How Do E-Bikes Measure Up? Analyzing Speed Differences and Network Impacts of São Paulo's Bikesharing System

Thiago Vinicius Louro[1,2] , Anna Beatriz Grigolon[1], Alejandro Tirachini[3,4], André Luiz Cunha[2], Karst T. Geurs[1]

[1] University of Twente, Enschede, Netherlands
[2] University of Sao Paulo, Sao Carlos School of Engineering, Sao Carlos, Brazil
[3] Technische Universität Braunschweig, Germany
[4] Universidad de Chile, Santiago, Chile

E-mail: t.v.louro@utwente.nl

ORCID IDs

Thiago Vinicius Louro: 0000-0002-1118-3119
Anna Beatriz Grigolon: 0000-0002-7044-2934
Alejandro Tirachini: 0000-0001-9911-1242
André Luiz Cunha: 0000-0002-0520-0621
Karst T. Geurs: 0000-0003-0918-8903

Abstract

Cycling speed is a pivotal factor in active mobility, influencing travel times, accessibility, and mode choice. However, empirical models of cycling speed, particularly for electric bicycles, remain limited. This study aims to examine the impact of infrastructure and topographic determinants on cycling speeds and network performance of electric bicycles, compared to conventional bicycles. Using a comprehensive GPS-based dataset from São Paulo's Tembici bikesharing system, we employ linear mixed-effects models to analyze detailed variations in cycling speeds at the tracking point level, and linear regression at the trip level to assess overall trip dynamics. Novel variables, such as car traffic data from Uber, are incorporated, along with interactions between road segment length and cycling infrastructure type. Results show that electric bicycles consistently achieve higher speeds than conventional bikes, especially on challenging terrains and longer trips, due to motor assistance. Car speed significantly reduces speed for both conventional and electric bicycles. Segregated cycling infrastructure significantly increases speeds, with the greatest benefits observed for e-bikes on steep gradients. At intersections and traffic lights, e-bikes show greater sensitivity in the segment-based analysis, but they recover more quickly at the trip level. E-bikes also led to a 14.2% improvement in the global efficiency of the cycling network, and our analysis of betweenness centrality showed that e-bikes change the importance of specific routes, making longer roads more critical. These findings underscore the importance of targeted investments in cycling infrastructure, particularly segregated lanes, to optimize e-bike performance, and support sustainable urban transport systems.

Keywords: Cycling, GPS data, Speed model, Multilevel mixed-effects model, Active travel

Declarations

Competing Interests: The authors declare no competing interests.

Funding: This work was supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001.

Acknowledgements: The authors would like to thank Tembici for providing data support.

Introduction

Speed models for road vehicles are essential to provide reliable travel time estimates which are relevant for the transport demand and accessibility models and transportation planning process (Clarry et al. 2019; Geurs & van Wee 2004). Despite the importance of travel times, the speed assigned to bicycles in transport models is often a single average value around 15 km/h (Highway Capacity Manual 2010; Saneinejad et al. 2012; Pereira et al. 2021), even though speed variations do exist due to differences in route elevation profiles, surface conditions, the presence of cycling facilities, cyclist fitness, among other factor (Meyer de Freitas & Axhausen 2024; Yan et al. 2024). In addition, the emergence of electric bicycles (e-bikes), adds even more variance to cycling speed profiles (Schleinitz et al. 2017). E-bikes enable cyclists to travel faster and with less physical effort, particularly on challenging terrain or longer routes. For this reason, e-bikes may attract a broader demographic to cycling, including those who might otherwise avoid it due to physical limitations or distance thresholds. Additionally, the increasing popularity of e-bikes within bikesharing systems presents a novel area for exploration, especially regarding their usage compared to conventional bikes (Xie et al. 2024). Understanding cycling speed dynamcs of regular and e-bikes can provide a clearer picture of their potential benefits, such as improved travel times, enhanced accessibility, and reduced physical effort for users.

The literature on cycling speed models can be classified in several ways. One classification relates to the data collection method, which falls into three categories: *i*) at fixed locations, *ii)* measuring the start and end time, and *iii*) using GPS data collected throughout the ride (Yan et al. 2024). Some studies analyze the speed for the entire trip (Schleinitz et al. 2018), while others analyze speeds at segments (Flügel et al. 2019; Romanillos & Gutierrez 2020), and others analyze the speed at every tracking point (Clarry et al., 2019). The analysis methods vary from simple descriptive analysis (Jensen et al., 2010), to OLS regression (Flügel et al. 2019), to multilevel linear mixed-effects models (Clarry et al. 2019; Yan et al. 2024), with the right model depending on the characteristics of the data. Linear Mixed Models (LMMs) are particularly advantageous when data have a hierarchical structure or involve repeated measurements, such as where observations are nested within trips or segments. LMMs allow for the inclusion of random effects, capturing both individual-level and group-level variations. This makes them more flexible and accurate compared to simpler models, which might overlook dependencies within the data, leading to biased results.

Existing studies highlight that cycling speeds are influenced by factors such as cycling infrastructure, topography, user characteristics, and bike type. The presence of cycling infrastructure significantly increases speeds (Clarry et al. 2019; Cubells et al. 2023; Flügel et al. 2019; Strauss & Miranda-Moreno 2017; Yan et al. 2024). Regarding topography, speeds are typically lower on uphill segments and higher on flat terrain or downhill segments. However, Yan et al. (2024) reported lower speeds even on downhill segments. Furthermore, the effect of cycling uphill is stronger than the positive effect of cycling downhill (Clarry et al. 2019; Flügel et al. 2019; Parkin & Rotheram 2010) as cyclists often brake while descending. E-bikes are generally faster than conventional bikes, and men tend to cycle faster than women. However, the speed gap between men and women is smaller among e-bike users (Flügel et al. 2019; Yan et al. 2024). Proximity to intersections also affects speeds, with lower speeds observed near intersections (Clarry et al. 2019; Strauss & Miranda-Moreno 2017; Yan et al. 2024). Additionally, speeds tend to be higher during the morning commuting and on arterial roads (Strauss & Miranda-Moreno 2017).

Despite the previous insights provided by cycling studies, several gaps remain. While prior research, such as Flügel et al. (2019) and Yan et al. (2019), has explored e-bikes, the overall body of work remains limited, first, in addressing how riders in electric bicycles interact with infrastructure, topography and the speed of cars next to them. Second, the general effects of electric bicycles on network performance also remains unexplored. Furthermore, many studies rely on small sample sizes, which restrict generalizability and hinder the development of robust models capable of capturing the full variability of cycling behavior. For example, studies often analyze data from a few dozen participants or short observation periods, limiting their applicability to broader populations.

Another significant gap lies in the lack of studies conducted in the Global South, where cycling networks are often underdeveloped and cycling usually remains as a marginal mode in terms of modal share. Research from cities or countries in the Global North often reflects more mature cycling infrastructure and/or a completely different pattern of city planning and development, while the contextual challenges of cities in the Global South—such as inconsistent infrastructure, high levels of car traffic and congestion, and different cycling cultures—remain understudied. These differences make it difficult to generalize findings across regions with vastly different cycling environments.

Additionally, while it is well-known that cycling speed profiles are context-dependent, few studies have explicitly analyzed how factors like infrastructure type, terrain, and weather interact with cycling network performance metrics, such as connectivity and efficiency. The impact of cycling speeds, especially those of e-bikes, on the overall performance of cycling networks has also been insufficiently explored. For instance, how e-bikes alter the importance of specific network links or enhance connectivity within fragmented networks remains unclear.

To address these gaps, this study examines how infrastructure and topography influence cycling speeds and assesses the impact of e-bikes on the performance of cycling networks. Using a large-scale GPS-based dataset from Tembici, a leading bikesharing provider in Brazil, we apply linear mixed-effects models to capture variations in cycling speeds at the tracking-point level (analysing local effects in short segments) and linear regression to analyze trip-level dynamics (analysing network-wide effects on entire cycling trips). Novel variables, such as car traffic data from Uber, are incorporated to enrich the analysis, as from previous studies there is no indication of what is the actual effect of car speeds on bicycle speeds. Interactions between factors—such as road segment length and cycling infrastructure type—are investigated to better understand the determinants of cycling speed for both conventional and electric bicycles. Finally, we evaluate the broader impacts of e-bikes on network efficiency, including their influence on the importance of specific network links, quantified using edge betweenness centrality.

Methodology

Study area and data

This study focuses on the urban area of São Paulo, Brazil. The dataset used in this paper comes from Tembici, a leading Brazilian bikesharing provider, obtained through an agreement with the University of São Paulo. Tembici operates a network of docking stations where users can rent and return bikes at any station. Recently, the company expanded its service by introducing electric bicycles (e-bikes) in São Paulo and other major Brazilian cities. These e-bikes feature a pedal-assist system that allows riders to reach speeds of up to 25 km/h, and their batteries are maintained through a swapping system. It includes trips made between January and July 2023, which corresponds to the period during which Tembici had GPS data available. It consists of two main sources: *i)* operational data, including attributes such as start and end stations, start and end times, trip duration, cyclist gender, trip purpose, and other relevant details, and *ii)* GPS data, consisting of the geographic coordinates (latitude and longitude) and corresponding timestamps recorded throughout each trip. Figure 1 shows a map of São Paulo, highlighting both the bikesharing stations and the cycling infrastructure. The network data comes from OpenStreetMap, the stations from the Tembici dataset that was provided, and the cycling network comes from the *Mapa de Infraestrutura Cicloviária* provided by *Companhia de Engenharia de Tráfego de São Paulo* (CET 2024), and consists of segregated (ciclovias) and non-segregated (ciclofaixas) cycle lanes.

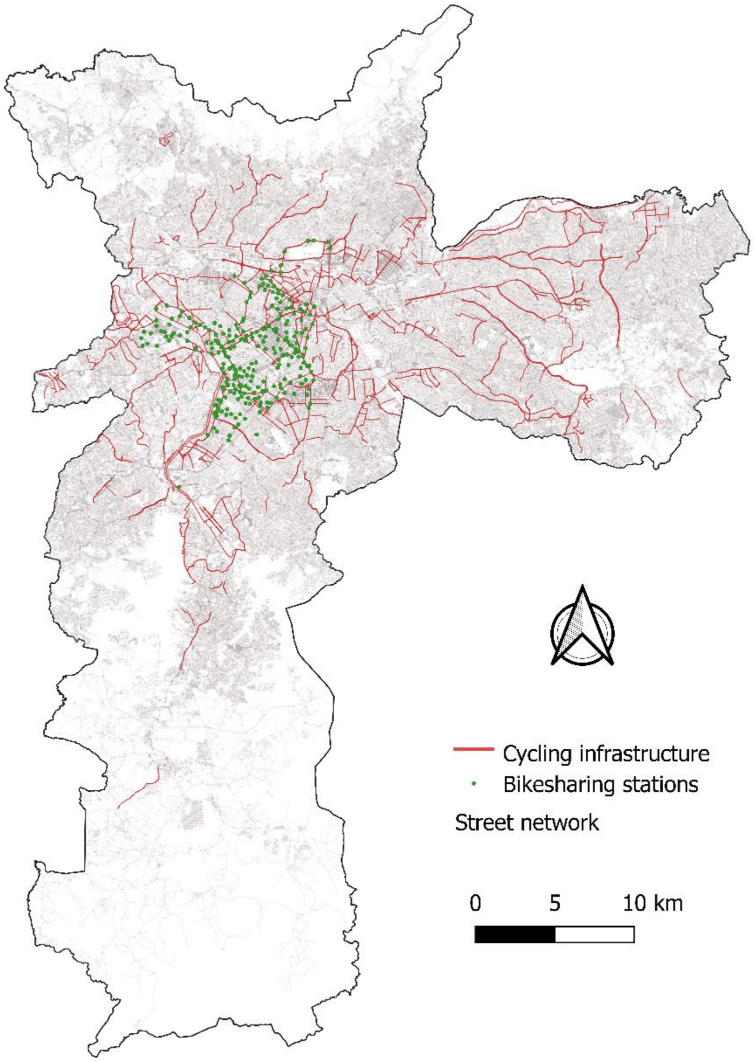

Figure 1. Distribution of cycling infrastructure and bikesharing locations in São Paulo

Most of the bikesharing stations are concentrated in the central and center-west areas of São Paulo, which include many high-income neighborhoods and central business districts. Equity concerns related to this distribution, for São Paulo and other cities, are discussed in (Fortes et al. 2024). Additionally, it is noticeable that the cycling infrastructure is not very dense around these stations, suggesting potential gaps in access to cycling infrastructure and network integration. The concentration of bikesharing stations in higher-income areas may introduce bias, as these regions might not represent the broader city's cycling patterns, particularly in lower-income neighborhoods where bike use and access to infrastructure may differ. Additionally, the sparse cycling infrastructure around the stations could influence cycling speeds and route choices, potentially leading to slower or less consistent speeds.

Preprocessing

The GPS data in the study comes from the GPS device on the user's smartphone and was generally recorded every 30 seconds, with some variations. To ensure consistency, we filtered out trips where the GPS frequency was greater than 30 seconds. Additionally, we limited our analysis to utilitarian trips, which are primarily made for commuting or other practical purposes, and to weekdays, ensuring that our models focus on typical everyday travel patterns. Utilitarian trips, as defined by the bikesharing provider, are direct trips with no stops (which could happen for services like food delivery) and are no bike tours (where the rider would unlock and return the bike to the same or nearby locations)

After integrating the operational and GPS datasets, we performed map matching of the GPS points to align them with the transportation network. We used ArcGIS Pro's SnapTracks tool for map matching, ensuring that the GPS points accurately corresponded to the road infrastructure. After matching the points to the network, we used the geosphere package in R (Hijmans et al. 2023) to apply the Haversine formula, which computes the distance between consecutive GPS points while accounting for the curvature of the Earth. Speed was calculated by dividing the distance between two points by the time difference between them. We retained only observations with speeds between 4 and 35 km/h. This step was taken to exclude speeds lower than typical walking speeds and to ensure that the dataset reflected realistic cycling conditions. Additionally, trips were filtered to include only segments longer than 30 meters, minimizing distortions in slope data that could arise from shorter segments. Finally, trips were limited to those with durations between 120 and 7,200 seconds.

The final dataset consisted of 1,119,993 observations, including 729,219 for conventional bikes and 390,774 for e-bikes, derived from a total of 76,870 trips (49,788 by conventional bike and 31,584 by e-bike). While it is possible that the same user made multiple trips, this information was not available, so each trip was treated as being made by a unique individual.

Variables

A range of variables were used to capture the factors influencing cycling speed at both the tracking point and trip level. The variables and their reference level are presented in Table 1.

Table 1 – Variables

Variable	Categories	Reference Category
Tracking point model		
Cyclist Gender	Female, Male, Other	Female
Road Class	Residential, Tertiary, Primary, Trunk, Secondary	Residential
Cycling Infrastructure	No cycling infrastructure, Cycle Lane, Cycle Path	No cycling infrastructure
Slope Category	Less than -5%, -5% to -2%, Between -2% and 2%, 2% to 5%, Greater than 5%	Between -2% and 2%
Time of Day	Early Morning, Morning Peak, Morning Off-Peak, Lunch,	Morning Off-

	Afternoon, Evening Peak, Night	Peak
Speed Limit	30 km/h, 40 km/h, 50 km/h, 60 km/h	40 km/h
Day of the Week	Monday, Tuesday, Wednesday, Thursday, Friday	Wednesday
Proximity to Intersections	No, Yes	No
Proximity to Traffic Lights	No, Yes	No
Length of Road Segment	-	-
Cyclist Speed (Previous Segment)	-	-
Uber Speed	-	-
Trip Distance	-	-
Trip model		
Trip distance	-	-
Percentage with cycling infrastructure	-	-
Number of intersections	-	-
Number of traffic lights		
Elevation difference	-	-
Bike Type		
Gender	Female, Male, Other	Female
Day of the Week	Monday, Tuesday, Wednesday, Thursday, Friday	Wednesday
Time of Day	Early Morning, Morning Peak, Morning Off-Peak, Lunch, Afternoon, Evening Peak, Night	Morning Off-Peak

For the tracking point model, the slope of the terrain was included as a categorical variable with five levels: Less than -5%, -5% to -2%, -2% to 2%, 2% to 5%, and Greater than 5%. On the other hand, for the trip level, we considered the net elevation between the first and last gps point. The elevation data was obtained from the NASA Shuttle Radar Topography Mission (2013), via the Open Topography DEM downloader plugin on QGIS (Win 2024), and the slope was then calculated using the osmnx package in Python (Boeing 2017).

Gender was provided directly by the Tembici dataset. While the dataset included three categories—female, male, and other—it was unclear whether the 'other' category represented non-binary individuals or those who chose not to disclose their gender. Therefore, the analysis focused solely on male and female cyclists, as these categories were more reliable in terms of data accuracy.

The network link length, road classification, and speed limits were obtained from Open-StreetMap. The day of the week, restricted to weekdays, was also included. Additionally, two bi-nary variables were created to capture proximity to intersections and traffic lights. In the tracking point model, these variables were generated by creating a 20-meter buffer around intersections and traffic lights, then overlaying GPS points to determine whether a cyclist was near these fea-tures. For the trip-level model, we used the exact count of intersections and traffic lights along the route. This method was not applied in the tracking point model due to the computational demands of extracting this data for all routes.

Cycling infrastructure was represented by a variable distinguishing among segregated cycling infrastructure (ciclovias), referred to as cycle lanes; on-road infrastructure (ciclofaixas), referred to as cycle paths; and routes with no infrastructure, with data obtained from the São Pau-lo municipality. To capture trip characteristics, total trip distance was included to reflect the full distance covered by a cyclist during a trip. Additionally, a categorical variable representing the hour of the day was incorporated, with time periods classified into seven categories: Early Morn-ing, Morning Peak, Morning Off-Peak, Lunch, Afternoon, Evening Peak, and Night. These catego-ries were developed based on observed patterns in the data, allowing for an analysis of speed variations according to the time of day.

Finally, external data sources were used to analyse further interactions with cycling speed. Uber speed data was matched with the dataset to provide the average hourly speed on road links during weekdays. The Uber speed dataset contained the average hourly speed for each segment of the São Paulo network over a single day, meaning it does not reflect the exact Uber speed at the time when the cyclist passed through each link. However, we considered this data valuable as a good proxy for the speed of motorized traffic experienced by cyclists. To account for dependencies between consecutive speed observations, a lagged speed variable was included to represent the speed at the previous GPS point. Table 2 presents descriptive statistics for cycling speeds by bike type across the categorical variables.

Table 2 – Descriptive statistics of cycling speeds for each categorical variable

Variable	Bike Type	N	Mean	Median	Min	Max	SD
Speed	Bike	685169	15.18	14.93	4.00	34.999	4.80
	E-bike	367812	17.46	17.70	4.00	34.996	5.14
Gender							
Female	Bike	145964	14.54	14.34	4.00	34.90	4.54
	E-Bike	69519	16.74	16.86	4.02	34.95	5.06
Male	Bike	479933	15.36	15.12	4.00	35.00	4.85
	E-Bike	265406	17.63	17.92	4.00	35.00	5.13
Other	Bike	59272	15.22	14.95	4.03	34.96	4.87
	E-Bike	32887	17.61	17.80	4.12	34.91	5.25
Grade Percent Category							
-2% to 2%	Bike	274423	15.08	14.84	4.00	34.96	4.68
	E-Bike	149597	17.48	17.77	4.03	35.00	5.03
-5% to -2%	Bike	118804	15.70	15.47	4.01	34.96	4.81
	E-Bike	63017	17.93	18.20	4.00	34.95	5.16
2% to 5%	Bike	110800	14.82	14.54	4.00	34.95	4.68
	E-Bike	58295	17.22	17.41	4.00	34.97	5.08
Above 5%	Bike	91488	15.06	14.88	4.01	35.00	4.97
	E-Bike	48966	17.43	17.71	4.08	35.00	5.22
Below -5%	Bike	89654	15.34	14.98	4.00	34.95	5.03
	E-Bike	47937	17.09	17.14	4.05	34.91	5.39
Day of week							
Monday	Bike	147152	15.16	14.93	4.01	34.96	4.75
	E-Bike	77914	17.45	17.66	4.00	34.97	5.09
Tuesday	Bike	145574	15.24	14.98	4.00	35.00	4.80

	E-Bike	91108	17.44	17.68	4.10	34.99	5.14
Wednesday	Bike	138369	15.20	14.96	4.00	34.96	4.79
	E-Bike	69220	17.50	17.75	4.02	35.00	5.16
Thursday	Bike	139731	15.14	14.88	4.00	34.96	4.82
	E-Bike	67042	17.46	17.70	4.00	35.00	5.17
Friday	Bike	114343	15.12	14.86	4.01	34.95	4.83
	E-Bike	62528	17.48	17.73	4.08	34.89	5.13
Hour category							
Early Morning	Bike	13787	15.83	15.58	4.07	34.53	4.70
	E-Bike	10980	18.83	19.20	4.60	35.00	4.99
Morning Out of Peak	Bike	107656	15.43	15.20	4.01	34.96	4.84
	E-Bike	53961	17.52	17.83	4.18	34.95	5.14
Morning Peak	Bike	184105	15.66	15.46	4.00	35.00	4.77
	E-Bike	99155	17.84	18.18	4.05	34.99	5.05
Lunch	Bike	42469	14.57	14.30	4.01	34.52	4.64
	E-Bike	21547	16.94	17.09	4.08	35.00	5.12
Afternoon	Bike	37942	14.55	14.21	4.08	34.74	4.70
	E-Bike	20461	17.19	17.33	4.15	34.84	5.23
Evening Peak	Bike	226904	14.88	14.64	4.00	34.96	4.79
	E-Bike	123483	17.09	17.24	4.00	34.94	5.15
Night	Bike	72306	15.03	14.69	4.01	34.95	4.82
	E-Bike	38225	17.63	17.80	4.18	34.95	5.16
Proximity to traffic light							
No	Bike	529658	15.17	14.91	4.00	35.00	4.81
	E-Bike	284882	17.48	17.71	4.00	34.99	5.14
Yes	Bike	155511	15.20	14.97	4.03	34.96	4.76
	E-Bike	82930	17.41	17.67	4.05	35.00	5.14
Proximity to intersection							
No	Bike	292642	15.46	15.16	4.00	34.96	4.82
	E-Bike	154157	17.91	18.20	4.03	34.99	5.06
Yes	Bike	392527	14.96	14.74	4.00	35.00	4.77
	E-Bike	213655	17.14	17.31	4.00	35.00	5.17
Road Type							
Primary	Bike	294209	15.46	15.19	4.01	34.96	5.02
	E-Bike	136695	17.78	18.09	4.08	34.95	5.29
Residential	Bike	59289	13.97	13.52	4.00	34.90	4.75
	E-Bike	38780	15.88	15.50	4.08	34.97	5.02
Secondary	Bike	33880	13.71	13.31	4.08	34.74	4.63
	E-Bike	18136	15.95	15.84	4.00	34.90	5.30
Tertiary	Bike	70948	15.02	14.76	4.01	34.96	4.75
	E-Bike	35660	17.18	17.21	4.13	34.99	5.15
Trunk	Bike	226843	15.38	15.29	4.01	35.00	4.45
	E-Bike	138541	17.86	18.27	4.00	35.00	4.87
Speed limit							

30	Bike	17342	13.86	13.54	4.01	34.46	4.14
	E-Bike	10194	16.34	16.21	4.35	34.42	4.82
40	Bike	100354	14.64	14.23	4.00	34.96	4.84
	E-Bike	55427	16.66	16.51	4.00	34.97	5.14
50	Bike	566307	15.31	15.10	4.00	35.00	4.79
	E-Bike	301653	17.65	17.97	4.00	35.00	5.13
60	Bike	1053	14.91	14.60	4.44	33.71	4.85
	E-Bike	480	17.15	16.93	4.44	34.82	5.26

Modelling

Cycling speed was analyzed at two levels: the tracking point level and the trip level. Analyzing cycling speed at both the tracking point level and the trip level provides a more comprehensive understanding of the factors influencing speed. The tracking point level analysis allows for a detailed examination of variations in speed at specific moments during a trip, distinguishing between electric and conventional bicycles and identifying nuanced differences in their behavior. By contrast, the trip level analysis focuses on the overall average speed for an entire trip, offering insights into the broader impact of bicycle type. Using both approaches ensures that both granular and aggregated perspectives are considered, enabling a more robust and well-rounded interpretation of the data.

At the tracking point level, separate linear mixed-effects models (LMMs) were developed for each bicycle type, allowing for a detailed examination of the coefficients specific to electric and conventional bicycles. Additionally, we built a model using the combined data from both conventional and electric bicycles to test for interaction effects and confirm whether the effects of bicycle type were significant. At the trip level, a single linear regression model was employed, incorporating interaction terms to investigate the effect of bicycle type on overall trip speed. We opted to use a single model at trip level due to the reduced number of variables, which makes it possible to effectively capture the relationship between bicycle type and overall trip speed.

Tracking point level

The dependent variable for this model was cycling speed (km/h), measured at each tracking point, which corresponds to the specific moments when a GPS location was recorded along the cyclist's route. This allows for a detailed analysis of speed variations throughout the trip. The variable was log-transformed to normalize its distribution and stabilize variance, ensuring that key assumptions of linear modeling—such as the normality of residuals and homoscedasticity—were met. Before estimating the model, continuous predictors were standardized to improve variable comparability and support model convergence.

For both conventional and electric bicycles, three models were estimated: we first fitted a null model to account for the hierarchical structure of the data, including random intercepts for each trip ID. This model served as a baseline to capture variability in cycling speed across individual trips. Next, we fitted a model with a random intercept for each trip, adding a range of fixed effects to explain cycling speed. This model allowed us to assess the individual contributions of these factors while accounting for differences across trips. Finally, we developed a model with random slopes for slope category, intersection proximity, and infrastructure type, allowing these effects to vary by trip. Additionally, interaction terms were included to explore relationships between slope category and gender, slope category and link length, slope category and infrastructure type, and infrastructure type and gender.

Trip level

For the trip level, we developed a linear regression model using trip speed as the dependent variable. The explanatory variables in the model included gender, the proportion of the route with cycling infrastructure, the number of intersections and traffic lights, net elevation (the difference in elevation between the first and last points), trip distance, time of day and day of the week.

Network analysis

To investigate the impacts of e-bikes on São Paulo's cycling network, a network analysis was conducted using two key metrics: global efficiency and edge betweenness centrality. Global efficiency (G_e) provides a measure of how effectively a network facilitates movement between all pairs of nodes. It is defined as the average inverse of the shortest path lengths between nodes, normalized by the total number of node pairs (Latora & Marchiori, 2001; Vragović et al., 2005):

$$G_e = \frac{1}{N(N-1)} \sum_{i \neq j} \frac{1}{d_{ij}} \tag{1}$$

where N is the total number of nodes, and d_{ij} represents the shortest path length between nodes i and j. The global efficiency values range from 0 to 1, where 0 represents a completely disconnected network and 1 represents perfect efficiency with direct connections between all nodes. In this study, edge weights were calculated based on the average travel times along network segments, which were derived from the speed models developed for both conventional and electric bicycles. The calculations were implemented using the igraph library (Csardi & Nepusz 2006) in R.

Edge betweenness centrality (E_{bc}) was used to identify critical links within the cycling network. This metric quantifies the importance of an edge by calculating the proportion of all shortest paths between node pairs that pass through it:

$$E_{bc} = \sum_{s \neq t} \frac{\sigma_{st(e)}}{\sigma_{st}} \tag{2}$$

where σ_{st} is the total number of shortest paths between nodes s and t, and $\sigma_{st(e)}$ represents the number of those paths that pass through edge e. Edges with higher betweenness centrality values are considered critical for maintaining connectivity and facilitating flow within the network. For this analysis, the NetworkX library in Python (Hagberg et al. 2008) was used to calculate edge betweenness centrality for both bicycle types. By comparing centrality values between e-bikes and conventional bicycles, it was possible to assess how the introduction of e-bikes may alter the importance of specific network links.

Results and discussion

Tracking point level

To evaluate the models' robustness, we performed 5-fold cross-validation, a method used to evaluate model performance and generalizability. The dataset was split into five equal parts or "folds." In each iteration, one-fold was used as the test set, while the remaining four were used for training the model. This process was repeated five times, ensuring each fold served as the test set once. The results were then averaged to provide a robust estimate of model accuracy and to minimize the impact of data variability.

The results indicated consistent performance across folds, with an average Root Mean Square Error (RMSE) of approximately 1.3 km/h for both the e-bike and conventional bicycle models. Diagnostic checks confirmed that residuals were normally distributed, variances were homoscedastic, and no spatial autocorrelation was detected, as assessed using variograms.

As previously explained, we estimated three models for both types of bikes, and the more complete models consistently provided the best fit to the data, as assessed by log-likelihood and AIC values. For this reason, we opted to present only the results of the models with main plus interaction effects (the third type of model described in Section 2.4.1) for both bicycle types, enabling a direct comparison of their outcomes. The significant difference between marginal and conditional R^2 values indicates that trip-level random effects explained considerably more variance than fixed effects alone. This pattern was consistent across both bicycle types, with conventional bicycles showing slightly higher R^2 values.

In both cases, since the dependent variable was log-transformed, coefficients are interpreted as percentage changes in speed relative to reference categories. For instance, a coefficient of 0.05 represents a 5% speed increase, while -0.05 indicates a 5% decrease. Table 1 presents the results for the models of conventional and electric bicycles.

Table 1 – Tracking point model results

Variable	**Conventional bicycle**			**Electric bicycle**		
	Coef.	**Std. Error**	**p-value**	**Coef.**	**Std. Error**	**p-value**
Intercept	2.294	0.004	< 2e-16	2.471	0.005	<2e-16
Length	0.017	0.001	< 2e-16	0.021	0.001	<2e-16
Proximity to intersection	-0.038	0.001	< 2e-16	-0.042	0.001	<2e-16
Proximity to traffic light	-0.003	0.001	0.001	-0.008	0.001	0.000
Uber Speed	-0.027	0.000	< 2e-16	-0.031	0.001	<2e-16
Speed previous section	0.127	0.000	< 2e-16	0.099	0.001	< 2e-16
Total Distance	0.030	0.001	< 2e-16	0.030	0.001	< 2e-16
Slope Category (base= -2% to 2%)						
Below -5%	0.022	0.003	0.000	0.012	0.005	0.022

-5% to -2%	0.001	0.003	0.651	-0.024	0.004	0.000
2% to 5%	-0.030	0.003	< 2e-16	-0.023	0.004	0.000
Above 5%	-0.069	0.003	< 2e-16	-0.053	0.005	< 2e-16
Gender (base= Female)						
Male	0.044	0.003	< 2e-16	0.053	0.004	< 2e-16
Road class (base= Residential)						
Tertiary	0.059	0.002	< 2e-16	0.052	0.003	< 2e-16
Primary	0.043	0.002	< 2e-16	0.052	0.003	< 2e-16
Trunk	0.070	0.002	< 2e-16	0.098	0.003	< 2e-16
Secondary	-0.023	0.003	< 2e-16	-0.007	0.003	0.029
Speed limit (base= 40 km/h)						
30	-0.307	0.003	0.000	-0.007	0.004	0.102
50	-0.009	0.002	0.000	-0.017	0.002	0.000
60	-0.017	0.010	0.001	-0.065	0.015	0.000
Day of week (base=Wednesday)						
Monday	-0.002	0.002	0.325	0.002	0.003	0.423
Tuesday	0.000	0.002	0.834	-0.002	0.003	0.555
Thursday	-0.007	0.002	0.002	-0.001	0.003	0.668
Friday	-0.006	0.002	0.015	0.006	0.003	0.036
Cycling Infrastructure (base = No infrastructure)						
Cycle Lane	0.033	0.003	< 2e-16	0.049	0.004	< 2e-16
Cycle Path	0.043	0.004	< 2e-16	0.054	0.006	< 2e-16
Hour Category (base = Morning Off Peak)						
Early Morning	0.047	0.005	< 2e-16	0.067	0.006	< 2e-16
Morning Peak	0.030	0.002	< 2e-16	0.020	0.003	0.000
Lunch	-0.027	0.003	< 2e-16	-0.016	0.004	0.000
Afternoon	-0.035	0.003	< 2e-16	-0.017	0.004	0.000
Evening Peak	-0.034	0.002	< 2e-16	-0.025	0.003	< 2e-16
Night	-0.018	0.003	0.000	0.009	0.004	0.010
Interaction terms						
Below -5%:Male	0.007	0.002	0.023	0.001	0.005	0.795
-5% to -2%:Male	0.004	0.002	0.139	0.008	0.004	0.034
2% to 5%:Male	0.004	0.002	0.111	0.006	0.004	0.092
Above 5%:Male	-0.003	0.002	0.033	-0.001	0.004	0.905
Below -5%:Other	0.013	0.002	0.008	-0.002	0.007	0.774
-5% to -2%:Other	0.011	0.002	0.010	-0.004	0.006	0.439
Below -5%:Length	0.067	0.002	< 2e-16	0.060	0.003	< 2e-16
-5% to -2%:Length	0.045	0.001	< 2e-16	0.041	0.002	< 2e-16
2% to 5%:Length	-0.009	0.001	0.000	0.024	0.002	< 2e-16
Above 5%:Length	-0.006	0.002	0.012	0.015	0.003	0.000
Below -5%:Cycle Lane	-0.012	0.003	0.000	-0.018	0.004	0.000

-5% to -2%:Cycle Lane	0.042	0.002	< 2e-16	0.077	0.003	< 2e-16
2% to 5%:Cycle Lane	0.011	0.002	0.000	0.015	0.003	0.000
Above 5%:Cycle Lane	0.108	0.003	< 2e-16	0.105	0.004	< 2e-16
Below -5%:Cycle Path	0.023	0.003	0.000	-0.022	0.005	0.000
-5% to -2%:Cycle Path	0.034	0.003	< 2e-16	0.047	0.004	< 2e-16
2% to 5%:Cycle Path	0.001	0.003	0.842	-0.005	0.004	0.283
Above 5%:Cycle Path	0.075	0.004	< 2e-16	0.037	0.005	0.000
Male:Cycle Lane	-0.007	0.003	0.031	-0.018	0.004	0.000
Male:Cycle Path	0.003	0.004	0.465	-0.002	0.006	0.713
Random intercept						
Trip variance		0.03			0.03	
Random slopes						
Proximity to intersection		0.00			0.00	
Slope Category						
Below -5%		0.01			0.02	
-5% to -2%		0.01			0.01	
2% to 5%		0.01			0.01	
Above 5%		0.01			0.02	
Cycling Infrastructure						
Cycle Lane		0.02			0.02	
Cycle Path		0.03			0.03	
Model Fit						
LL		-89945			-56285	
AIC		180079.40			112761	
R2m		0.145			0.134	
R2c		0.411			0.387	

The baseline speeds, estimated from the model, are 9.9 km/h for conventional bikes and 11.8 km/h for e-bikes. These speeds represent the predicted averages at the reference levels of the model's variables or when continuous predictors are at their mean values. Cycling speeds were significantly influenced by slope, with notable differences between regular bikes and e-bikes. For conventional bikes, slopes below -5% were associated with a modest speed increase of 2.2% (0.2 km/h) compared to flat terrain. In contrast, slopes between -5% and -2% did not produce a statistically significant change in speed, suggesting that cyclists might control their acceleration on moderate slopes, potentially prioritizing safety and comfort. Uphill gradients, however, had a more pronounced slowing effect: slopes between 2% and 5% reduced speeds by 3.0%, approximately 0.3 km/h, while steeper inclines above 5% had the most substantial decrease, reducing speeds by 6.9% (0.7 km/h), and this finding is also confirmed by previous literature (Arnesen et al. 2020; Flugel et al. 2019; Clarry et al. 2019; Yan et al. 2024). This more pronounced effect of cycling uphill when compared to cycling downhill was also found by Parkin and Rotheram (2010).

For e-bikes, the relationship between slope and speed was more nuanced. On slopes below -5%, speeds increased by only 1.1% (0.14 km/h), a smaller effect than observed for regular bikes. Interestingly, moderate downhill slopes between -5% and -2% resulted in a surprising 2.4% (0.28 km/h) reduction in speed, suggesting a non-linear interaction between slope and speed for e-bikes. This counterintuitive finding may reflect deliberate speed reduction by riders due to safety concerns or the influence of e-bike motor dynamics, which could limit acceleration on such gradients. Uphill slopes followed a pattern similar to regular bikes, with speeds decreasing by 2.3% (0.3 km/h) on slopes between 2% and 5% and by 5.3% (0.6 km/h) on steeper inclines above 5%. The smaller reductions for e-bikes suggest that motor assistance helps mitigate some of the physical effort required to cycle uphill.

Gender also influenced cycling speeds, though the degree of influence varied by bike type and terrain. For conventional bikes, males rode on average 4.4% (0.44 km/h) faster than females, confirming findings from other studies (Boufous et al. 2018; Yan et al. 2024). For e-bikes, the gender gap in speed was slightly larger, with males riding 5.3% (0.6 km/h) faster than females. This aligns with findings by Flugel et al. (2019), who reported a comparable 5% speed difference between male and female e-bike riders, although in their study e-bikes reduced the difference, which did not happen in our study.

As for the interaction between gender and slope, while it revealed some variations, these effects were generally minor. For conventional bikes, on slopes below -5%, male riders experienced a modest additional speed increase of 0.7% (0.07 km/h) compared to females. On the other hand, on slopes between -5% to -2% and 2% to 5%, the interaction effects were not statistically significant. Interestingly, on steep uphill gradients (above 5%), male riders showed a slight relative reduction in speed compared to females, with a decrease of 0.3% (0.03 km/h). These findings suggest that, while males generally rode faster than females on conventional bikes, the interaction between gender and terrain was minimal, with only small adjustments based on slope conditions. As for the e-bikes, the only significant interaction was on gradients between -5% and -2%, where male riders achieved slightly higher speeds than females, with an additional increase of 0.8% (0.1 km/h).

Regarding cycling infrastructure, the presence of cycle lanes increased speeds by 3.3% (0.3 km/h) for conventional bikes and 4.9% (0.6 km/h) for e-bikes. Cycle paths provided an even greater boost, with speed increases of 4.3% (0.4 km/h) for conventional bikes and 5.40% (0.6 km/h) for e-bikes. Other studies also found that cycling infrastructure increases speeds (El-Geneidy et al. 2007; Flügel et al. 2019, Cubells et al. 2023), and the same result, infrastructure without physical separation being more beneficial regarding speed increases than segregated paths, was also found by Yan et al. (2024) in the Netherlands. Although the context in Brazil is completely different than in the Netherlands, some of the reasons mentioned by Yan et al. (2024) could also apply here, such as cyclists being pressured by vehicular traffic, more experienced cyclists using cycle paths or the ability to overtake other cyclists when the infrastructure is not segregated. Therefore, cofounders might be playing a role in our estimations. On the contrary to conventional bikes, riders on e-bikes gain more speed on segregated cycle tracks relative to unsegregated cycle lanes, suggesting that the effect on different levels of infrastructure segregation might be significantly different for riders on e-bikes vs riders on conventional bikes. The available data is not sufficient to determine if the observed differences are due to differences in cycling behavior, differences by the type of bicycle, or (more likely) by a combination of both effects.

The interactions between slope categories and infrastructure were almost all significant. For conventional bikes, slopes below -5% with cycle lanes were associated with a 1.3% decrease in speed, likely reflecting increased caution. In contrast, on slopes between -5% and -2%, cycle lanes led to a 4.2% speed increase. On slopes between 2% and 5%, cycle lanes contributed to a modest 1.2% speed increase, while on slopes above 5%, they had the most pronounced impact, increasing speeds by 10.8%, which shows that cycle lanes had an important role in supporting cyclists on challenging terrain. A similar pattern emerged for e-bikes, though with some differences in magnitude. On slopes below -5%, the presence of cycle lanes was associated with a 1.8% reduction in speed, while for slopes between -5% and -2%, speeds increased by 7.7%. On slopes between 2% and 5%, e-bikes saw a speed gain of 1.5%, while steep uphill slopes exhibited the largest benefit, with speeds increasing by 10.5%. Cycle paths also enhanced speeds across all slope categories but generally had weaker effects than cycle lanes.

For both bike types, the benefits of cycle paths were less pronounced on steeper gradients, suggesting that their design or placement may offer less support compared to cycle lanes. Overall, these results underscore the significant role of cycling infrastructure in enhancing cycling performance, particularly on challenging terrain. Cycle lanes, in particular, provided consistent and substantial benefits, especially on moderate downhill and steep uphill slopes. This superior performance may be attributed to the fact that they are segregated from traffic, which minimizes conflicts.

Road type and speed limits also significantly influenced cycling speeds. For conventional bikes, tertiary roads increased speeds by 5.9%, and primary roads increased speeds by 4.3%. The most substantial increase was observed on trunk roads, where speeds were 7.0% higher than the reference. In contrast, secondary roads were associated with a slight speed reduction of 2.3%. These trends (in line with those previously reported by Strauss & Miranda-Moreno, 2017) likely reflect variations in traffic flow, road design, and cyclists' perceptions of safety. For e-bikes, tertiary roads increased speeds by 5.2%, while primary roads resulted in similar gains. Trunk roads provided the largest increase, with speeds increasing by 9.8%. Secondary roads, however, showed a minor speed reduction of 0.7%, suggesting a more favorable performance on higher-class roads.

For conventional bikes, roads with a 60 km/h speed limit resulted in a 1.7% reduction in speed, while a 50 km/h limit led to a smaller reduction of 0.9%. In stark contrast, roads with a 30 km/h limit led to a substantial 30.7% reduction in speed. This sharp drop suggests that cyclists on conventional bikes may perceive these roads as particularly constrained, perhaps due to higher traffic density, or road design features that require them to slow down considerably. For e-bikes, the effects were different. At a 60 km/h limit, e-bikes exhibited a larger reduction in speed (6.5%), indicating greater sensitivity to higher speed limits, possibly because faster cruising speeds make cyclists more cautious. However, at lower limits, such as 50 km/h and 30 km/h, the reductions were smaller (1.7% and 0.7%, respectively), with the effect at 30 km/h not being statistically significant ($p = 0.1$).

Regarding the time of day, conventional bikes had higher speeds during early morning hours and progressively declined throughout the day. E-bikes had a very similar pattern, with the only difference being the night period, where speeds had a slight increase of 0.1%. These results suggest that early morning hours provided optimal conditions for cycling, possibly due to lower traffic volumes and was also found in previous literature (Clarry et al. 2019).

Proximity to intersections and traffic lights also had a significant influence on cycling speeds. For conventional bikes, proximity to intersections resulted in a 3.8% reduction in speed, similar to what was reported in the literature (Clarry et al. 2019; Strauss and Miranda-Moreno 2017; Yan et al. 2024), while proximity to traffic lights led to a smaller but significant reduction of 0.3%. Similar results were also found by Clarry et al. (2019). For e-bikes, the impact of proximity to intersections and traffic lights was slightly stronger. E-bikes were 4.2% slower near intersections, while traffic lights had a smaller but notable effect, causing a 0.8% reduction in speed. These findings suggest that, while both types of cyclists adjust their speeds similarly in these contexts, the reductions are more pronounced for e-bikes. In general, for a vehicle moving, delays due to stops depend on the cruising speed and on the acceleration and deceleration rates. E-bikes have larger cruising speeds (which push to larger stopping delays), and potentially larger acceleration and deceleration rates (which would push to lower stopping delays). Our results show that the e-bike effect on acceleration/deceleration, if it exists, does not compensate for the fact that e-bikes have larger cruising speeds, and therefore, stopping, for any reason, implies a larger delay than that for conventional bikes. Therefore, e-bikes could benefit more from straightforward paths with fewer intersections and traffic lights.

Regarding link length, conventional bikes had a 1.7% increase in speed for each standard deviation increase in road segment length, while e-bikes experienced a slightly higher benefit, with a 2.1% speed increase per standard deviation in road segment length. This highlights the fundamental role of well-connected and uninterrupted cycling routes in allowing riders to maintain higher, more consistent speeds. For trip distance, both conventional and e-bikes had a 3% increase in speed for each standard deviation increase in trip distance, suggesting that riders adjusted their pace based on the anticipated length of their journey, and similar results were found by Clarry et al. (2019).

As for the interaction between road segment length and slope, all of them were significant. For conventional bicycles, longer road segments were associated with higher speeds on downhill slopes, with a 6.7% increase in speed on slopes below -5%, and 4.5% for slopes between -5% and -2%, for each standard deviation increase in link length. In contrast, on uphill slopes, longer segments were linked to reductions in speed. For slopes between 2% to 5%, speed decreased by 0.9%, while for slopes above > 5%, the decrease was smaller at 0.6%, also for each standard deviation increase in link length.

For e-bikes, a different pattern was observed. On slopes below -5%, each standard deviation increase in road segment length was associated with a 6.0% increase in speed, while on slopes between -5% and -2%, the increase was slightly smaller at 4.1%. For uphill gradients, in contrast with conventional bikes, e-bikes shown a positive association between segment length and speed, with a 2.4% increase observed on slopes between 2% and 5% and a 1.5% increase on slopes above 5%. Both bicycle types showed increased speeds on longer road segments with downhill slopes, although e-bikes had slightly smaller gains compared to conventional bikes. On uphill segments, the patterns diverged significantly. For conventional bicycles, speeds decreased on longer uphill stretches, likely reflecting the increased physical effort required. In contrast, e-bike speeds increased on longer positive slopes, suggesting that motor assistance enabled riders to maintain or even accelerate their pace despite the gradient. This positive association underscores the advantage of motor assistance in mitigating the impact of slope on cycling effort.

Uber speeds also significantly impacted cycling speeds. For conventional bikes, each standard deviation increase in Uber speeds was associated with a 2.7% decrease in cycling speed, while for e-bikes, the reduction was slightly larger at 3.1%. Lastly, for both conventional and electric bicycles, the random intercept for baseline cycling showed a moderate variability in baseline speed across trips, with a variance of 0.03 for both bike types. The random slopes for slope categories, proximity to intersections and cycling infrastructure also showed modest variability for both bike types. As for the correlations among random effects, for conventional bikes, a strong negative correlation (-0.6 for conventional and -0.7 for e-bikes) between the random intercept and the cycle lane category suggested that trips with higher baseline speeds experienced less benefit from cycle lanes, indicating that the effect of cycling infrastructure was smaller for faster cyclists.

Interaction terms for the grouped data set

In order to observe how bike type moderated the relationship between speed and other factors, highlighting differences in responses between e-bikes and conventional bikes, we opted to run another model using the full dataset, using interactions between the bike type and the explanatory variables.

Table 2 – Interactions results

Interaction	**Coef.**	**Std. Error**	**p-value**
E-bike:Below -5%	-0.0131	0.0021	0.0000
E-bike:-5% to -2%	0.0005	0.0018	0.7563
E-bike:2% to 5%	0.0095	0.0022	0.0000
E-bike:Above 5%	0.0070	0.0047	0.0013
E-bike:Cycle lane	0.0146	0.0047	0.0019
E-bike:Cycle path	0.0040	0.0067	0.0019
E-bike:Male	0.0034	0.0048	0.4776
E-bike:Length	0.0090	0.0008	< 2e-16
E-bike:Tertiary	-0.0037	0.0033	0.2655
E-bike:Primary	0.0087	0.0038	0.0238
E-bike:Trunk	0.0224	0.0039	0.0000
E-bike:Secondary	0.0148	0.0042	0.0004
E-bike:30	0.0033	0.0055	0.5464
E-bike:50	-0.0081	0.0031	0.0083

E-bike:60	-0.0274	0.0180	0.1268
E-bike:Intersection	-0.0047	0.0013	0.0003
E-bike:Traffic light	-0.0042	0.0014	0.0027
E-bike:Total distance	0.0074	0.0017	0.0000
E-bike:Early Morning	0.0203	0.0078	0.0090
E-bike:Morning Peak	-0.0079	0.0035	0.0227
E-bike:Lunch	0.0134	0.0053	0.0110
E-bike:Afternoon	0.0184	0.0056	0.0010
E-bike:Evening Peak	0.0114	0.0035	0.0013
E-bike:Night	0.0288	0.0047	0.0000
E-bike:Uber speed	-0.0039	0.0007	0.0000
E-bike:Cycle lane:Male	-0.0056	0.0051	0.2844
E-bike:Cycle path:Male	-0.0056	0.0075	0.4555

For slope categories, slopes below -5% had a smaller effect on e-bikes compared to conventional bikes, which means that conventional bikes increased more their speed in this slope category, while the effect of slopes between 2% and 5% and above 5% was positive for e-bikes, which indicates that e-bikes performed better.

Gender did not significantly interact with bike type ($p = 0.48$ for male, $p = 0.66$ for other), indicating that the effect of using an e-bike on speed was consistent across gender categories. The effect of cycling infrastructure was higher for e-bikes, particularly with cycle lanes, as evidenced by the positive interaction (coefficient = 0.02). In contrast, the interaction with cycle paths was not significant ($p = 0.55$).

The effect of road type was bigger for e-bikes across most road types, except on tertiary roads, where no significant difference was observed. For speed limits, a significant difference was observed for 50 km/h, with the e-bikes decreasing more their speed than conventional bikes. For the hour categories, e-bikes had a better performance than conventional bikes, except for the morning peak. In all other periods of the day, the interaction term for electric bikes was positive.

The interaction effects for uber speed were also significant. Although both conventional and electric bikes slowed down as Uber speeds increased, the effect was more pronounced for e-bikes. Proximity to intersections and to traffic lights had a significant negative effect on e-bikes.

Spatial visualization

In order to showcase the difference in speed between the two bicycle types, Figure 2 shows the difference on average speed per link between each bike type for a part of the network.

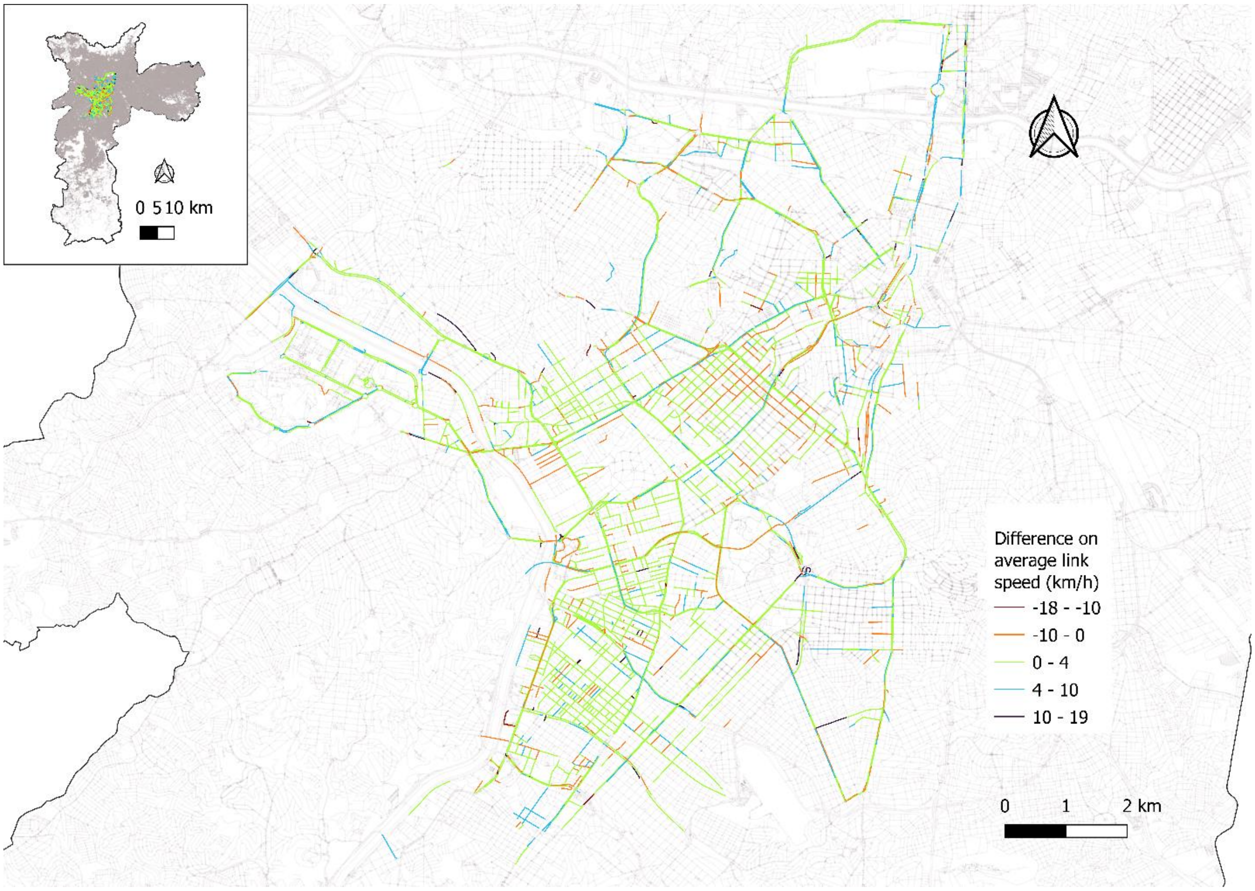

Figure 2. Difference on average link speed between conventional and electric bicycles in São Paulo

Most of the links had a 0-4 km/h speed advantage for e-bikes compared to conventional bikes, indicating that e-bikes offer a moderate performance improvement across the area on average. However, the map also shows pockets of the city where e-bikes had a more substantial 4-10 km/h speed advantage. Additionally, for a small number of links, the speed difference was negative, which means the average speed of conventional bikes was higher than e-bikes. It is also notable that the links with larger e-bike speed advantages are likely located on longer arterial roads, which are probably connected to a smaller number of intersections and other interruptions, enabling e-bikes to reach higher speeds.

Also, to showcase the impact of electric bicycles on the network, we measured the global efficiency of the transportation network for both conventional and electric bicycles. Using the average travel times as edge weights, the global efficiency was 0.0521 for conventional bikes and 0.0595 for e-bikes. This represents a 14.2% improvement in global efficiency for the e-bikes, reflecting its reduced travel times and enhanced overall connectivity compared to the conventional mode.

To further analyze the impact of electric bicycles on the network, we evaluated the difference in edge betweenness centrality between conventional and electric modes of travel. Figure 3 illustrates the edges within the 10th quantile of differences in betweenness centrality between electric and conventional bicycles. These edges represent the top 10% of changes, indicating where the introduction of electric bicycles has the most significant impact on network flow.

With the exception of a few edges located in the denser zone on the right side of the map, the majority of the top 10% of edges with the highest difference in betweenness centrality are concentrated along longer, likely arterial roads. This suggests that the introduction of e-bikes increases the reliance on these longer roads. This highlights how e-bikes may shift the critical pathways in the network toward major roads. These findings could inform infrastructure prioritization, such as enhancing safety and capacity along these arterial routes to better accommodate the growing use of e-bikes.

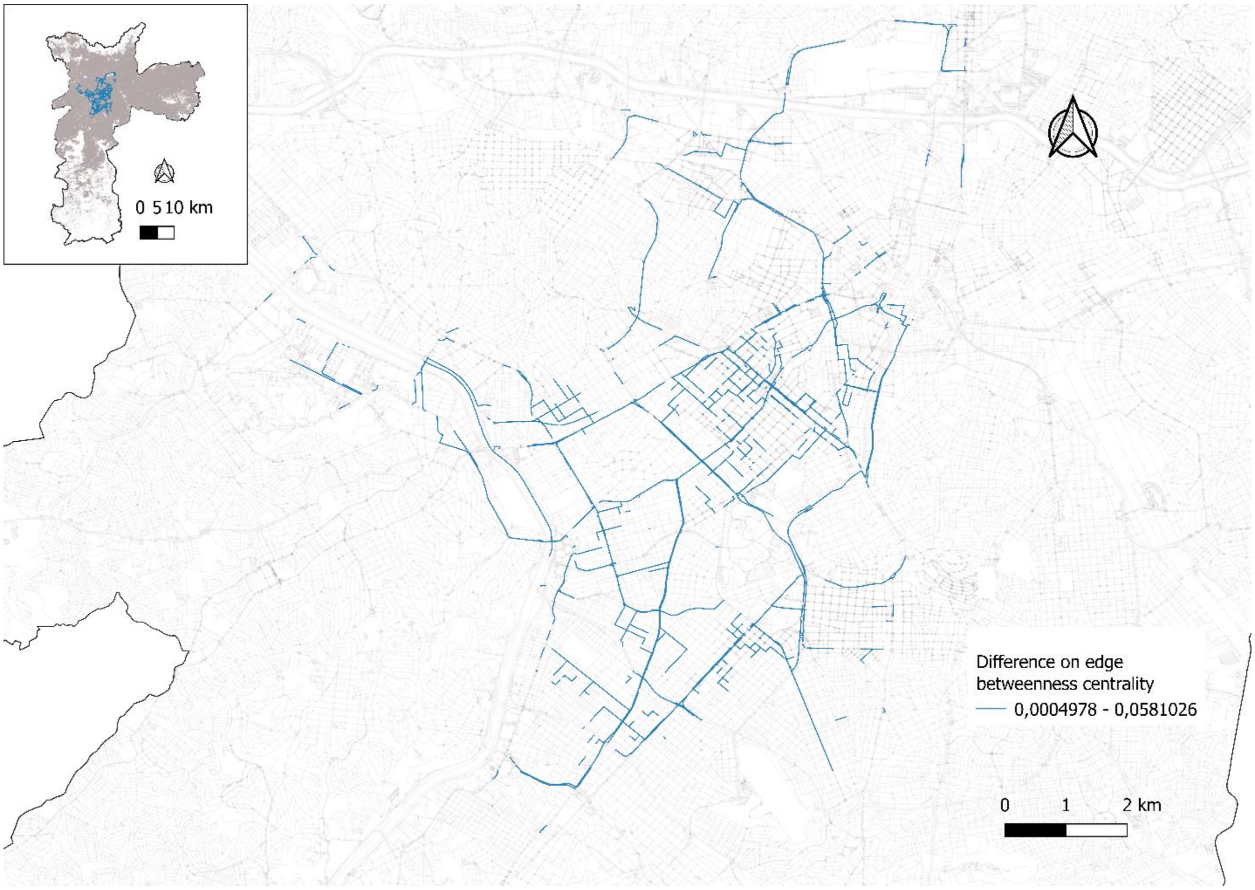

Figure 3. Difference on edges betweenness centrality between conventional and electric bicycles in São Paulo

Trip level

To examine the factors influencing cycling speed, we also developed a linear regression model with the average trip speed as the dependent variable. To account for differences between bike types, we included interaction terms between key variables and bike type, capturing potential variations in how these factors affect speed. Table 2 presents the detailed results of the estimated model for average trip speed.

Table 2 – Model for average trip speed

Variable	Coef.	Std. Error	p-value
Intercept	2.420	0.020	< 2e-16
Number of intersections	-0.003	0.000	< 2e-16
Number of traffic lights	-0.011	0.005	0.027
Elevation difference	-0.002	0.000	< 2e-16
Trip distance	0.000	0.000	< 2e-16
Percentage with cycling infrastructure	0.001	0.000	0.000
Bike Type			
Electric	0.185	0.052	0.000
Gender			
Male	0.070	0.012	0.000
Day of week			
Monday	-0.006	0.012	0.637
Tuesday	0.022	0.015	0.143
Thursday	-0.001	0.014	0.920
Friday	-0.013	0.014	0.343
Hour Category			
Early morning	0.165	0.028	0.000
Morning peak	0.014	0.014	0.326
Lunch	-0.005	0.025	0.837
Afternoon	-0.012	0.026	0.628
Evening peak	-0.034	0.015	0.025
Night	0.037	0.018	0.035
Interaction terms			
E-bike:Male	-0.113	0.033	0.001
E-bike:Percentage with cycling infrastructure	0.000	0.000	0.400
E-bike:Number of intersections	0.002	0.001	0.051
E-bike:Number of traffic lights	0.019	0.012	0.109
E-bike:Elevation difference	0.001	0.001	0.380
E-bike:Early morning	-0.001	0.069	0.989
E-bike:Morning peak	-0.019	0.038	0.613
E-bike:Lunch	-0.023	0.072	0.748
E-bike:Afternoon	-0.023	0.077	0.759
E-bike:Evening peak	0.014	0.042	0.742
E-bike:Night	0.021	0.045	0.634
E-bike:Trip distance	0.000	0.000	0.589
Model Fit			
Adjusted R-squared		0.3396	
Residual standard error		0.1974	
p-value		< 2.2e-16	

The model explained approximately 34% of the variance in cycling speeds (adjusted $R^2 = 0.3$) and showed an overall significant fit ($p < 0.001$). The baseline speed, estimated from the model, is 11.3 km/h. The speed of e-bikes was 18.5% (2.1 km/h) higher than those of conventional bikes, a substantial increase. Furthermore, while electric bicycles were associated with higher speeds compared to conventional bicycles, this advantage varied by gender, since males cycled 7.0% (0.8 km/h) faster than female cyclists. The interaction between gender and bike type was significant and negative (-11.3%, 1.3 km/h) indicating that men had a substantially smaller speed increase with electric bicycles when compared to women. This shows that, although men were faster cyclists than women in average, electric bikes reduce gender-based speed disparities at the trip-level.

For the percentage of dedicated cycling infrastructure, a one percentage point increase in the cycling infrastructure along the route led to a 0.01% increase in trip speed. This means that if 50% of a route had cycling infrastructure, the trip speed would increase by approximately 4.9% (0.6 km/h) compared to a route with no cycling infrastructure. Additionally, the interaction between bike type and percentage of cycling infrastructure was not significant, meaning that effect of this variable was the same for both bike types.

As for the number of intersections, each additional (non-signalized) intersection reduced speed by 0.3%. Similarly, traffic lights also had a negative impact, with each additional traffic light reducing speed by 1.1%. The interaction between electric bikes and intersections was marginally significant ($p = 0.05$), suggesting that while intersections generally reduce speed, the reduction may be slightly less pronounced for e-bikes at the trip level. However, it is important to note that results from the tracking point models indicate the opposite trend: intersections had a stronger negative effect on e-bikes compared to conventional bikes. This apparent contradiction likely arises due to differences in analysis levels. At the segment level, the immediate stop-start behavior of e-bikes near intersections shows a sharper speed loss, reflecting their higher cruising speeds and greater deceleration. By contrast, the trip-level model averages speeds across the entire trip, where motor assistance allows e-bikes to recover and compensate for these localized slowdowns.

Regarding the net elevation, each meter of elevation was associated with a 0.2% reduction in speed, reflecting the physical challenge of uphill cycling. And, in this case, no difference was found between bike types.

As for time of day, the speed was higher during early morning hours and progressively declined throughout the day, and the effect was the same for both bike types. Day-of-week variations showed no significant effects. Trip distance, however, was positively associated with speed, with each additional kilometer increasing speed by 0.08%. Additionally, the interaction between trip distance and bike type was not significant ($p = 0.6$), suggesting that speed adjustments related to distance were similar for both conventional and electric bikes.

Both the tracking point and trip-level models confirmed that e-bikes achieved higher speeds than conventional bicycles. Similarly, both models highlighted the influence of gender, with males cycling faster than females. In the tracking point model, males were 4.4% and 5.3% faster for conventional and e-bikes, respectively, while in the trip-level the difference was 7.0% between males and females for conventional bikes. The key difference between the models was on the impact of e-bikes, with the trip-level model suggesting that e-bikes helped reduce gender-based speed disparities. This suggests that, although males still maintain some advantages over females when using e-bikes, this difference dissipates over the entire trip, possibly due to interruptions such as intersections or other factors, resulting in a similar average speed for both genders.

Regarding cycling infrastructure, both models identified a positive impact on speed. However, in the tracking point model, the impact was greater for e-bikes, while in the trip-level model, the effect was closer for both bike types. A notable difference was observed in the impact of terrain: e-bikes outperformed conventional bikes in the tracking point model, but no such difference was found in the trip-level model. This discrepancy is likely related to interruptions during the trip, which diminish the advantage of e-bikes. Additionally, reductions in speed near intersections and traffic lights were more pronounced for e-bikes in the tracking point model, whereas the trip-level model showed that e-bikes experienced less speed loss at intersections compared to conventional bikes. This is likely due to the higher deceleration and subsequent higher acceleration of e-bikes.

Conclusions

In this study, speed models for conventional and electric bicycles were developed, using extensive GPS data from over 75,000 trips made by users of the Tembici bikesharing system in São Paulo. By employing linear mixed-effects models at the tracking-point level and regression models at the trip level, the research provides a comprehensive analysis of cycling speeds across various spatial, demographic, and environmental contexts. The inclusion of separate models for conventional and electric bicycles made it possible to highlight the unique dynamics of these bike types, addressing a key gap in urban mobility research. We added relevant explanatory variables that were not addressed previously in the literature, such as the speed of cars, as well as interactions between key variables.

Our findings show that uphill segments reduced speeds, but e-bikes experienced smaller reductions than conventional bikes. At the tracking point level, we found that males cycled faster than females on both bike types, whereas at the trip level, e-bikes nearly eliminated the gender gap. Cycling infrastructure consistently increased speeds for both bike types, with e-bikes benefiting more. The most significant benefits of cycling infrastructure were observed on steep gradients. Additionally, higher-class roads and lower speed limits were generally associated with increased cycling speeds. For both bike types, speeds were higher in the morning and decreased throughout the day. Proximity to intersections and traffic lights negatively impacted speed for both bike types. Trip distance and link length were positively associated with cycling speeds, but e-bikes benefited more than conventional bikes from longer links and trips. Additionally, we found that longer uphill segments reduced speeds for conventional bikes but increased speeds for e-bikes, and also that higher Uber speeds were associated with lower cycling speeds. E-bikes also increased the global efficiency of the network by 14.2%, and edge betweenness centrality analysis revealed that longer roads, likely arterial routes, gained greater importance with the use of e-bikes.

In conclusion, this study provides a nuanced understanding of cycling speeds in São Paulo by examining the distinct dynamics of conventional and electric bicycles across various spatial, demographic, and environmental contexts. The findings highlight the significant role of cycling infrastructure, road characteristics, and slope in shaping cycling performance. Electric bicycles demonstrate advantages on challenging terrains and longer trips, and their ability to reduce physical effort and narrow the gender speed gap underscores their potential to support more inclusive and sustainable urban transportation systems. These findings emphasize the need for cycling infrastructure, particularly segregated lanes, which are especially effective in optimizing e-bike performance. Future research should explore the interactions between environmental, cultural, and behavioral variables and e-bike speeds, as well as their broader integration into urban transportation systems across diverse geographical contexts. Comparative analyses of cities with varying cycling cultures and infrastructure configurations would further substantiate these findings and inform evidence-based planning for e-bike adoption in rapidly urbanizing regions.

References

Arnesen, P., Malmin, O.K., Dahl, E.: A forward Markov model for predicting bicycle speed. Transportation 47, 2415–2437 (2020). https://doi.org/10.1007/s11116-019-10021-x

Boeing, G.: OSMnx: New methods for acquiring, constructing, analyzing, and visualizing complex street networks. Comput. Environ. Urban Syst. 65, 126–139 (2017). https://doi.org/10.1016/j.compenvurbsys.2017.05.004

Boufous, S., Hatfield, J., Grzebieta, R.: The impact of environmental factors on cycling speed on shared paths. Accid. Anal. Prev. 110, 171–176 (2018). https://doi.org/10.1016/j.aap.2017.09.017

Clarry, A., Faghih Imani, A., Miller, E.J.: Where we ride faster? Examining cycling speed using smartphone GPS data. Sustain. Cities Soc. 49, 101594 (2019). https://doi.org/10.1016/j.scs.2019.101594

Csardi, G., Nepusz, T.: The igraph software package for complex network research. InterJ. Complex Syst. 1695 (2006).

Cubells, J., Miralles-Guasch, C., Marquet, O.: Gendered travel behaviour in micromobility? Travel speed and route choice through the lens of intersecting identities. J. Transp. Geogr. 106, 103502 (2023). https://doi.org/10.1016/j.jtrangeo.2022.103502

Flügel, S., Hulleberg, N., Fyhri, A., Weber, C., Ævarsson, G.: Empirical speed models for cycling in the Oslo road network. Transportation 46, 1395–1419 (2019). https://doi.org/10.1007/s11116-017-9841-8

Fortes, L.M., Giannotti, M., Soares de Freitas, F.: Gender, class and race uneven access to bike systems across five Brazilian cities. Cities 148, 104822 (2024). https://doi.org/10.1016/j.cities.2024.104822

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

Hagberg, A., Swart, P.J., Schult, D.A.: Exploring network structure, dynamics, and function using NetworkX. Los Alamos National Laboratory (2008). https://www.osti.gov/biblio/960616

Hijmans, R.J., Williams, E., Vennes, C., Hijmans, M.R.J.: geosphere: Spherical trigonometry for geographic applications. Comprehensive R Archive Network (2023). https://CRAN.R-project.org/package=geosphere

Jensen, P., Rouquier, J.-B., Ovtracht, N., Robardet, C.: Characterizing the speed and paths of shared bicycle use in Lyon. Transp. Res. Part D Transp. Environ. 15, 522–524 (2010). https://doi.org/10.1016/j.trd.2010.07.002

Latora, V., Marchiori, M.: Efficient Behavior of Small-World Networks. Phys. Rev. Lett. 87, 198701 (2001). https://doi.org/10.1103/PhysRevLett.87.198701

Meyer de Freitas, L., Axhausen, K.W.: The influence of individual physical capabilities for cycling adoption: Understanding its influence and mode-shift potentials. Transp. Res. Part A Policy Pract. 185, 104105 (2024). https://doi.org/10.1016/j.tra.2024.104105

Parkin, J., Rotheram, J.: Design speeds and acceleration characteristics of bicycle traffic for use in planning, design and appraisal. Transp. Policy 17, 335–341 (2010). https://doi.org/10.1016/j.tranpol.2010.03.001

Pereira, R.H.M., Saraiva, M., Herszenhut, D., Braga, C.K.V., Conway, M.W.: r5r: Rapid Realistic Routing on Multimodal Transport Networks with R5 in R. Findings (2021). https://doi.org/10.32866/001c.21262

Romanillos, G., Gutierrez, J.: Cyclists do better. Analyzing urban cycling operating speeds and accessibility. Int. J. Sustain. Transp. 14, 448–464 (2020). https://doi.org/10.1080/15568318.2019.1575493

Saneinejad, S., Roorda, M.J., Kennedy, C.: Modelling the impact of weather conditions on active transportation travel behaviour. Transp. Res. Part D Transp. Environ. 17, 129–137 (2012). https://doi.org/10.1016/j.trd.2011.09.005

Schleinitz, K., Petzoldt, T., Franke-Bartholdt, L., Krems, J., Gehlert, T.: The German Naturalistic Cycling Study – Comparing cycling speed of riders of different e-bikes and conventional bicycles. Saf. Sci. 92, 290–297 (2017). https://doi.org/10.1016/j.ssci.2015.07.027

Schleinitz, K., Petzoldt, T., Gehlert, T.: Risk compensation? The relationship between helmet use and cycling speed under naturalistic conditions. J. Saf. Res. 67, 165–171 (2018). https://doi.org/10.1016/j.jsr.2018.10.006

Strauss, J., Miranda-Moreno, L.F.: Speed, travel time and delay for intersections and road segments in the Montreal network using cyclist Smartphone GPS data. Transp. Res. Part D Transp. Environ. 57, 155–171 (2017). https://doi.org/10.1016/j.trd.2017.09.001

Transportation Research Board: Highway Capacity Manual. 5th edn. (2010).

Vragović, I., Louis, E., Díaz-Guilera, A.: Efficiency of informational transfer in regular and complex networks. Phys. Rev. E 71, 036122 (2005). https://doi.org/10.1103/PhysRevE.71.036122

Xie, Y., Smart, M., Noland, R.B.: Powering bikeshare in New York City: Does the usage of e-bikes differ from regular bikes? Transp. Plan. Technol. (2024). https://doi.org/10.1080/03081060.2024.2341301

Yan, H., Maat, K., van Wee, B.: Cycling speed variation: A multilevel model of characteristics of cyclists, trips and route tracking points. Transportation (2024). https://doi.org/10.1007/s11116-024-10477-6

Win, K.N.: OpenTopography DEM Downloader (Version 3.0) [QGIS plugin]. https://plugins.qgis.org/plugins/OpenTopography-DEM-Downloader/ (2024).