



# Exploring the time geography of public transport networks with the `gtfs2gps` package

Rafael H. M. Pereira<sup>1</sup> · Pedro R. Andrade<sup>2</sup> · João Pedro Bazzo Vieira<sup>1</sup>

Received: 10 May 2022 / Accepted: 16 November 2022

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022

## Abstract

The creation of the General Transit Feed Specification (GTFS) in the mid-2000s provided a new data format for cities to organize and share digital information on their public transport systems. GTFS feeds store geolocated data on public transport networks, including information on routes, stops, timetables, and service levels. The GTFS standard is now widely adopted by thousands of transport authorities and a wide variety of software applications for different purposes, including trip planning, timetable creation and accessibility analysis. Yet, there is still a lack of tools to parse GTFS data when the objective is to analyze the complex spatial and temporal patterns of public transport systems. This paper presents `{gtfs2gps}`, a new general-purpose computational tool to easily process static GTFS data that allows one to analyze the space–time trajectories of public transport vehicles at fine spatial and temporal resolutions. `{gtfs2gps}` is an open-source R package that employs parallel computing to convert GTFS feeds from relational text files into a trajectory data table, similar to GPS records, with the timestamps of vehicles in every trip. This paper explains the package functionalities and demonstrates how `{gtfs2gps}` can be used to articulate key concepts in time geography to explore and visualize the spatial and temporal patterns of public transport networks. We also present a case study looking at how `{gtfs2gps}` can be used to examine socioeconomic and spatial–temporal inequalities in access to public transport, providing key information to monitor cities’ progress toward the Sustainable Development Goals. The paper is accompanied by a computational notebook in R Markdown to support reproducibility of the results in this paper and to replicate the analysis for other contexts where GTFS data are available. Given the widespread use of GTFS, `{gtfs2gps}` opens new possibilities for researchers to examine the time geography of public transport systems in urban areas across the globe.

**Keywords** GTFS · Time geography · Public transport · Transport networks · `gtfs2gps`

---

✉ Rafael H. M. Pereira  
rafael.pereira@ipea.gov.br

Extended author information available on the last page of the article

## 1 Introduction

The digitization of modern society has substantially changed how we understand the relationships between people, places, and mobility patterns. This has been no different when it comes to how we plan and organize public transport services in our cities. The creation of the General Transit Feed Specification (GTFS) standard in the mid-2000s (Antrim and Barbeau 2017) provided a new data format for cities to organize and share digital information on their public transport systems. The popularity of GTFS is due to its simplicity and machine-readability. GTFS feeds are now widely adopted by thousands of transport agencies worldwide,<sup>1</sup> and a wide variety of commercial and open-source software have been developed on top of the GTFS format to perform trip planning (e.g., Google Maps, Moovit), routing and accessibility analysis (OpenTripPlanner, R5), and timetable creation and management (TransLoc, Trillium). Yet, there are still just a handful of tools to parse GTFS data in a simple and convenient way (Herszenhut et al. 2021; Poletti et al. 2021), in particular when the objective is to analyze the complex spatial and temporal patterns of public transport systems as well as the space–time trajectories of vehicles at fine spatial and temporal resolutions.

This short paper presents `{gtfs2gps}`, an R package to convert static GTFS feeds from relational text files into a trajectory data table, similar to GPS records, allowing one to keep track and analyze the space–time properties of vehicles, trips, routes, and transport systems. We introduce `{gtfs2gps}` functionalities and demonstrate how they can be used to process GTFS data and articulate key concepts in time geography to examine the spatial and temporal patterns of public transport systems. Time geography provides a conceptual framework for investigating how human activities and mobility patterns are physically constrained or shaped by space and time (Miller 2017). The original development of time geography in the 1970s by Hägerstrand (1970) was a feat of imagination that allowed scientists to understand the space–time fabric of human mobility at a time when there was so little data and computing power available. Since then, time geography has been increasingly applied to study animal movement ecology (Demšar et al. 2015) and, more commonly, human behavior movement embedded in transport networks (Long and Nelson 2013; Neutens et al. 2011).

The growing use of time geography has been fueled by the increasing availability of GPS trajectory data from vehicles and smart phones (Song and Miller 2012; Yang et al. 2017, 2020) as well as public transport in GTFS format (Lee and Miller 2018, 2020). It has also been made possible due to the remarkable advances in geographic data science and open-source development seen in recent years (Singleton and Arribas-Bel 2021), with the growing recognition of open-source software

---

<sup>1</sup> GTFS feeds of several cities can be downloaded from repositories such as <https://www.transit.land> or <https://transitfeeds.com>.

development as a vital part of academic research to support open reproducible science (Arribas-Bel et al. 2021; Rey 2009; Wolf et al. 2021). This paper falls into a growing literature developing open-source applications for geographical analysis for transport planning (Lovelace 2021). The {gtfs2gps} tool presented here provides a simple way to overcome the technical challenge that is to take flat GTFS files and unravel the space–time patterns of public transport networks, thus helping to better represent the time geography of Hägerstrand’s imagination. This paper has a practical purpose to show how {gtfs2gps} can be used to visualize the space–time trajectories and patterns of public transport systems. For a more in-depth exploration of the theoretical and mathematical underpinnings of time geography concepts, see the works of Miller (2005, 2017) and Sui (2012).

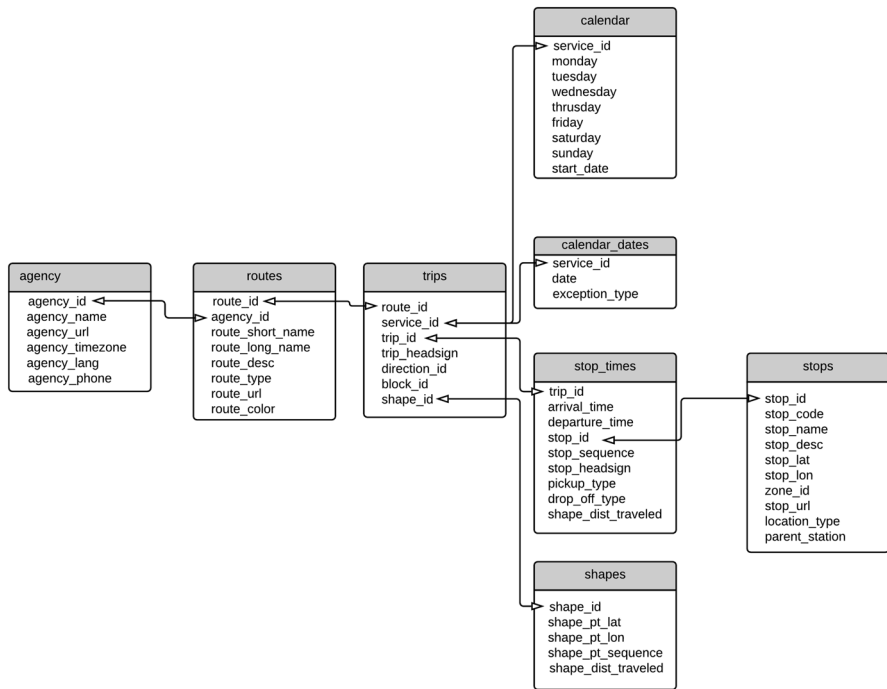
To support reproducibility (see Páez 2021; Brunsdon and Comber 2021), this paper is accompanied by a computational notebook in R Markdown that allows one to easily reproduce the results presented in this paper and replicate the analysis and data visualizations for other contexts where GTFS data are available.

Apart from this introduction, this paper is organized as follows: Section 2 presents the basic structure of the GTFS data format. Section 3 gives a general explanation of the main algorithm behind the core function of the {gtfs2gps} package. Section 4 presents a few examples and data visualizations that illustrate how one can analyze GTFS data with {gtfs2gps} to shed light on the time geography of public transport systems. Finally, Sect. 5 presents some final remarks with reflections on the time geography of public transport networks. The computational notebook to reproduce the analysis and visualization in this study is available as supplementary material, also available as a GitHub repository: [https://github.com/ipeaGIT/gtfs2gps-time\\_geography](https://github.com/ipeaGIT/gtfs2gps-time_geography).

## 2 GTFS data

The General Transit Feed Specification (GTFS) is a standard data format to store and organize data on the infrastructure and planning services of public transport systems. A GTFS feed is a zipped archive that gathers multiple text files with geolocated data and detailed information on routes, stops, timetables, and service levels (Fig. 1). Each file in a GTFS feed is a plain text file with standard columns and formats. Similar to a relational database, key columns link the information between routes, trips, timetables, stops, etc. These key columns are highlighted as the end-points of the arrows in Fig. 1. GTFS is an open and collaborative data specification format.

The “stops.txt” file, for example, brings the spatial coordinates of all the stops and stations, which are uniquely identified by the “stop\_id” column. The “routes.txt” file defines every unique route in the public transport system with the column `route\_id`, and brings information such as the transport mode (column ‘route\_type’) and the transport agency (‘agency\_id’) running each particular route. Meanwhile, the “trips.txt” file brings information for every trip in the system. Every trip has a unique id (‘trip\_id’) and it is associated with several other information such as the days of the week when that trip occurs (specified by the ‘service\_id’ column) and



**Fig. 1** Schema of the general transit feed specification (GTFS) data format

the spatial trajectory if that trip (specified by the ‘shape\_id’ column). The detailed information of the weekly time table of the service is presented in detail in the “calendar.txt” file, while spatial coordinates of the trip trajectory are presented in the “shapes.txt” file. The “stop\_times.txt” file, in turn, brings information on what are the sequence of stops served by each trip at what are the expected arrival and departure times from each stop. The frequency of each trip can often be derived from the “stop\_times.txt” file but some GTFS feeds also include a “frequencies.txt” file with that information. It is the key columns in each file that allow {gtfs2gps} and other computational packages such as gtfstools (Herszenhut et al. 2021), r5r (Pereira et al. 2021) and OpenTripPlanner to associate and connect information between routes, trips, stops, timetables, etc. Figure 1 only illustrates a few of the most important files and columns but more information about other files and extensions of the GTFS standard can be found at <https://gtfs.org/>.

This description applies mostly to GTFS static feeds, and not to GTFS real-time, an extension of the GTFS format that uses automatic vehicle locating data to provide real-time information about current arrival and departure times of vehicles. A caveat about using static GTFS data is that they are based on information on scheduled services. Consequently, the information on this type of dataset typically does not consider the variability of travel time estimates due to unplanned factors such as a temporary decrease in road capacity, traffic crashes, or nonrecurrent congestion

levels. Recurrent congestion, nonetheless, is normally accounted for by transport authorities in the planning of timetables.

The relative simplicity of the GTFS format, its ability to capture numerous entities in space and time efficiently in a small number of files, has contributed to its popularity among transport agencies. The widespread use of the GTFS standard has also helped to create a common ground for developers, researchers, and planners to create and share computational tools to process and analyze public transport networks in multiple contexts across the globe. In recent years, several computational tools have been developed to manipulate GTFS data, such as *gtfstools* (Herszenhut et al. 2021) *tidytransit* (Poletti et al. 2021), or to perform routing and accessibility analysis, including *opentripplanner* (Morgan et al. 2021) and *r5r* (Pereira et al. 2021). More recently, Dardas et al. (2022) developed a geospatial workflow to calculate different reliability metrics and speed performance of public transit systems based on real-time GTFS feeds. The `{gtfs2gps}` package complements these existing tools currently by introducing tools to convert static GTFS feeds from relational text files into structured GPS-like data tables that can be more easily handled, processed, and analyzed.

### 3 The `{gtfs2gps}` package

While the simplicity of the GTFS data format makes it easily machine-readable, the structure of relational tables is less human-friendly. The package `{gtfs2gps}` converts GTFS feeds from relational text files into a data table similar to GPS records with timestamps that represent the trajectories of vehicles. The package is developed in R, an open-source programming language, and it leverages parallel computing for efficient computation <https://github.com/ipeaGIT/gtfs2gps>.

The main function of the package is named ``gtfs2gps``, after the package. The function has two main inputs, a GTFS zipped file and the maximum spatial resolution of the output. The function executes the following steps:

1. *Spatial sampling of points along route shapes* for each transport route in the GTFS, the function samples several spatial points along the route shape to guarantee that the distance between two consecutive points will never be greater than the spatial resolution set by the user. By default, the spatial resolution is set to 100 m. Finer resolutions lead to more accurate representation of vehicles' spatial trajectory at the cost of computation performance.
2. *Snap stops to route shapes* oftentimes, the geographical coordinates of stops in GTFS feeds are not precise. To address this issue, the function snaps public transport stops to the nearest point in the route shape, keeping the ordered sequence of stops.
3. *Calculate network distance between stops* next, the function uses the data generated in the two previous steps to calculate the network distance between each pair of consecutive stops and the total distance of each trip. This is done considering the Euclidean distance between consecutive points sampled along the route in

step one. Using finer spatial resolutions in step one lead to more accurate distance estimates along the street network.

4. *Calculate trip duration* trip duration is calculated based on the timetable recorded in the GTFS, indicating the expected time of arrival and departure of the vehicle for each trip at each stop. The function calculates the travel duration between each pair of consecutive stops and for the whole trip.
5. *Calculate trip speeds* next, we use the data generated in steps (3) and (4) to calculate the vehicle speed between each pair of consecutive stops for each trip. The package assumes each trip presents a constant average speed between two consecutive stops, but it captures how speed values can vary along the route and at different times of the day based on the arrival and departure times from the GTFS input data. Oftentimes, though, GTFS feeds only have information about the departure time from the first stop and the arrival time at the last stop of each trip. In these cases, we assume the trip will have a constant average speed, which is automatically calculated considering the total length and duration of the trip. While this is a reasonable simplification for most trips, it might generate additional data inaccuracies for routes entering and/or leaving highly congested segments.
6. *Interpolate space–time positions* finally, the {gtfs2gps} interpolates the space–time position of each vehicle in each trip at the spatial resolution set by the user by considering the network distance and speed between stops. In other words, the function estimates the timestamp when each vehicle will be at each of the spatial points sampled along with the route shape, created in step 1.

The ‘gtfs2gps()’ function can run this sequence of steps in parallel on all route shapes and trips present in a GTFS feed. The output of the function is a trajectory data table with a similar structure to traditional GPS data sets. The Code block 1 shows how a GTFS feed can be processed to generate a table with GPS-like records using the ‘gtfs2gps()’ function, and how the output looks like. In this paper, we use a GTFS feed published by the Urban Transport Metropolitan Agency (EMTU) of São Paulo, Brazil, valid from June/2019 until January/2020. This GTFS feed is available as supplementary material for this paper. For this article, we used {gtfs2gps} version v2.0-2.

**Code block 1: Load libraries and convert a GTFS data feed into a data table with GPS-like format.**

```
# Load gtfs2gps other library
library(gtfs2gps)

# address of gtfs file
gtfs_file <- system.file("extdata/saopaulo.zip", package = "gtfs2gps")

# convert gtfs to gps-like data table
df <- gtfs2gps::gtfs2gps('gtfs.zip', spatial_resolution = 100, parallel = TRUE)
head(df$shapes)
```

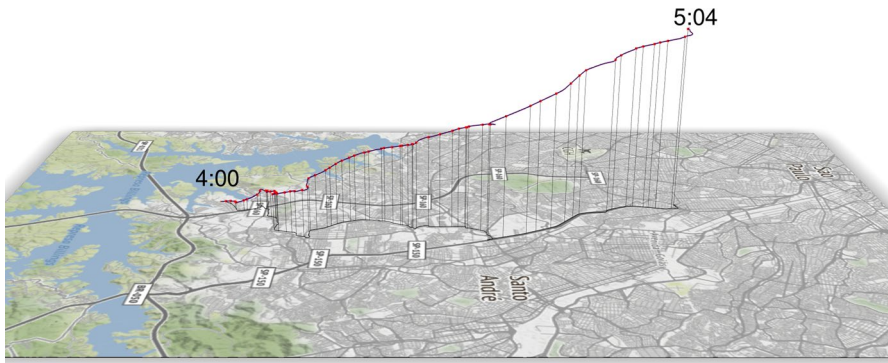
	shape_id	shape_pt_lat	shape_pt_lon	shape_pt_sequence
1:	52421	-23.52212	-46.56379	1
2:	52421	-23.52157	-46.56276	2
3:	52421	-23.52129	-46.56222	3
4:	52421	-23.52124	-46.56212	4
5:	52421	-23.52108	-46.56183	5
6:	52421	-23.52265	-46.56161	6

```

  id shape_id trip_id trip_number route_type shape_pt_lon shape_pt_lat departure_time stop_id
1: 1 50784 407E-10-0#1 1 3 -46.44303 -23.63370 00:00:00 770007651
2: 2 50784 407E-10-0#1 1 3 -46.44290 -23.63336 00:00:26 <NA>
3: 3 50784 407E-10-0#1 1 3 -46.44285 -23.63315 00:00:42 <NA>
4: 4 50784 407E-10-0#1 1 3 -46.44270 -23.63269 00:00:52 770007650
5: 5 50784 407E-10-0#1 1 3 -46.44269 -23.63263 00:00:53 <NA>
6: 6 50784 407E-10-0#1 1 3 -46.44275 -23.63234 00:01:00 <NA>
  stop_sequence dist cumdist cumtime speed
1: 1 0.000000 [m] 0.00000 [m] 0.00000 [s] 5.558455 [km/h]
2: NA 39.873149 [m] 39.87315 [m] 25.82432 [s] 5.558455 [km/h]
3: NA 24.678450 [m] 64.55160 [m] 41.80762 [s] 5.558455 [km/h]
4: 2 52.793554 [m] 117.34515 [m] 52.01429 [s] 18.620850 [km/h]
5: NA 6.587615 [m] 123.93277 [m] 53.28788 [s] 18.620850 [km/h]
6: NA 33.248587 [m] 157.18135 [m] 59.71588 [s] 18.620850 [km/h]
```

## 4 The time geography of public transport networks

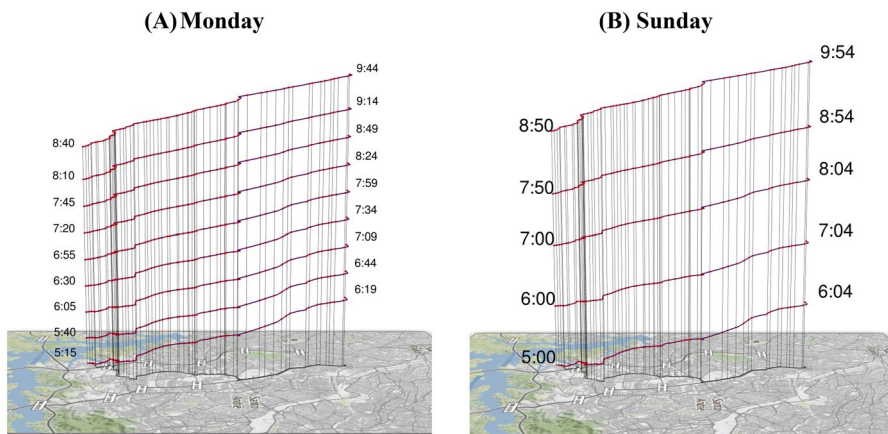
Geovisualization is an important element in time geography research (Kwan, 2004). Package {gtfs2gps} makes it possible to visualize and explore transit trajectories in ways that were not easily done before. The output of the {gtfs2gps} allows one to examine the individual space–time path of a public transport trip using a 3-D space–time aquarium, as originally proposed by Hägerstrand (Kwan 2000). Figure 2 shows the space–time trajectory of a given bus trip in the public transport system of São Paulo. In this figure, time is represented on a third orthogonal axis, while bus stops are represented with red dots. In traditional public transport systems with fixed routes, vehicles should follow a fixed itinerary with a sequence of stops according to a predetermined level of service. The boarding/alighting of passengers can be understood as ‘fixed activities’ that occur at public transport stations. These stations are sparsely distributed in space and time, where vehicles are expected to arrive at and depart from according to a planned timetable recorded in the GTFS feed. In this sense, it is possible to say that, from a transport planning perspective, the programmed boarding/alighting of passengers for a given trip acts as a set of



**Fig. 2** Space–time path of a single public transport trip in São Paulo, Brazil

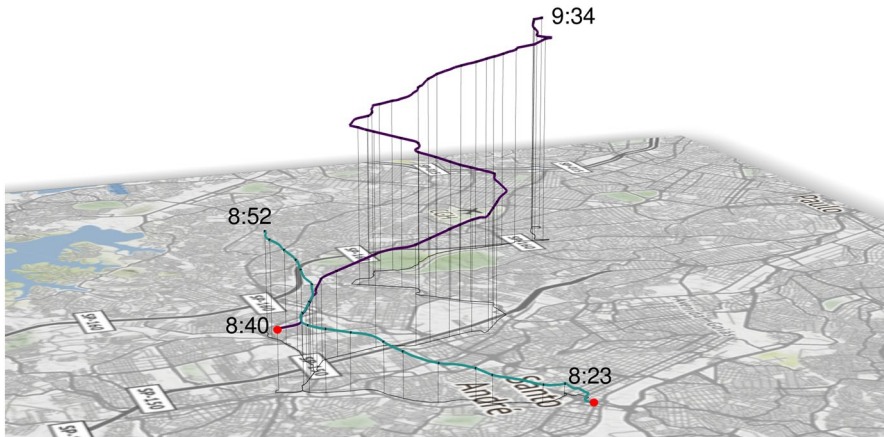
‘space–time anchors’ with predetermined timing and limited durations at relatively few discrete locations in space and time. The duration of boarding time is equivalent to a ‘stationary activity time’. Figures 2, 3 and 4 are presented as static images, but the code to create these figures allows for interactive visualization of the maps. The code necessary to replicate all the analyses and maps presented in this study is available as a computational notebook in R Markdown from the GitHub repository [https://github.com/ipeaGIT/gtfs2gps-time\\_geography](https://github.com/ipeaGIT/gtfs2gps-time_geography).

Each public transport route consists of multiple trips departing at different times of the day. This trip frequency set by transport agencies can also be visualized in a 3-D space–time aquarium with a sequence of space–time paths, one for each trip, stacked over time. Figure 3 shows the space–time paths of the trips of a same route between 7:00 am and 10:00 am on Monday and Sunday to illustrate how differences in frequency levels for a given route can be visualized. Vertical spacing between each departure time in the z-axis gives us a compelling visual understanding of



**Fig. 3** Space–time paths of multiple trips of a public transport route between 5 and 10 am in São Paulo, Brazil





**Fig. 4** Bundling of space–time paths of routes/trips that arrive at a selected bus stop within a 10-min window. São Paulo, Brazil. Note: Red points indicate the start position of each trip (color figure online)

how trip frequencies vary throughout the day (Fig. 3A) and across days of the week (Fig. 3B).

Another key concept in time geography is the notion of *path bundling*, which refers to the spatial and temporal intersection or convergence of two or more space–time paths (Miller 2005, 2017). Path bundling articulates the potential co-location of agents to jointly participate in a given activity. After parsing the GTFS data generated by {gtfs2gps}, one can query for example all the bus trips that would allow two or more individuals to meet at a given stop within a certain time window. In this example, we find trips arriving at a selected bus stop within a 10 min time window, between 08:41 am and 08:50 am. This type of query tells us at what time each individual would have to start their trips and which public transport routes they should take to meet at an agreed stop within that given time window. The bundling of the space–time paths of these trips is shown in Fig. 4.

## 5 Application

In this section, we illustrate how the {gtfs2gps} package can be used to address a particular research question: How do public transport service levels differ for different income groups at different times of the day?

Providing access to sustainable transport systems is instrumental in promoting economic development and making cities more inclusive and sustainable (Banister 2002; Holden et al. 2018). Recognizing this, the United Nations included among its Sustainable Development Goals (SDGs) the target to provide convenient access to public transport with special attention to the needs of those in vulnerable situations (United Nations 2015). As local governments increasingly incorporate the SDGs into their policy agendas, we need to grapple with the question of how ‘convenient’ access to public transport should be defined.

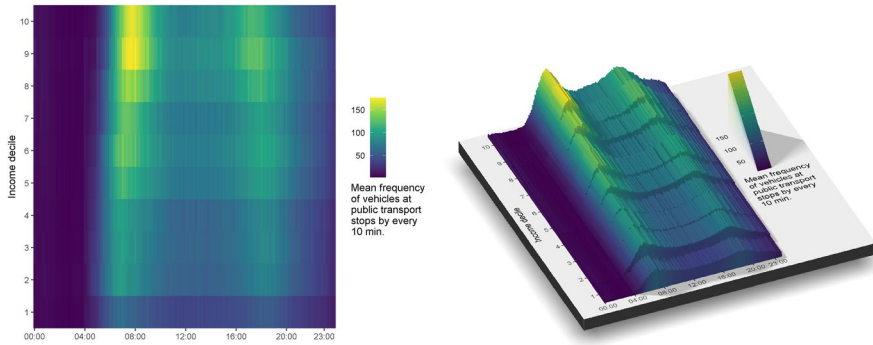
Defining minimum levels of public transport provision is an open-ended question in the transportation equity literature (Lucas et al. 2016; Pereira and Karner 2021). Nonetheless, a general recommendation by UN-Habitat (2018) is that a definition of ‘convenient’ access to public transport should, whenever possible, be based on at least two aspects. The first one is spatial proximity. The UN-Habitat (2018) recommends that people’s distance to public transport stops should be shorter than 500 m (considering either Euclidean or network distance). The second aspect is the temporal availability of public transport, which speaks directly to the level of service frequency and how it impacts people’s waiting times. Here, there is no specific frequency level recommended by UN-Habitat or agreed upon in the literature. Whatever threshold values of proximity and service frequency are chosen to define convenient access, understanding how service levels differ for different socioeconomic groups and at different times of the day becomes critical from an equity point of view.

Here, we demonstrate how the {gtfs2gps} package can help researchers and policymakers examine the socioeconomic and spatial–temporal inequalities in access to public transport services. As a case study, we analyzed the GTFS data of the public transport system of São Paulo (Brazil) on a typical business day in October 2019. Data on the spatial distribution of the population and socioeconomic characteristics come from the latest census of the Brazilian population, conducted in 2010. Population and socioeconomic data were spatially aggregated over an H3 hexagonal grid at resolution 8, where each cell has an area of 0.74 km<sup>2</sup>; this is approximately the size of a typical city block. Each hexagonal cell was classified by income decile based on the average household income per capita of its resident population.

The first step of the data analysis was to run the ‘gtfs2gps()’ function, converting the GTFS of São Paulo into a trajectory data table. This allowed us to calculate the frequency of public transport vehicles at every single public transport stop at 10-min intervals throughout the day. Next, we used a spatial intersection operation to determine the hexagonal cells in which each stop falls into, and associate each stop with the income level of the population in its closest proximity. The code used in these operations is presented in the supplementary material.

By merging these intermediary results, we can calculate the average frequency of public transport services at stops that intersect hexagons of different income levels across different times of the day. These estimates are presented in 2D and 3D in Fig. 5, where the *Y* axis presents information on income levels, the *X* axis represents the time of the day divided into 10-min intervals, and values in the *Z* axis and color vary according to public transport frequency levels.

As expected, the public transport services in São Paulo are planned with higher frequency at the morning peak (between 6 and 9 am) and afternoon peak (between 5 and 7 pm) to cope with higher passenger demand during these periods. However, the level of service frequency is remarkably higher among higher-income groups (deciles 9–10). During the morning peak time, the average service frequency in the 20% richest areas was 4.904 vehicles per hexagon; this is 28% higher than the average frequency provided to low-income regions between the first and fifth income deciles (3823 vehicles per hexagon). This difference is more striking at the morning peak, but it persists throughout the day, except for late evenings (between midnight



**Fig. 5** 2D and 3D representation of the average frequency of public transport services in regions of different income levels across different times of the day, São Paulo, 2019

and 4 am) when there are virtually no services operating. Additionally, we can also see that the peak service in low-income areas (deciles 1–4) starts slightly earlier than the peak service for high-income groups. This is possibly planned to accommodate the fact that low-income communities in São Paulo generally live in urban peripheries and have to leave early to work/study due to significantly longer commute times (Pereira and Schwanen 2015).

These results show remarkable inequalities in the spatial and temporal access to public transport services in São Paulo. It also shows how these inequalities are persistent through most of the day, particularly during business hours when public transport is critical for low-income communities to access employment opportunities, education, and health services. This kind of analysis is fundamental to inform public debates about how cities should set their policy targets of convenient access to public transport, and how such targets should account for the interplay between socioeconomic, spatial, and temporal inequalities. The `{gtfs2gps}` package makes it very easy for researchers and policymakers to conduct this and other types of analysis that need to consider in detail the time geography of public transport systems.

## 6 Discussion

In this paper, we illustrated how the `{gtfs2gps}` package can convert GTFS feeds from relational text files into a structured trajectory data table, allowing one to examine the spatial–temporal patterns of public transport systems. In this sense, the `{gtfs2gps}` package can help researchers to visualize the space–time paths of public transport vehicles and articulate key concepts in time geography in the context of big data and transport analytics.

It is important to note that there is a difference between traditional applications of time geography to study human mobility and the application of time geography to cast light on the spatial–temporal organization of public transport systems, as presented in this paper. The framework of time geography is commonly used to analyze the mobility patterns of humans, who have agency about the possible activities

they engage in and on how they navigate transport networks. Because of this agency, concepts such as potential path area and space–time prism become critical to reflect on issues of uncertainty and potentiality underlying human mobility. In conventional public transport systems, by contrast, fixed transit routes are designed in such a way that vehicles on each trip have to abide by a predetermined space–time path scheduled by a transport agency.

While the spatial paths of public transport routes are fixed, the temporal paths might vary due to traffic or weather conditions, for example. In this case, there is no uncertainty about the spatial trajectory of vehicles. There are only travel time uncertainties about whether the transport services will be able to follow the timetables according to the schedule. This would be equivalent to considering the boarding of passengers at public transport stops as spatially fixed and temporally flexible activities (Charleux 2015). As a result, some concepts from time geography such as space–time prism and potential path area might not be directly applicable to the study of conventional public transport systems. Further research could investigate how concepts like potential path area (PPA) could be adapted to capture the notion of potential time window to study mobility patterns where there are only temporal uncertainties related to service reliability. This research agenda could benefit from a growing literature developing methods to enrich GTFS feeds with data from smart cards and onboard GPS devices that record historical public transport speeds (Arbex and Cunha 2020; Wessel et al. 2017).

The `gfs2gps` package is intended as a general-purpose tool to create public transport data outputs that open up possibilities for multiple applications based on GTFS data. A recent study has used `{gfs2gps}` to compare the distributions of bus stop spacings across cities in the USA (Pandey et al. 2021), a research question with important implications for service convenience and accessibility (Wu and Levinson 2021). Nonetheless, this type of output could have much broader applicability. For example, `{gfs2gps}` applications could be used to create interactive data visualizations, simulate and visualize changes in public transport networks or even compare the time geography of public transport systems across different urban areas. The authors of this study are currently developing a new computational method to estimate public transport emissions from GTFS data using the output of the `gfs2gps` package.

Recent innovations in the management and planning of public transport data open up new opportunities to apply and develop time geography concepts, but they also raise new challenges for the development of computational tools to organize and process large swaths of data. The story of GTFS illustrates how a global standard data format can create a common ground for transport agencies, researchers, and software developers to build new transport planning tools in different parts of the world and learn from one another. It also creates new opportunities for researchers to delve into and visually explore the time geography of public transport systems. We hope the `{gfs2gps}` package can be helpful in this endeavor.

## References

- Antrim A, Barbeau SJ (2017) Opening the door to multimodal applications: creation, maintenance and application of GTFS data (No 17-03702). Article 17–03702. Transportation research board 96th annual meeting transportation research board. <https://trid.trb.org/view/1438473>
- Arbex R, Cunha CB (2020) Estimating the influence of crowding and travel time variability on accessibility to jobs in a large public transport network using smart card big data. *J Transp Geogr* 85:102671. <https://doi.org/10.1016/j.jtrangeo.2020.102671>
- Arribas-Bel D, Alvanides S, Batty M, Crooks A, See L, Wolf L (2021) Urban data/code: a new EP-B section. *Environ Plan: Urban Anal City Sci* 48(9):2517–2519. <https://doi.org/10.1177/23998083211059670>
- Banister D (2002) *Transport planning* (2 edition.). London u.a.: Spon: Routledge
- Brunsdon C, Comber A (2021) Opening practice: supporting reproducibility and critical spatial data science. *J Geogr Syst* 23:477–496. <https://doi.org/10.1007/s10109-020-00334-2>
- Charleux L (2015) A modification of the time-geographic framework to support temporal flexibility in ‘fixed’ activities. *Int J Geogr Inf Sci* 29(7):1125–1143. <https://doi.org/10.1080/13658816.2015.1009464>
- Dardas A, Hall B, Salter J, Hosseini H (2022) A geospatial workflow for the assessment of public transit system performance using near real-time data. *Trans GIS* 26(4):1642–1664. <https://doi.org/10.1111/tgis.12942>
- Demšar U, Buchin K, Cagnacci F, Safi K, Speckmann B, Van de Weghe N, Weiskopf D, Weibel R (2015) Analysis and visualisation of movement: an interdisciplinary review. *Mov Ecol* 3(1):5. <https://doi.org/10.1186/s40462-015-0032-y>
- Hägerstrand T (1970) What about people in regional science? *Pap Reg Sci Assoc* 24(1):6–21. <https://doi.org/10.1007/BF01936872>
- Herszenhut D, Pereira RHM, Andrade PR, Bazzo J (2021) gtfstools: general transit feed specification (GTFS) editing and analysing tools (1.0.0.9) [R]. *R Found Stat Comput*. <https://doi.org/10.5281/zenodo.5703236>
- Holden E, Linnerud K, Banister D, Wierling A (2018) *The imperatives of sustainable development: needs, justice, limits*. Routledge, Taylor & Francis Group, London
- Kwan M-P (2000) Interactive geovisualization of activity-travel patterns using three-dimensional geographical information systems: a methodological exploration with a large data set. *Transp Res Part c: Emerg Technol* 8(1–6):185–203
- Kwan M-P (2004) GIS methods in time-geographic research: geocomputation and geovisualization of human activity patterns. *Geogr Ann: Ser B Hum Geogr* 86(4):267–280
- Lee J, Miller HJ (2018) Measuring the impacts of new public transit services on space-time accessibility: an analysis of transit system redesign and new bus rapid transit in Columbus, Ohio, USA. *Appl Geogr* 93:47–63. <https://doi.org/10.1016/j.apgeog.2018.02.012>
- Lee J, Miller HJ (2020) Robust accessibility: Measuring accessibility based on travelers’ heterogeneous strategies for managing travel time uncertainty. *J Transp Geogr* 86:102747. <https://doi.org/10.1016/j.jtrangeo.2020.102747>
- Long JA, Nelson TA (2013) A review of quantitative methods for movement data. *Int J Geogr Inf Sci* 27(2):292–318. <https://doi.org/10.1080/13658816.2012.682578>
- Lovelace R (2021) Open source tools for geographic analysis in transport planning. *J Geogr Syst*. <https://doi.org/10.1007/s10109-020-00342-2>
- Lucas K, Mattioli G, Verlinghieri E, Guzman A (2016) Transport poverty and its adverse social consequences. *Proc Inst Civil Eng: Transp* 169(6):353–365. <https://doi.org/10.1680/jtran.15.00073>
- Miller HJ (2005) A Measurement theory for time geography. *Geogr Anal* 37(1):17–45. <https://doi.org/10.1111/j.1538-4632.2005.00575.x>
- Miller HJ (2017) Time geography and space–time prism. *Int Encycl Geogr: People Earth Environ Technol* 1–19
- Morgan M, Young M, Lovelace R, Layik H (2021) Opentripplanner for R: setup and connect to “Open-TripPlanner” (0.4.0.99999) [R]. *R Found Stat Comput*. <https://zenodo.org/record/3558311>
- Neutens T, Schwanen T, Witlox F (2011) The prism of everyday life: towards a new research agenda for time geography. *Transp Rev* 31(1):25–47. <https://doi.org/10.1080/01441647.2010.484153>
- Páez A (2021) Open spatial sciences: an introduction. *J Geogr Syst* 23:467–476. <https://doi.org/10.1007/s10109-021-00364-4>

- Pandey A, Lehe L, Monzer D (2021) Distributions of bus stop spacings in the United States. Findings. <https://doi.org/10.32866/001c.27373>
- Pereira RHM, Karner A (2021) Transportation equity. In: Vickerman R (ed) International encyclopedia of transportation, vol 1. Elsevier Ltd, United Kingdom, pp 271–277
- Pereira RHM, Saraiva M, Herszenhut D, Braga CKV, Conway MW (2021) r5r: rapid realistic routing on multimodal transport networks with R<sup>5</sup> in R. *Transp Find*. <https://doi.org/10.32866/001c.21262>
- Pereira RHM, Schwanen T (2015) Commute time in Brazil (1992–2009): differences between metropolitan areas, by income levels and gender. Discussion Paper No. 192. Ipea, Institute for Applied Economic Research. [https://repositorio.ipea.gov.br/bitstream/11058/5140/1/DiscussionPaper\\_192.pdf](https://repositorio.ipea.gov.br/bitstream/11058/5140/1/DiscussionPaper_192.pdf)
- Poletti F, Herszenhut D, Buckley T, Padgham M, Noriega-Goodwin D (2021) tidytransit: read, validate, analyze, and map GTFS feeds (1.2.0) [R]. *R Found Stat Comput*, <https://github.com/r-transit/tidytransit>
- Rey SJ (2009) Show me the code: spatial analysis and open source. *J Geogr Syst* 11(2):191–207. <https://doi.org/10.1007/s10109-009-0086-8>
- Singleton A, Arribas-Bel D (2021) Geographic data science. *Geogr Anal* 53(1):61–75. <https://doi.org/10.1111/gean.12194>
- Song Y, Miller HJ (2012) Exploring traffic flow databases using space–time plots and data cubes. *Transportation* 39(2):215–234. <https://doi.org/10.1007/s11116-011-9343-z>
- Sui D (2012) Looking through Hågerstrand’s dual vistas: towards a unifying framework for time geography. *J Transp Geogr* 23:5–16. <https://doi.org/10.1016/j.jtrangeo.2012.03.020>
- UN-Habitat (2018) SDG 11+ metadata: a guide to assist national and local governments to monitor and report on SDG goal 11+ indicators; UN-Habitat: Nairobi, Kenya. <https://www.local2030.org/library/60/SDG-Goal-11-Monitoring-Framework-A-guide-to-assist-national-and-local-governments-to-monitor-and-report-on-SDG-goal-11-indicators.pdf>
- United Nations U.N. (2015) Transforming the World: the 2030 agenda for sustainable development; United nations general assembly; United Nations, New York
- Wessel N, Allen J, Farber S (2017) Constructing a routable retrospective transit timetable from a real-time vehicle location feed and GTFS. *J Transp Geogr* 62:92–97. <https://doi.org/10.1016/j.jtrangeo.2017.04.012>
- Wolf LJ, Fox S, Harris R, Johnston R, Jones K, Manley D, Tranos E, Wang WW (2021) Quantitative geography III: future challenges and challenging futures. *Prog Hum Geogr* 45(3):596–608. <https://doi.org/10.1177/0309132520924722>
- Wu H, Levinson D (2021) Optimum stop spacing for accessibility. *Eur J Transp Infrastruct Res* 21(2):1–18. <https://doi.org/10.18757/ejtir.2021.21.2.4794>
- Yang L, Kwan M-P, Pan X, Wan B, Zhou S (2017) Scalable space-time trajectory cube for path-finding: a study using big taxi trajectory data. *Transp Res Part B: Methodol* 101:1–27. <https://doi.org/10.1016/j.trb.2017.03.010>
- Yang L, Zhang F, Kwan M-P, Wang K, Zuo Z, Xia S, Zhang Z, Zhao X (2020) Space-time demand cube for spatial-temporal coverage optimization model of shared bicycle system: a study using big bike GPS data. *J Transp Geogr* 88:102861. <https://doi.org/10.1016/j.jtrangeo.2020.102861>

**Publisher’s Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

## Authors and Affiliations

Rafael H. M. Pereira<sup>1</sup>  · Pedro R. Andrade<sup>2</sup>  · João Pedro Bazzo Vieira<sup>1</sup> 

<sup>1</sup> Ipea - Institute for Applied Economic Research, Brasília, Brazil

<sup>2</sup> INPE - National Institute for Space Research, São José dos Campos, Brazil