee, 2004), and if transit ridership reflects the direct utility people derive from transit services (Merlin et al., 2021), we should expect that the accessibility provided by a transit route could be a good predictor of its passenger demand. We examine the association between route-level accessibility and ridership using both an exploratory correlation analysis and a multivariate regression model. Since there are many factors that affect accessibility and ridership, we applied a multivariate regression considering other factors that are found in the literature. To run the multivariate regression model, we propose a Directed Acyclic Graph (DAG) to map the causal paths of the key variables that influence route-level transit ridership. The construction of the DAG is important to make explicit the assumptions behind the regression model, and to guide which covariates should be controlled for if we want to find unbiased estimates of the association between accessibility and ridership. The next subsections present in more detail the data and methods used.

5.1. Data

Table 1 lists the data sets used in this study. Population and income data comes from the 2010 Census, the latest demographic data

Table 2
Parameters used in routing.

Parameter	Value
Maximum travel time (door-to-door)	1 h (transit)
Active mode speeds	3.6 km/h (walking) 12 km/h (bicycle)
Maximum walking distance to access or egress the transit network	1000 m
Time of suboptimal routes	10 min

available. Employment data includes all formal jobs registered in 2018 in an administrative record database organized by the Ministry of Economy. Public transit data in the General Transit Feed Specification (GTFS) format was used to do routing analysis and calculate travel time matrices (details in Section 5.2). The smart card data and the origin–destination survey data are combined to determine the ridership of bus routes, as explained in the next section.

5.1.1. Ridership data

Transit ridership data collected through smart card records was provided by the city's transportation planning agency (ETUFOR). We used data from October 2018 to calculate average ridership on each bus route during the morning peak (between 6am and 8am) of business days. We focused on this particular period because this is when passenger ridership is most concentrated.

Although most ridership count data is captured through smart cards, the number of passengers boarding bus routes that depart from a few terminals are not registered in the case of terminals with controlled entry. This is because once a passenger taps in to enter this type of terminal, there is no need to validate the ticket when boarding a bus. As a consequence, there is no information on which bus routes passengers take after entering these terminals. This is the case for seven of nine transit terminals in Fortaleza. To get around this limitation, we used data from a boarding survey conducted in 2018 at the terminals as part of the city's latest travel survey (Pasfor, 2019). This survey asked passengers at terminals with controlled entry which routes they take when departing from those terminals at different times of the day. Assuming that the passenger demand data collected in this survey is representative of the daily average ridership, we added this average demand of routes departing from those terminals to their average ridership information collected from smart cards.

Ridership data for the two subway lines in Fortaleza were not available, so it was not possible to include this transport mode in the analysis. Since only a small share of commuting trips by transit were done by subway – 0.4% according to the latest household travel survey (Pasfor, 2019) - we believe the exclusion of this mode from the analysis should not significantly affect our results.

5.2. Routing and accessibility analyses

In this section we detail the method used to estimate the proposed route-based accessibility indicator. In the first step, OpenStreetMap road network, topography and GTFS data are used to estimate travel times between the centroids of a spatial grid of hexagons. We use a spatial grid that corresponds to the global H3 index at resolution 9 (UBER, 2018). At this resolution, each cell has an area of approximately 0.10 km^2, which is equivalent to an average of three city blocks in the context of Fortaleza, allowing for an analysis at high spatial resolution without compromising computational tractability.

These travel time estimates were calculated using r5r, an open computational package for routing multimodal transportation routing in R (Pereira et al., 2021). The detailed_intineraries() function of the package allows one to estimate door-to-door travel time estimates for any origin–destination pairs considering both the optimal and multiple suboptimal paths that use different transit route alternatives. This is possible because the function uses an extension to the McRAPTOR routing algorithm, for which detailed description of the base algorithm can be found in (Delling et al., 2015).

After calculating the travel time matrices from all-to-all considering multiple bus route alternatives, we can select for each bus route the trips between origin–destination pairs for which that route would have been used in at least some part of the journey. For each of these multiple travel time matrices at hand (one for each bus route), we use Eq. (1) to calculate the active accessibility to jobs from each hexagonal cell. Next, we followed Eq. (2) to calculate for each travel time matrix the average accessibility of hexagonal cells weighted by their population, which gives us the accessibility conveyed by each bus route in the transport system.

The parameters used to calculate travel times are summarized in Table 2. The route analysis considered 6 a.m. as the single departure time to capture the level of service and accessibility levels during the morning peak. Ideally, several routings would have been carried out with different departure times. However, this was not possible due to computational limitations. In addition, for this exercise, we only considered suboptimal routes that arrive at the most 10 min after the arrival of the optimal route, considering that a 10-min delay would still be tolerable for most passengers.

We calculated accessibility using both cumulative and gravitational decay functions. This was done to make sure we have a robust comparison between the indicator proposed in this paper and other route-based accessibility metrics (Diab et al., 2020a; Karner, 2018; Liu and Miller, 2023) considering the two most popular decay functions used in the literature. In the case of the cumulative measure, the travel time threshold considered was 60 min, a value similar to the average commuting time by transit in Fortaleza (59 min) according to the latest household travel survey (Pasfor, 2019).

Meanwhile, in the case of the gravity decay function, we calibrated the impedance function considering the distribution of travel times from transit trips observed in the household travel survey. Impedance values were estimated using a combined impedance

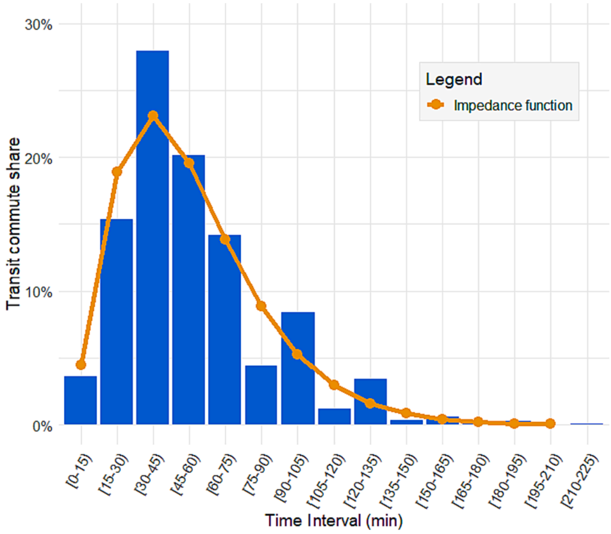

Fig. 3. Calibrated impedance function (yellow curve) and distribution of travel times from public transit trips (bars) in Fortaleza, 2019. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

function (de Ortúzar et al., 2011), represented in Eq. (4), which presented a better fit to the data distribution. This impedance function represents people's abilities and preferences at the same time. On the one hand, it reflects how a person's ability to access opportunities diminishes with longer journeys that consume more travel time. On the other hand, it also reflects preferences because, for short trips under 15 min, some people might prefer to use active modes of travel, which account for almost 40% of daily trips in Fortaleza according to the latest household survey. As a robustness check, we also conducted our analysis considering an exponential-type gravitational impedance function calibrated with the same data. The calibrated Beta value was −0.00038 and the conclusions of our analysis remained the same.

Fig. 3 shows the calibrated decay curve with the corresponding parameter values and the distribution of observed travel times. The calibrated parameters were $a = -86.91$, $b = 2.1$ and $c = -0.058$.

$$f(c_{ij}) = exp(a) \times t_{ij}{}^{b} \times exp(c \times t_{ij}) \tag{4}$$

where $f(c_{ij})$ = travel impedance factor; a = standardizing constant; b = adjustment parameter; c = adjustment parameter; t_{ij} = travel time from origin i to destination j.

5.3. Analysis of the association between accessibility and ridership

A first exploratory analysis is conducted by using Pearson's correlation between the ridership of bus routes and route-level accessibility estimates calculated using our proposed indicator, derivative metric proposed by Liu and Miller (2023) and the indicators proposed by Karner (2018) and Diab et al. (2020a). The analysis was conducted considering all bus routes in the transit system of Fortaleza, and for three different types of routes separately: trunk routes with terminals, trunk routes without terminals, and feeder routes. Trunk routes with terminals are those routes connecting two terminals or which connect terminals to the city center. Trunk routes without terminals are those which connect neighborhoods to the city center but do not connect to any terminal. Finally, feeder routes were defined as those connecting neighborhoods to terminals or to other neighborhoods without connecting to terminals or the city center. The motivation for these segmented analyses is to test the extent to which the results are consistent for different types of routes with different operational characteristics and connectivity patterns.

Next, we further examine the association between accessibility and ridership using a multivariate regression to estimate how transit ridership is associated with the different route-based accessibility metrics while controlling for other covariates. To select the covariates included in the regression, we built a Directed Acyclic Graph (DAG) that maps the potential causal paths of relevant variables found in the literature about the relationship between accessibility and determinants of transit ridership (Section 5.4). The construction of a DAG makes explicit the assumptions in the data analysis about which variables are directly or indirectly related to one another, and it allows for the identification of those control variables that are essential to estimate the unbiased effects of the variable of interest, and those variables that could introduce bias if included (Huntington-Klein, 2021).

5.4. Directed Acyclic Graph (DAG) of accessibility and ridership

Several studies have analyzed the determinants of transit ridership (Balcombe et al., 2004; Boisjoly et al., 2018; Taylor et al., 2009; Taylor and Fink, 2012). A more detailed review of these studies can be found in Miller et al. (2018). These determinants are divided

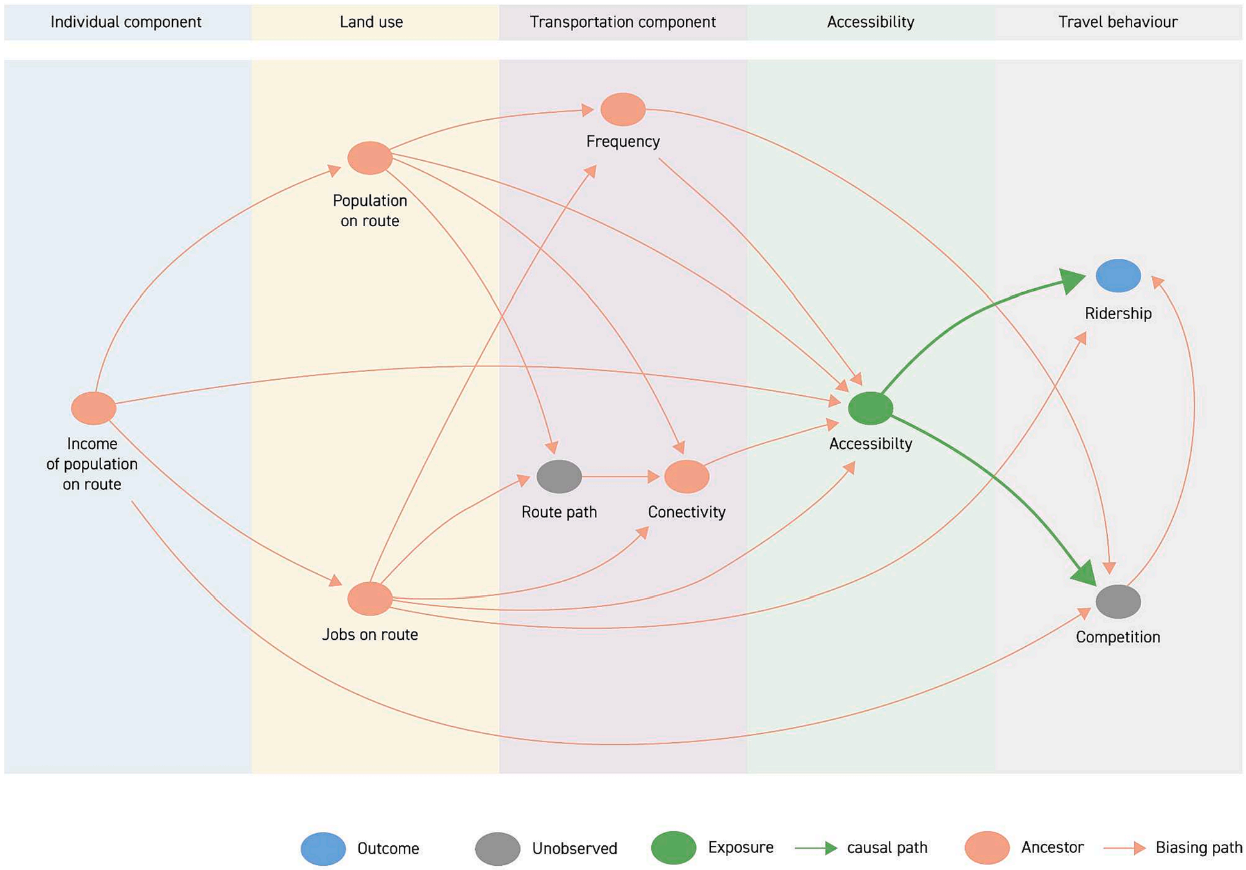

Fig. 4. Directed Acyclic Graph of the causal paths between transit route-level accessibility and ridership.

into external and internal factors. External factors are those outside the direct management of planning agencies, such as the spatial distribution of population, income, and jobs. Internal factors, on the other hand, are those that agencies can directly control, such as service frequency, network coverage. More recently, some of these studies have started to consider accessibility as an important variable in explaining or predicting passenger demand and/or mobility (Bree et al., 2020; Cui et al., 2022; Diab et al., 2020a; Merlin et al., 2021; Siqueira, 2020). However, efforts to systematize the determinants of transit ridership from a causal point of view are still incipient (Taylor et al., 2009).

Fig. 4 depicts the proposed causal diagram between transit route-level accessibility and ridership. The ridership of a transit route is directly impacted by the accessibility conveyed by that route (Merlin et al., 2021). This is largely because passengers may select bus routes based on how much a route can help them reach desired opportunities. Routes that are able to connect larger numbers of people to larger numbers of opportunities are expected to attract more passengers.

The ridership of a transit route is also directly impacted by the competition between the route and other alternatives. Although we do not measure the competition variable in this study, it is meant to capture peoples' travel choices both in terms of alternative transit routes and transport modes. When people are deciding between bus routes, for example, the frequency of services have a direct influence on their decision, particularly when the trajectory of routes overlap spatially (Diab et al., 2020a). This competition is also directly affected by the population's income levels because of how income determines the affordability of transit services and the availability of private vehicle alternatives (Schwanen et al., 2002; Siqueira, 2020). Sometimes, the opportunities that a transit route may access have a direct impact on ridership, even if it is a low service route. This can occur, for example, in commuter routes because work trips are mandatory. Those commuting for work might have no choice but to use a particular commuter route (Merlin et al., 2021). As a result, a transit route may have some ridership even if it does not provide high levels of service, connectivity, or access.

Meanwhile, route-level accessibility is influenced by different transport, land use and individual components, in line with the broad components of accessibility proposed by Geurs and van Wee (2004). Under the transport component, the accessibility conveyed by a route is directly influenced by its fare cost, service frequency, operational speed and its connectivity to the rest of the network, which is at least in part a result of its spatial itinerary. This accessibility is also directly affected by the land use component, which includes the spatial distribution of population and opportunities (such as jobs). This involves direct effects because of the proximity between a route with population and opportunities, and also indirect effects because of how the spatial distribution of population and opportunities influence the supply of transit services in terms of network configuration and service frequency. As a result, the people benefited and

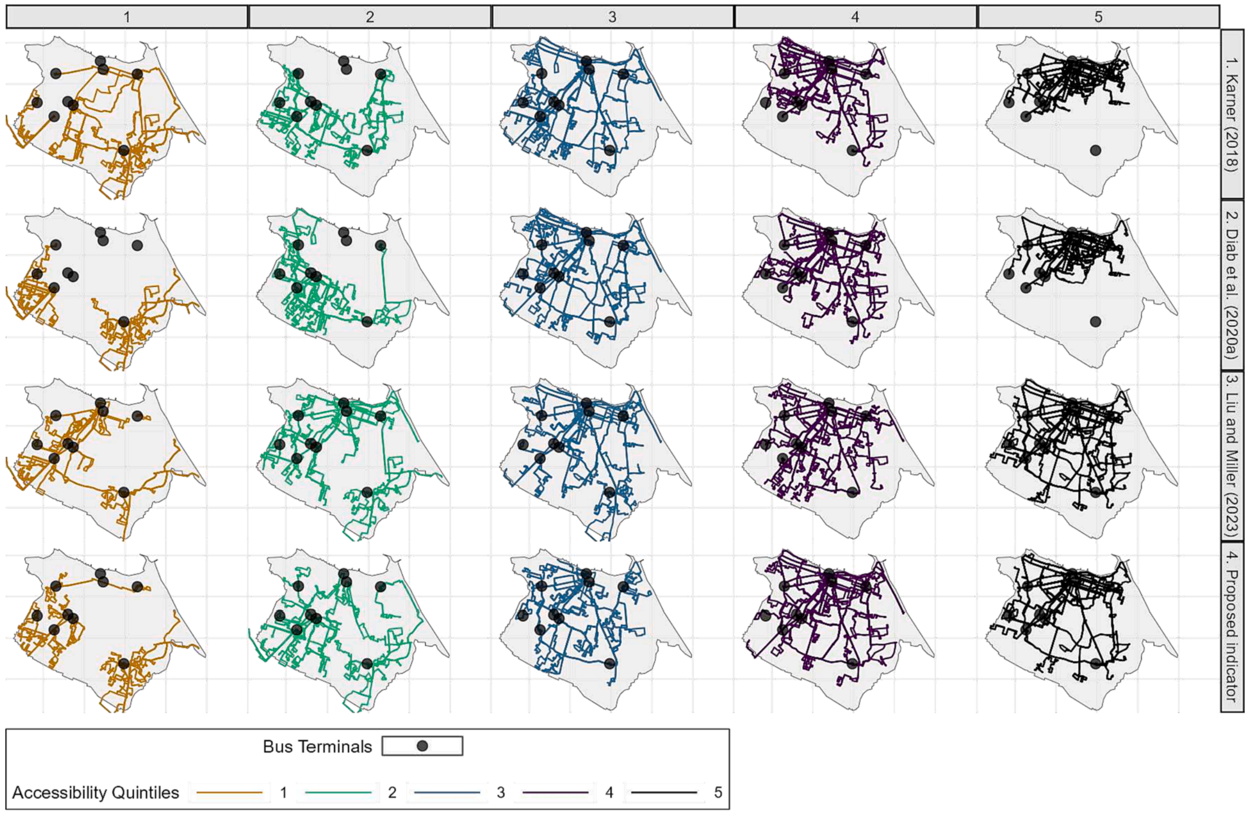

Fig. 5. Quintiles of route accessibility for the different indicators.

the jobs accessed determine the supply and degree of service, indirectly impacting ridership (Merlin et al., 2021).

Finally, the accessibility conveyed by a route is also influenced by a component of individuals' characteristics. Even though we do not account for these characteristics in the present paper, this component involves various factors that impact people's ability to use transit systems. This includes income (due to fare costs and affordability issues mentioned above) but also other aspects such as age and various types of physical disabilities and whether vehicles are built following universal design standards.

It is also important to highlight a few points. This DAG aims to illustrate the most immediate processes of ridership determinants. As a result, some dynamics are not considered here. One of these is the endogenous influence of ridership on the planning of transit supply (Diab et al., 2020b; Taylor et al., 2009). Another aspect is the dynamic impact of accessibility on the spatial distribution of population and activities as a result of locational choices. Individuals and firms may prefer to locate in areas with better transit accessibility (Cao et al., 2008; Cervero, 2007). Additionally, the competition variable could also consider people's perceptions of comfort, convenience and safety of different transport mode and transit route alternatives (Rahman et al., 2020; Vredin Johansson et al., 2006). This perspective certainly includes experiences with service level of routes, such as crowding and reliability (Brakewood et al., 2015; Cats et al., 2016; Drabicki et al., 2023; Tang and Thakuriah, 2012). As a result, current ridership may have an impact on future perceptions of competition across routes. Lastly, the DAG proposed here does not include the potential effects of the built environment or urban form on route-level accessibility and ridership, but this could be incorporated in future versions (Aston et al., 2021; Ewing and Cervero, 2010, 2001).

Obs. In the vocabulary of the DAG literature, the arrows that represent the direct and indirect causal paths between one's variable of interest (the exposure variable) and the outcome variable are denominated as "causal paths". All the other paths are referred to as "biasing paths" because they can transmit bias for estimating the effect of the exposure on the outcome.

5.5. *Regression model specification*

Based on the DAG proposed in Fig. 4, we set out to specify the model to estimate the causal effect of interest, which is the total effect of accessibility on ridership. The diagram shows different sources of potential biases caused by variables that affect both accessibility and ridership through different indirect pathways (backdoor effect), such as the jobs and frequency variables. These variables are called confounding variables in the causal inference literature (Huntington-Klein, 2021) because they are not part of the causal pathway between accessibility and ridership but can create a false association between them. Based on path analysis, it is possible to identify all possible confounding pathways on the diagram and find the appropriate variables which should be controlled for in order

Table 3
Pearson Correlation coefficients between route-level accessibility and ridership considering different route-level accessibility measures, impedance functions and types of bus routes.

Impedance function	Indicator	Type route			Entire system
		Feeder	Trunk, with terminal	Trunk, no terminal	
Cumulative	Karner (2018)	0.02	−0.05	−0.17	0.14*
	Diab et al. (2020a)	0.03	−0.14	−0.20	0.17*
	Liu and Miller (2023)	0.43*	0.45*	0.41*	0.49*
	Proposed	0.44*	0.41*	0.79*	0.55*
Gravitational	Karner (2018)	0.02	−0.06	−0.18	0.13*
	Diab et al. (2020a)	0.03	−0.15	−0.21	0.17*
	Liu and Miller (2023)	0.43*	0.34*	0.41*	0.43*
	Proposed	0.44*	0.40*	0.79*	0.55*

* $p < 0.05$ (all remaining values are not significant).

to obtain unbiased and valid estimates of causal relationships of interest (Shrier and Platt, 2008; Textor et al., 2017). Based on the DAG proposed above, the minimal sufficient adjustment sets (Shrier and Platt, 2008; Textor et al., 2017) suggest there are two model specifications to estimate the total effect of accessibility on ridership removing potential biases: (i) frequency, jobs and population or (ii) frequency, jobs and income. We have used both specifications in our regression analyses. For the sake of brevity, the results for the first set are presented in the body of the text, while the results for the second set of variables are shown in the Appendix A.

The covariates were measured as follows. We measured routes' frequency as the number of trips per hour between 6 a.m. and 8 a.m. reported in the GTFS data for weekdays. All the socio-economic variables (jobs, population and income) were defined from the hexagons that intersect a 500-m buffer of the routes' itinerary. The 500-m buffer is intended to take into account users who can access the route on foot. Thus, the employment and population variables represent the total number of jobs and residents within this area nearby the route, respectively. Finally, the income variable is represented using the same criteria as the average per capita income in the vicinity of the route.

We should note that while the accessibility indicator can be correlated to some covariates, there is not perfect collinearity between the independent variables. This collinearity between independent variables leads to large variance of the estimated coefficient of accessibility, but it does not bias the point estimates of the coefficients. To compare the relationship between accessibility and ridership across the specified indicators, we used log-transformed variables so the results can be interpreted as elasticities. Because Karner's indicator (2018) can take on negative values due to a normalization procedure of accessibility scores, we decided to ignore this normalization step given that it is not essential for the purposes of our analysis.

6. Results

Fig. 5 displays maps of the bus routes of Fortaleza divided by quintiles of accessibility levels for each indicator, allowing for a spatial assessment of the routes in Fortaleza. Because the use of cumulative and gravitational decay functions yielded very similar results, we have opted to show only the maps based on the cumulative function for the sake of brevity. As shown in the figure, local routes located in the peripheral regions of the city with weaker connectivity to the transit system tend to convey the lowest accessibility levels (1st quintile), as expected. However, our proposed indicator and the derivative metric allow us to identify a few routes that connect different neighborhoods to the city center at the lowest level of accessibility, presumably due to their low-frequency services.

The results from the four indicators show that bus routes with high accessibility levels are typically those directly connected to the city center or with stops in densely populated areas. However, only the proposed indicator and the derivative accessibility metric consider how a route's connectivity to the rest of the transit system impacts its accessibility. Nonetheless, the effect of connectivity is underestimated in the derivative indicator due to the fact that suboptimal routes are not taken into account. This can lead to many routes being suboptimal in areas with a high supply of public transport and therefore being assigned low accessibility values. This helps us understand why there is a greater concentration of routes with low accessibility (first three quintiles) near the city center for the derivative indicator. This makes our proposed indicator relatively better able to measure route accessibility and identify routes that, despite being located in the southern outskirts of the city and far from the city center, offer high levels of accessibility due to their connections to trunk routes or terminals.

Table 3 shows the correlation estimates between route-level accessibility and transit ridership for all types of bus routes. The results show that both cumulative and gravity impedance functions produce very similar outcomes. We also observe that the indicator proposed in this study exhibits a strong positive association with route ridership, both for the entire system (with a correlation of around 0.55) and for each route type (ranging from 0.44 to 0.79). The derivative accessibility metric also has good correlation levels, although slightly inferior to the proposed indicator. Specifically, when evaluating trunk routes without terminals, the proposed indicator exhibited higher correlation. Conversely, the other accessibility indicators presented non-singificative correlations. Particularly the indicators of Diab et al. (2020) and Karner (2018), which presented negative correlations for certain route types. These negative correlations are counter-intuitive and could be attributed to the limitations of these indicators in measuring the accessibility provided by each particular route, instead of focusing on the overall accessibility of locations along the route (as discussed in Section 2). It is worth noting that the negative correlations were more pronounced for trunk routes without terminals, which intersect with several feeder routes.

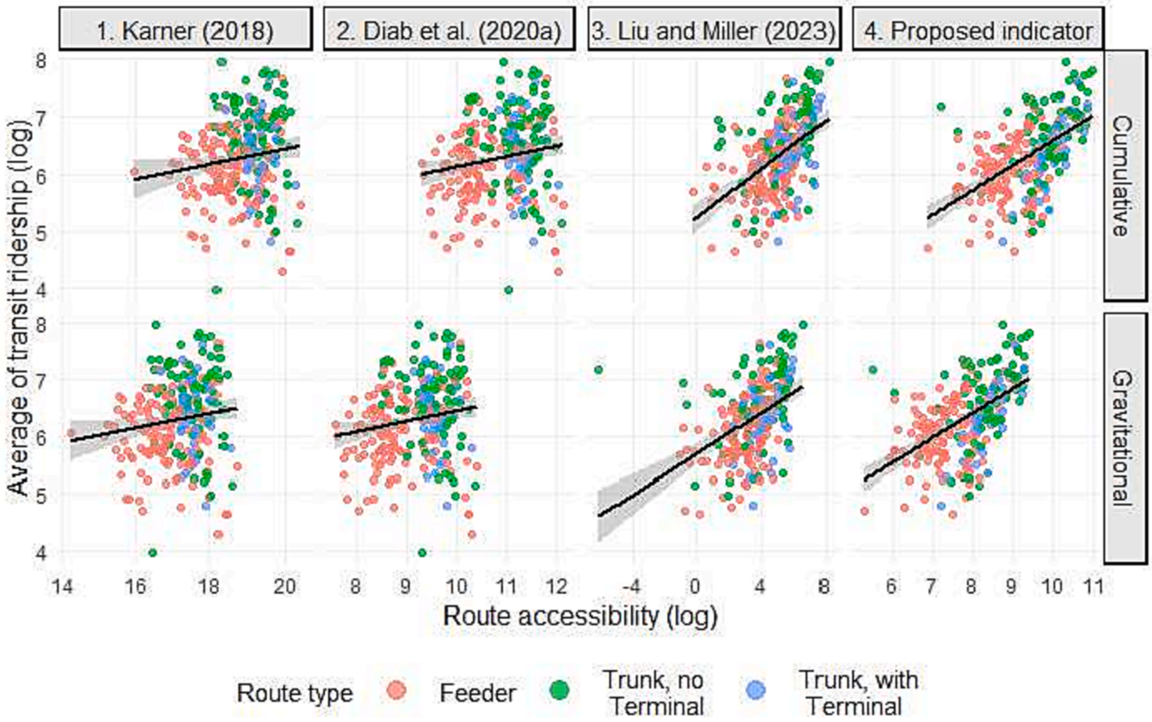

Fig. 6. Scatter plot between accessibility and ridership for the different indicators obs. Black line represents the linear regression line that minimizes the squared perpendicular distance to all data points. The shaded area indicates the 95% confidence interval.

Table 4
Summary of multivariate regression of the effect of accessibility on ridership considering different route-level accessibility metrics.

Indicator	Karner (2018)	Diab et al. (2020a)	Liu and Miller (2023)	Proposed indicator
Constant	**2.614*****	**2.982*****	**−0.120**	**−0.141**
	(0.661)	(0.649)	(0.628)	(0.632)
Route accessibility	**−0.387*****	**−0.682*****	**0.088*****	**0.318*****
	(0.042)	(0.068)	(0.021)	(0.079)
Route Frequency	**1.015*****	**1.056*****	**0.865*****	**0.801*****
	(0.052)	(0.052)	(0.058)	(0.063)
Jobs on route	**0.264*****	**0.542*****	**0.073**$^{+}$	**−0.042**
	(0.038)	(0.054)	(0.039)	(0.053)
Population on route	**0.397*****	**0.173****	**0.233****	**0.138**$^{+}$
	(0.062)	(0.062)	(0.072)	(0.083)
Num. Obs.	245	245	245	245
R^2	0.714	0.730	0.642	0.640
R^2 Adj.	0.710	0.725	0.636	0.634
AIC	194.5	180.9	249.9	251.1

*$p < 0.05$.
$^{+}$ $p < 0.1$.
** $p < 0.01$.
*** $p < 0.001$.

Fig. 6 shows the scatter plots illustrating the association between route-level ridership numbers and accessibility level calculated for each indicator and route type. The solid lines show the ordinary least squares regression lines, included for illustrative purposes. This figure shows route-level accessibility calculated with the proposed indicator presents the strongest association with ridership. Again, this result is robust regardless of impedance function considered.

The proposed indicator and the derivative accessibility metric are also better able to distinguish accessibility levels between types of bus routes. The results with our indicator capture a general pattern where feeder routes tend to have the lower accessibility levels, followed by trunk routes with and without terminals. Meanwhile, the indicator proposed by Karner (2018) was not able to differentiate between the accessibility levels across various route types, and the indicator by Diab et al. (2020a) was only able to distinguish accessibility levels between feeder and trunk routes. The result found for Diab et al. (2020a) indicator may be attributed to the fact that this measure is merely an average of the accessibility of census tracts. As a result, it yields high accessibility values for routes that traverse through the city's central area (where job opportunities are abundant) and low accessibility values for routes that pass through low-density regions, such as feeder routes. Again, the fundamental issue with Diab et al. (2020a) and Karner (2018) indicators is that they calculate route accessibility by averaging accessibility values across geographic units without considering which route is taken, which results in assigning a default system accessibility level for all routes. This approach overlooks the distinctive features of each route, such as the served population, operational efficiency, and connectivity.

Finally, Table 4 presents the results of the multivariate regression model. Because the model specification was selected considering accessibility as our variable of interest, the coefficients of covariates should be interpreted with caution as they do not necessarily reflect unbiased estimates (Keele et al., 2020). In addition, it is important to note that although there is a small variation in the measures of fit between the models, for the purposes of our study, the significance and magnitude of the accessibility coefficient is what matters the most as we are particularly interested in examining the effect of accessibility on transit ridership and not on the overall model prediction performance.

As indicated by the exploratory analyses above, the results of the association between accessibility and ridership are extremely similar when using either cumulative or gravity impedance functions. For the sake of brevity, we only show the regression results for the analysis calculated using the cumulative function. Furthermore, these results are statistically identical to the results obtained when we use the alternative model specification suggested by the DAG to estimate the total effect of accessibility on ridership (see Appendix A). This suggests the assumptions of the DAG and our analysis are robust, given that we arrive at the same results even when using regressions with alternative model specifications.

The results of the regression model show that the coefficient of the route-level accessibility varies substantially between the four accessibility measures. For the Karner (2018) and Diab et al. (2020a) indicators, the values of the coefficients are statistically significant but negative, which is rather unexpected. As these indicators do not restrict the route considered in their calculation process, they can inadvertently indicate that an inefficient route (e.g. due to low frequency and operational speed) provides high accessibility simply because it has some overlap with more efficient routes. It is notable in Fig. 6 that these indicators are the ones for which some feeder routes have the highest accessibility.

Meanwhile, the proposed indicator and the derivative accessibility metric present positive and statistically significant coefficients. Using the indicator proposed in this article, the results suggest that, when comparing transit routes with similar frequency, and average job and population densities, a route with a 10% higher accessibility would have on average approximately 3.18% higher ridership. By contrast, a 10% increase in the marginal accessibility metric is associated with an increase of only approximately 0.8% in a route's ridership. This seems to suggest that our proposed indicator can more sensitively capture the association between accessibility and ridership at the route level. This is likely because, as discussed in Section 2, the derivative measure underestimates the accessibility of suboptimal routes.

It becomes easier to grasp these coefficients once we understand how transit networks could provide different accessibility levels despite having similar frequency, or average jobs and population density. This can occur because of how route-level accessibility and its effect on ridership are influenced by two key factors. First, the potential benefit generated by a route is largely affected by its connectivity to the rest of the transit network. Routes that are more connected are generally more attractive as they can better expand passengers' access to other routes and places and thus cater to the needs of more potential passengers. Second, although the regression model controls for the average density of jobs and population near each route, the accessibility provided by a transit route is fundamentally affected by how these densities are spatially distributed along its path. For example, transit routes can have similar average population and employment densities and still provide quite different accessibility levels depending on whether they connect suburban neighborhoods to a more distant central business district (CBD) or serve populations closer to the CBD. In other words, a route can provide relatively low access if most jobs are located farther away from the population it serves, or if job opportunities are more concentrated upstream the transit route.

7. Conclusions

In this paper, we proposed a novel route-based accessibility metric. This indicator presents a summary measure of the access to opportunities that each route of a public transit system provides to the population. Differently from other similar indicators found in the literature, the route-level accessibility measure proposed in this study simultaneously accounts for land use patterns (i.e. spatial distribution of population and activities) and for route-specific characteristics in terms of its spatial configuration and stop-spacing, operational performance (speed and frequency) and its spatial and temporal connectivity to the rest of the transit network. Based on a case study in the city of Fortaleza (Brazil), we calculated the level of access to jobs provided by each bus route considering the proposed indicator and three other indicators found in the literature using two types of impedance functions, cumulative and gravitational. Finally, we compared to what extent the results of these indicators would be associated with transit ridership, using an exploratory correlation analysis and a multivariate regression model informed by a directed acyclic graph (DAG) that seeks to map the relevant causal paths mediating the relationship between route-level accessibility and transit ridership.

The results show that the route-level accessibility measure proposed in this paper can better differentiate the level of accessibility between different types of routes than other indicators, with more substantial differences when compared to the indicators proposed by Karner (2018) and Diab et al. (2020a). Our findings also show that, after controlling for route frequency and average jobs and population density near each route, our proposed indicator presents substantially stronger association with transit ridership, regardless of the impedance function considered. Based on this indicator, our findings suggest that a 10% increase in the access to jobs provided by a transit route is associated with approximately 3.18% increase in ridership. These findings suggest that the proposed indicator can more sensitively capture the association between accessibility and ridership at the route level even after controlling for confounding variables.

Compared to the indicators of Karner (2018) and Diab et al. (2020a), these results emerge largely due to two main characteristics that differentiate the accessibility measure proposed in this paper. First, the proposed indicator is able to capture accessibility benefits of using each particular transit route rather than a simple aggregation of the accessibility levels from the places along the route. This is possible because of how the routing analysis used to calculate the proposed accessibility metric allows one to account for route-specific characteristics such as the operational performance and connectivity of each route. A second feature that differentiates our route-level accessibility measure from the indicators used by Karner (2018) and Diab et al. (2020a) is the fact that it accounts for the population that could potentially benefit from a transit route both directly and indirectly. This allows our indicator to more adequately reflect both the downstream and upstream connectivity of transit routes with the rest of the transit network. On the whole, these characteristics could explain why the accessibility measure proposed in this paper has a substantially stronger association with transit ridership.

There are also important differences between the proposed indicator and the derivative accessibility metric. While the derivative measure focuses on quantifying the additional accessibility brought about by a given route, we direct our attention towards assessing the average accessibility benefit the route provides. Although this distinction may seem subtle, it bears significant implications concerning our methods and results. By emphasizing the average accessibility benefit, our proposed measure presents a more nuanced view of the performance and benefits of transit routes, accounting for their overall contribution to accessibility rather than just the incremental gains they provide.

We understand that the derivative accessibility approach can be more appropriate if one's aim is to assess the extent to which transportation projects impact accessibility levels. It is reasonable to consider that a new transport investment is only worthwhile if it enables individuals to access additional opportunities. Nonetheless, if one's aim is to examine what factors have greater impact on transit ridership, and how ridership levels could change as a result of interventions on the existing transit network, the route-level access measure we proposed could be more appropriate. This is largely because the derivative accessibility indicator may introduce certain biases by underestimating the importance of sub-optimal routes.

The new route-level accessibility measure presented in this study has a few limitations. The current version of the indicator considers that the various legs of a trip (e.g. waiting time, transfer time or time in the vehicle) have the same weight, i.e., the user chooses the best route based only on the total travel time. This may not be true since passengers may weigh differently the impedance of each step. The consequence of this may be a reduction in the association between accessibility and ridership. Another limitation is that the current version of the indicator does not take into consideration different modes of transport (e.g. bus, subway, trains, etc.) nor the monetary cost of trips and affordability issues, which could also be incorporated in future work. Further work could investigate the robustness of the indicator in other contexts with different transport network configurations, such as grid-like and predominantly radial networks.

The findings of this study suggest that the proposed indicator could help advance research methods to investigate the determinants of transit ridership. A route-level accessibility metric could also provide new insights for equity analyses that examine how much accessibility different routes convey to different socio-economic and demographic groups. From a policy perspective, this indicator could also help transportation agencies to measure the accessibility benefits conveyed by public transit routes and its associations with transit ridership.

For example, future work can use the indicator to examine how changes in land use patterns (e.g., by the construction of housing developments, zoning changes) could impact the accessibility of different transit routes and changes in transit ridership as a consequence. Similarly, the indicator could also be used by transportation agencies to simulate for example how different scenarios of changes to land use patterns (e.g. changes in population or jobs densities) and to the transit network (e.g. changes in the frequency of particular routes, the redesign of existing routes or creation of new ones) could affect route-level accessibility and what impact on their ridership could be expected.

In this sense, the new indicator could also make an important contribution to transport planning policy by generating relevant information to inform the operational planning of public transit systems. As the indicator generates accessibility estimates that directly concern the unit of analysis and planning of transit networks, it is expected that this indicator could help practitioners integrate the concept of accessibility into the day-to-day work of planning transportation systems.

CRediT authorship contribution statement

João Lucas Albuquerque-Oliveira: Conceptualization, Methodology, Data curation, Software, Writing – original draft, Visualization, Formal analysis. **Francisco Moraes Oliveira-Neto:** Conceptualization, Methodology, Writing – review & editing, Supervision. **Rafael H.M. Pereira:** Conceptualization, Methodology, Data curation, Software, Writing – review & editing, Supervision.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Table 5
Summary of the multivariate regression of the effect of accessibility on the number of passengers, considering different accessibility metrics at the route level and including income as a covariate.

Indicator	Karner (2018)	Diab et al. (2020a)	Liu and Miller (2023)	Proposed indicator
Constant	**3.604*****	**3.033*****	**0.941***	**0.158**
	(0.635)	(0.452)	(0.428)	(0.414)
Route accessibility	**−0.307*****	**−0.735*****	**0.094*****	**0.369*****
	(0.049)	(0.077)	(0.022)	(0.072)
Route Frequency	**1.076*****	**1.103*****	**0.914*****	**0.807*****
	(0.061)	(0.056)	(0.064)	(0.071)
Jobs on route	**0.483*****	**0.652*****	**0.273*****	**0.055**
	(0.047)	(0.050)	(0.051)	(0.077)
Average income on route	**−0.169+**	**0.051**	**−0.210***	**−0.137**
	(0.090)	(0.088)	(0.094)	(0.095)
Num.Obs.	245	245	245	245
R2	0.629	0.687	0.598	0.609
R2 Adj.	0.622	0.682	0.592	0.603
AIC	258.7	216.7	278.0	271.0

**p < 0.01.
\+ p < 0.1.
* p < 0.05.
*** p < 0.001.

Appendix A

Table A1

References

Adler, M.W., van Ommeren, J.N., 2016. Does public transit reduce car travel externalities? Quasi-natural experiments' evidence from transit strikes. J. Urban Econ. 92, 106–119. https://doi.org/10.1016/j.jue.2016.01.001.

Arriagada, J., Munizaga, M.A., Guevara, C.A., Prato, C., 2022. Unveiling route choice strategy heterogeneity from smart card data in a large-scale public transport network. Transp. Res. Part C Emerg. Technol. 134, 103467 https://doi.org/10.1016/j.trc.2021.103467.

Aston, L., Currie, G., Delbosc, A., Kamruzzaman, Md., Teller, D., 2021. Exploring built environment impacts on transit use – an updated meta-analysis. Transp. Rev. 41, 73–96. https://doi.org/10.1080/01441647.2020.1806941.

Balcombe, R., Mackett, R., Paulley, N., Preston, J., Shires, J., Titheridge, H., Wardman, M., White, P., 2004. The demand for public transport: a practical guide.

Beaudoin, J., Farzin, Y.H., Lin Lawell, C.-Y.-C., 2015. Public transit investment and sustainable transportation: a review of studies of transit's impact on traffic congestion and air quality. Res. Transp. Econ. 52, 15–22. https://doi.org/10.1016/j.retrec.2015.10.004.

Bertolini, L., le Clercq, F., Kapoen, L., 2005. Sustainable accessibility: a conceptual framework to integrate transport and land use plan-making. Two test-applications in the Netherlands and a reflection on the way forward. Transp. Policy 12, 207–220. https://doi.org/10.1016/j.tranpol.2005.01.006.

Boisjoly, G., El-Geneidy, A., 2016. Daily fluctuations in transit and job availability: a comparative assessment of time-sensitive accessibility measures. J. Transp. Geogr. 52, 73–81. https://doi.org/10.1016/j.jtrangeo.2016.03.004.

Boisjoly, G., El-Geneidy, A.M., 2017. How to get there? A critical assessment of accessibility objectives and indicators in metropolitan transportation plans. Transp. Policy 55, 38–50. https://doi.org/10.1016/j.tranpol.2016.12.011.

Boisjoly, G., Grisé, E., Maguire, M., Veillette, M.-P., Deboosere, R., Berrebi, E., El-Geneidy, A., 2018. Invest in the ride: a 14 year longitudinal analysis of the determinants of public transport ridership in 25 North American cities. Transp. Res. Part Policy Pract. 116, 434–445. https://doi.org/10.1016/j.tra.2018.07.005.

Brakewood, C., Macfarlane, G.S., Watkins, K., 2015. The impact of real-time information on bus ridership in New York City. Transp. Res. Part C Emerg. Technol. 53, 59–75. https://doi.org/10.1016/j.trc.2015.01.021.

Bree, S., Fuller, D., Diab, E., 2020. Access to transit? Validating local transit accessibility measures using transit ridership. Transp. Res. Part Policy Pract. 141, 430–442. https://doi.org/10.1016/j.tra.2020.09.019.

Cao, X., Mokhtarian, P.L., Handy, S.L., 2008. Examining the impacts of residential self-selection on travel behavior: methodologies and empirical findings.

Cats, O., West, J., Eliasson, J., 2016. A dynamic stochastic model for evaluating congestion and crowding effects in transit systems. Transp. Res. Part B Methodol. 89, 43–57. https://doi.org/10.1016/j.trb.2016.04.001.

Cervero, R., 2007. Transit-oriented development's ridership bonus: a product of self-selection and public policies. Environ. Plan. Econ. Space 39, 2068–2085. https://doi.org/10.1068/a38377.

Cui, B., DeWeese, J., Wu, H., King, D.A., Levinson, D., El-Geneidy, A., 2022. All ridership is local: accessibility, competition, and stop-level determinants of daily bus boardings in Portland. Oregon. J. Transp. Geogr. 99, 103294 https://doi.org/10.1016/j.jtrangeo.2022.103294.

Ortúzar, J. de D., Willumsen, L.G., Consultancy, L.W., 2011. Modelling Transport. SPH P 608.

Delling, D., Pajor, T., Werneck, R.F., 2015. Round-based public transit routing. Transp. Sci. 49, 591–604. https://doi.org/10.1287/trsc.2014.0534.

Diab, E., DeWeese, J., Chaloux, N., El-Geneidy, A., 2020a. Adjusting the service? Understanding the factors affecting bus ridership over time at the route level in Montréal, Canada. Transportation 48, 2765–2786. https://doi.org/10.1007/s11116-020-10147-3.

Diab, E., Kasraian, D., Miller, E.J., Shalaby, A., 2020b. The rise and fall of transit ridership across Canada: understanding the determinants. Transp. Policy 96, 101–112. https://doi.org/10.1016/j.tranpol.2020.07.002.

Drabicki, A., Cats, O., Kucharski, R., Fonzone, A., Szarata, A., 2023. Should I stay or should I board? Willingness to wait with real-time crowding information in urban public transport. Res. Transp. Bus. Manag. 100963 https://doi.org/10.1016/j.rtbm.2023.100963.

El-Geneidy, A., Levinson, D., 2022. Making accessibility work in practice. Transp. Rev. 42, 129–133. https://doi.org/10.1080/01441647.2021.1975954.

Ewing, R., Cervero, R., 2001. Travel and the built environment: a synthesis. Transp. Res. Rec. 1780, 87–114. https://doi.org/10.3141/1780-10.

Ewing, R., Cervero, R., 2010. Travel and the built environment. J. Am. Plann. Assoc. 76, 265–294. https://doi.org/10.1080/01944361003766766.

Farber, S., Fu, L., 2017. Dynamic public transit accessibility using travel time cubes: comparing the effects of infrastructure (dis)investments over time. Comput. Environ. Urban Syst. 62, 30–40. https://doi.org/10.1016/j.compenvurbsys.2016.10.005.

Fransen, K., Farber, S., Deruyter, G., De Maeyer, P., 2018. A spatio-temporal accessibility measure for modelling activity participation in discretionary activities. Travel Behav. Soc. 10, 10–20. https://doi.org/10.1016/j.tbs.2017.09.002.

Geurs, K.T., van Wee, B., 2004. Accessibility evaluation of land-use and transport strategies: review and research directions. J. Transp. Geogr. 12, 127–140. https://doi.org/10.1016/j.jtrangeo.2003.10.005.

Huntington-Klein, N., 2021. The Effect: An Introduction to Research Design and Causality | The Effect.

IBGE, 2020. Fortaleza (CE) | Cidades e Estados | IBGE [WWW Document]. URL https://www.ibge.gov.br/cidades-e-estados/ce/fortaleza.html (accessed 2.28.23).

Karner, A., 2018. Assessing public transit service equity using route-level accessibility measures and public data. J. Transp. Geogr. 67, 24–32. https://doi.org/10.1016/j.jtrangeo.2018.01.005.

Keele, L., Stevenson, R.T., Elwert, F., 2020. The causal interpretation of estimated associations in regression models. Polit. Sci. Res. Methods 8, 1–13. https://doi.org/10.1017/psrm.2019.31.

Liu, L., Miller, H.J., 2023. Accessibility derivative: measuring the accessibility contribution of public transit routes. In: Presented at the 2023 AAG Annual Meeting.

Manaugh, K., Badami, M.G., El-Geneidy, A.M., 2015. Integrating social equity into urban transportation planning: a critical evaluation of equity objectives and measures in transportation plans in North America. Transp. Policy 37, 167–176. https://doi.org/10.1016/j.tranpol.2014.09.013.

Martens, K., 2019. Why accessibility measurement is not merely an option, but an absolute necessity. In: Silva, C., Bertolini, L., Pinto, N. (Eds.), Designing Accessibility Instruments. Routledge, Routledge, New York, NY, pp. 37–51. https://doi.org/10.4324/9781315463612-4.

Merlin, L.A., Singer, M., Levine, J., 2021. Influences on transit ridership and transit accessibility in US urban areas. Transp. Res. Part Policy Pract. 150, 63–73. https://doi.org/10.1016/j.tra.2021.04.014.

Miller, E.J., Shalaby, A., Ehab, D., Kasraian, D., 2018. Canadian Transit Ridership Trends Study.

Niehaus, M., Galilea, P., Hurtubia, R., 2016. Accessibility and equity: an approach for wider transport project assessment in Chile. Res. Transp. Econ. 59, 412–422. https://doi.org/10.1016/j.retrec.2016.05.003.

Pasfor, 2019. PASFOR | Plano de Acessibilidade Sustentável de Fortaleza [WWW Document]. URL https://www.pasfor.com.br/ (accessed 3.7.23).

Pereira, R.H.M., Banister, D., Schwanen, T., Wessel, N., 2019. Distributional effects of transport policies on inequalities in access to opportunities in Rio de Janeiro. J. Transp. Land Use 12. https://doi.org/10.5198/jtlu.2019.1523.

Pereira, R.H.M., Saraiva, M., Herszenhut, D., Braga, C.K.V., Conway, M.W., 2021. r5r: Rapid realistic routing on multimodal transport networks with R [5] in R. findings. https://doi.org/10.32866/001c.21262.

Rahman, F., Mazumder, R.J.R., Shamsul Kabir, Md., Hadiuzzaman, Md., 2020. An exploratory analysis of factors affecting comfort level of work trip chaining and mode choice: a case study for Dhaka City. Transp. Dev. Econ. 6, 11. https://doi.org/10.1007/s40890-020-0095-3.

Raveau, S., Guo, Z., Muñoz, J.C., Wilson, N.H.M., 2014. A behavioural comparison of route choice on metro networks: time, transfers, crowding, topology and socio-demographics. Transp. Res. Part Policy Pract. 66, 185–195. https://doi.org/10.1016/j.tra.2014.05.010.

Schwanen, T., Dijst, M., Dieleman, F.M., 2002. A microlevel analysis of residential context and travel time. Environ. Plan. Econ. Space 34, 1487–1507. https://doi.org/10.1068/a34159.

Shrier, I., Platt, R.W., 2008. Reducing bias through directed acyclic graphs. BMC Med. Res. Methodol. 8, 70. https://doi.org/10.1186/1471-2288-8-70.

Silva, C., Larsson, A., 2018. Challenges for accessibility planning and research in the context of sustainable mobility (Discussion paper No. 2018/07). International Transport Forum Discussion Papers. https://doi.org/10.1787/8e37f587-en.

Silva, C., Bertolini, L., te Brömmelstroet, M., Milakis, D., Papa, E., 2016. Accessibility instruments in planning practice: bridging the implementation gap. Transp. Policy 53, 135–145. https://doi.org/10.1016/j.tranpol.2016.09.006.

Siqueira, M.F., 2020. Metodologia de análise dos determinantes da demanda por transportes no paradigma da inferência causal.

Tang, L., Thakuriah (Vonu), P., 2012. Ridership effects of real-time bus information system: a case study in the City of Chicago. Transp. Res. Part C Emerg. Technol. 22, 146–161. https://doi.org/10.1016/j.trc.2012.01.001.

Taylor, B.D., Fink, C.N.Y., 2012. Explaining transit ridership: What has the evidence shown? Transp. Lett. 5, 15–26. https://doi.org/10.1179/1942786712Z.0000000003.

Taylor, B.D., Miller, D., Iseki, H., Fink, C., 2009. Nature and/or nurture? Analyzing the determinants of transit ridership across US urbanized areas. Transp. Res. Part Policy Pract. 43, 60–77. https://doi.org/10.1016/j.tra.2008.06.007.

Textor, J., van der Zander, B., Gilthorpe, M.S., Liśkiewicz, M., Ellison, G.T.H., 2017. Robust causal inference using directed acyclic graphs: the R package 'dagitty'. Int. J. Epidemiol. dyw341 https://doi.org/10.1093/ije/dyw341.

UBER, 2018. H3: Uber's Hexagonal Hierarchical Spatial Index [WWW Document]. Uber Blog. URL https://www.uber.com/blog/h3/ (accessed 3.7.23).

van Wee, B., 2016. Accessible accessibility research challenges. J. Transp. Geogr. 51, 9–16. https://doi.org/10.1016/j.jtrangeo.2015.10.018.

Vredin Johansson, M., Heldt, T., Johansson, P., 2006. The effects of attitudes and personality traits on mode choice. Transp. Res. Part Policy Pract. 40, 507–525. https://doi.org/10.1016/j.tra.2005.09.001.