GEORGIA DOT RESEARCH PROJECT 22-30

FINAL REPORT

BIKEWAYSIM: USING CHANGES IN REVEALED IMPEDANCE TO ASSESS POTENTIAL BENEFITS OF NEW CYCLING INFRASTRUCTURE

OFFICE OF PERFORMANCE-BASED MANAGEMENT AND RESEARCH

600 WEST PEACHTREE STREET NW ATLANTA, GA 30308

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No.: FHWA-GA-24-2230	2. Government Accession No.: N/A	3. Recipient's Catalog No.: N/A
4. Title and Subtitle: BikewaySim: Using Changes in Revealed Impedance to Assess potential Benefits of New Cycling Infrastructure		5. Report Date: October 2024 6. Performing Organization Code: N/A
7. Author(s): Reid Passmore (https://orcid.org/0000-0001-6602-2702); Kari Watkins, Ph.D. (https://orcid.org/0000-0002-3824-2027); Randall Guensler, Ph.D. (https://orcid.org/0000-0003-2204-7427)		8. Performing Organization Report No.: 22-30
9. Performing Organization Name and Address: Georgia Institute of Technology, 790 Atlantic Drive, Atlanta, GA 30332-0355, Phone: (206) 250-4415, Email: kari.watkins@ce.gatech.edu		10. Work Unit No.: N/A 11. Contract or Grant No.: PI#0019312
12. Sponsoring Agency Name and Address: Georgia Department of Transportation, Office of Performance-based Management and Research, 600 West Peachtree St. NW, Atlanta, GA 30308		13. Type of Report and Period Covered: Final; Oct. 2022 – Oct. 2024 14. Sponsoring Agency Code: N/A
15. Supplementary Notes: Conducted in cooperation with the U.S. Department of Transportation, Federal Highway Administration.		

16. Abstract:

This research proposes a framework for assessing the impacts of new and existing cycling infrastructure using minimum impedance routing. Impedance represents the relative difficulty of cycling taking into consideration travel time, exposure to automobiles, hills, and the provision of cycling infrastructure. An all-paths network is created from OpenStreetMap data, and data on traffic volumes, vehicle speeds, the number of vehicle lanes, elevation, and bicycle facilities are reconciled with the all-paths network. Map-matched cycling GPS traces from the Cycle Atlanta app are cleaned and then used to calibrate link and turn impedance functions that consider cyclists' preferences for link and turn attributes. The calibration process uses stochastic optimization techniques to maximize the overlap between the map-matched and impedance paths. The calibrated link and turn impedance functions are then applied to the 250 square mile metro Atlanta study area to assess the impact of 38 planned bicycle facilities for 3.6 million trips from the Atlanta Regional Commission Activity Based Model. Both the minimum travel time and impedance paths are calculated for the existing network and compared against the minimum impedance paths for the future network with the planned bicycle facilities. The results were then processed to create metrics and visuals on trip impedance reduction, percent detour, change in link betweenness centrality, impedance reduction contribution, and bike sheds. The calibrated impedance factors and framework are then applied to Savannah, GA to assess the impacts of 57 proposed bicycle facilities. This demonstrated that the framework and calibrated impedance factors could likely be used throughout the state of Georgia to assist GDOT, MPOs, and cities in assessing and prioritizing cycling infrastructure projects.

17. Keywords: Bicycling, Impedance, Shortest Path, Route Choice, Accessibility, Routing	18. Distribution Statement: No Restriction		
19. Security Classification (of this report): Unclassified	20. Security Classification (of this page): Unclassified	21. No. of Pages: 192	22. Price: Free

Form DOT 1700.7 (8-72) Reproduction of completed page authorized

GDOT Research Project No. 22-30

Final Report

BIKEWAYSIM: USING CHANGES IN REVEALED IMPEDANCE TO ASSESS PERCEIVED SAFETY BENEFITSOF NEW CYCLING INFRASTRUCTURE

By

Randall Guensler, Ph.D.
Professor – School of Civil and Environmental Engineering

Kari Watkins, Ph.D.
Associate Professor – School of Civil and Environmental Engineering

Reid Passmore

Graduate Research Assistant

Georgia Tech Research Corporation

Contract with
Georgia Department of Transportation

In cooperation with
U.S. Department of Transportation
Federal Highway Administration

October 2024

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	$candela/m^2$	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	$candela/m^2$	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

* SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.
(Revised March 2003)

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

Commented [BD1]: Chapter numbers appear off

The lack of cycling infrastructure is a major deterrent to the use of bicycle transportation in the United States. Planners and engineers currently lack a comprehensive set of tools that can be used to assess and communicate how new and improved cycling infrastructure would potentially improve cycling mobility and accessibility. Without these tools, cycling infrastructure may be built ad hoc or where it is politically convenient, instead of where it would be the most effective, thus reducing the likelihood that bicycles will be used for transportation. This research outlines a framework for assessing new and existing cycling infrastructure through minimum impedance routing, which is calibrated through stochastic optimization techniques on a dataset of monitored cycling GPS traces. Impedance represents the relative difficulty of cycling along different routes, taking into consideration travel time, road grade, exposure to traffic, turn movements, characteristics of cycling infrastructure provided, etc.

The framework is implemented by first developing an all-paths network (containing all viable streets and non-motorized paths for cycling) from OpenStreetMap data. Then, data on traffic volumes, vehicle speeds, the number of vehicle lanes, elevation, and bicycle facilities are joined to the all-paths network. Once joined, links where cycling is not viable (Interstates, Interstate access ramps, sidewalks, parking aisles, driveways, etc.) are removed. In the case of sidewalks, it should be acknowledged that cyclists do often ride on sidewalks if they find it uncomfortable to ride on the adjacent road (Barajas, 2021; Chaloux & El-Geneidy, 2019; Marshall et al., 2017). However, sidewalks were infeasible to include for the analyses of this report because the GPS data are not accurate enough to determine if a cyclist was on the road or the sidewalk. In the research study area, the all-

Commented [BCD2]: Is sidewalk riding able to be removed from the CycleAtlanta GPS traces? If not, then cyclists may be choosing to ride on the sidewalk but would be reported as being on the road in this model, which would give the road an inaccurately low impedance.

Note that I would want to consider sidewalk riding in the model as it is a common for riders to do, especially for lower income riders who are riding on hostile roads out of necessity.

paths network consisted of 76,912 links, 66,151 nodes, and 222,764 link-to-link turn opportunities.

Processed and filtered cycling GPS traces from the Cycle Atlanta app were map-matched to the all-paths network. More than 1,900 trips, across 586 cyclists, were successfully map-matched to the all-paths network. These data were used to calibrate link and turn impedance functions for minimum impedance routing. The impedance calibration process uses stochastic optimization techniques to maximize the similarity between the map-matched and impedance routes. At the time this report was completed, the best-performing combination of impedance attributes included: the presence of a multi-use path, the presence of a bike lane, the number of lanes, whether the link's average grade was above 4%, and the presence of an unsignalized turn across a major road. The researchers are working on calibrating additional impedance factors that should be considered (e.g., bicycle infrastructure separation, pedestrian volumes on multi-use paths, user characteristics, etc.).

The calibrated link and turn impedance functions reported herein were then applied to the 250 square mile metro Atlanta study area to assess the impact of 38 planned bicycle facilities for 127,682 unique origin-destination pairs (representing 3.6 million trips across from the Atlanta Regional Commission Activity Based Model TIP Amendment Six 2030 model run). The minimum travel time and minimum impedance routes were calculated for the existing network and compared against the minimum impedance routes for the future network that contained the planned bicycle facilities. The results were then processed to create metrics and visuals on trip impedance reduction, percent detour, change in link betweenness centrality, impedance reduction contribution, and bikesheds.

The calibrated impedance factors and framework were then applied to Savannah, GA to assess the impacts of 57 proposed bicycle facilities. The success of this demonstration indicates that the framework and calibrated impedance factors are likely to be transferrable throughout the state of Georgia to assist GDOT, MPOs, and cities in assessing and prioritizing cycling infrastructure projects.

The open source BikewaySim repository contains scripts for creating an all-paths network from OpenStreetMap data, reconciling other data sources with the all-paths network, map matching cycling GPS traces, calibrating link and turn impedance functions, and using the framework to evaluate new and improved cycling infrastructure.

CHAPTER 1. INTRODUCTION

According to the 2017 National Household Travel Survey (NHTS), about 46% of all single occupancy motor vehicle trips in the U.S. are three miles or less, yet the NHTS estimates that only 1% of all trips are performed by bicycle (Federal Highway Administration, 2017). This mismatch is suspected to be in large part due to the real and perceived danger faced by cyclists when they share the same right-of-way as motor vehicles (Buehler & Pucher, 2021). According to GDOT's crash data, 30 cyclists died on Georgia roads in 2020 (Georgia Department of Transportation, 2022). Despite only accounting for 0.02% of all crashes, cyclists represent 1.85% of fatalities in Georgia. Past estimates of cycling fatalities per 100 million kilometers cycled show the U.S. well above peer European countries like the U.K., the Netherlands, and Germany (Buehler & Pucher, 2021).

Cyclists are particularly vulnerable because their chances of surviving if struck by a car dramatically decrease when vehicle speeds exceed 30 mph (Groeger, 2016). Separated cycling facilities are one such proposed solution to make cycling safer and more accessible. The literature suggests that separated cycling facilities can decrease crash risk (DiGioia et al., 2017), and people, regardless of cycling frequency or ability, state that they would be more willing to try cycling on roads with bicycle facilities that provide a buffer or separation from car traffic (Watkins et al., 2019).

With modern, Dutch-inspired, bicycle facility design guidelines from NACTO, the technical guidance is available to develop high-quality, separated bicycle facilities. However, given limited funding for the construction of active transportation

infrastructure, cities, MPOs, and/or state departments of transportation (DOTs), such as the Georgia Department of Transportation (GDOT), need tools to highlight the location and degree of separation with which these facilities would provide the greatest increase in the mobility and accessibility of cycling (i.e., bike ability) so that alternatives that stand to have the greatest impact are implemented.

To that end, this research developed the BikewaySim analytical framework for assessing how new and improved bicycle infrastructure reduces impedance for, and improves mobility and accessibility of, cycling. BikewaySim is an open-source repository that is primarily coded in Python. BikewaySim contains several modules for downloading the necessary data to perform most of its basic functions, making it deployable for any study area in the state of Georgia as well as most metropolitan areas in the United States with supplementary data from other state DOTs. The BikewaySim code repository contains scripts for creating an all-paths network from OpenStreetMap data, reconciling other data sources with the all-paths network, map matching cycling GPS traces, calibrating link and turn impedance functions, and assessing route impedance for new, improved, and existing cycling infrastructure.

The contents of this report are split into six additional chapters:

- Chapter 2 presents a comprehensive literature review of cycling route choice, identifying prior models designed to represent cyclists' revealed preferences and tradeoffs for route attributes (i.e., route length, proportion of the route on bicycle facilities). Findings from this cycling route choice literature review informed the development of the cycling impedance functions used in BikewaySim. In addition, existing measures for cycling accessibility are discussed, as these

measures are incorporated into Bikeway Sim's bicycle facility assessment framework.

- Chapter 3 details the development of a cycling-specific, all-paths network, which formed the main data structure for developing and applying Bikeway Sim's impedance functions and assessing the impacts of new cycling facilities. Bikeway Sim's main network structure is derived from OpenStreetMap data, which are publicly available. However, additional data from local, regional, state, and federal government agencies had to be reconciled with the OpenStreetMap data to introduce basic data on the number of lanes, speed limits, and other infrastructure elements not present in OSM. This chapter also discusses the QA/QC procedures required to use an open-source network that does contain errors and proposes future work and recommendations for further developing cycling all-paths networks and collecting relevant attribute data.
- Chapter 4 describes Bikeway Sim's cycling impedance measure and presents a method for calibrating cycling impedance functions from cycling GPS traces through stochastic optimization. The method is carried out using GPS traces collected from the Cycle Atlanta smartphone app between 2012-2016. After these GPS traces had been cleaned and map-matched to the network developed in Chapter 3, a total of 1,792 traces were used to develop initial impedance functions for the impedance calibration process.
- Chapter 5 demonstrates Bikeway Sim's ability to assess the impacts of new bicycle facilities. Using the initial impedance functions developed in Chapter 4, 38 planned bicycle facilities in the City of Atlanta were assessed for 3.6 million

trips across 127,000 origin-destination pairs taken from the Atlanta Regional Commission's TIP Amendment Six Activity Based Model run. Metrics for trip impedance reduction and the additional minutes of travel due to impedance are visualized at the traffic analysis zone level to communicate the broad impacts of the planned cycling facilities. The "link betweenness centrality" metric is used to visualize network flow changes as a result of impedance and the planned cycling facilities. Lastly, the impedance reduction impact of each planned cycling facility is calculated, and the planned cycling facilities are then ranked in order of increasing impedance reduction impact per mile. An additional visualization, in the form of a bike shed, is also presented to show the localized impact of cycling impedance and new cycling infrastructure. Analytical results are discussed, and recommendations are made for how planners and engineers can use BikewaySim results in the bicycle facility planning process.

- Chapter 6 replicates the analyses performed in Chapter 5 for Savannah, GA to demonstrate the transferability of the BikewaySim analytical framework and the calibrated impedance functions from in Chapter 4. A total of 57 planned bicycle facilities in the Savannah metropolitan area were assessed for 17,856 work trips across 15,275 origin-destination pairs (derived from the US Census Bureau's Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics data). This chapter uses the complete workflow of the BikewaySim repository from developing and processing a cycling all-paths network to assessing the impacts of proposed cycling facilities.

- Finally, Chapter 7 concludes the report with additional discussion of the research findings, recommended applications of BikewaySim in state and regional transportation planning, the current limitations of this research, future work designed to refine the impedance calibration process, and the need to collect many more cycling GPS traces to better understand cyclist travel behavior.

CHAPTER 2. LITERATURE REVIEW

Cycling route choice research seeks to develop behavioral inferences about how cyclists choose a route between an origin (point A) and destination (point B). Route choice is typically framed as a random utility problem, where cyclists choose amongst a set of feasible routes based on the attributes of each route (Ben-Akiva & Bierlaire, 2003). These route choice models demonstrate the tradeoffs that cyclists make between various route attributes across a feasible set of routes. Route attributes include but are not limited to the travel time of a route, length of a route, the number of turns (and type of turns) along a route, the distribution of road grade along a route, the total elevation gain of a route, the length of the route on bicycle facilities, the length of the route on major arterials, etc.

Cycling route choice research at this level originated in the early 2010s, once smartphones made it feasible to collect volunteered GPS traces of bicycle activity on a large scale. Before the 2010's, researchers relied primarily on stated preference survey data to assess the impact of route attributes on cyclist choices. This report examines 19 cycling route choice studies that estimated multinomial logit models (MNL) for route selection, using activity data collected from cycling GPS traces. These studies were conducted primarily in North America and Western Europe in the following cities/regions: Davis, CA, USA (Fitch & Handy, 2020), San Francisco, CA, USA (Fitch & Handy, 2020; Hood et al., 2011), Seattle, WA, USA (Chen et al., 2017), Portland, OR, USA (Broach et al., 2012), Eugene, OR, USA (Zimmermann et al., 2017), Atlanta, GA, USA (Misra, 2016; Misra & Watkins, 2018), Phoenix, AZ, USA (Khatri et al., 2016; Shah & Cherry, 2021), Waterloo, ON, CA (Casello & Usyukov, 2014), Copenhagen, DK (Prato et al., 2018; Skov-Petersen et al., 2018), Amsterdam, NL (Koch & Dugundji,

2021; Ton et al., 2017), the Netherlands (Bernardi et al., 2018), Zurich, CH (Meister et al., 2022; Menghini et al., 2010), and Tel Aviv, IL (Ghanayim & Bekhor, 2018).

Generally speaking, these models have shown that the following attributes significantly influence cycling route choice, all decreasing the utility of a route and the likelihood of a route being chosen: increasing distance (Broach et al., 2012; Ghanayim & Bekhor, 2018; Hood et al., 2011; Khatri et al., 2016; Menghini et al., 2010); the number of turns, sometimes separated by type of turn (Broach et al., 2012; Hood et al., 2011; Khatri et al., 2016); and the proportion of a routes on steep hills, measured by average percent grade, percent grade categories, or cumulative elevation gain (Broach et al., 2012; Hood et al., 2011; Meister et al., 2022; Prato et al., 2018). Meanwhile, the presence of cycling infrastructure (regardless of type) increases the utility of a route (Broach et al., 2012; Hood et al., 2011; Khatri et al., 2016; Menghini et al., 2010).

Models have also found that a variety of land-use attributes also influence route choice, although the directionality of the utility is not always consistent across these studies and respective models (Ghanayim & Bekhor, 2018; Koch & Dugundji, 2021; Prato et al., 2018). Additional operating environmental variables also influence route choice, such as: adjacent traffic volumes (Broach et al., 2012; Khatri et al., 2016), the number of traffic signals (Broach et al., 2012; Khatri et al., 2016; Koch & Dugundji, 2021), the number stop signs (Broach et al., 2012; Hood et al., 2011), and the presence or absence of street lighting (Chen et al., 2017).

Koch and Dugundji (2021), Broach et al. (2012), and Prato et al. (2018) found that additional traffic signals decreased the utility of a route, likely due to the additional delay incurred, but Shah et al. (2021) and Khatri et al. (2016), both using data from bikeshare

users in Phoenix, AZ, found that traffic signals added utility to a route. The authors attributed this to Phoenix's gridded street network design, where cyclists riding on local streets will inevitably have to cross a major road at an unsignalized intersection. Similarly, a Copenhagen study also found that traffic signals added utility to a route, which the authors suspected aided cyclists when crossing major roads (Skov-Petersen et al., 2018). The Portland study found that all-way stop sign intersections had less dis-utility than signalized intersections, and that both straight and left turn movements across high-volume roads at unsignalized intersections decreased the utility of a route more than turns at signalized intersections (Broach et al., 2012).

While increasing route distance decreases the utility of a route, the extent of this decrease appears to vary by trip purpose and study location. In Portland, commuters preferred more direct trips than non-commuters (Broach et al., 2012). Additionally, cyclists in places with high cycling mode shares and extensive cycling networks tended to choose the shortest route, potentially supporting the assertion indicating that these cyclists don't have to travel far to find cycling infrastructure to their destination (Bernardi et al., 2018; Fitch & Handy, 2020; Ton et al., 2017). This variation in the degree of dis-utility exists for hills as well. In Hood et al. (2011) and Broach et al. (2012), women showed a higher avoidance of hills, and Meister et al. (2022) showed that cyclists on e-bikes were less impacted by hills.

Preferences for bicycle infrastructure also varied by type of infrastructure, study area location, and the cyclists' gender. It would seem logical to assume that preferences for cycling infrastructure would be hierarchical according to the infrastructure's separation from traffic (i.e., in order of increasing utility: bike lanes, separated bike lanes, and multi-

use paths), but this was not always the case. While Fitch and Handy (2020) found this to be true in San Francisco, it was not found to be true in Davis, where bike lanes and multi-use paths were equally preferred. The authors suggest that this is due to the lack of a connected network of cycling infrastructure in San Francisco, compared to Davis's extensive interconnected infrastructure that has continuously expanded and evolved over a period of more than 50 years (Fitch & Handy, 2020; Schultheiss et al., 2018). However, these differences may also arise from interaction effects between infrastructure design and vehicle traffic volumes and speeds, or other specific design elements beyond cycling network connectivity. Similarly, Ton et al. (2017) found that cyclists in Amsterdam had no strong preferences for separated cycling infrastructure, which the authors attributed to Dutch road design standards that only recommend mixed traffic when automobile speeds are low (i.e., the separated infrastructure is predominantly inherent). While bicycle lanes in other studies were preferred, bicycle lanes in Portland were not a significant route attribute, except when they were present on bridges (Broach et al., 2012). Rather, multi-use paths were the most preferred followed by bicycle boulevards in Portland. The authors attributed this observation to the fact that bicycle lanes in Portland are often located on busy arterials that can be easily avoided by using Portland's extensive network of bicycle boulevards. Lastly, several studies showed that women had higher preferences for cycling infrastructure than men (Meister et al., 2022; Misra & Watkins, 2018).

In assessing the literature on bicycle route choice, it is important to remember that each study is unique, with different participant demographics, different potential tolerances for mixed traffic amongst cyclists studied, different data collection methods, different and evolving local infrastructure design and availability, different operating environment

conditions (e.g., network signalization, traffic volumes, speeds, heavy-duty truck activity, etc.), and even different environmental conditions (e.g., heat, humidity, and rain). Hence, it is difficult to compare the magnitudes of cyclists' preferences across studies. That is, each resulting model reflects local user demographics, road designs, land use, topography, multimodal connectivity, miles/kilometers of dedicated cycling infrastructure, and road network complexity.

Cycling route choice models are built on GPS traces collected from monitored bicycle trips, and these trips may not be representative of all people who currently cycle or those who would potentially cycle (especially in cities with a low cycling mode share), and these models use different network datasets that have varying availability of attributes that may also be coded differently (e.g., elevation as cumulative elevation gain or average percent grade). Studies that have estimated multiple MNL models also found differences in preferences between males and females, commuters and non-commuters, regular and new bikeshare users, and across persons (Broach et al., 2012; Fitch & Handy, 2020; Khatri et al., 2016; Koch & Dugundji, 2021; Misra, 2016).

MNL model coefficients are also notoriously difficult to interpret. Results from cycling route choice models are often interpreted by converting model coefficients into average marginal rates of substitution (MRS). MRS values represent how much more of, or less of, an attribute a cyclist would trade for a one-unit increase in another attribute. In several cases, the MRS is calculated with respect to the distance coefficient, so that a change in an attribute is seen as a percent distance reduction or addition to a trip. Broach et al. (2012) expresses this as an equivalent percent change in distance, as shown in Equation 1. Some models are calibrated in value of distance (VoD) space so that the coefficients

themselves already represent the percent change in distance from a 1-unit increase in an attribute (Meister et al., 2022; Prato et al., 2018).

$$\text{Equivalent \% } \Delta\text{distance} = \left(e^{\Delta_{\text{attribute}}\frac{\beta_{\text{attribute}}}{\beta_{\ln(\text{dist})}}} - 1\right) * 100$$

Equation 1: Equivalent Percent Change in Distance

Commented [BD3]: Fix formatting for equations (typ)

Equation 1 finds the "equivalent percent change in distance" given the coefficients from a cycling route choice model and a change in a route attribute. A negative value corresponds to a reduction in equivalent percent change in distance. As noted earlier, most existing cyclist route choice models are developed and calibrated in VoD. This is significantly different than most vehicle route choice models, which are developed and calibrated using travel time and value of time (VoT), but cyclists and pedestrians generally do not face traffic-induced congestion delay. Hence, value of distance can generally be converted to value of time using average cyclist speeds (for various demographic user groups).

Using the coefficients from the Broach et al. (2012) non-commute model, a one-unit increase in the proportion of a route on a bike path (i.e. a route with no bike paths compared to a route that is entirely on bike paths), corresponds to an equivalent percent change in distance of -26%:

$$(\Delta_{\text{aaaaaaaaaaaaaaaa}} = 1.0, \beta\beta_{\text{ppoaappppppaaaauppppp ppoo aaaabbaa ppaaaah}} = 1.57, \beta\beta_{\ln(\text{ddaaddaa})} = -5.22).$$

In other words, for the example above, a cyclist would prefer to ride up to 26% further or longer than the shortest distance or travel time route on average, if it meant taking a route

entirely composed of bike paths. To further illustrate using the example above, a typical cyclist would perceive a 1.0-mile ride (i.e., a 7.5-minute ride, assuming an average cycling speed of 8.0 miles per hour) on a route that is entirely composed of multi-use paths as being only 0.76 miles in length or 5.7 minute in duration (Table 1).

A positive marginal rate of substitution corresponds to an increase in equivalent percent change in distance (and travel time). Using the coefficients from the Broach et al. (2012) non-commute model again, an additional turn per mile corresponds to an equivalent percent change in distance of 7.4%:

$$(\Delta_{\text{aaaaaaaaaaaaaaaa}} = 1.0,\ \beta\beta_{\text{nnaaaaappdd ppaaaa mmaamaa}} = -0.371, \beta\beta_{\ln(\text{ddaaddaa})} = -5.22).$$

A route with one more turn per mile is equivalent in utility to a route that is 7.4% longer. In other words, a cyclist would avoid adding an additional turn per mile unless it reduced the trip distance or travel time by 7.4% or more on average. To further illustrate, a cyclist would perceive a 1.0-mile route (7.5-minutes) with one turn per mile to be 1.07 miles in length (8.3 minutes) on average (Table 1).

Table 1: Example Equivalent Percent Change in Distance Calculations

Example	Equivalent %Δ distance for $\Delta_{\text{aaaaaaaaaaaaaaaa}} = 1$	Actual Distance (Actual Travel Time)	Perceived Distance (Perceived Travel Time)
Route entirely on bike paths	-26%	1.0 miles (7.5 minutes)	0.76 miles (5.7 minutes)
Route with one turn per mile	7.4%	1.0 miles (7.5 minutes)	1.07 miles (8.3 minutes)

Values derived from Broach et al. (2012)

MRS values have the potential to standardize model results across cities, but it remains unclear if MRS comparisons across cities are statistically valid. Fitch and Handy (2020) also found significant variation in the MRS between persons even within the same study

area, so it seems inadvisable to compare MRS values at face value across these studies, given that the variability in participant demographics, participant experience, local infrastructure design, and local transportation network operating conditions is so large.

It is also not clear whether cycling route choice models can predict route choice effectively. Out of the 19 cycling route choice papers studied, only a few performed some form of out-of-sample model verification to test the predictive power of the estimated route choice model. Hood et al. (2011) validated their model using a holdback sample (without including the chosen route) and found that the route choice model was able to predict the chosen route exactly for 15% of the holdback sample. Meister et al. (2024) compared two route choice models, a link-based route choice model and the shortest path, the researchers and found that the shortest path outperformed the two route choice models in route prediction.

Even if a reliable cycling route choice model can be developed and calibrated for a region, integrating a cycling route choice model into the regional TDM may prove to be very difficult. Route choice models are complex and require modelers to generate a large set of potential routes within the model, from which each choice will be made. Recent efforts in the cycling literature have instead focused on translating findings from route choice models into link impedance that can be used directly in TDM shortest path modeling. Existing methods include creating stress modification factors based on MRS values (Lowry et al., 2016) and scaling route choice utility function coefficients by the shortest distance route so that they can be used in an impedance function (Broach & Dill, 2016).

This research took an alternative approach that would avoid the need to develop cycling route choice models, by calibrating an impedance function using cycling GPS traces. Impedance function calibration is treated as an optimization problem, where the dissimilarity between the chosen routes and the modeled routes from the impedance function is minimized (see Schweizer et al., 2020). This approach is described further in Chapter 5.

ASSESSING THE IMPACTS OF NEW CYCLING INFRASTRUCTURE

This section of the literature review focuses on research that has used cycling route choice utility models or cycling impedance estimates to quantify the potential impacts of new cycling infrastructure. There were only a few papers identified, which highlights the research gap between estimating cycling route choice models and applying cycling route choice models to cycling infrastructure planning.

Hood et al. (2011) developed a route choice model for San Francisco and demonstrated how the model could be used to quantify utility increases resulting from new cycling infrastructure. However, due to the stochastic nature of their choice set algorithm, their research occasionally found a decrease in utility for infrastructure improvements, which was counterintuitive (Nassir et al., 2014).

Research by Lowry et al. (2016) interpreted MRS values from previous route choice models and converted them to impedance for shortest path routing purposes. Their research employed two impedance functions, one for links and one for turns. Link impedance was modified by a stress factor that was quantified from road link attributes, which included consideration of the number of lanes, speed limit, road grade, and provision of cycling infrastructure. The turn impedance considered the number of lanes,

the speed limit, the presence of a traffic signal, and the presence of a bicycle intersection treatment (e.g., bike boxes, protected intersections, etc. (NACTO, 2014)) for the street that was crossed as well as the turn type (i.e., left, straight, and right).

Using the limited studies present in the literature that assessed cycling infrastructure improvements using cycling route choice models and more general literature that assessed cycling network connectivity, four metrics to assess the impacts of cycling infrastructure improvements were identified:

- **Directness:** Directness refers to the physical distance reduction provided by an infrastructure improvement (reduced circuity). Increasing the directness of cycling networks makes cycling more attractive by shortening the trip distance to places people want to go (CROW, 2009; FHWA, 2018). Thus, using change in trip directness can be a way of prioritizing projects (Cabral et al., 2019). Studies in places with high cycling ridership like Davis, CA and Amsterdam, NL have lower detour rates likely due to their highly-connected network of cycling infrastructure which provides for direct and safe routes (Fitch & Handy, 2020; Ton et al., 2017).
- **Impedance:** Infrastructure improvements are usually designed to reduce cyclists' travel costs by reducing travel time and a variety of perceived costs associated with the trip (e.g., physical exertion costs, convenience costs, experience costs, safety costs, etc.). Cyclists are assumed to want to reduce their overall travel costs; hence, infrastructure improvements can be prioritized by finding the projects that provide the greatest decrease in impedance for the highest number of origin-destination pairs (Hood et al., 2011; Nassir et al., 2014). Impedance can be

in units of time or money but can also be unitless. Impedance is relative, meaning that once a choice is being made, the relative difference in impedance across alternatives is more meaningful than the impedance value itself.

- **Betweenness Centrality:** Betweenness centrality refers to how often a link appears across all shortest paths (Alattar et al., 2021; Freeman, 1977; Newman, 2008). A high betweenness centrality indicates that a transportation network link provides an important connection for many trips. Cooper (2017) also showed that betweenness centrality could be calibrated with targeted cycling counts to predict ridership across a network. Cycling infrastructure improvements on links with high betweenness centrality may further enhance cycling travel, and cycling facilities connected to these links are more likely to be used (Lowry et al., 2016). Changes in betweenness centrality can be used to assess the addition or removal of cycling infrastructure, restrictions on what links can be traversed, and changes in impedance. For example, if cycling infrastructure designed to reduce link impedance is added to high-stress links, betweenness centrality will likely increase for these links. If the impedance reduction is large enough, these links will appear in more shortest path routes.
- **Destination Accessibility:** Destination accessibility refers to whether people can reach key destinations (e.g. hospitals, grocery stores, dentists, restaurants, workplaces) within specified time or cost tolerances. This method requires the implementation of thresholds that define whether a trip is considered feasible. Lowry et al. (Lowry et al., 2016) created a basket of key destinations and assessed whether every trip origin had access to at least two of each type of key

destination. The overall cycling trip distance allowed was two miles and stress tolerances were created to reduce the candidate bikeshed for trip origins. Infrastructure improvements can then be evaluated by seeing how they increase access to destinations. A similar approach was taken by the online People for Bikes Bicycle Network Analysis (BNA) tool (People for Bikes, n.d.).

LITERATURE REVIEW SUMMARY

While there is an extensive body of literature on cycling route choice for different populations in different regions, there are fewer studies that demonstrate how to best apply the estimated models to planning new cycling infrastructure. To address this gap, this paper presents a framework for modeling the impacts of new cycling infrastructure. Before any technical analyses could be done, various GIS data representing streets, bicycle facilities, elevation, and more had to be reconciled to form a cycling all-paths network for impedance calibration and routing purposes.

CHAPTER 3. BIKEWAYSIM NETWORK DEVELOPMENT

This chapter begins with a discussion about what will be defined as the "all-paths" cycling network for trip routing, followed by how the all-paths network is derived from OpenStreetMap. The all-paths network developed for this project is then summarized. The graph structure and basic data schema for BikewaySim are then described. The project study area is introduced, followed by details describing the various aspects of the BikewaySim network development and reconciliation with external data sources: acquiring and processing OpenStreetMap data into graph format, reconciling road attributes from other data sources (traffic signals, speed limits, number of lanes, traffic volumes, and bicycle infrastructure), and, lastly, adding elevation data from DEM and LiDAR sources. Parts of this chapter have been derived from Passmore et al. (2021) and further expanded.

ROUTABLE ALL-PATHS NETWORKS AND OPENSTREETMAP

Critical to performing any of the technical analyses in this project was the development of a routable all-paths network specific to cycling. Routable all-paths networks are spatial databases containing the attributes and centerline geometry of every legally traversable transportation asset. A routable all-paths network can be represented as a graph, which is necessary for performing network analyses and shortest-path calculations.

All-paths networks may include public roads, Interstates, sidewalks, bicycle paths, parking lot aisles, service roads, and more. Traditional travel demand models use simplified road networks containing only Interstates, arterials, and major collectors while

excluding local residential roads, service roads, and non-motorized paths. According to NCHRP 08-36 Task 141, less than 30% of the MPOs surveyed used an all-paths network for their travel demand model (RSG, 2019). Without local roads and non-motorized paths, these simplified networks cannot properly model bicycle routing. Bicyclists can make use of almost every public road (Interstates excluded), but unlike motor vehicles, they can also travel on paved trails through parks, wide sidewalks designated for cycling, cycletracks, and links where they may briefly have to walk their bicycle (staircases, sidewalks, etc.). As the literature review on cycling route choice modeling also suggests, cyclists may prefer taking local streets and non-motorized paths to avoid exposure to high-speed motor vehicle traffic.

Simplified regional transportation in travel demand modeling networks are geared towards modeling congested travel times using an iterative traffic assignment step. Increasing the network size to incorporate all local roads would exponentially increase computational time. However, bicycle traffic is assumed to be less impacted by congestion, so shortest path calculations for bicycle travel do not need to be run iteratively. An all-paths network for cycling, used in conjunction with the simplified travel demand model motor vehicle network can facilitate both types of routing as needed. For example, the San Diego Association of Governments (SANDAG) activity-based model switches network resolutions as needed, using a more refined network for the first and last quarter mile and a simplified network for the intermediate sections (RSG, 2019). The routing tiles utilized by the Valhalla routing engine also facilitate transitions between network resolutions to cut down on computational time (Valhalla, n.d.).

In previous work, the research team identified three street network datasets that could be used for cycling routing: the Atlanta Regional Commission's (ARC) Activity-Based Model (ABM) network, the proprietary HERE Streets network, and the open source and community-built OpenStreetMap (OSM) network (Passmore et al., 2021). The research team identified OSM as the most comprehensive and widely available all-paths network dataset. Municipal network datasets such as the ARC network are not always available in sufficient detail for a city or region, and proprietary network datasets such as the HERE network require the purchase of a license to utilize and cannot be shared outside of an organization, limiting the repeatability of research findings. In contrast, OSM is open source and retrievable for any region desired, though it may be less complete for rural areas where there are fewer contributors.

While containing a similar total centerline length of streets to the HERE and ARC dataset, OSM carried substantially more non-motorized paths and service roads which may be needed for cycling routing. Additionally, key bicycling trails and bicycle infrastructure were simply not present in ABM or HERE, as both networks had been primarily intended for automobile routing.

However, the research team also found that OSM incorporates an excessive number of attributes and the data completion rate for these attributes is very low. Figure 1 shows attribute completion for the 293 attributes attached to OSM road features in the network used in Passmore et al. (2021), weighted by the length of the feature. These attributes include speed limit (13% complete), number of lanes (16% complete), and street parking (0.1% complete). This means that while OSM typically contained more centerline miles of features than the comparison networks, the OSM road features didn't have many

attached attributes about speed limit, the number of lanes, or traffic volume attributes; all of which have been shown to influence cycling routing as mentioned in Chapter 2.

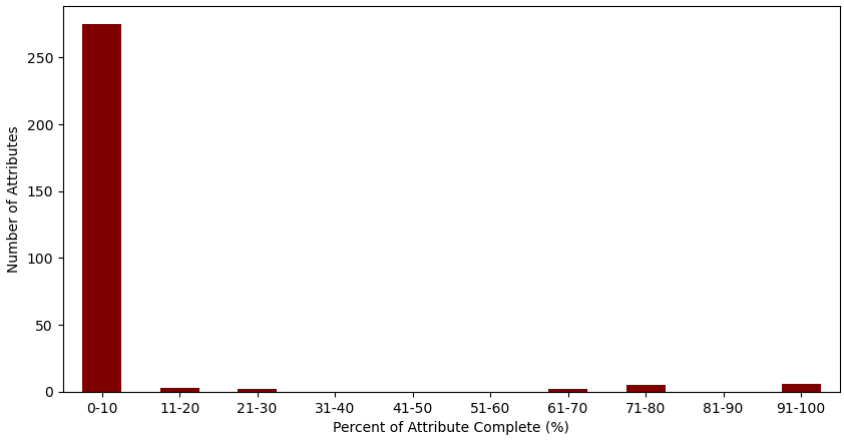

Figure 1: OpenStreetMap Attribute completion weighted by road length and rounded to the nearest percent (N = 293)

It is important to remember that OSM is a map of the world contributed primarily by volunteers. OSM can be edited by anyone, and while there are some conventions for "tagging" features in OSM, there may be multiple ways to tag or draw a feature that would all be technically correct (*Tags*, n.d.). Additionally, users can create new tagging and drawing conventions as desired.

For instance, sidewalks can be drawn separately or represented as an attribute of the road adjacent to the sidewalk. A similar inconsistency occurs with cycletracks. For example, this two-way cycletrack separated by flex posts in Figure 2 is represented as a separate feature (white line) in the 2023 OSM data, but as recently as 2022 it was represented as an attribute on the adjacent road (red line).

OSM data consistency and quality will likely always be a challenge, but one advantage of the OSM data structure is its traceability. OSM edits are public, and the history of features can be inspected to know how the features have changed over time. This offers an advantage over municipal and proprietary data that is often shared in ESRI shapefile or geodatabase format, where features are not typically tracked in a similar manner.

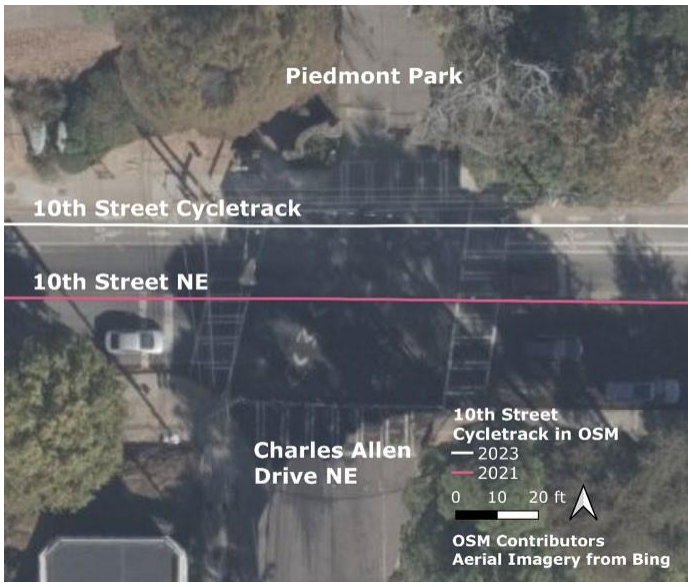

Figure 2: 10th Street cycletrack representation in 2023 and 2021

Because OSM data do not generally contain complete road attributes, supplementary data from other network datasets and asset inventories must be reconciled with OSM data. A scripted reconciliation process to add supplementary data from the HERE roads and ARC networks was created in Passmore et al. (2021). These reconciliation scripts have been expanded upon in this project. However, before discussing the reconciliation process, the BikewaySim graph structure and attribute schema are described.

BIKEWAYSIM GRAPH STRUCTURE AND ATTRIBUTE SCHEMA

To allow for shortest path calculations, BikewaySim uses a directed multi-graph structure to represent all of the transportation network that is legally traversable by bicycle with relevant attribute data, subject to the availability of that data in the specified study area. Graphs can be used to represent the presence and connectivity of highways, local roads, bike paths, sidewalks, and other transportation infrastructure. Graphs are routable, meaning that they can be used to find the optimal path between any two points using a shortest path algorithm, such as Dijkstra's algorithm (Dijkstra, 1959). This property gives graphs a wide range of applications. They are used in trip planning apps such as Google Maps and Waze for wayfinding, and they are also used in traffic assignment in transportation demand modeling.

Graphs are formed as a network of links and nodes (outside of the transportation community, links are also referred to as edges or arcs, and nodes are also referred to as vertices). As shown in Figure 3, links (the white lines) represent streets, sidewalks, etc., while nodes (the white circles) represent the intersections of these features. Transportation links are directional because the presence of one-way streets and because some attributes, such as elevation, depend on the direction of travel. This generally means that two-way streets are represented with two links, one for each direction of travel, while one-way streets are represented with only one link (unless contra-flow travel is permitted). Lastly, the BikewaySim graph is a multi-graph because there are instances where two links share the same start and ending node and thus need to be represented with a unique link ID. This may occur when a residential street leaves a main street briefly only to return to the main street at the next intersection.

Loops, though present in the transportation network (e.g., cul-de-sacs), were not considered in this project because they are not needed in shortest path processes. When these were encountered, they were removed.

Figure 3: Streets represented as links and intersections represented as nodes

The BikewaySim network was stored in the form of node and edge lists that can be saved in ".gpkg" format so that they can be visualized in GIS software. A node list is a table that identifies the intersection or terminus nodes of the network, their N attributes, and their geometry. Table 2 shows the generalized schema of the BikewaySim node list.

Table 2: Basic Schema of Node List

Node ID	Node Attribute 1	…	Node Attribute N	Node Impedance	Geometry
…	…	…	…	…	…

Commented [BD4]: Is the "..." field in this schema supposed to represent fields not shown? (Typ other schema tables)

An edge list identifies the different links of a network, their unique IDs, their associated start and end nodes, whether they represent a reversed link, their N attributes, and their geometry attributes. Table 3 shows the generalized schema of the BikewaySim edge list.

Table 3: Basic Schema of Edge List

Link ID	Node A	Node B	Reverse Link	Link Attribute 1	…	Link Attribute N	Link Impedance	Geometry
…	…	…	…	…	…	…		…

Generally, graphs as-is don't allow for the consideration of turns or turn costs in shortest path routing. To enable this, a pseudo dual graph is constructed (Winter, 2002). Figure 4, reproduced from Winter (2002), visualizes this process. The black lines with arrows and the white circles represent the pseudo dual graph while the gray lines and circles represent the original graph. Essentially, the new graph is created from the original graph by using the original graph's nodes as links and using the original graph's links as nodes. This new graph, the complete pseudo dual graph (Figure 4a), is then pruned to remove U-turns and restricted turns (Figure 4b). This does increase the size of the graph used for shortest path routing which increases computational time.

Figure 4: a) the complete pseudo-dual graph
b) the restricted pseudo-dual graph (reproduced from Winter 2002)

Turns movements are classified as left, straight, right, or U-turns/sharp left/rights based on the change in azimuth between the two links involved. Azimuth measures the change in angle between two points relative to North. A link starting with a node at (0,0) and (0,1) would have an azimuth of zero, and a link starting with a node at (0,0) and (1,1) would have an azimuth of 45 degrees. Each link's azimuth was normalized to 180 degrees to account for instances where the geometric directions of the two links don't align. So, a link with a starting node at (1,0) and an ending node at (0,0) that would normally have an azimuth of 270 degrees would have a new azimuth of 90 degrees.

The difference in Azimuths between the two links gives an amount from 0 to 360. The turns were classified using the angles in Table 4. While this approach has certainly been used before, the literature did not reveal any recommendations for alternative or preferred degree ranges.

Table 4: Turn Classification by Azimuth Difference

Turn Type	Azimuth Difference (Degrees)
Left	>= 210 and <= 270
Straight	< 30 and > 150
Right	>= 30 <= 150
U-Turn or Sharp Left/Right	> 150 and < 210

When performing shortest path routing in BikewaySim, Dijkstra's algorithm is run using the pseudo dual graph, rather than the normal graph. To facilitate the transfer from the starting node to the pseudo dual graph to the ending node, directed connector links are added. The connector links going from a node to the pseudo dual graph are given a cost equal to the connecting link. The connector links going from the pseudo dual graph to a node on the graph are given an initial weight of zero because the impedance of this link has already been incorporated in the pseudo dual graph. Table 5 shows the BikewaySim pseudo dual graph edge list schema.

Table 5: Example schema for pseudo dual graph edge list

Source Link ID	Source Reverse Direction	Target Link ID	Target Link Direction	Target Link Cost	Turn Attribute 1	…	Turn Attribute 2	Turn Cost	Total Cost
…	…	…	…	…	…	…		…	…

Graphically, the turns can be visualized using the midpoints of the links. Figure **5** shows two routes. The one on the left did not have turn costs while the one on the right did. The black lines on each show the path taken from the starting node (lower left corner of each) to the ending node (upper right corner). The turns are represented as the colored line segments connecting the link midpoints (orange circles). When turn costs are applied that

discourage left turns, the route on the right takes fewer turns. These turn costs are demonstrative; the actual turn costs are calibrated in Chapter 4.

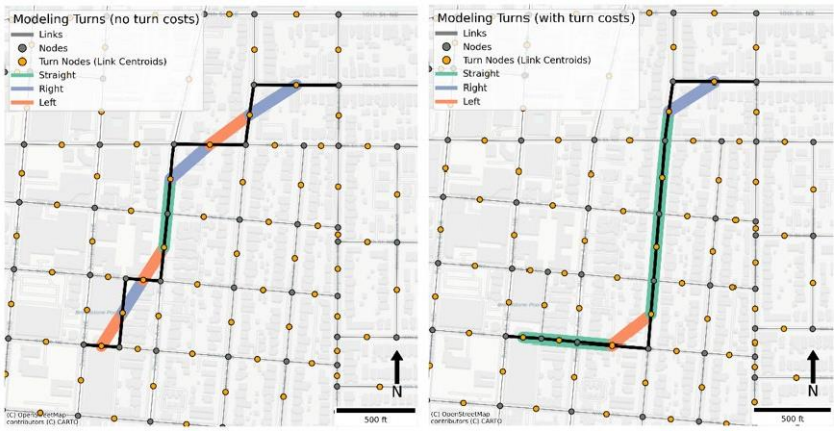

Figure 5: Example of routing that incorporates turn costs

BIKEWAYSIM STUDY AREA

This study employed a 250 square-mile area within the Atlanta Metro Region. As shown in Figure 6, the study area is cross-hatched and is delimited by Interstate 285, which is colored in red. The Interstates and various city boundaries are shown for reference. This study area contained a large portion or all of the cities of Atlanta, Decatur, East Point, College Park, Hapeville, Brookhaven, Avondale Estates, and Chamblee. The white space on the eastern side of the study area is unincorporated Dekalb County.

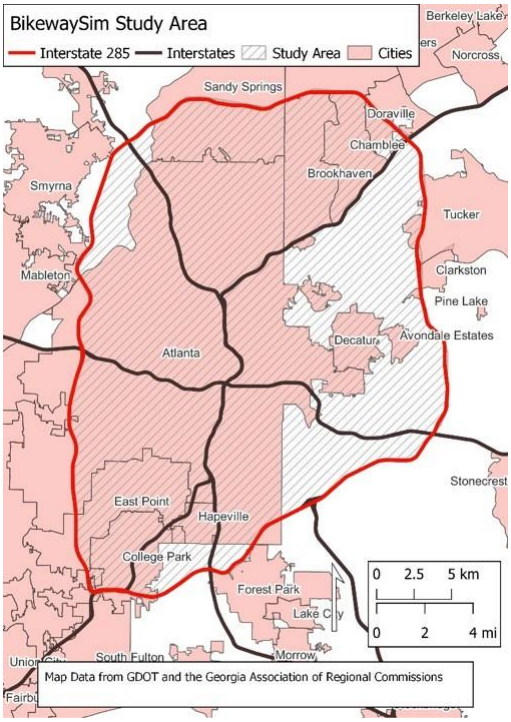

Figure 6: Map of the Project Study Area

This study area was delineated to encompass the majority of the CycleAtlanta GPS trace data (Chapter 4), which was needed for impedance calibration (Chapter 5), and because I-285 functions as an informal boundary between urban and suburban metro Atlanta. Using the city limits of Atlanta alone is insufficient because many bike trips cross into the City of Decatur or unincorporated DeKalb County.

There are a variety of land uses across the study area, ranging from the dense high-rise-filled Midtown and Downtown Atlanta to the small downtown subarea of Decatur. However, most of the study area comprises single-family homes with non-gridded street networks. The university campuses of Georgia Tech and Emory, each with an

interconnected network of wide sidewalks that are used as multi-use paths, are also within the study area. This area also contains a variety of urban on-street bicycle infrastructure as well as major regional bicycle trails such as the Atlanta BeltLine and Stone Mountain Trail. The development of the study area network using OSM and supplementary data is described in the section that follows.

ACQUIRING AND PROCESSING OPENSTREETMAP DATA

OSM data were acquired from the website, Geofabrik, which prepares state, regional, and national extracts of OSM data (Geofabrik, n.d.). Geofabrik archives these extracts regularly (daily for the current month, monthly for the current year, and yearly for at every January 1st since 2014). The January 1st, 2023 extract for the state of Georgia was chosen for this study. The 2023 version was chosen over older versions because it was assumed that the available OSM tags and data were more complete, consistent, and of higher quality than earlier versions. Using an archived version of OSM, rather than the current version, ensures the repeatably of the analyses conducted in this study. However, there is a tradeoff in that the network of 2023 will not always match up exactly with older networks due to new links being added or modified because of changes in the built environment.

The archived OSM data is downloaded in protocol-binary format (PBF) and was masked by the study area boundaries, filtered to only include transportation assets (OSM contains building footprints, points of interest, and more), and converted to XML format using the command line tool, Osmium Tool (*Osmium Tool*, n.d.). Both shell (.sh) and batch (.bat) scripts, for Linux/MacOS and Windows machines respectively, were prepared to script

these steps. After this initial processing step, there were approximately 101,911 unique OSM features totaling 7,016 miles in length.

Raw OSM features are not routable because features often extend through intersections, but features in OSM that intersect with other features at the same elevation will have a shared node that allows the features to be broken down into routable links. The OSM data was converted to a routable network graph using OSMnx (Boeing, 2017). As an example, Figure 7 shows Myrtle Street from North Avenue to 8th Street. In the OpenStreetMap extract, this street is stored as it is shown in the figure, meaning that it spans across multiple intersections, crosswalks, and driveways. The OSMnx Python module breaks this feature down so that each line segment that makes up Myrtle Street is represented as a link. Then, it removes interstitial nodes (i.e., nodes with degree two that only connect to two links of the same OSM ID) and returns a set of links that can be used for routing. The processed links retain the OSM ID that identifies what attributes to assign to the new links, and new reference nodes are assigned based on the relevant OSM nodes. For BikewaySim's routing and map matching processes, a temporary ID is assigned to each of the new links in sequential order from zero. The length of the new links are re-calculated based on the new geometry so that link travel times can be updated accordingly.

Manual edits to the routable network graph version of OSM should be limited as these edits will not be easily re-mergeable into the current or the archived extract of OpenStreetMap. If a manual edit is needed, that edit should be made on the archived extract through JOSM, a Java based OpenStreetMap editor. JOSM can also help facilitate adding any manual edits to the most current version of OSM, if those edits are in

accordance with OSM's open data license, so as to improve the network for everyone that uses OSM.

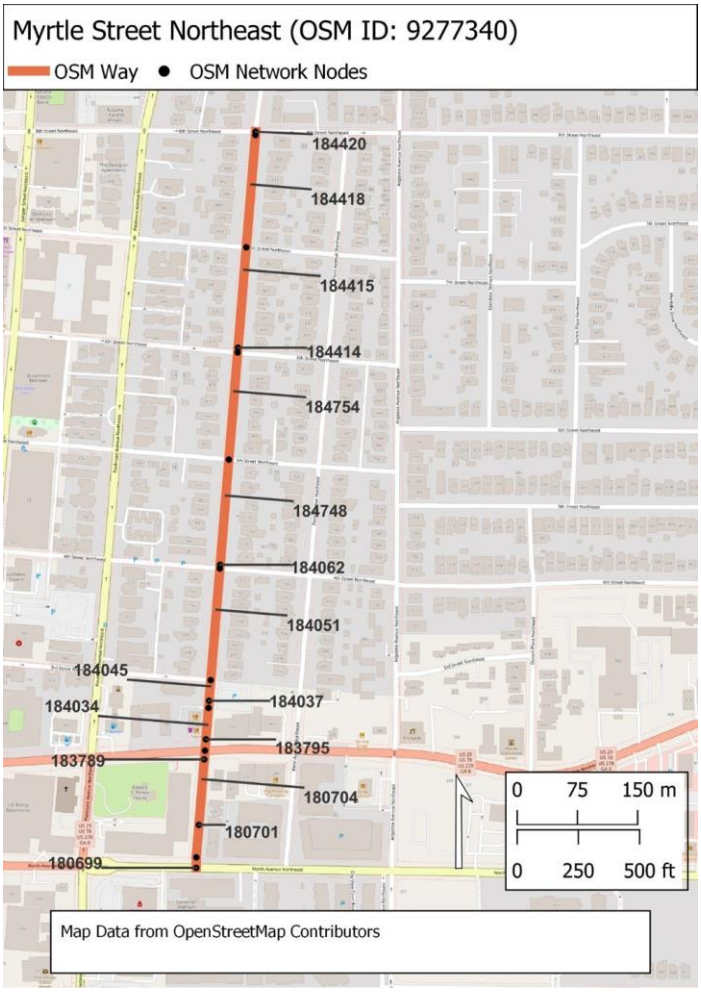

Figure 7: Network breakdown example for Myrtle Street

Once converted to network graph format, there were a total of 197,082 links (totaling 6,887 miles in length) and 150,882 nodes. Around 1,805 links were dropped because they

looped into themselves. The raw OSM links were still retained in a separate database as they are used in subsequent processes to streamline network reconciliation and joining of data to link features.

These links were then classified using the available OSM tags into nine different classifications: restricted access road, no access or private, sidewalks and crosswalks (unmarked and marked), pedestrian paths, bicycle paths, parking aisles and driveways, service roads, and roads. These categories were collectively exhaustive. A description of each category is given in Table 6.

Table 6: Link Categories

Bicycle:	Off-street paved paths where cycling is designated and cycletracks that are drawn separate from the road network
Construction or Proposed:	Features that are under construction or being planned
No Access or Private:	Private and/or gated roads that cannot be accessed by the general public
Parking or Driveway:	Streets used by vehicles primarily for parking and access that typically dead end or provide redundant connectivity
Pedestrian:	Off-street paved paths where cycling is allowed but the path is not explicitly designated for cyclists
Restricted Access:	Interstates and Interstate Access Ramps where cycling is not permitted
Road:	All public roads where cycling is permitted
Service:	Alleyways and entrances that may offer some enhanced connectivity for cyclists
Sidewalk or Crosswalk:	Sidewalks and crosswalks where cycling is not permitted

This classification step was necessary to identify links that bicyclists could not legally or practically traverse such as Interstates and inaccessible private roads but also to remove excess links that don't contribute meaningful connectivity to bicycling like parking aisles in park lots and dead-end driveways. It also aids in assisting the reconciliation steps as

traffic signals don't need to be added to parking aisles or bike paths unless they are directly connected to a road. The number and total length of features for each category are listed in Table 7.

Table 7: Number and Total Length of Links by Link Category

Link Category	Number of Links	Total Miles of Links		Used in Bicycle Routing
Road	65,292	3078	(45%)	Yes
Service	45,583	955	(14%)	No*
Parking or Driveway	31,673	852	(12%)	No*
No Access or Private	15,088	606	(9%)	No
Sidewalk or Crossing	22,950	546	(8%)	No*
Restricted Access Road	2,864	459	(7%)	No
Pedestrian	11,019	243	(3%)	Yes
Bike Facility	2,552	125	(2%)	Yes
Construction or Proposed	61	9	(0%)	No
Total	197,082	6,877	(100%)	

*Unless a link has been exempted due to evidence of bicycle travel from CycleAtlanta data or through the research team's local knowledge

"Road" links make up the largest portion of the network, both in link count and in mileage. "Service", "Parking or Driveway", and "No Access or Private" links make up the next largest percent (35%). "Sidewalk or Crossing" links make up about 8%. "Pedestrian" and "Bike" links, which should represent the non-motorized, off-street, transportation network. From these numbers alone, it's clear that cycling in Atlanta requires using roads with motor vehicles to get to most places.

The last column of Table 7 indicates whether that link was used for bicycle routing. "No" with an asterisk indicates that exceptions were made for certain links that were used in the CycleAtlanta data. Sidewalks and crossings, unless providing connectivity between an off-street cycletrack or multi-use path and a road, were not considered for routing. Sidewalk riding, while illegal for adults in many jurisdictions of the United States, does

occur in cases where a cyclist feels that it is safer than cycling on the road (Barajas, 2021; Chaloux & El-Geneidy, 2019; Marshall et al., 2017). However, GPS trace data alone are not accurate enough to indicate whether a cyclist was on the road or an adjacent sidewalk. Including sidewalks can lead to ambiguous map matching and routing results where a cyclist is seen jumping between the sidewalk and the road due to their parallel connectivity. While sidewalks can be included as an attribute on roads, OpenStreetMap data on sidewalks in many areas are incomplete (in the ITS4US project, the GDOT team found that 70% of sidewalks were missing from OSM).

The sidewalk network for the study area is shown in Figure 8. Most of the mapped sidewalks in OSM are in the center of Atlanta. However, as the ITS4US team found in Gwinnet County, there are lots of OSM sidewalks in central areas (such as Lawrenceville) but few sidewalks as you move into the suburban areas. A quick inspection of the BikewaySim study area indicates that there are lots of missing sidewalks in OSM that show up in satellite imagery. Because the network is incomplete, it is not feasible to include sidewalks as a variable at this time. But as the network becomes more complete, future research should investigate whether sidewalks (or even informal desire lines) play a role in routing decisions for cyclists, particularly on large arterials without cycling infrastructure.

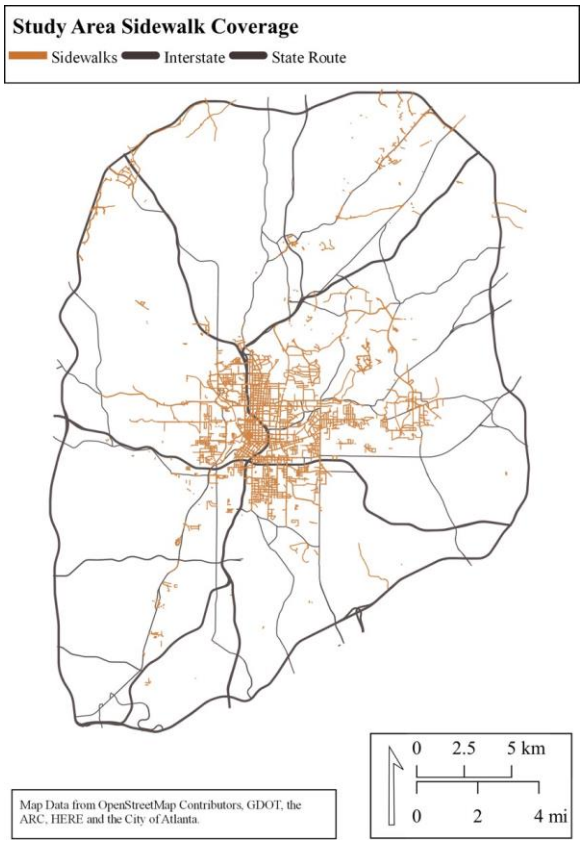

Figure 8: Sidewalk coverage in the study area

Figure 9 shows the study area network where the links are symbolized by link type. While there are some continuous sections of off-street bicycle/pedestrian paths, particularly in the center of the study area, it confirms that most travel requires cyclists to share the road with motor vehicles.

After identifying the link types, the next step was to classify the bicycle infrastructure present in OSM. Research shows that for metropolitan areas, OSM data on bicycle

facilities are comparable with municipal data on bicycling facilities (Ferster et al., 2020). After some initial inspections, the research team determined that the study area's OSM data on bicycle facilities matched well with municipal data from the City of Atlanta and the Atlanta Regional Commission on bicycle facilities (Atlanta Regional Commission, 2022; City of Atlanta Department of Public Works, 2024).

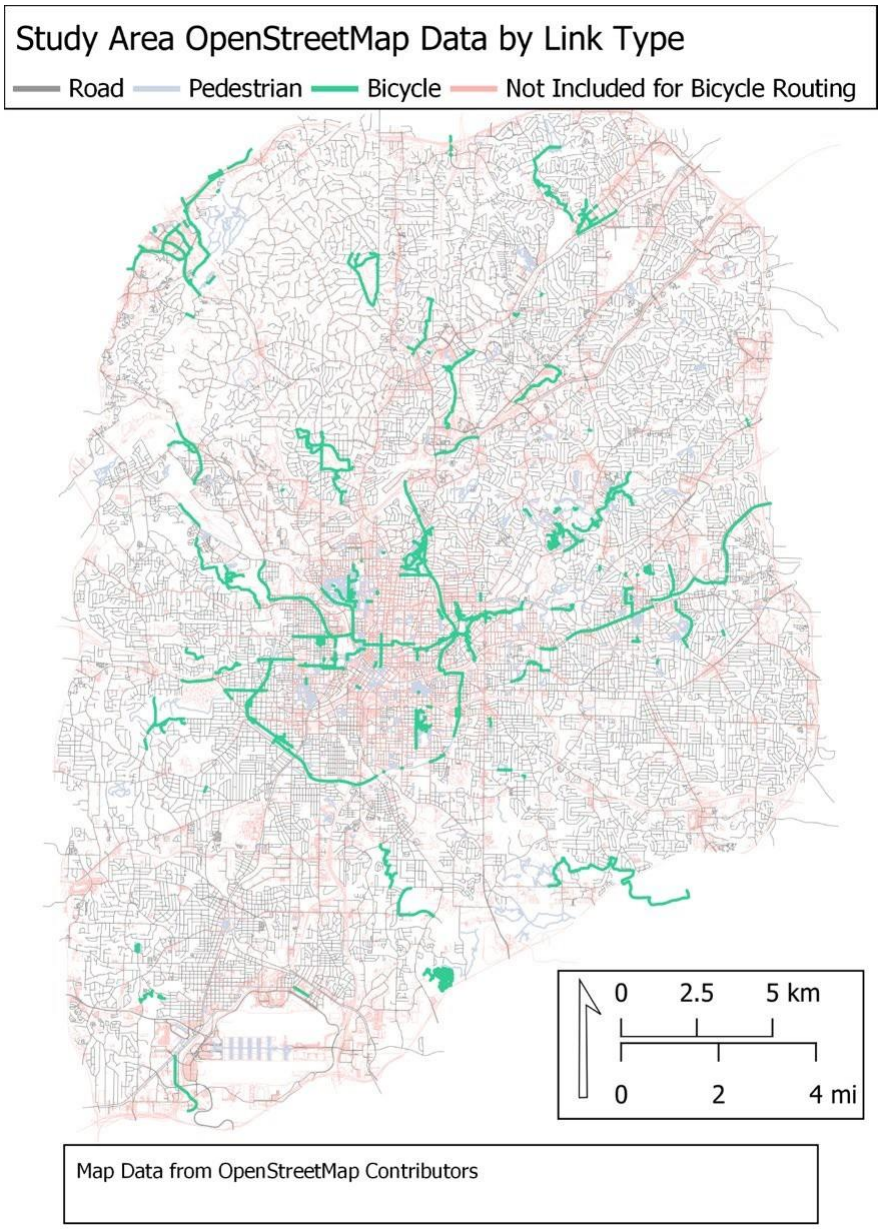

Figure 9: Link types in the study area

OSM has various tags for representing bicycle infrastructure. In general, cycling infrastructure can be found using the "cycleway" key or if it's a dedicated cycling path "highway=cycleway." The "cycleway" tag can be used in conjunction with more detailed tags to specify the type of bicycle infrastructure by direction. Each of the following cycling facility types was assigned to each link in a facility forward and facility reverse columns that represented the type of facility present based on the direction a cyclist was traveling. Facilities were assigned by direction because there are several roads with "climbing lanes," where a bike lane will be present in the uphill direction of a road but not the downhill direction. Six types of infrastructure were classified in the OSM data (see Table 8).

Table 8: Bicycle facility types

Facility Type	Example
No Facility	
Shared Use Marking (Sharrow) or Marked Route	
Conventional Bicycle Lane	
Buffered Bicycle Lane	
Cycletracks or Protected Bicycle Lanes	
Multi-Use Path	

One challenging aspect of this process was differentiating between cycletracks and multi-use paths. Cycletracks in OSM are often represented as separate links (Figure 2) which is similar to how multi-use paths are represented. As such, both the 10th Street cycletrack and the eastside section of the Atlanta BeltLine trail are marked as "cycleways" in OSM. In a few instances, cycletracks were distinguished from multi-use paths based on whether pedestrians were allowed (e.g., foot travel was not allowed on the 10th Street cycletrack but was allowed on the Atlanta BeltLine).

The OSM data also already contained node attributes for traffic signals. These node attributes were joined to the pseudo dual graph to indicate whether a turn movement passed through a signalized intersection. However, the current set of signals in OSM needs to be reconciled with, and expanded to match, more comprehensive GDOT sources in the future. Upon a brief inspection, it was clear that there were traffic signals in the OSM data that were not in the GDOT data, and vice versa.

Once the network was retrieved and processed, the next step was to add road attribute data from other sources.

NETWORK RECONCILIATION

OSM data for this study area did not contain the relevant road attribute data for explaining cycling impedance factors such as speed limits, the number of vehicle lanes, or motor vehicle traffic volumes. Instead, several other data sources were imported to fill these road attributes. These data sources are shown in Table 9.

Table 9: Attribute Data Sources

Data	Source	Data Type
Bicycle Facilities	OSM, ARC, CoA	Links
Bridges/Tunnels	OSM	Links
Elevation	USGS	Raster, Point Cloud
Link and Node Geometry	OSM	Link, Node
Link Directionality	OSM	Link
Number of Vehicle Lanes	HERE, GDOT	Links
Speed Limit	HERE	Links
Traffic Signals	OSM, GDOT	Nodes
Turn Restrictions	OSM	Link-to-link

HERE data were licensed for use in this project and cannot be shared, but data from USGS and GDOT were publicly available. The BikewaySim repository contains scripts and instructions for downloading both sources. USGS DEM and LiDAR data are available across the United States, but GDOT data are only available within the state of Georgia. An equivalent data source would need to be identified when using BikewaySim outside of Georgia.

The BikewaySim code contains a set of scripts that assist in reconciling these various datasets with OSM (Passmore, 2024). There isn't one particular series of steps that works for all the data sources, but generally, the reconciliation process follows the decision tree shown in Figure 10 for link features. Node, elevation raster, and elevation point cloud features were reconciled with a different process.

First, OSM was filtered based on the type of added data. For instance, speed data were only needed for OSM roads, so the OSM links were filtered to only include roads. Similarly, for bicycle facilities data from other sources, the OSM links were filtered to include roads with bicycle facilities and off-road bicycle facilities.

Second, the resolution of the added data is assessed to see if the data should be added to the full OSM link or to the truncated network OSM links. Using the full OSM link reduces the number of possible matches to consider later in the reconciliation process, so when working with data that are not at network resolution (i.e., links span across intersections), it's better to use the full OSM features instead. Once this has been decided, the OSM links are then buffered and intersected with the data to be added.

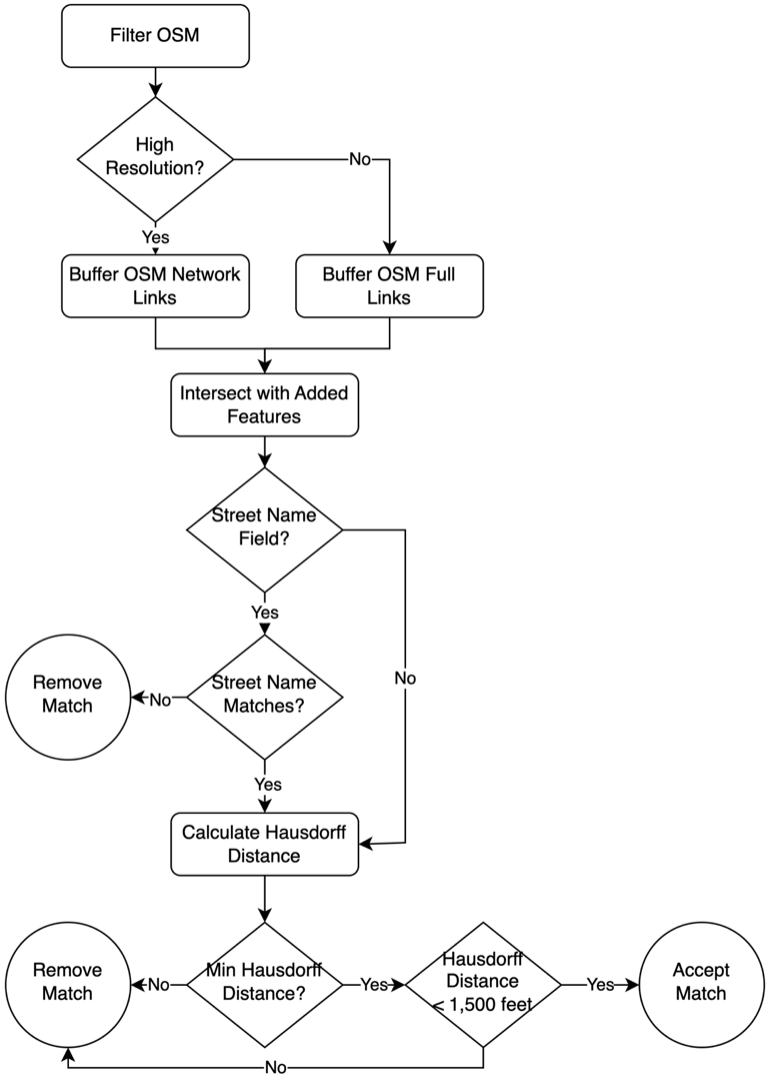

Figure 10: Generalized reconciliation process for link features

If the added features contain a street name field, then the street names are compared between the OSM and the added data. The street names for both OSM and the added data had all capital letters lowercased, punctuations removed, street suffixes (e.g., Drive, Avenue, Northeast, etc.) removed, and street suffix abbreviations (e.g., Dr., Ave., N.E., etc.) removed. The street names were then compared to see if at least one element of the street name was present in both the OSM data and the added data. For example, consider a candidate match where the OSM name was 10^{th} Street Northeast, and the added data name was 10^{th} St. N.E. The OSM name would be converted to "10^{th}" and the added data name would be converted to "10^{th}". Because at least one word of the street name is present in both, the name check would return "True". Matches where the name check returned "False" could be removed. In cases where either feature didn't have a street name, these matches were skipped.

Next, the Hausdorff distance between the intersected feature and the original OSM feature was calculated for each match (Min et al., 2007). The Hausdorff distance measures the similarity between two-line features. The higher the Hausdorff distance, the less similar the line features. This metric was used because it accounted for both the length and orientation of the OSM feature and the intersect match feature. This was especially helpful when a street in the other dataset matched to an intersecting street in the OSM dataset. Once the Hausdorff distance was calculated for all matches, only the candidate matches with the minimum Hausdorff distance per OSM feature and a Hausdorff distance of less than 1,500 feet were accepted as matches. The reconciled data was stored independently of the network data during the reconciliation process so that all the reconciled data could be added at the end. Depending on the data added, some

variations in the reconciliation process were made, so the next sections will describe specific steps taken to reconcile each added data source.

GDOT Link Data

Attributes on traffic volume and the number of lanes were available through GDOT (Georgia Department of Transportation, n.d.). Both datasets had link geometry derived from GDOT's 2021 road inventory. However, the links in the datasets were segmented differently because GDOT's road inventory uses a linear referencing system to split the roads into segments.

The OSM data were filtered to only include roads since traffic volume and the number of lanes are both road features. Because links from both data sources were not in network graph format, the OSM full links were used over the OSM network links. The GDOT data did not contain street names, so only the Hausdorff distance check was used to match features.

The GDOT traffic data contain traffic volumes on all public roads in Georgia, where some link traffic volumes were measured from permanent counters or traffic studies and other link traffic volumes had been calculated from nearby counts or given a default value. For this paper, the field for annual average daily traffic (AADT) was used to calculate average daily truck traffic %, by dividing the AADT of combined-unit trucks field by the AADT field.

After the Hausdorff distance check, 18,578 of the 20,176 features of the GDOT data were matched to 8,935 of the 12,017 features of the OSM data.

The number of lanes data contains the total number of vehicle lanes, ignoring center turn lanes, left/right turn lanes, and shoulders. The data set did not indicate the number of lanes per direction, so a three-lane one-way road would have the same number of lanes as a two-way road with two lanes in one direction and one lane in the other direction.

After the Hausdorff distance check, 18,551 of the 20,176 features of the GDOT data were matched to 8,935 of the 13,410 features of the road full link OSM data.

HERE Link Data

The HERE link data contained attributes for vehicle speed and the number of lanes. HERE had to first be filtered to remove links that represented restricted access roads or weren't public roads.

The HERE speed attributes were categories describing the "general speed of a road based on the posted or legal speed." In the future, probe speeds, derived from the methods used in Arias et al., (2021) should be used. The HERE speed categories were simplified into four categories, shown in Table 10, as there were few roads above 40 miles per hour that cyclists were allowed on. The speed limit category ordering was also flipped, so that the ascending category number corresponds to the ascending speed.

Table 10: HERE Speed Categories Simplified

Original HERE Speed Category	Description	Simplified Speed Category
1	> 80 MPH	4
2	65-80 MPH	4
3	55-64 MPH	4
4	41-54 MPH	4
5	31-40 MPH	3
6	21-30 MPH	2
7	6-20 MPH	1
8	< 6 MPH	1

HERE has three fields for describing the number of lanes on a link. The first two attributes are used in a few scenarios to specify the number of travel lanes per direction, usually when there is an odd number of lanes. The last attribute used on all of the HERE data is categorical and describes the number of lanes per direction (see dictionary in Table 11). Like the GDOT lanes data, the HERE lanes data ignored center turn lanes, left/right turn lanes, and shoulders.

Table 11: HERE's Lane and Speed Fields

Field	Description	Data Type	Possible Values
To Lanes	The number of lanes in the forward direction	Integer	0 to 8
From Lanes	The number of lanes in the reverse direction	Integer	0 to 8
Lane Category	The number of lanes	Enumerated	1: one lane per direction 2: two or three lanes per direction 3: four or more lanes per direction

For conflation, the OSM data were filtered to only include roads, given that traffic volume and the number of lanes only apply to road features. HERE links are at the network resolution, so the OSM network links were used for reconciliation over the full links. The HERE data did contain street names, so both the street name check and the Hausdorff distance check were used to accept/reject candidate matches.

The reconciliation process with the HERE data yielded 314,862 initial candidate matches. After the street name and Hausdorff distance checks, 58,349 of the 65,292 HERE links were matched to 35,863 of the 73,866 features of the road full link OSM data.

Bicycle Facilities Data from the City of Atlanta and the Atlanta Regional Commission

This section details how bicycle facilities were added to the network with the approximate date of opening for public use, so that bicycle activity traces could be assigned properly to the evolving network. The impedance calibration process (Chapter 5) project uses cycling GPS traces recorded between 2012-2016; hence, the facilities available in the 2023 OSM data had to be coded to show what infrastructure was available at the time each CycleAtlanta trip was made. One way to accomplish this is to utilize an OSM history file. One of the advantages of OSM is that all changes are tracked, meaning that one can retrieve the instance of OSM on any date (starting from the inception of OSM in the early 2000s). However, there is likely a gap between when bicycle facilities are installed and when they get added to OSM by contributors. For this reason, the research team avoided using OSM history files. Future research should investigate this gap to help approximate when cycling infrastructure was installed.

Fortunately, the available municipal data from the Atlanta Regional Commission and the City of Atlanta contained installation year data for each facility. These dates were brought over to OSM through a similar process to the reconciliation done with HERE/GDOT in the previous sections. Further manual efforts may be needed to add more precise installation dates from research and discussions with City of Atlanta staff and local advocacy organizations. These more precise installation dates could potentially help calibrate more accurate impedance factors using existing cycling GPS traces and as cycling GPS trace data collection continues into the future.

Elevation and Grade from USGS DEM and LiDAR Data

Cyclists' avoidance of steep/hilly routes has been repeatedly verified in the cycling route choice literature (Broach et al., 2012; Hood et al., 2011), yet OSM does not carry elevation attributes. This is because elevation attributes would need to be re-calculated whenever there is a modification in the network (i.e., splitting a link up into smaller segments or adding a new one), and elevation attributes can be represented in countless ways (average grade, length of grade between 4% and 6%, total ascent, net gain).

Elevation attributes can be added to OSM by sampling elevations from a digital elevation model (DEM). A DEM is a "representation of the bare ground (bare earth) topographic surface of the Earth excluding trees, buildings, and any other surface objects" (USGS, n.d.-b). Practically speaking, DEM files are geo-referenced rasters where the raster pixel values correspond to coordinates and elevations. Elevation attributes were added to network links by sampling the elevation from the DEM at specified intervals along the network links.

The resulting sampled elevation data are often noisy, as network links may not follow the surface of the earth (e.g., tunnels, bridges, underpasses, overpasses, cut and fill) and link centerlines may not fall within the correct raster cell (Liu et al., 2018); both of these can result in unrealistic spikes or dips in the elevation profile of a road. To remove outliers and smooth the elevation data, this research followed the procedure outlined in Liu et al. (2018) but added a pre-processing component for re-sampling bridge elevations from USGS LiDAR data.

One-by-one meter resolution DEM data were acquired through the USGS's TNM Access API (USGS, n.d.-a). To begin the elevation sampling process, points for sampling elevation data had to be created from the OSM links. These sampling points were created from the raw OSM links instead of the network OSM links because the longer raw links would carry more sampled elevation data to use for outlier detection and smoothing. The network used in Liu et al. (2018) had much longer links and didn't contain local roads and non-motorized paths like in the OSM network.

For each raw OSM link, a sampling point was generated every 10 meters (the same distance used in Liu et al. (2018)) along the link's geometry starting from the first point. When a link was less than 10 meters, the first and last points were used as the sampling points. This applied to 4,881 out of the 67,594 raw OSM links or 29,600 km of the total 7,803,312 km. These short roads were mostly made up of OSM footway features or residential roads. In cases where the last point on a link was after the last 10-meter segment (e.g. last point at 25m), the last point was not included as a sampling point.

The sampling points were then overlaid with the DEM and assigned elevations using the data from the DEM raster cell that they occupied. After the DEM sampling, links that were tagged as bridges in OSM were re-sampled using LiDAR data on bridge deck elevations. LiDAR data were also acquired through the USGS's TNM Access API. LiDAR data are stored as point clouds where each point has a latitude, longitude, elevation, and classification. For USGS LiDAR data, points representing bridge decks are assigned the value "17" for classification. The LiDAR data for the study area was filtered to only points representing bridge decks to minimize the amount of LiDAR data that needed to be stored. The sampling points of the tagged bridges were buffered by 20

meters and intersected with the bridge deck LiDAR points. If a sampling point did not intersect with any bridge deck LiDAR points, then the existing sampled elevation was kept. If a sampling point did intersect with a bridge deck LiDAR point, then the sampling point was reassigned an elevation equal to the average of all the intersecting bridge deck LiDAR points.

To gain an intuition for the process, consider the profile of the link representing the 5th Street Northwest bridge across Interstate 85 and Interstate 75 shown in Figure 11. The plot on the far left shows the elevation profile of the link, where the sampled DEM data is represented with circles. The red and green points indicate where the road grade on the link exceeded 20% (green) or fell below -20% (red) road grade.

In the elevation profile, the link starts with a steep descent into what appears to be a valley before rising rapidly. The two maps on the right (one showing a satellite view and the other map data) contextualize the link's location. On each map, the directional arrows on the link show the link's direction respective to the elevation profile. In the satellite view, it appears that the steep descent roughly corresponds with where the bridge starts over the Interstate, and the steep ascent corresponds with where the bridge ends. These sampled elevations from the DEM represent the elevation at the Interstate level.

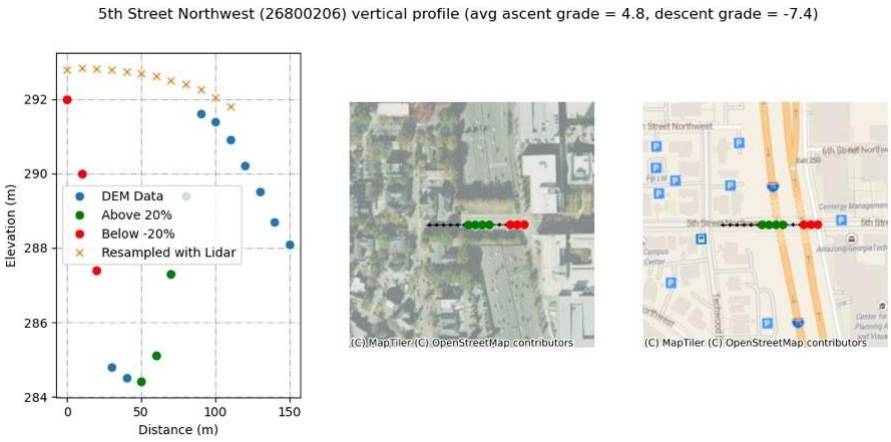

Figure 11: 5th Street Northwest with elevation sampled from LiDAR bridge decks

After re-sampling the elevations with the bridge deck LiDAR points, new LiDAR elevations are added to the elevation profile (orange "X" marks). The new points "fill in" the valley that was present using just the DEM elevations and now the elevation profile of the bridge is more accurate.

After sampling elevation from DEM and LiDAR, the next step was to remove noisy elevation points by identifying unrealistic road grade changes between the sampled points. The segment grade between sampled points was calculated for each segment in a link using Equation 2. For each segment, the difference in elevation between the first point ($ee_{pp+1} - ee_{pp}$) and the last point $ee_{pp+1} - ee_{pp}$ was divided by the length of the segment (ll). Most of the time, the length of the segment was 10 meters unless the link was less than 10 meters.

$$ssee sssseessss\ ssgggggee\ (\%) = \frac{ee_{pp+1} - ee_{pp}}{ll} * 100$$

Equation 2: Segment grade formula

Once the segment grades had been calculated for all of the links, a segment grade threshold corresponding to the links' road type was used to removed "noisy" points. For all segments in a link that exceeded the specified segment grade threshold, the ending point of the segment was removed, and the segment grades were recalculated. This process was repeated until no segments exceeded the specified segment grade threshold. If this process removed too many points, then only the elevation of the first and last points was used, and the link was flagged for further inspection. These links either need new sampled points from the bridge LiDAR data, are the grade specified, or might be too short of a link.

Segment grade thresholds were specified by link type to prevent the removal of steep segment grades that were accurate. Liu et al. (2018) set established segment grade thresholds based on road design standards (8% for arterials and highways and 15% for local roads). These thresholds were used but an additional, higher threshold of 20% was applied to smaller local roads and non-motorized paths. Some non-motorized paths were unpaved hiking trails that more regularly exceeded 20% grade.

Some of the roads in the study area were simply steep, so using a lower threshold could incorrectly remove reasonable elevations. For example, Mary George Avenue Northwest (Figure 12) had many segments that exceeded 15% grade, but those grade segments reflect the actual conditions and are not due to sampling noise.

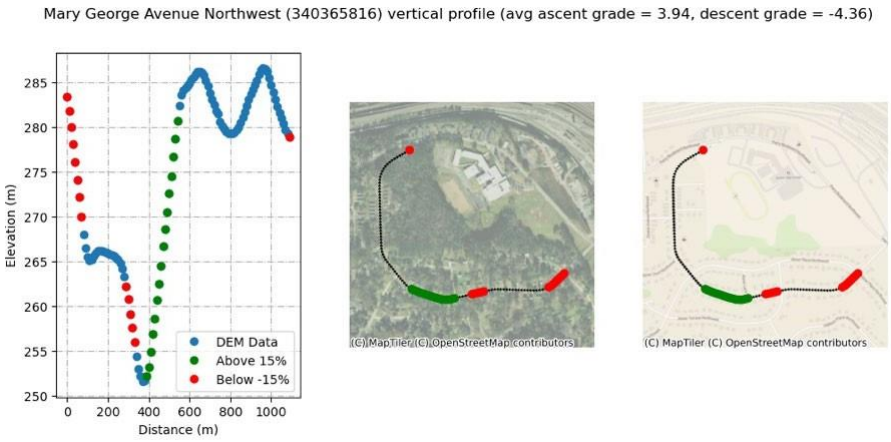

Figure 12: Vertical profile and map views of Mary George Avenue Northwest

The layout of Figure 12 is the same as Figure 11. The elevation profile (the plot on the far left) shows that the link starts with a steep descent, levels off before descending again, and then hits a low point before a rapid ascent back up. The map view (on the far right) shows that there are no rivers or built environment features that would cause these grades, and a field visit by the researchers confirmed that this road was that steep.

The AASHTO Guide states that multi-use paths should not exceed 5% grade (AASHTO Task Force on Geometric Design, 2012), but in reality, there were several multi-use paths, such as the Lionel Hampton PATH Trail (Figure 13), that contained segments exceeding 5% grade. Using a 5% threshold for multi-use trails would have removed valid points.

The layout of Figure 13 is the same as Figure 11 except the point threshold was lowered from 20% to 5%. The elevation profile (the plot on the far left) shows that the link starts

with a descent below -5% grade before flattening out and then continuing with a descent below -5% grade.

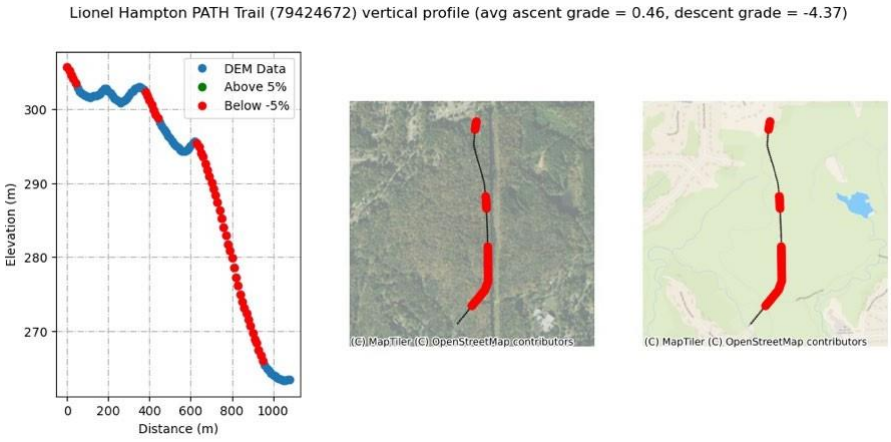

Figure 13: Vertical profile of the Lionel Hampton PATH Trail

After the point removal step, point elevations that were removed are replaced using linear interpolation between the remaining points. Then, a cubic smoothing spline with a smoothing parameter of 0.5 was fit to the links. The fitted splines were then applied to calculate elevations for the OSM network links. For each OSM network link the start and end points were projected onto the corresponding raw OSM link and the distance along the line from the start of the raw OSM link and the projected points was calculated. These distances were then input into the corresponding fitted cubic spline to calculate a network link-specific elevation profile, where elevation values were predicted every 10 meters. The link-specific elevation profiles were then aggregated to calculate the link's total ascent, total descent, average ascent grade, and average descent grade. Average ascent and descent grades were calculated using Equation 3 and Equation 4, respectively.

$$ggaaeeggggssee\ ggssaaeessss\ ssgggggggee\ (\%) = \frac{ssttssggll\ \ ggssaaeessss}{ssttssggll\ \ lleesssssh}$$

Equation 3: Average ascent grade (%)

$$ggaaeeggggssee\ ggeessaaeessss\ ssgggggggee\ (\%) = \frac{ssttssggll\ ggeessaaeessss}{ssttssggll\ \ lleesssssh}$$

Equation 4: Average descent grade (%)

In future work, the researchers will create link elevation attributes designed to better capture the effect of the elevation profile for use in impedance. For example, it may prove worthwhile to calculate the length of each segment across several grade bins (e.g., length of link less than -5% grade, between -5% and +5% grade, and greater than 5% grade).

FINAL NETWORK

The final network size was 153,157 links (306,314 directed links), 123,628 nodes, and 485,706 directed turns. This includes all connected OSM links, even the ones that may not be traversable by bike. The full network was not used for the analyses in the later chapters of this report as it was typically filtered to only include the links with a type of "Road", "Bicycle", or "Pedestrian" that were specified for bicycle routing (Table 6) which would bring the network size to 76,912 links (144,458 directed links), 66,151 nodes, and 222,764 link-to-link turn opportunities. This filtered network will be used to describe the makeup of the link attributes. The filtered network had road attributes on the type of bicycle facility, the AADT, truck traffic percent, the number of lanes per direction, speed, traffic signals, and elevation. The attribute makeup of these final links is described in the next sections.

Lanes

After the GDOT and HERE lane attribute data had been reconciled to the network, the attribute values of both were merged. The GDOT lanes data were converted to a lane category based on the logic in Table 12, and the HERE lanes data were modified slightly to account for the data in the "from lanes" and "to lanes" link attributes as shown in Table 13.

Table 12: Lane Category Assignment for GDOT

GDOT Number of Lanes	One-way	New Lane Category
1	True	1
2	True	2
3+	True	3
1, 2	False	1
3, 4	False	2
5+	False	3

Table 13: Lane Category Assignment for HERE

Lane Category	Maximum of the "To Lanes" and "From Lanes" Columns	New Lane Category
1	0	1
2	0	2
3	0	3
Any value	1	1
Any value	2	2
Any value	3+	3

Once the GDOT and HERE lane attributes were comparable, the attribute with the lowest lane category was retained after inspecting some of the mismatched lane attributes. All non-road links were given a lane category of zero, and roads with a cycletrack or multi-use path were also given a lane category of zero.

Figure 14 shows the network symbolized by the number of lanes (links with a value of zero or one were grouped). Most of the links had one lane per direction, which makes

sense as most of the roads were residential. The roads with two or more lanes per direction are mostly clustered in Midtown and Downtown Atlanta. There are a lot of large one-way streets in this area designed to move vehicle traffic.

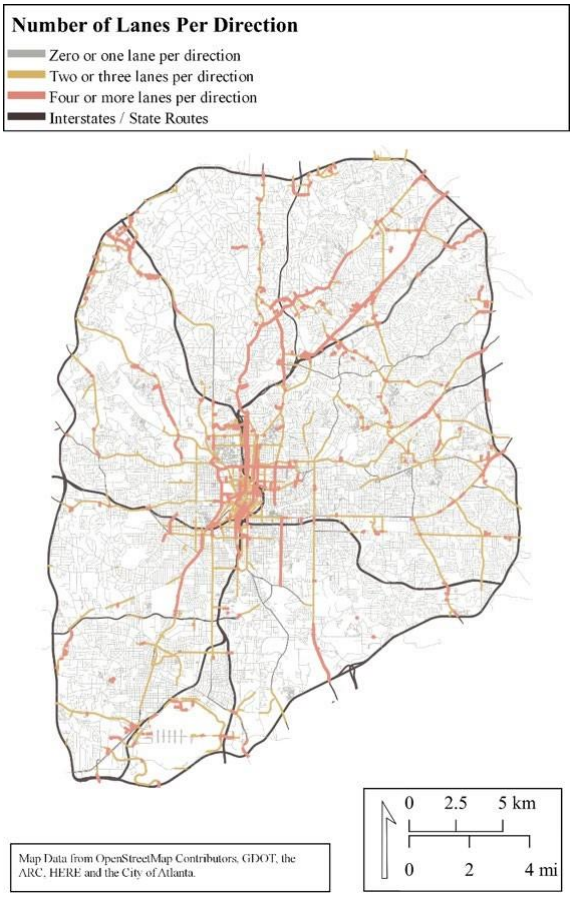

Figure 14: Number of lanes per direction

Average Annual Traffic Volume (AADT)

Figure 15 shows the AADT for the network links. Most links in the study area, being residential, had traffic volumes below 5,000. The roads with higher traffic volumes appear to be located in the Northern half of the study area. Midtown and Downtown Atlanta has a large concentration of roads between 10,000 and 30,000 AADT. Peachtree Road, Northside Drive, and Buford Highway are all above 30,000 AADT.

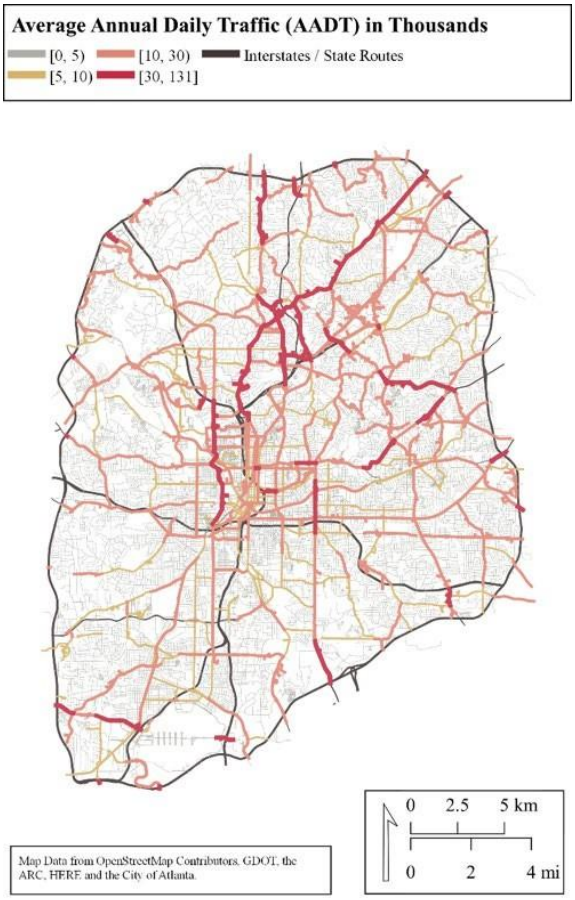

Figure 15: Average annual daily traffic

Percent Truck Traffic

Figure 16 shows the percent of truck traffic for the network links. Most of the links in the study area had below 1% truck traffic volume. Roads with higher truck traffic volume are scattered throughout the study area. Some areas of note include the Norfolk Southern Inman Yard in the area directly west of Midtown, and the access roads along Interstate-85 in the northeast quadrant of the study area. Due to the spatial distribution of percent truck traffic, it may not be practical to calibrate link impedance for truck traffic.

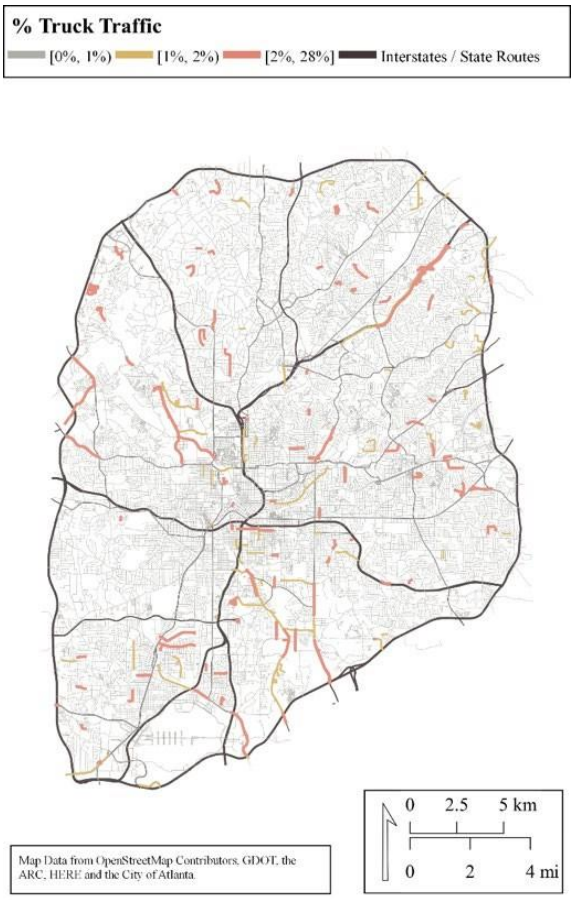

Figure 16: Percent truck traffic

Speed Category

The HERE speed category data are visualized in Figure 17. Most links in the study area had speed limits below 30 miles per hour. Links with speed limits between 31 and 40 miles per hour were sprinkled throughout this study area, but there were only a few links with speeds above 40 miles per hour.

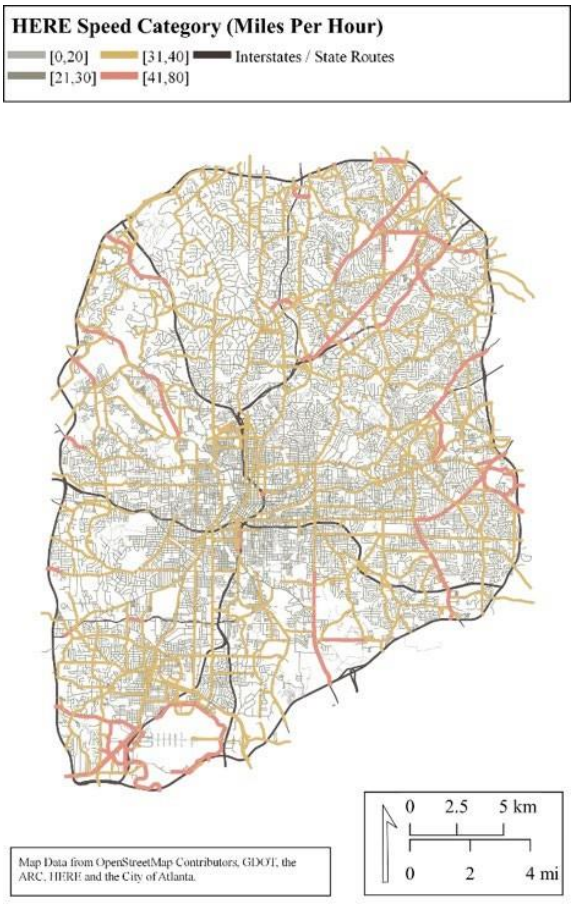

Figure 17: HERE speed categories

Elevation / Percent Grade

Figure 18 shows the spatial distribution of link average percent grade. Many links were above 4% grade, indicating that the study area was hilly. This also likely means that cyclists may not have many options for avoiding hills.

Commented [BCD5]: Why was 4% used as the cut off? There's a big difference between ~4% grade hills and hills in the 6-10% grade range.

In future research, it would also be interesting to see how grade impacts impedance for ebikes/escooters vs regular bikes. Given that ebikes give riders the option to "smooth out" hills and are becoming more and more popular, grade may not be as big of a concern as it once was.

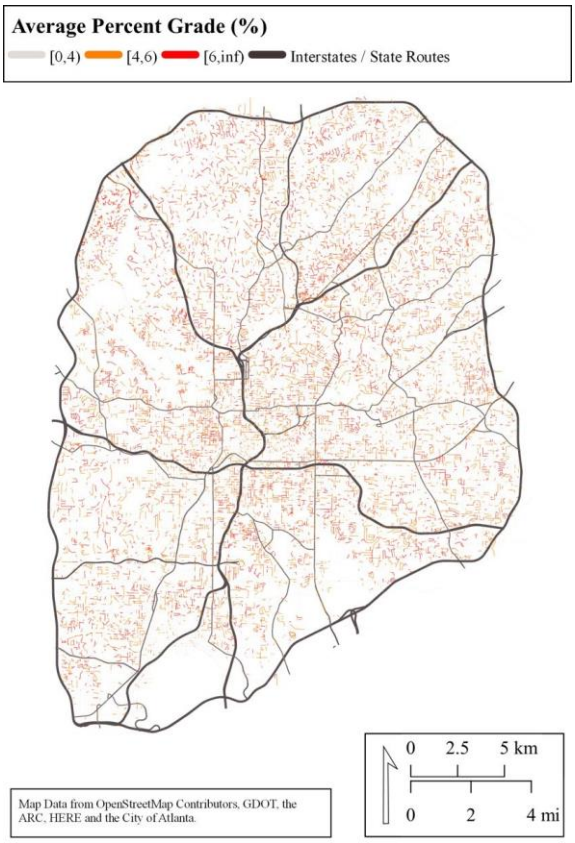

Figure 18: Average road grade

Cycling Facilities

Figure 19 shows the spatial distribution of different types of cycling infrastructure in the study area. Most of the facilities in the study area were either traditional painted bike lanes or multi-use paths.

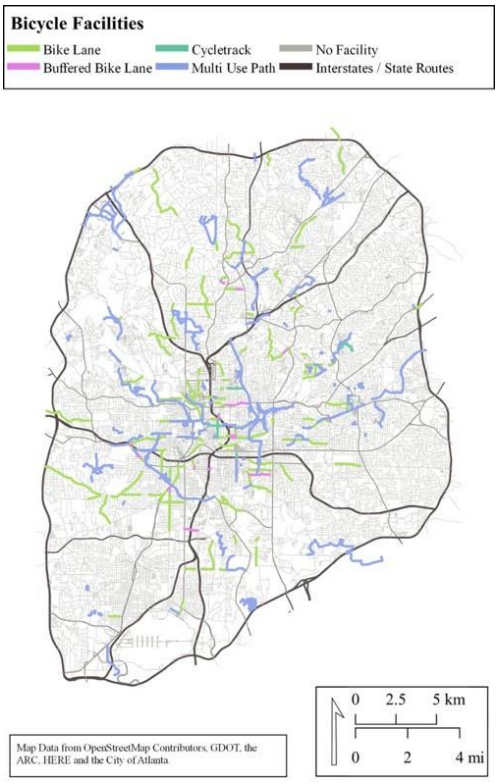

Figure 19: Study area bicycle facilities

FIELD DATA COLLECTION FOR CYCLING INFRASTRUCTURE FEATURES

The research team has created standardized bicycle facility inspection forms that are applicable to most contexts in which bicycle infrastructure is expected to be present in the United States. These inspection forms are designed to identify the presence of a bicycle infrastructure feature and capture pertinent design and condition elements that are required to assess a bicycle facility's conformance with federal policy and design guidelines, such as the bicycle facility's width. In addition, these forms can be used to record information useful for impedance value assignment to route links, such as the

volume and speed of vehicular traffic adjacent to the bicycle facility. The forms are designed with clarity in mind, such that an inspector with minimal background knowledge can complete bicycle facility inspections. Likewise, the forms are constrained to one page per bicycle facility, to ensure ease of use in the field. Potential use cases for these forms include municipal transportation asset management teams compiling information regarding the infrastructure within their jurisdiction, non-governmental activists seeking to collect data on the state of bicycle facilities in areas of interest, and/or researchers seeking to more accurately assign impedance values to bicycle facility links.

Inspection Form Development

Standardized infrastructure inspection forms are useful for identifying the condition of built infrastructure and conformance with relevant design standards. The authors have previous experience in the design and/or use of standardized infrastructure inspection forms in the field, including inspection forms for sidewalks, curb ramps, curb cuts, crosswalks, and bus stops (Reichard, 2023; Reichard et al., 2022). This expertise has contributed to an understanding of general inspection form design principles and best practices.

Well-designed inspection forms are general enough to be applicable in a broad range of contexts. For example, it would be preferable to have one bicycle facility inspection form that is applicable to all categories of bicycle facility, as opposed to separate forms for each facility type. This eliminates the need for the user of the form to have sufficient knowledge to assign a given bicycle facility to its correct category and identify the corresponding form for inspection. For ease of use in the field, the manual data entry sections of an infrastructure inspection form should also be contained to one side of one

8.5 ” x 11” sheet of paper, to avoid having to repeatedly rearrange papers on a clipboard. Likewise, ideal infrastructure inspection forms should be easy to understand by non-practitioners, such that a user could accurately perform a field inspection with minimal prior knowledge or training. This greatly lowers the barriers to entry for performing inspections. One way to do this is to include instructions and brief definitions of terms on the back side of the inspection form sheet, allowing users to quickly access information that may help clarify questions they may have. It is also very helpful to test draft form designs by having relatively inexperienced inspectors (e.g., interns) use the forms in the field and iteratively update the form based on their feedback.

In developing an inspection form, one needs to start by identifying the pertinent variables that need to be collected in the field, and therefore included on an inspection form. The research team begins by identifying relevant federal policies that constrain the design of the infrastructure in question. For example, in the case of sidewalks, design is primarily constrained by provisions within the Americans with Disabilities Act (ADA). Thus, any sidewalk inspection form should record data pertinent to ensuring ADA compliance (e.g., sidewalk slope, cross slope, disjoint sizes, etc.). If, after the inclusion of all variables necessary to assess compliance with federal policy, there is room on the form to include additional variables, a secondary priority is to include any additional variables needed to assess compliance with federal design standards or best practices published by a federal agency. A tertiary priority is to include any additional variables that are necessary for assessing compliance with best practices published by non-federal agencies, such as prominent advocacy groups. Field inspection forms should also include metadata, such as the date and time of the inspection, location of the inspection, and name(s) of the

inspector(s), to aid in processing and archiving inspection data. Finally, it is useful to include space for inspectors to include miscellaneous field notes, to account for unique conditions that warrant special consideration.

It is often difficult to create a single form that is generally applicable in a broad range of contexts and simple enough to be usable by inexperienced inspectors, while also including a comprehensive set of variables (and enumerations). As such, compromises are sometimes necessary. In the case of bicycle facility inspections, the research team opted to create two inspection forms; a general bike facility inspection form, and a bike facility conflict point inspection form. This delineation made sense because there are many different types of bicycle facility conflict points (e.g., intersections, bus stops, loading zones, mid-block crosswalks, etc.) that have unique design requirements and considerations. Accounting for each of these considerations on the generalized bicycle facility inspection form would have come at the expense of other useful data that contained on the bike facility inspection form.

Bicycle Facility Inspection Forms

The bicycle facility inspection form contains the following data fields:

Metadata

- Inspector(s) - This field contains the name(s) of the inspectors filling out the form in the field. This is useful information when there are multiple teams of inspectors performing inspections.
- Inspection date - This field contains the date the inspection was performed. This information is important because some of the other data included on the form may

be affected by the date and time of the inspection, such as observed vehicular and pedestrian volumes concurrent with the bicycle facility.

- Inspection time - This field contains the time the inspection was performed. This information is important because some of the other data included on the form may be affected by the date and time of the inspection, such as observed vehicular and pedestrian volumes adjacent to the bicycle facility.
- Map location - This field contains the location of the bicycle facility on a map. This field can be completed prior to commencing data collection, to assist in directing field inspectors to the location in need of inspection. Alternatively, this field can be completed after data collection is completed. The location helps third parties spatially visualize the location of the bicycle facility in question within the context of the transportation network.
- Photo(s) - This field contains photo(s) of the bicycle facility. This field can be completed prior to the commencement of data collection (e.g., using street view video or still imagery) to assist in directing field inspectors to the location in need of inspection. Alternatively, this field can be completed after data collection is complete, using photos captured by the inspectors to better document the condition of the bicycle facility.
- Bike Facility ID - This field contains the ID of the bicycle facility (often tied to a reference system, such as the OSM transportation network). This is useful if the inspection is part of a larger coordinated inspection effort and specific facilities have been assigned identification codes.

- Zone/Neighborhood - This field contains the zone or neighborhood of the bicycle facility. This is useful for any data analysis that involves comparing a group of bicycle facilities across sub-regions using collected data.
- Latitude - This field contains the latitude of the bicycle facility. This is useful for pinpointing the precise location of the inspection.
- Longitude - This field contains the longitude of the bicycle facility. This is useful for pinpointing the precise location of the inspection.
- Nearest Street - This field contains the name of the nearest street. This is useful for quickly getting a sense of the location of the bicycle facility and is a good failsafe if the latitude and/or longitude are recorded incorrectly.
- Nearby Landmark - This field contains the name of the nearest land use landmark. This is useful for quickly getting a sense of the location of the bicycle facility and is a good failsafe if the latitude and/or longitude are recorded incorrectly.

Right of Way Elements

- Bike Lane Alignment (Motor Vehicles) - This field denotes the alignment of the bicycle facility with respect to motor vehicle traffic. Valid entries would include left of traffic, right of traffic, center of traffic, mixed with traffic, or not applicable (if the bicycle facility in question is not associated with a roadway).
- Bike Lane Alignment (Pedestrians) - This field denotes the alignment of the bicycle facility with respect to pedestrian traffic. Valid entries would include left of traffic, right of traffic, center of traffic, mixed with traffic, or not applicable (if the bicycle facility in question is not associated with a walkway).

- Type of Horizontal Separation (Motor Vehicles) - This field notes the type of horizontal separation used to separate bicycle traffic from motor vehicle traffic. Common horizontal separation types include striping or the use of a buffer zone. Alternatively, there may be no horizontal separation used.
- Type of Horizontal Separation (Pedestrians) - This field notes the type of horizontal separation used to separate bicycle traffic from pedestrian traffic. Common horizontal separation types include striping or the use of a buffer zone. Alternatively, there may be no horizontal separation used.
- Horizontal Separation Condition (Motor Vehicles) - This field denotes that condition of the horizontal separation between bicycle traffic and motor vehicle traffic, generally graded as good, fair, poor, or absent. A freshly painted separation stripe would likely be classified as being in good condition, whereas a faded or otherwise damaged stripe would likely be classified as fair or poor depending on the extent of the damage.
- Horizontal Separation Condition (Pedestrians) - This field denotes that condition of the horizontal separation between bicycle traffic and pedestrian traffic, generally graded as good, fair, poor, or absent. A freshly painted separation stripe would likely be classified as being in good condition, whereas a faded or otherwise damaged stripe would likely be classified as fair or poor depending on the extent of the damage.
- Buffer Width (Motor Vehicles) - This field contains the width of the buffer zone separating bicycle traffic from motor vehicle traffic (width=0 if there is no buffer zone present). Buffer zones are known to increase rider comfort, and presumably

a wider buffer zone results in a higher degree of rider comfort. Rider comfort has implications for the impedance value associated with bicycle facility links.

- Buffer Width (Pedestrians) - This field contains the width of the buffer zone separating bicycle traffic from pedestrian traffic (width=0 if there is no buffer zone present). The inclusion of a buffer zone is known to increase rider comfort, and presumably a wider buffer zone results in a higher degree of rider comfort. Rider comfort has implications for the impedance value associated with bicycle facility links.
- Type of Vertical Separation (Motor Vehicles) - This field denotes the type of vertical separation between bicycle traffic and motor vehicle traffic. Vertical separations include any method of separation that has a vertical component. Common examples of vertical separations include plastic delineators, planters, parking curbs, parked cars, or a median. The type of vertical separation used has implications on rider comfort, and therefore has implications on impedance values applied to bicycle facilities.
- Type of Vertical Separation (Pedestrians) - This field denotes the type of vertical separation between bicycle traffic and pedestrian traffic. Vertical separations include any method of separation that has a vertical component. Common examples of vertical separations include plastic delineators, planters, parking curbs, parked cars, or a median. The type of vertical separation used has implications on rider comfort, and therefore has implications on impedance values applied to bicycle facilities.

- Vertical Separation Condition (Motor Vehicles) - This field denotes that condition of the vertical separation between bicycle traffic and motor vehicle traffic, generally graded as good, fair, poor, or absent. Freshly installed plastic delineators would likely be classified as being in good condition, whereas broken or otherwise damaged plastic delineators would likely be classified as fair or poor depending on the extent of the damage.
- Vertical Separation Condition (Pedestrians) - This field denotes that condition of the vertical separation between bicycle traffic and pedestrian traffic, generally graded as good, fair, poor, or absent. Freshly installed plastic delineators would likely be classified as being in good condition, whereas broken or otherwise damaged plastic delineators would likely be classified as fair or poor depending on the extent of the damage.
- Shared Lane Markings (Motor Vehicles) - This field denotes the presence or absence of shared lane markings (often referred to as sharrows) indicating that bicycles and motor vehicles are expected to share the same right of way. The presence or absence of shared lane markings has implications for rider comfort.
- Shared Lane Markings (Pedestrians) - This field denotes the presence or absence of shared lane markings (often referred to as sharrows) indicating that bicycles and pedestrians are expected to share the same right of way. The presence or absence of shared lane markings has implications for rider comfort.
- Shared Lane Marking Condition (Motor Vehicles) - This field denotes the condition of the shared lane markings, often grades as good, fair, poor, or absent. Freshly painted shared lane markings would likely be graded as being in good

condition, whereas faded or otherwise damaged shared lane markings may be graded as being in fair or poor condition, depending on the extent of the damage.

- Shared Lane Marking Condition (Pedestrians) - This field denotes the condition of the shared lane markings, often grades as good, fair, poor, or absent. Freshly painted shared lane markings would likely be graded as being in good condition, whereas faded or otherwise damaged shared lane markings may be graded as being in fair or poor condition, depending on the extent of the damage.

Adjacent Infrastructure and Operating Conditions

- Adjacent Street Parking - This field indicates if there is street parking adjacent to the bicycle travel lane. Adjacent street parking can improve rider comfort if it is used as a form of vertical separation between bicycle traffic and motor vehicle traffic. However, street parking can also be a detriment to rider comfort if there is a risk of car doors opening into the bicycle travel lane, or if the adjacent street parking creates conflict points between bicycle traffic and vehicles attempting to park.
- Motor Vehicle Volume - This field denotes the volume of motor vehicles on the adjacent or concurrent roadway. A higher volume of motor vehicles corresponds with decreased rider comfort. It is likely to be impractical for inspectors to reliably assess traffic volumes while in the field, so it is acceptable to report the traffic volume as high, medium, or low. Subsequent post processing can bring in measured/sampled AADT values, in lieu of field estimates.
- Motor Vehicle Speed - This field denotes the travel speed of motor vehicles on the adjacent or concurrent roadway. A higher prevailing traffic speed corresponds

to decreased rider comfort. It may be impractical for inspectors to assess precise traffic speeds while out in the field, so it is acceptable to roughly estimate the speed within 5-10 miles per hour. Subsequent post processing can bring in measured/sampled floating car speed values, in lieu of field estimates.

- Pedestrian Volume - This field denotes the volume of pedestrians on the adjacent or concurrent roadway. A higher volume of pedestrians corresponds with decreased rider comfort (particularly if bicycles and pedestrians share the same right of way). It may be impractical for inspectors to assess precise pedestrian volumes while out in the field, so it is acceptable to report the traffic volume as high, medium, or low.

General Elements

- Bike Route Signage - This field denotes the presence or absence of signage indicating that the facility is a bicycle route. Bicycle route signage increases rider comfort.
- Signage Visibility - This field denotes the visibility of bicycle route signage, typically graded as good, fair, poor, or absent. The level of rider comfort gained from the presence of bicycle route signage is partially dependent on the visibility of these signs.
- Bicycle Facility Directionality - This field denotes the directionality of the bicycle facility traffic (one way or two way).
- Directional Separation Type - This field denotes the type of separation used to separate two-way traffic on bicycle facilities, such as striping, a buffer zone, or

plastic delineators. The type of separation (or lack thereof) used has impacts on rider comfort.

- Lane Width (per direction) - This field denotes the width of each direction of the bicycle lane. Wider lanes are correlated with increased rider comfort.
- Shared Bike / Bus ROW - This field denotes if the bicycle facility is a combined bus and bike lane, which would have impacts on rider comfort.
- Running Slope - This field denotes the slope of the bicycle lane in the direction of travel. Steep uphill climbs are more difficult to traverse than level terrain or downhill slopes.
- Cross Slope - This field denotes the slope of the bicycle lane perpendicular to the direction of travel (the banking or tilt of the bicycle facility). A substantial cross slope (beyond a normal surface crown) may lead to difficulties in traversing the bicycle facility.
- Surface Material - This field denotes the material of the bicycle facility surface. Asphalt and concrete surfaces are generally preferred over brick or cobblestone surfaces. Dirt and gravel surfaces are also occasionally encountered, mainly on side paths.
- Surface Condition - This field denotes the condition of the surface, graded on a scale of good, fair, or poor. A freshly paved road would likely be graded as being in good condition, whereas a very old road with potholes would likely be graded as being in fair or poor condition, depending on the extent of the damage. Surface conditions impact rider comfort.

- Colored Surface - This field indicates if the surface of the bicycle facility is colored differently (typically green) from adjacent right of way. Colored surfaces may increase rider comfort.
- Potholes Present - This field indicates the presence or absence of potholes. These data provide further insight into the condition of the surface. Potholes decrease rider comfort.
- Cracking Present - This field indicates the presence or absence of surface cracking. These data provide further insight into the condition of the surface. Substantial cracking can decrease rider comfort.
- Debris Present - This field indicates the presence or absence of debris within the bicycle facility (e.g., accumulated leaves or trash that cause some or all of the bicycle facility to become impassable). These data provide further insight into the condition of the surface. Substantial debris can decrease rider comfort.
- Rutting Present - This field indicates the presence or absence of rutting. These data provide further insight into the condition of the surface. Rutting decrease rider comfort.
- Loose Materials Present - This field indicates the presence or absence of loose materials on the surface of the bicycle facility. Loose materials, such as a dirt surface that is not sufficiently compacted, or a severely damaged asphalt surface, decrease rider comfort.
- Gutter Width - This field indicates the width of a gutter that runs adjacent to a bicycle facility. There are design guidelines specifying maximum allowable slopes for gutters adjacent to bicycle lanes.

- Gutter Depth - This field indicates the depth of a gutter that runs adjacent to a bicycle facility. There are design guidelines specifying maximum allowable slopes for gutters adjacent to bicycle lanes.
- Field Notes - This field allows data inspectors to record any additional pertinent information that has not been captured by any of the above fields.

Conflict Point Inspection Forms

The bicycle conflict point inspection form contains many of the same fields, including Inspector(s), Inspection Date, Inspection Time, Map Location, Photo(s), Bike Facility ID, Zone/Neighborhood, Latitude, Longitude, Parallel Street Name, Nearby Landmark Name, and Field Notes. The bicycle facility conflict point inspection form also contains the following fields that are not present on the general bicycle facility inspection form:

General Elements

- Parked Cars within 20-feet - This field indicates if there are parked vehicles within 20 feet of the conflict point, which is prohibited in most circumstances.
- Conflict Point Visibility - This field grades the visibility at the conflict point, on a scale of good, fair, or poor. A clear view of the conflict point with no visual obstructions would likely be rated as good, while an obstructed view would likely be graded as fair or poor, depending on the severity of obstruction.
- Vehicles Yield to Bicycles Signage - This field denotes the presence or absence of signage indicating that motor vehicles should yield to bicycles. These signs should typically be installed at conflict points between bicycles and motor vehicles.

Mid-block Pedestrian Crossing Conflict Points

- Bicycles Yield to Pedestrians Signage / Markings - This field denotes the presence or absence of signage or markings indicating that bicycles should yield to pedestrians at the conflict point. These signs or markings should typically be installed at pedestrian-bicycle conflict points, particularly if the conflict point is mid-block. This field is recorded as n/a if the conflict point in question is not a mid-block pedestrian crossing conflict point.
- Crosswalk Present - This field denotes the presence or absence of a crosswalk at the mid-block pedestrian crossing conflict point. Federal design standards stipulate that a crosswalk should be present at mid-block conflict points. This field is recorded as n/a if the conflict point in question is not a mid-block pedestrian crossing conflict point.
- Crosswalk Condition - This field grades the condition of the crosswalk, typically on a scale of good, fair, poor, or absent. A fully ADA compliant crosswalk would likely be graded as good, while a crosswalk with faded paint or otherwise violates ADA design standards would likely be graded as fair or poor, depending on the severity of the condition. This field is recorded as n/a if the conflict point in question is not a mid-block pedestrian crossing conflict point.
- Raised Crosswalk - This field denotes the presence or absence of a raised crosswalk at the mid-block pedestrian crossing conflict point. Raised crosswalks improve pedestrian safety by prompting vehicles and bicycles to slow down before crossing the crosswalk. This field is recorded as n/a if the conflict point in question is not a mid-block pedestrian crossing conflict point.

Mid-Block Transit Stop Conflict Points

- Total Duration of Conflicts (minutes per hour) - this field denotes the total duration in which a transit vehicle is expected to be occupying the bicycle lane at the mid-block bus stop conflict point. The FHWA recommends that mid-block transit stop conflict points be eliminated (e.g., create a bulb out for the transit vehicle to load and unload without blocking the bicycle lane), if the bicycle lane will be blocked by transit vehicles for more than five minutes each hour. This field is recorded as n/a if the conflict point in question is not a mid-block transit stop conflict point.
- Bus Only Pavement Markings - This field denotes the presence or absence of bus only pavement markings. These pavement markings are optional at mid-block bus stop conflict points. This field is recorded as n/a if the conflict point in question is not a mid-block transit stop conflict point.
- Bicycles Yield to Bus Signage / markings - This field denotes the presence or absence of signage or markings indicating that bicycles should yield to buses at the conflict point. These signs and markings are optional, but best practices indicate that they should be installed. This field is recorded as n/a if the conflict point in question is not a mid-block transit stop conflict point.

Loading Zone / Accessible Parking Mid-Block Conflict Points

- Crosswalk Present - This field denotes the presence or absence of a crosswalk at the mid-block pedestrian crossing conflict point. Federal design standards stipulate that a crosswalk should be present at accessible street parking locations

that conflict with a bicycle facility. This field is recorded as n/a if the conflict point in question is not a mid-block pedestrian crossing conflict point.

- Crosswalk Condition - This field grades the condition of the crosswalk, typically on a scale of good, fair, poor, or absent. A fully ADA compliant crosswalk would likely be graded as good, while a crosswalk with faded paint or otherwise violates ADA design standards are graded as fair or poor, depending on the severity of the defect. This field is recorded as n/a if the conflict point in question is not a mid-block pedestrian crossing conflict point.
- Bicycle Lane Width Reduction (ft) - This field denotes the magnitude of bike lane width reduction associated with the conflict point. A reduction in lane width may have impacts on rider comfort. If there is no lane width reduction, this field is recorded as reduction with = 0.

Intersection Conflict Points

- Intersection Vehicle Directions - This field indicates the number of directions from which vehicles enter the intersection (e.g., three-way intersection, four-way intersection, etc.) This information is helpful for contextualizing the conflict point.
- Adjacent Street Vehicle Travel Lanes - This field denotes the number of motor vehicle travel lanes on the street adjacent to or concurrent with the bicycle facility. This information is helpful for contextualizing the conflict point.
- Crossing Street Vehicle Travel Lanes - This field denotes the number of motor vehicle travel lanes on the street intersecting with the bicycle facility. This information is helpful for contextualizing the conflict point.

- Traffic Control Type - This field denotes the type of traffic control or signalization used at the intersection (e.g., signalized intersection, 4-way stop, 1-way stop, yield controlled, etc.). This information is helpful for predicting rider comfort, rider safety, and travel time.
- Buffer Width - This field indicates the width of the buffer zone (if one exists) between bicycle traffic and motor vehicle traffic queuing at an intersection. This has impacts on the visibility of bikers to vehicles making a turn.
- Two-Stage Bicycle Left Turn Present - This field denotes the presence or absence of a two-stage bicycle left turn traffic pattern at the intersection. In a two-stage bicycle left turn, left turning bicycles are required to cross the intersection going straight, stop at the opposite side of the intersection, and wait for a signal change to complete the turn. This traffic pattern can improve safety but requires riders to wait for an extra signal-cycle.
- Bicycle Box Present - This field denotes the presence or absence of a bicycle box at an intersection. Bicycle boxes allow bicycles to queue to the left or right of motor vehicle traffic, which is useful for riders attempting to turn.
- Bicycle Box Condition - This field grades the condition of the bicycle box, on a scale of good, fair, poor, or absent. A freshly painted and easily visible bicycle box would likely be graded as being in good condition, while a faded or otherwise damaged bicycle box would likely be graded as being in fair or poor condition, depending on the extent of the damage.
- Signal Cycle Length - This field indicates the length of the signal cycle, in seconds. This information is useful for predicting travel time.

- Bicycle Specific Signals - This field denotes the presence or absence of traffic signals specific to bicycle traffic. The inclusion of bicycle signals can improve safety and rider comfort but may not be warranted in areas with low traffic volume.
- Bicycle Lane Clearly Marked Before Intersection - This field indicates if the bicycle lane is clearly marked on the section of the facility before entering the intersection.
- Bicycle Markings in Intersection - This field denotes the presence or absence of bicycle lane markings within the intersection. It is often helpful to provide pavement markings within the intersection to remind turning vehicles to be aware of riders.
- Bicycle Marking Condition - This field denotes the condition of bicycle markings within the intersection, graded on a scale of good, fair, poor, or absent. Freshly painted markings that are easily visible would likely be graded as being in good condition, while faded markings would likely be graded as being in fair or poor condition, depending on the extent of the damage.
- Turning Vehicles Cross Over Bicycle Lane - This field indicates if motor vehicles making a turn (typically right) are required to cross over the bicycle lane prior to making the turn. This creates a secondary conflict point between bicycles and motor vehicles.
- Merge Area Length - This field notes the length of the merge area between bicycle traffic and motor vehicle traffic, if one exists.

- Merge Area Distance to Intersection - This field notes the distance between the start of the merge area and the start of the intersection, if one exists.

Inspection Form Testing

Preliminary versions of the inspection forms described in the previous sections such that each form could capture the field information necessary to assess conformance with bicycle facility design standards. The Bicycle Facility Inspection form is presented in Appendix A and the Bicycle Conflict Point Inspection form is presented in Appendix B.

The forms were field tested by assigning undergraduate students who performed inspections of several distinct bicycle facilities throughout the Georgia Institute of Technology Atlanta campus. The undergraduates recruited for this testing had previously taken an introductory transportation engineering class but did not have any other prior experience with transportation engineering. It is reasonable to assume that any other users of this form would have approximately as much, if not more, background knowledge on the design of bicycle facilities.

The primary focus of this testing was to improve the ease of understanding of the forms. As such, the undergraduate inspectors were accompanied by one experienced graduate student inspector. The graduate student inspector's role was to note the aspects of the form design that were difficult to intuitively understand. Through this process, the designs of the forms were iteratively updated.

The study route included a variety of separate bicycle facility contexts, including separated bicycle facilities, on-street bike lanes, signed routes, unmarked routes, and side paths. Likewise, the route contained several different types of conflict points, including

standard intersections, mid-block crosswalks, and bus stops. It was necessary to test the form in a broad array of contexts to ensure that the form makes sense regardless of facility type.

Case Study Locations

The bicycle facility inspection forms were tested at the following locations on the Georgia Institute of Technology's Atlanta campus (see Figure 20):

- Bicycle Facility 1 (red) - PATH Parkway. The section of the PATH Parkway through Georgia Tech's campus is a dedicated bikeway fully separated from pedestrian and motor vehicle traffic, other than at intersections.
- Bicycle Facility 2 (orange) - Hemphill Avenue Northwest. This street transitions from an on-street non-protected bicycle lane to a signed route indicating that bicycles and motor vehicles share the road.
- Bicycle Facility 3 (yellow) - Eco-Commons. This is a side path that is separated from pedestrian and motor vehicle traffic.
- Bicycle Facility 4 (green) - State Street Northwest. This street contains an on-street protected bike lane.
- Bicycle Facility 5 (blue) - Atlantic Drive Northwest. This street is off-limits to motor-vehicle traffic and has no physical delineation between bicycle traffic and pedestrian traffic.
- Conflict Point 1 - Three way stop controlled intersection containing on-street non-protected bike lanes and a bike box for left turning bicycles.
- Conflict Point 2 - Mid block pedestrian crosswalk intersecting with two motor vehicle travel lanes and one on-street unprotected bicycle lane.

- Conflict Point 3 - Side path intersecting two motor vehicle travel lanes and two protected bicycle lanes.
- Conflict Point 4 - Signalized 4-way intersection that requires right-turning vehicles to merge through the protected bicycle lane.
- Conflict Point 5 - Signalized 4-way intersection between a two-lane motor vehicle lane road and a bicycle / pedestrian road.

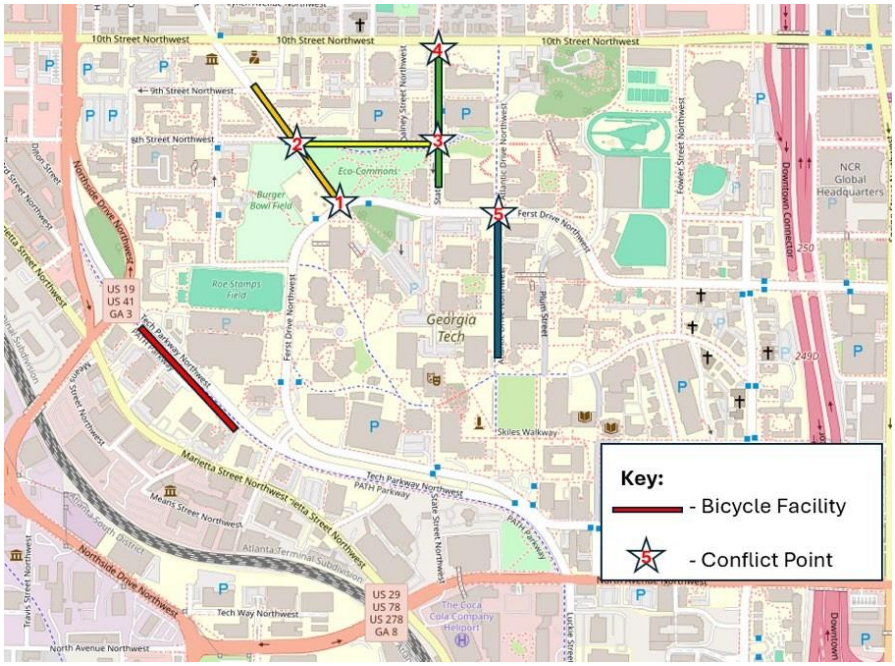

Figure 20: Bicycle Facility Inspection Form Test Locations

The form designs were iteratively improved through continuous field testing in Atlanta, Georgia. Updates were made to each form to improve practicality, clarity, conciseness, and to optimize the types and amounts of data contained on each form. The forms were

tested by teams of undergraduate students, with one graduate student supervisor present to receive real-time feedback and to provide guidance as necessary. The manuscript describing the development and implementation of these bicycle facility and bicycle conflict point inspection forms has been submitted to the 2025 Transportation Research Board Annual Meeting call for papers (see Reichard et al., 2024).

In 2024, Georgia Tech's ITS4US team research team (partners with GDOT on the $11 million FHWA ITS4US project) developed a number of video-based flythrough inspection applications for sidewalk assets (performing inspections to gather the same information on field inspection forms, but within a Web-based app). The manual field inspection forms for sidewalks, ramps, and curb cuts are very similar to the bicycle facility inspection forms prepared for this project. The ITS4US research team has also developed a number of automated machine vision sidewalk asset feature identification algorithms that could also be adapted to bicycle infrastructure (which would even use the same vehicle-based video collected for the ITS4US project). The team believes that a modest follow-on GDOT project could adapt the online video inspection and machine vision tools for use in developing the feature set and asset design parameters that were used in impedance algorithms presented in this report.

CHAPTER 4. CYCLEATLANTA DATA PROCESSING AND MAP MATCHING

The impedance calibration process requires GPS traces from monitored bicycle trips that show what routes were chosen by users. These traces are typically difficult to acquire because software development is required to collect the traces, and resources are required to deploy the software. Existing traces often cannot be acquired from third parties due to privacy concerns and aggregated data does not have the detailed routing information required. Fortunately, the research team had access to data from a prior research project collected through a publicly available smartphone app known as CycleAtlanta (Misra et al., 2014). The app was operational in Atlanta between 2012 – 2016. CycleAtlanta was a branched software version of CycleTracks, the app used by Hood et al. (2011) to collect GPS traces. Anyone could download the app to their smartphone, create an account, and record their trips. When creating an account, the app would ask users to provide sociodemographic information, their rider type, cycling frequency, and cycling experience level. When recording trips, the app would ask users to identify their trip purpose. A cycling route choice model was developed from these data and presented in a Ph.D. dissertation by Dr. Aditi Misra (Misra, 2016).

An in-depth analysis of the trip and user characteristics of the CycleAtlanta data can be found in Misra (2016). This report will reproduce and discuss some of the analyses and figures shown in Misra (2016) but will not go in as much depth to avoid retreading this work. Both the data cleaning and map matching processes have been updated since Misra (2016), so these processes are described in the next two sections.

DATA CLEANING

The CycleAtlanta data are composed of three different data tables: traces, trips, and users. The traces table contains the second-by-second GPS position and related data for every trip that was logged, including the trip ID, timestamp, latitude, longitude, altitude, instantaneous speed, horizontal accuracy, and vertical accuracy. Ephemeris data were not provided with the GPS data, which limited the options for filtering out potentially erroneous single-second data points in the subsequent cleaning steps.

The trips table contains the trip characteristics of each trip, including the trip ID, user ID, trip type, and trip description. The user table contains metadata on the device ID, user creation date, and the app version as well as user characteristics such as user ID, email, age, gender, income, ethnicity, home zip code, school zip code, work zip code, cycling frequency, rider history, and rider type.

These data were cleaned using the series of steps listed in Table 14. The first step of the cleaning process was to load all three data tables and cross-reference them against each other using the user ID and the trip ID. User IDs that weren't present in both the users and trips data tables as well as trip IDs that weren't present in both the traces and trips tables were removed. Users that didn't have trips or traces may have been users that never recorded a trip. Trips that didn't have traces may have been canceled or misrecorded trips.

The next step was to clean the traces data table. First, duplicate GPS points with the same timestamp were dropped. Then, trips with fewer than 300 records (or five minutes' worth of recorded data) were dropped. While the CycleAtlanta app generally records second-by-second GPS data, there are instances where there are large gaps between GPS records.

Occasionally, there would be so many large gaps that traces became too sparse to map match. These sparse trips were removed by removing trips with fewer than 300 GPS records, which corresponds to five minutes' worth of recorded data.

Remaining trips that had at least one point outside of the study area (introduced in Chapter 3) were removed. Duplicate trips where two or more trips had the same trace data but different trip IDs were also removed.

Any remaining trips that had more than a 2-minute gap between recorded points were trimmed. Misra (2016) selected a 2-minute pause to represent the longest feasible wait time at a traffic signal. Trips with long pauses could be trips where the user performed a trip chain. At the moment, these trips are just trimmed to represent the first segment, but these recorded points will be investigated in the future to see if they should be retained as separate trips.

The instantaneous speed column was then used to check for trips that may have involved some travel in a car or transit vehicle. A trip was removed if there were at least 60 GPS points with instantaneous speeds above 30 miles per hour (i.e., a person sustained at least a minute of travel at 30 miles per hour). It's feasible for a bike to break 30 miles per hour, but it isn't generally feasible to maintain that speed very frequently. An upper bound of 100 miles per hour was also set to account for any unreasonable short-duration speeds caused by satellite data errors. The filter was modified so that a trip would only be dropped if it contained 60 points between 30 and 100 miles per hour.

Horizontal accuracy data was then used to drop points with low-accuracy readings. High horizontal accuracy values, measured in feet, were associated with lower accuracy. After

inspecting many trips, it was clear that some trips had a high average horizontal accuracy despite appearing to be accurate on the map. Because of this, GPS points were only removed if they were at least four standard deviations above a trip's mean horizontal accuracy. Because the previous cleaning steps dropped individual data points, the 300-records filter was re-imposed to remove trips that had now lost too many records in the cleaning process.

Finally, the email field in the user table was used to find duplicate users. These users may have registered multiple times, gotten a new phone, or forgot their old login. In addition to email, the researchers inspected the origins and destinations of the potential duplicate users to confirm that they were indeed duplicates. Combining the duplicate users required additional processing if the person had entered different user characteristics for each login. When this occurred, only non-null values were retained and rules were set in place for age (retain the higher age) income (retain the higher income), cycling frequency (retain the highest frequency), rider history (retain the highest experience), and rider type (retain the most confident rider type). Ultimately, the cleaned CycleAtlanta data contained 33,936,817 points, representing 20,273 trips, across 972 users, as shown in Table 14.

Table 14: CycleAtlanta data cleaning steps

Cleaning Step	GPS Points	GPS Points Dropped	Trips	Trips Dropped	Users	Users Dropped
Load all data	58,631,361	0	34,043	0	1,718	0
Cross-reference all data	58,568,491	42,950	28,187	5,840	1,483	235
Drop duplicate GPS points	48,588,441	9,999,970	28,203	0	1,488	0
Drop trips with fewer than 300 records	48,223,615	364,826	25,472	2,731	1,336	152
Drop trips with at least one point outside of the study area	37,793,055	10,430,560	22,046	3,426	1,064	272
Drop duplicate trips	35,444,448	2,348,607	20,497	1,549	1,063	1
Drop points recorded after a two-minute pause	34,850,695	593,753	20,497	0	1,063	0
Drop trip if 60 points have a speed above 30 mph but below 100 mph	34,226,290	624,405	20,344	153	1,060	3
Drop points when horizontal accuracy is four standard deviations above the mean horizontal accuracy	33,952,302	273,988	20,344	0	1,060	0
Drop trips with fewer than 300 records	33,936,817	15,485	20,272	71	1,058	2
Consolidate duplicate users	33,936,817	0	20,272	0	971	86

Figure 21 shows the number of trips across CycleAtlanta's lifespan, from October 2012 to September 2016. There were two periods of data collection, with a gap present from February 2014 and July 2014 that could have been during an outage or an update of the app. The first period of data collection, from October 2012 to January 2014 had a higher number of trips recorded per month than did the second data collection period, from August 2014 to September 2016. Figure 22 shows the creation date of users across the

same period which reveals that the number of newly registered users fell precipitously after the first period of data collection.

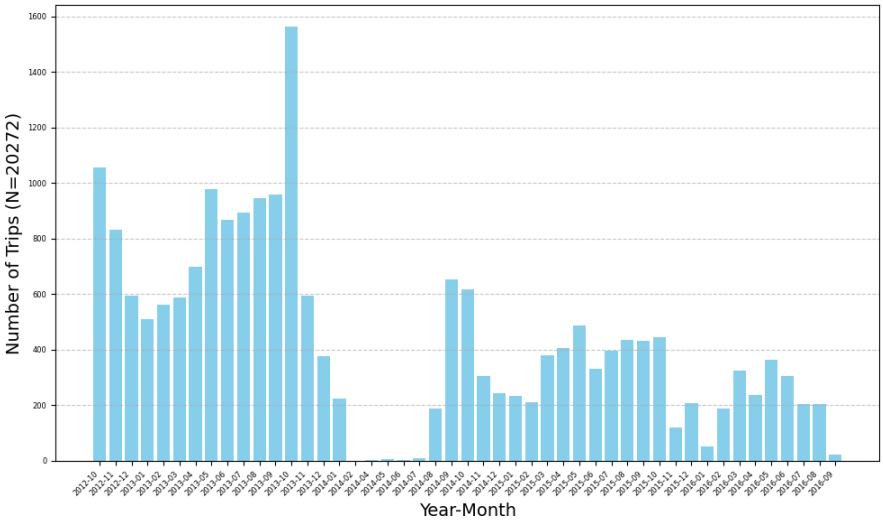

Figure 21: Number of CycleAtlanta trips over time

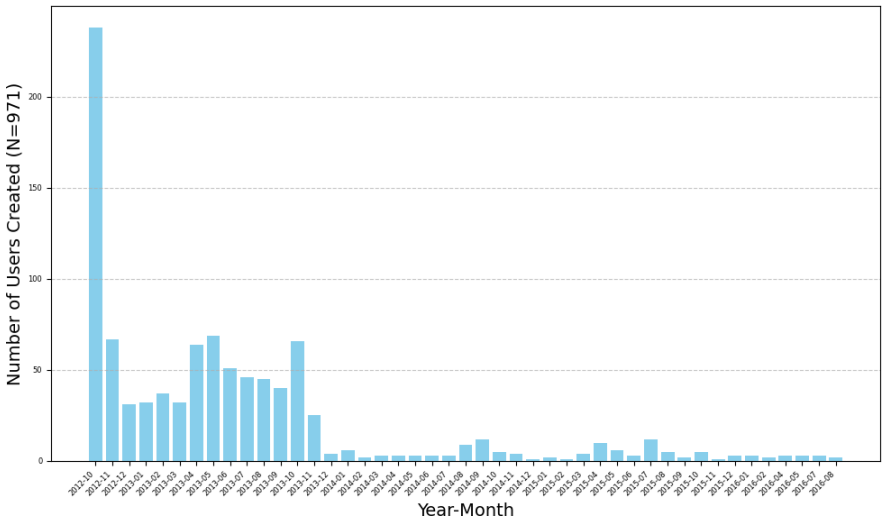

Figure 22: Number of new CycleAtlanta users over time

IDENTIFYING REDUNDANT TRIPS

With the data cleaned, the next step was to identify redundant trips in the data. A redundant trip is where two or more trips take approximately the same route, regardless of the direction between the two points. For example, consider a user logging their regular commute between home and work. If they recorded this trip multiple days and never deviated from their route, then that would be considered a redundant trip. If they took the same route between home and work as from work to home, that would also be a redundant trip. Redundant trips were identified to reduce the number of traces that needed to be map-matched and fed into the impedance calibration process.

To identify redundant trips, each user's origins and destinations were clustered using the density-based spatial clustering of applications with noise (DBSCAN) algorithm (Ester et

al., 1996), and then the trajectory similarity, measured through the Fréchet distance, was calculated for trips between the same points (Alt & Godau, 1995).

Each user's origins and destinations were retrieved by extracting the first and last coordinates recorded for each trip. The origins and destinations, combined into a set of points, were clustered using the DBSCAN algorithm implementation in the sci-kit learn Python module with an epsilon setting of 1,500 feet and a minimum cluster size of two (Scikit-learn, n.d.). This algorithm groups closely packed points together, which works well for origins and destinations from the same user where unique origins and destinations should be spread out and similar ones should be closely packed. The epsilon parameter, which controls the maximum distance allowed between points within a cluster, was adjusted to cluster origins and destinations that were spread out across college campuses or shopping centers. This also helped to account for GPS noise. Trip origins/destinations that weren't clustered were labeled as unique. Figure 23 shows an example on Georgia Tech campus where all the points are clustered despite being on different parts of campus. The average number of unique OD pairs per user was four, while the average number of trips per user was 21.

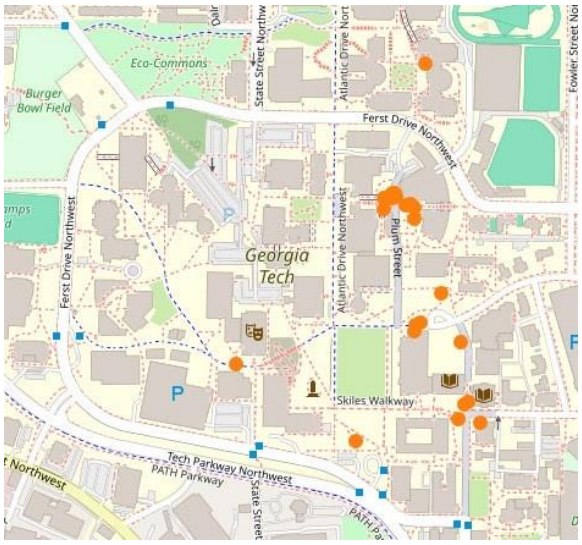

Figure 23: Clustered origins and destinations on Georgia Tech campus

After clustering, trips with an origin and destination that belonged to a cluster (excluding cases in which the origin and destination cluster belonged to the same cluster), were further evaluated to calculate the trajectory similarity between trips. Trajectory similarity was measured using Fréchet distance, because this metric considers both the location and ordering of points along a line. After testing, trips with a Fréchet distance of 500 feet or less were deemed to be similar trips. Essentially, this means that when moving two points along two lines representing two trips, the maximum allowed distance between the two points as they move along their respective lines cannot exceed 500 feet.

If this distance was exceeded, then the two trips were considered to have taken distinct routes. These distinct routes are referred to as “trip patterns” and each trip pattern is given a unique identifier specific to the user.

Figure 24 shows two sets of lines (one set colored in dark blue and the other set colored in light blue) that represent different trip patterns between points. Here, there is a deviation between routes between the intersections of North Avondale Road / South Candler Street and North Arcadia Avenue / East Ponce de Leon Avenue. The research teams' local knowledge suggests that this a common diverging point for bicyclists choosing between the more direct, but more stressful, North Avondale Road to the South, or the less direct, but less stressful, series of streets that connect to the Stone Mountain Path route to the North. The trip doesn't only just deviate at this section, but it also likely deviates further to the west, because there are two roads separated by train tracks.

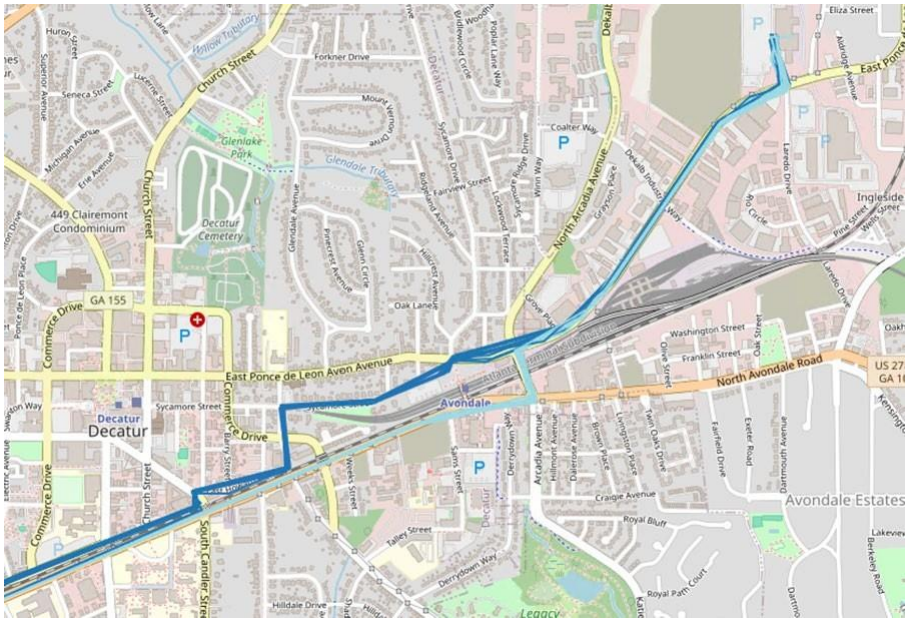

Figure 24: Example of two sets of trip patterns

The process for assigning a "trip pattern" ID was as follows. For each user, and then for each origin-destination pair (including both directions), each trip was assessed in an iterative process. Starting with the first trip in the origin-destination pair group, the Fréchet distance with all the other trips with the same origin-destination pair was calculated. If a trip was in the reverse direction (i.e., the first trip was from cluster 1 to cluster 2, but the second trip was from cluster 2 to cluster 1), the sequence of coordinates was flipped before calculating the Fréchet distance. All trips with a Fréchet distance of less than 500 feet were assigned to this first trip. Once a trip has been assigned, it was never re-assigned. The next trip is checked to see if it has already been assigned to an existing trip pattern. If not, a new trip pattern ID was generated for that trip, and the Fréchet distance was calculated to all the remaining unassigned trips. This process continued until all trips had been assigned a trip pattern ID.

To improve the computation time associated with calculating the Fréchet distance, the trajectories were simplified using the Ramer-Douglas-Peucker algorithm, which is a recursive algorithm for simplifying the shape of lines (Douglas & Peucker, 1973). After assigning trip patterns, the average number of unique trips per user fell from 19 to 12.

In future work, the research team plans to investigate the cause of these different trip patterns. One possible cause could be users taking a direct route to work and a more "scenic" or leisurely route back home after work. These route changes shouldn't be a result of trip chaining, because the origin and destination of the trips are the same (trip chaining would show up in the data as multiple trips). In particular, it would be especially interesting to know if the different trip patterns can be used to detect changes in routing from new bicycle facilities.

FILTERING THE DATA

After assigning the trip pattern IDs, the data were filtered for use in the map matching and impedance calibration processes. First, trips that were labelled as exercise trips by app users were removed. These trips are not likely to be destination-oriented, and the impedance calibration process requires an assessment of origin-destination choice. Next, trips that were less than 1.0 mile or greater than 10.0 miles were removed. Short trips had little opportunity for detouring from the shortest route and long trips are likely not representative of the typical cyclist. Next, trips with a mean speed of less than 5.0 miles per hour or greater than 16.0 miles per hour were removed. Excessively slow trips may need further processing to investigate if they were walking trips or perhaps included stops representing trip chains that were not identified in the cleaning process. Trips with high average speeds are also likely not representative of typical cyclists (e-bikes were not as common between 2012 and 2016 as they are today).

After applying these filters, loop trips and redundant trips were removed. Loop trips were identified by having the same origin and destination cluster ID. Unique trips were identified by the presence of an un-clustered origin or destination. For trips with a labeled origin and destination, unique trips were identified as trips for a single trip pattern ID that had a frequency of one. For the remaining redundant trips, the most frequently occurring trip pattern (the mode) was retained. If there were multiple modes, the lowest numerical trip pattern ID was retained. Of the trips with that trip pattern ID, the lowest numerical trip ID was used as the representative GPS trace for map matching and impedance calibration. The result of each filtering step on the number of trips and users is shown in

Table 15. In the end, there were 2,978 trips and 721 users for the map matching process in the next section.

Table 15: Filtering steps

Filter Step	Trips	Trips Dropped	Users	Users Dropped
Start	20,272	0	971	0
Remove exercise trips	17,562	2,710	830	141
Remove trips above 10 miles or less than 1 mile	15,953	1,609	792	38
Remove trips with average speed above 16 mph per hour or below 5 mph	15,276	677	742	50
Remove loop trips	14,028	1,248	721	21
Remove redundant trips	2,978	11,050	721	0

MAP MATCHING

Map matching is the process of determining the most likely series of streets/paths given the recorded GPS trace. Figure 25 shows an example of the GPS points (black dots) and the map-matched route (red polyline). This is not always a straightforward process because GPS data are noisy.

The state of the practice method for map matching is the Hidden Markov Model (HMM). For a detailed overview of this approach, see Newson and Krumm (2009), and for another study using HMM see Meister et al. (2021). The goal of HMMs in map matching is to find the most likely sequence of "states" (network links) given a set of "observations" (GPS records). HMMs are limited by road connectivity, which works well for preventing unrealistic link transitions between links that aren't directly connected (e.g., bridges and underpasses). However, informal shortcuts, such as the one shown in Figure 26, will cause the HMM to fail to find the correct path.

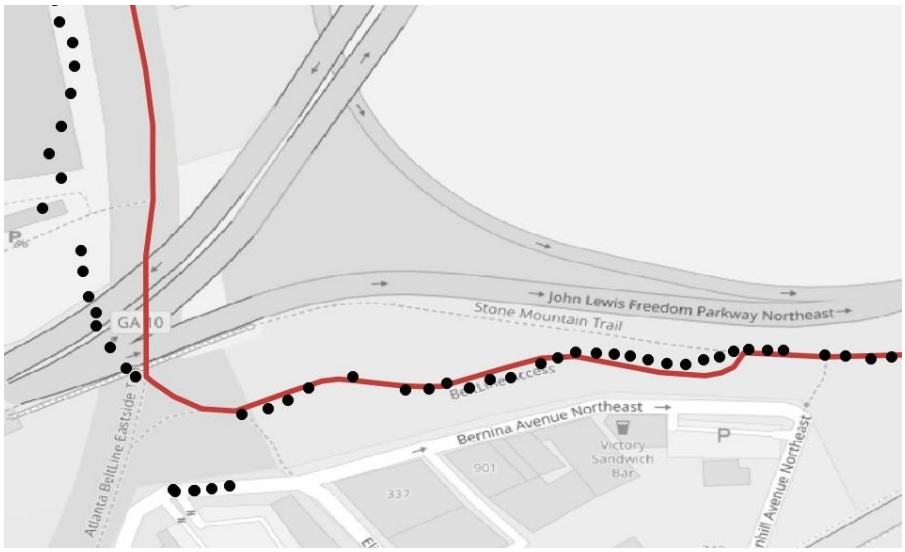

Figure 25: Example of a map-matched trace

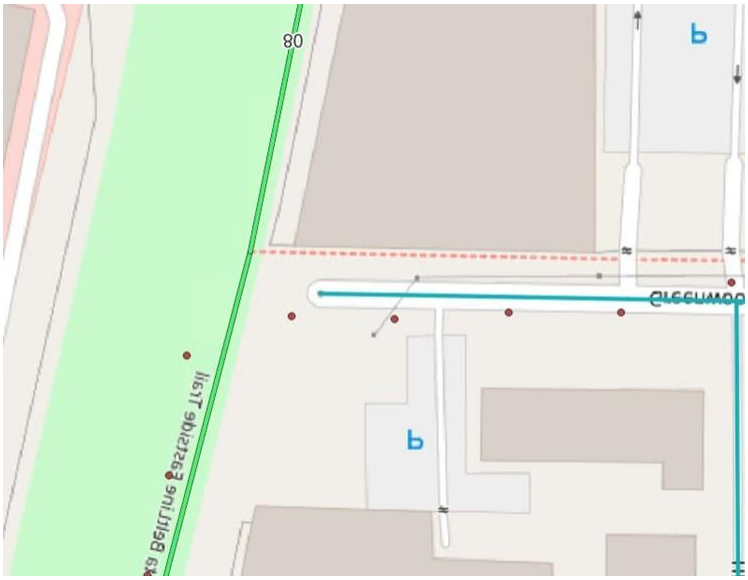

Figure 26: Example of failed map match due to informal shortcut

Great care had to be taken to filter the type of links allowed for map matching, particularly in the case when links were adjacent to each other. The GPS data collected from CycleAtlanta was not precise enough to know whether a cyclist was riding on a road or the sidewalk adjacent to the road. So, sidewalks, service roads, and parking lot aisles were removed, to favor matching to road, pedestrian, and bike links. Any isolated networks (i.e., disconnected links) created after filtering the network were also removed to ensure that there was only one fully connected graph. In the future, select links will be added back or drawn in as needed to improve map matching results.

This study used the HHM implemented in the Leuven matching Python package (Meert & Verbeke, 2018). This package adds the ability to model non-emitting states; these are links that don't match to a specific GPS point but improve the match by connecting to links that do match to GPS points. This is especially helpful for trips with fewer points or areas where there is a high density of links, because HMM requires each GPS point to match to a link.

To reduce the computational time for map matching, the spacing between points was reduced to every 30 feet. Out of the 2,978 trips from the cleaned and filtered dataset, 1,936 trips had at least 90% of their GPS points matched to links. The 1,042 trips that failed to be matched to routes will be investigated in future research. Of the unmatched trips, 388 trips failed to match entirely, and 654 only partially matched. These trips may have used shortcuts that weren't on the map or may have started within a large parking lot that was far enough away to not match to any nearby network nodes. For this report, the subset of 1,936 trips that were matched was sufficient for demonstrating the impedance calibration process.

MAP MATCHING RESULTS

As stated earlier, 1,936 (65%) of the 2,978 trips in the filtered data were successfully map-matched. This section provides descriptive statistics for the user and trip characteristics of the map-matched trips. When examining the composition of the CycleAtlanta users, Misra (2016) confirmed that the users were not representative of the local area and were overrepresented by white men. On whether the CycleAtlanta users were representative of the general Atlanta cycling population, it wasn't clear as demographic data on the cyclist population of Atlanta were not available.

This report uses the CycleAtlanta data to calibrate impedance factors, but the research team acknowledges that the CycleAtlanta userbase may not be representative of Atlanta's current or historic cycling population. Atlanta has seen extensive expansions in its bicycle network since the app was discontinued. In the future, it is recommended that new traces are collected to see if the routing preferences of cyclists have changed since then.

User Characteristics

Of the 721 users in the filtered data, 586 users were represented in the map-matched data. Figure 27 shows the distribution of the number of trips recorded per user. As with the full dataset, certain users recorded more trips than others, but most users recorded fewer than five unique trips. Only 26 users recorded more than 10 unique trips.

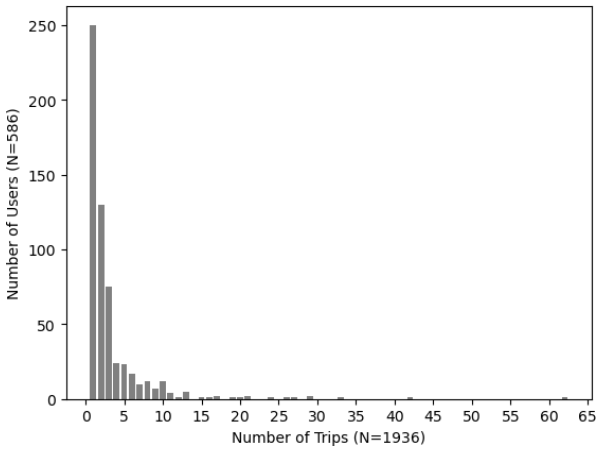

Figure 27: Trips per user in the map matched data

Table 16 shows the trip characteristics for the map-matched data. Most of the trips were commute trips, followed by social trips.

Table 16: Trip type distribution for the map matched CycleAtlanta traces

Field	Possible Values	Trips after map matching (N=1,936)
Trip Type	Commute	859
	Errand	173
	Other	88
	School	85
	Shopping	109
	Social	456
	Work-Related	166
	NULL	0

Table 17 shows the composition of the user characteristics. For each user characteristic, there were about 100 users that did not provide the requested information. As in Misra (2016), the users with map-matched data are predominantly white and male which is not

representative of the population demographics of the study area. As Dr. Misra discussed in her PhD thesis, it is difficult to assess the representativeness of the CycleAtlanta data of the people that cycle in Atlanta (that is, while we do know population demographics, we really do not have a good handle on regional or sub-regional demographics for those that bicycle for transport).

Regardless, the data collection method, a smartphone app (in the early 2010s), combined with the recruitment methods contributed to CycleAtlanta users being more likely to be tech-savvy, white, younger, and more active cyclists. This leaves out cyclists that did not have a smartphone at the time, which likely left out lower-income cyclists, particularly ones that rely on their bicycle for transportation and must ride regardless of the built environment. Despite this limitation, research has shown that even more confident cyclists still go out of their way for cycling infrastructure and report that they would prefer more separation from automobiles (Watkins et al., 2019).

Table 17: User characteristics in the map matched data

User Characteristic	Possible Values	Users after map matching (N=586)	Trips after map matching (N=1,936)
Gender	Female	124 (21%)	456 (24%)
	Male	335 (57%)	1,120 (58%)
	NULL	127 (22%)	360 (19%)
Age	18-24	43 (7%)	106 (5%)
	25-34	247 (42%)	934 (48%)
	35-44	108 (18%)	303 (16%)
	45-54	52 (9%)	188 (10%)
	55+	15 (3%)	54 (3%)
	NULL	121 (21%)	351 (18%)
Income	Less than $20,000	43 (7%)	168 (9%)
	$20,000 - $39,999	72 (12%)	230 (12%)
	$40,000 - $59,999	50 (9%)	170 (9%)
	$60,000 - $74,999	50 (9%)	259 (13%)
	$75,000 - $99,999	62 (11%)	217 (11%)
	$100,000 or greater	117 (20%)	409 (21%)
	NULL	192 (33%)	483 (25%)
Ethnicity	White	380 (65%)	1,419 (73%)
	African American	13 (2%)	28 (1%)
	Asian	15 (3%)	35 (2%)
	Native American	2 (0%)	2 (0%)
	Pacific Islander	0 (0%)	0 (0%)
	Multi-racial	16 (3%)	39 (2%)
	Hispanic/Mexican/Latino	21 (4%)	56 (3%)
	Other	7 (1%)	13 (1%)
	NULL	132 (23%)	344 (18%)
Cycling Frequency	Daily	119 (20%)	501 (26%)
	Several times per week	144 (25%)	578 (30%)
	Several times per month	45 (8%)	92 (5%)
	Less than once a month	25 (4%)	70 (4%)
	NULL	253 (43%)	272 (14%)
Rider History	Since childhood	249 (42%)	994 (51%)
	Several years	151 (26%)	514 (27%)
	One year or less	58 (10%)	130 (7%)
	Just trying it / just started	14 (2%)	26 (1%)
	NULL	114 (19%)	272 (14%)
Rider Type	Strong and fearless	105 (18%)	507 (26%)
	Enthused and confident	226 (39%)	765 (40%)
	Comfortable, but cautious	143 (24%)	399 (21%)
	NULL	112 (19%)	265 (14%)

Commented [BCD6]: I am concerned about using a data set that is mostly white cyclists and skews towards young, affluent male riders. If this model is to be used to select bike projects it needs to reflect the needs of all users.

Additionally, is it possible to break this table out by trips rather than users? Based on Figure 25, the majority of users only logged 1-3 trips, but a few logged 5-45. Depending on who the more consistent reporters were this data may be even more skewed than it appears.

CHOSEN ROUTE ANALYSIS

In this section, the spatial distribution of the 1,936 trips that were successfully matched are examined. First, Figure 28 shows the distribution of trip distances. The average trip distance was 3.9 miles. Note that the range of trip distances was impacted by the filtering steps, as trips that were less than a mile or greater than 10 miles were removed.

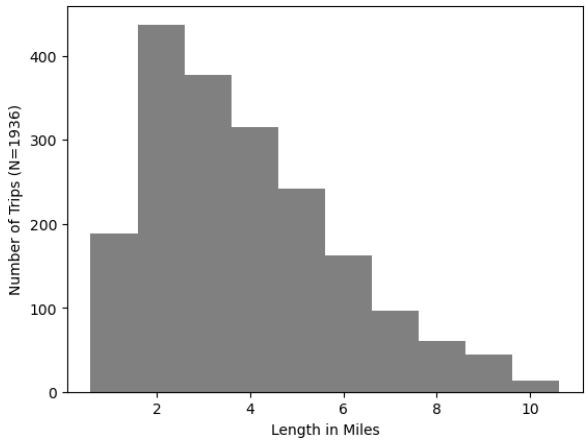

Figure 28: Distribution of trip distances

Figure 29 shows the number of unique users observed on each link in the network, where the thicker links indicate a higher number of users. A buffered version of the bicycle facilities in the area is shown for reference. From Figure 29, the CycleAtlanta users appear to be concentrated around the perimeter of the Old Fourth Ward neighborhood. The roads that make up this perimeter are 10th Street NE to the North, the Eastside Atlanta BeltLine trial to the East, Freedom Parkway Tail and Edgewood Avenue to the South, and Peachtree Street to the West. Some other standout routes include the Stone Mountain Path out to Decatur, Hosea L Williams Drive Southeast, and North Highland. It

appears that many of the highest volume routes are also near cycling infrastructure, with notable exceptions being Highland Drive, Howell Boulevard, and Peachtree Road.

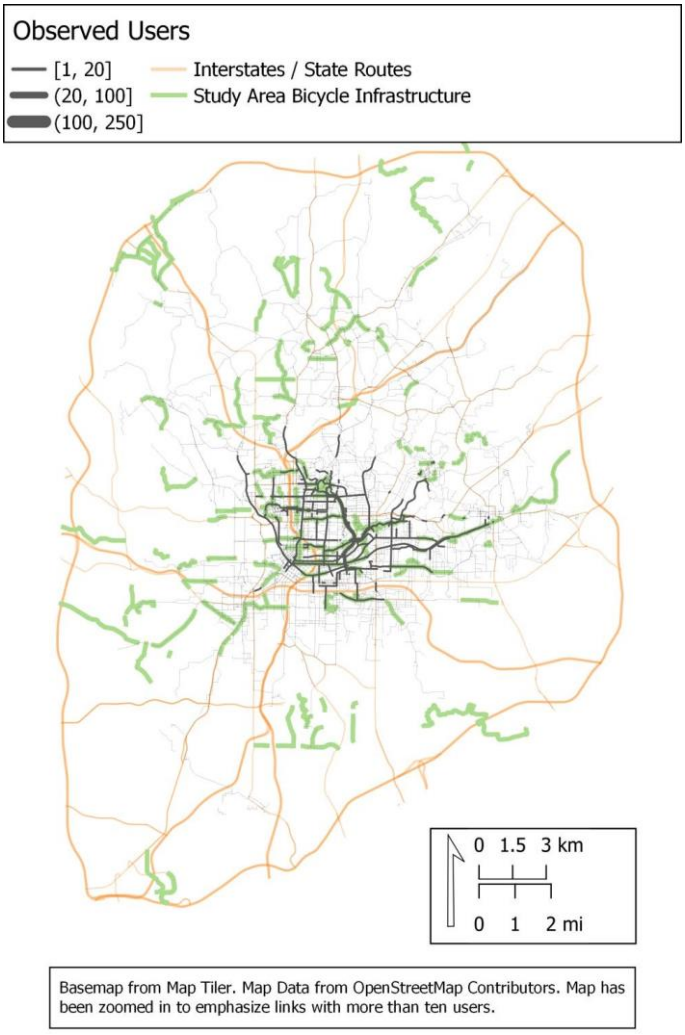

Figure 29: Unique users observed per link

Commented [BCD7]: A couple comments here:
-What percentage of trips were taken on bike facilities vs no bike facilities? Was there good variety to build an accurate model?
-Were all the bike facilities shown here present from 2012-2016 when the Cycle Atlanta data was collected?
-Why is there a big gap in the data on the Eastside Trail?
- Consider making the green highlighting less thick or making the lines on bike infrastructure green to make this map more legible.

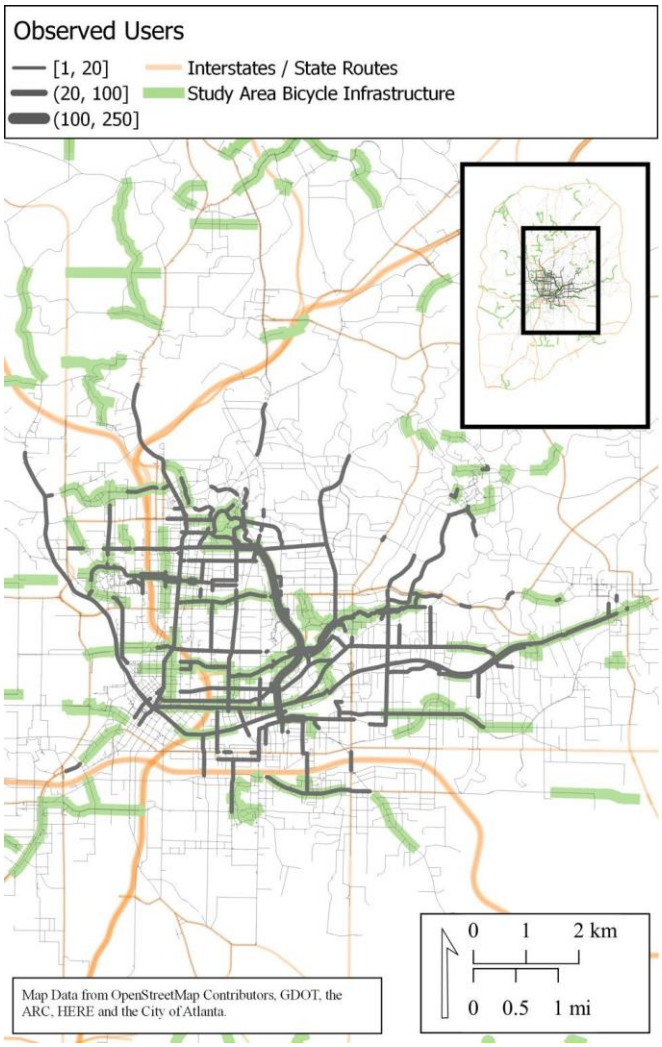

Figure 30: Zoomed in view of the links with the highest number of users

Table 18 shows the descriptive statistics for the chosen route characteristics. These statistics describe all of the trips, and the percentage of route variables are not weighted by trip distance to prevent longer trips from influencing the statistics. The difference in

the mean and median, in addition to the standard deviation, is large for several route characteristics. This likely hints that preferences for facilities varied a lot, but it may also indicate that some link attributes may not have been available for certain trips. For instance, the percentage of the route on bike lanes and multi-use paths variables show large differences in the mean and median, which could be due to many trips not having proximal access to these facilities.

The road attribute characteristics generally follow expectation in that most trips use roads with lower speed limits, fewer lanes, and lower traffic volumes. Again, this could be due to the layout of these facilities (e.g., there a fewer roads with high speed limits, many lanes, and/or high traffic volumes), but this likely also shows that cyclists have some preference for these facilities. Additionally, most travel occurred on roads (mean of 80%) rather than bicycle or pedestrian links, highlighting that at the time of the CycleAtlanta study, few of the recorded cyclists had access to separated cycling infrastructure for the majority of their travel.

Table 19 shows the descriptive statistics of the difference in route attribute composition between the chosen path and the shortest path. The shortest path attribute is subtracted from the chosen path attribute, so values that are positive indicate that the *chosen route* had a higher value. Values that are negative indicate that the *shortest route* had a higher value. Comparing the two helps demonstrate how cyclists avoid portions of the shortest route. For instance, the bike lane and multi-use path variables are positive, indicating that cyclists were often deviating from the least travel time path to access bicycle infrastructure. The speed limit, number of lanes, and the traffic volumes values are as

expected too, with negative values for higher speeds, more lanes, and higher traffic volumes.

Table 18: Descriptive statistics for chosen route characteristics

Route Characteristic	Mean	Standard Deviation	Min	25%	Median (50%)	75%	Max
Length (miles)	3.6	2	0.3	1.9	3.1	4.8	10.6
Travel Time (minutes)	24	13	2	13	21	32	70
Ascent (feet)	209	130	1	107	182	278	751
Oneway (%)	**11**	**15**	**0**	**1**	**6**	**15**	**100**
State Route (%)	**9**	**17**	**0**	**0**	**0**	**10**	**100**
Bicycle Link (%)	18	24	0	0	3	33	100
Pedestrian Link (%)	1	4	0	0	0	0	49
Road Link (%)	80	25	0	64	94	100	100
Service Link (%)	1	2	0	0	0	1	17
0 - 30 miles per hour (%)	**71**	**28**	**0**	**52**	**80**	**97**	**100**
31 - 40 miles per hour (%)	**28**	**28**	**0**	**2**	**20**	**46**	**100**
41+ miles per hour (%)	**0**	**2**	**0**	**0**	**0**	**0**	**53**
One lane per direction (%)	**67**	**32**	**0**	**44**	**76**	**96**	**100**
Two lanes per direction (%)	**27**	**27**	**0**	**3**	**19**	**43**	**100**
3+ lanes per direction (%)	**6**	**14**	**0**	**0**	**0**	**4**	**100**
0 – 4,000 AADT (%)	**50**	**30**	**0**	**27**	**51**	**74**	**100**
4,001 – 10,000 AADT (%)	**24**	**22**	**0**	**4**	**20**	**38**	**100**
10,000+ AADT (%)	**25**	**28**	**0**	**1**	**14**	**43**	**100**
[0%,4%) (%)	93	6	54	90	94	98	100
[4%, 6%) (%)	5	5	0	1	4	8	37
[6%, inf) (%)	2	2	0	0	1	2	21
Bike Lane (%)	10	15	0	0	5	15	95
Cycletrack (%)	1	4	0	0	0	0	60
Multi-Use Path (%)	16	23	0	0	1	32	100
*Eastside Beltline	9	18	0	0	0	6	100

Bolded variables only applied to road links, *Percentage of multi-use path that was the BeltLine Eastside trail

Table 19: Difference in route attributes (chosen - shortest)

Route Characteristic	Mean	Standard Deviation	Min	25%	Median (50%)	75%	Max
Length (miles)	0.6	0.9	0	0.1	0.3	0.8	8.9
Travel Time (minutes)	4	6	0	1	2	6	59
Ascent (feet)	28	66	-207	-2	10	48	552
Oneway (%)	-1	14	-100	-4	0	3	100
State Route (%)	-6	18	-88	-7	0	0	71
Bicycle Link (%)	9	20	-75	0	0	15	88
Pedestrian Link (%)	0	4	-19	0	0	0	49
Road Link (%)	-9	20	-89	-16	-1	0	75
Service Link (%)	0	2	-10	0	0	0	17
0 - 30 miles per hour (%)	4	25	-100	-5	0	16	100
31 - 40 miles per hour (%)	-4	24	-96	-16	0	5	96
41+ miles per hour (%)	0	3	-66	0	0	0	25
One lane per direction (%)	5	24	-100	-2	1	14	100
Two lanes per direction (%)	-4	22	-99	-12	0	3	100
3+ lanes per direction (%)	-2	10	-86	0	0	0	66
0 – 4,000 AADT (%)	4	26	-100	-9	0	17	100
4,001 – 10,000 AADT (%)	3	20	-100	-3	0	12	99
10,000+ AADT (%)	-7	24	-99	-19	0	2	88
[0%,4%) (%)	2	6	-23	-1	0	4	37
[4%, 6%) (%)	-1	5	-31	-3	0	1	25
[6%, inf) (%)	0	3	-19	-1	0	0	16
Bike Lane (%)	2	11	-54	0	0	4	70
Cycletrack (%)	0	4	-49	0	0	0	40
Multi-Use Path (%)	9	19	-75	0	0	14	85
*Eastside Beltline	5	14	-74	0	0	0	75

Bolded variables only applied to road links, *Percentage of multi-use path that was the BeltLine Eastside trail

CYCLEATLANTA PROCESSING SUMMARY

In this section, the process by which the CycleAtlanta data were cleaned, filtered, and map-matched was described. Trip and user characteristics of the map-matched data were discussed, and the spatial distribution of the map-matched trips was examined. The researchers plan to delve deeper into the existing map-matched data and improve the map

matching results. However, the existing map-matched traces were of high enough quality to calibrate cycling impedance factors.

CHAPTER 5. DERIVING CYCLING IMPEDANCE FROM GPS TRACES

To calculate the shortest path, BikewaySim relies on impedance functions. Impedance, expressed in minutes, represents the friction associated with travel (i.e., travelers want to minimize impedance). In typical motor vehicle simulation models, impedance incorporates travel time, parking costs, and tolls/fares. In contrast, cycling impedance incorporates the added costs, relative to the total travel time, that cyclists incur when exposed to motor vehicle traffic, elevation, or other link characteristics.

Chapter 2's literature review on route choice modeling suggests that the average marginal rate of substitution (MRS) for a unit increase in one route attribute with respect to distance could be used to create cycling impedance functions. However, given the variability in MRS within and across study areas (Fitch & Handy, 2020), differences in model specifications across study areas, the poor predictive power of these models compared to the shortest distance path (Meister et al., 2024), and the overhead required to implement these models in the travel demand planning process, an alternative method is needed.

The ideal alternative is an impedance function that is directly estimated from GPS traces. Unlike route choice models, impedance functions are directly applicable to large-scale routing services (e.g., Google Maps, Apple Maps, Valhalla, etc.) that already use custom impedance factors for routing (Dutta, 2020; Glennon, 2015). In addition, impedance functions could be integrated into existing travel demand modeling workflows.

Schweizer et al. (2016) appear to have been the first to utilize this type of approach using the impedance function (referred to as a cost function in Schweizer et al. (2016)) shown in Equation 5.

$$aa_{aa}(\beta\beta, xx_{aa}) = \beta\beta' xx_{aa} + \alpha\alpha$$

Equation 5: Link Cost Function used in Schweizer et al. (2016)

Where aa_{aa} is the cost of link gg, $\beta\beta$ is a vector of coefficients to be calibrated, xx_{aa} is a vector of link attributes, and $\alpha\,\alpha$ is an arbitrary non-zero constant that is fixed at a value of one. The values of the coefficients in $\boldsymbol{\beta}\,\boldsymbol{\beta}$ were estimated by minimizing the objective function (Equation 6).

Commented [BD8]: Formatting: missing equations

$$zz(\beta\beta) = -\frac{\sum_{\forall ii}\sum_{\forall aa\ a} LL_{aa}\delta\delta_{aaaa}(\beta\beta)}{\sum_{\forall ii}\sum_{aa\in aa_{ii}} LL_{aa}}$$

Equation 6: Objective function from Schweizer et al. (2016)

The numerator ($\sum_{\forall ii}\sum_{\forall aa} LL_{aa}\delta\delta_{aaaa}(\beta\beta)$) measures the total length (LL_{aa}) of links shared between the chosen and modeled routes for all trips (ii). The term, $\delta\delta_{aaaa}$, is one if link gg is contained in both the chosen route (gg_{ii}) and the modeled route (qq_{ii}) (i.e. $gg \in gg_{ii} \cap gg \in qq_{ii}$) and zero otherwise. The denominator ($\sum_{\forall ii}\sum_{aa\in aa_{ii}} LL_{aa}$) is the total length of all the links contained in the chosen route for all the trips. Together the numerator and denominator measure the total overlap of the modeled routes and the chosen routes. A value of zero means that the modeled routes overlapped with 0% of the chosen routes, and a value of one means that the modeled routes overlapped completely (100%) with the chosen routes. Longer trips have a greater influence on the objective function than shorter trips. The total overlap is

maximized but the objective function is minimized, so the total overlap is then multiplied by negative one.

The minimization of the objective function is not straightforward. The objective function requires a set of modeled routes, and the modeled routes are generated from performing Dijkstra's algorithm on a set of origin-destination pairs using a graph with link impedance determined by Equation 5. Dijkstra's algorithm requires non-negative link costs ($\mathbb{a}_{\mathbb{a}}(\boldsymbol{\beta}\boldsymbol{\beta}) \geq 0 \; \forall \; \mathbb{g}$), so the objective function by extension is non-continuous (i.e., it is undefined where the values of $\boldsymbol{\beta}\boldsymbol{\beta}$ result in a negative link cost).

Schweizer et al. (2016) simulated the objective function surface with two beta parameters, one for the length of a link and one for the presence of an exclusive bikeway, in Figure 31. The surface reveals a valley with multiple local minima, so Hessian and gradient-based optimization methods would fail to find the global minimum.

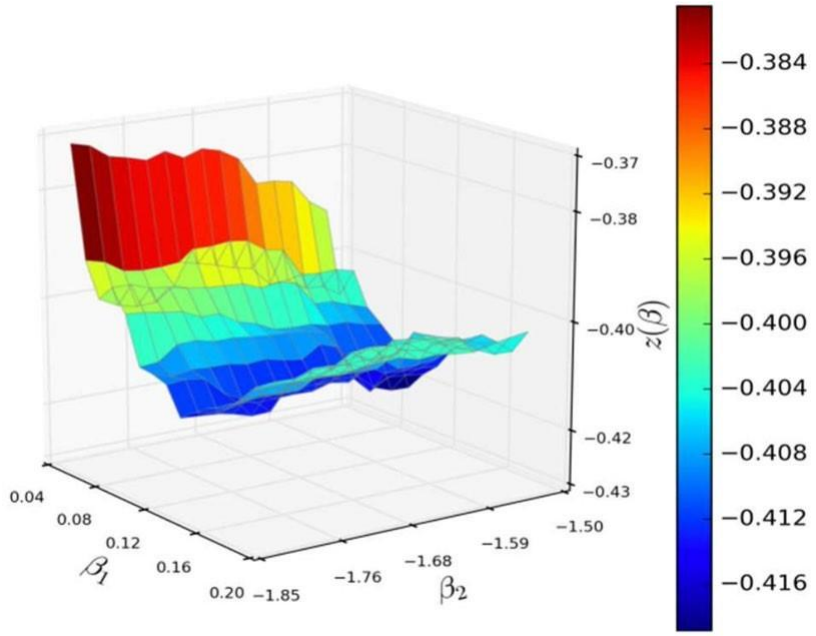

Figure 31: Illustration of the surface of the objective function using a two-dimensional beta vector (reproduced from Schweizer et al. (2016))

Schweizer et al. 2016 used bounded stochastic optimization methods to select parameters and calculate link costs. Three optimization methods, diverging evolution (DE), particle swarm optimization (PSO), and covariance matrix adaptation (CMA) were tested, and PSO was shown to perform the best. For a brief description of these methods, refer to Schweizer et al. 2016. Essentially, each method uses a population of "agents" to search the objective function surface. Each agent tests a unique vector of coefficients ($\boldsymbol{\beta\beta}$), where the coefficients are randomly selected from a bounded search space. Each agent returns an objective function value, and the $\boldsymbol{\beta\beta}$ of the agent with the best objective function value is used to adjust the selection of the next vector of coefficients for all the other agents.

Schweizer et al. 2016 used PSO to calibrate the attribute vector for several years of cycling GPS trace data collected in Bologna, Italy. Several link attributes were tested, including: "exclusive bikeways, bikeways with pedestrian access, road width, maximum allowed speed, low-priority roads, presence of traffic lights, number of entering roads at the entry node of the link and the number of incoming roads from the right side." However, the "exclusive bikeway" attribute (i.e., cycletracks and separated bicycle facilities) was the only attribute that provided a noticeable overlap increase compared to the shortest distance route prediction. Additionally, Schweizer et al. 2016 found that attribute vectors above two dimensions did not improve the overlap and in some cases reduced it.

Commented [BCD9]: What is an "exclusive bikeway"? Does that include unprotected bike lanes or is it only protected/separated facilities?

Ultimately, they settled on two attribute vectors. The first, shown in Equation 7, incorporates link length and the presence of an exclusive bikeway. The second, shown in Equation 8, incorporates link length and link length interacted with the presence of an exclusive bikeway.

$$\underset{aa}{xx^{CC}} = [\underset{aa}{LL}, \underset{aa}{BB}]'$$

Equation 7: Attribute vector with link length and exclusive bikeway presence

Commented [BD10]: Missing equations

$$\underset{aa}{xx^{PP}} = [\underset{aa}{LL}, \underset{aa}{LL}\,\underset{aa}{BB}]'$$

Equation 8: Attribute vector with link length and link length interacted with exclusive bikeway presence

The improvement in overlap from the shortest path prediction ranged from 15-31% where the overlap value ranged from 0.32 to 0.38 for the shortest path and 0.39 to 0.47 for the modeled routes. Equation 8 generally outperformed Equation 9. The improvement in

prediction from the shortest path was larger for longer trips than shorter trips (< 0.3 miles).

Two limitations of this approach were the lack of consideration for turn attributes and the limited number of attributes in the link cost function. This research sought to address these gaps as well as build upon this work. The main contributions of this research are in adding consideration of turn variables; repeating the methodology for a city in the U.S. with limited cycling infrastructure, fewer cyclists, and different built environment features than an Italian city; and lastly, applying the calibrated impedance factors to assess the impacts of new cycling facilities.

IMPEDANCE FUNCTIONS

Impedance functions are used to find the user-preferred route between any two points. The impedance functions minimize cyclists' exposure to high-speed motor vehicle traffic, delay at intersections, and/or steep hills, while at the same time not deviating too far in distance and time from the route with the least travel time. Consider the two routes presented in Figure 32. The route in black minimizes travel time while the purple route minimizes impedance based on link attributes. The black route proceeds along a four-lane arterial road with no cycling infrastructure for most of the way (Figure 33b). Meanwhile, the purple route has a lower impedance due to the presence of cycling infrastructure, namely the Atlanta BeltLine (Figure 33a), a multi-use path with no motor vehicle traffic. Hence, many users will prefer to take the longer, but safer, less stressful. and more enjoyable route.

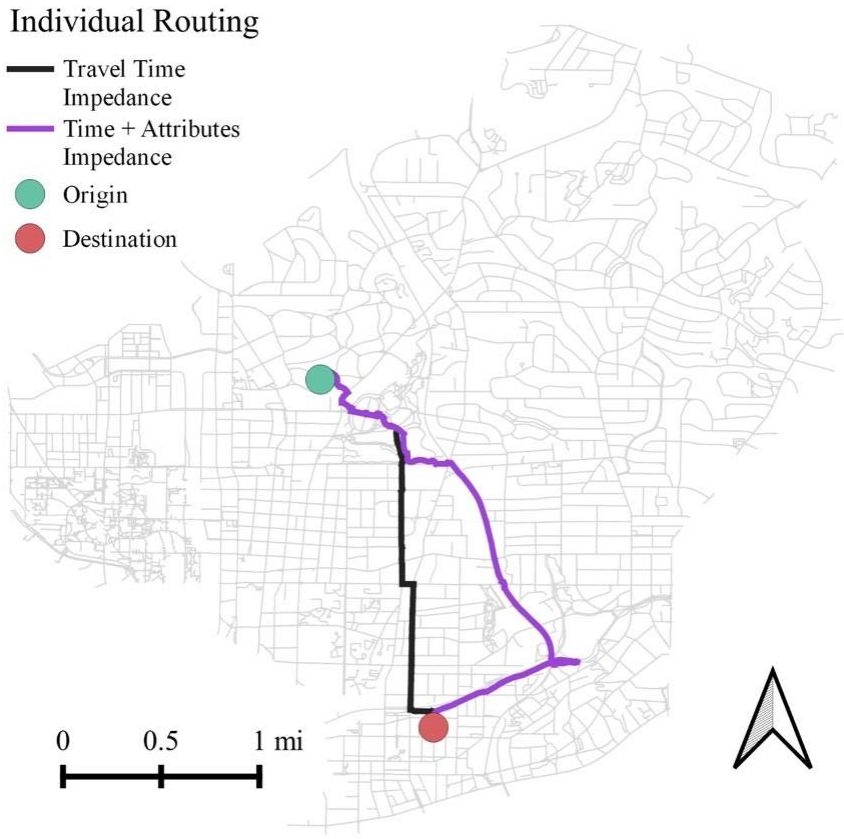

Figure 32: Example of Cycling Impedance Routing (purple) vs. Shortest Path Routing (black)

Figure 33: a) Boulevard b) Atlanta BeltLine

In the above scenario, the impedance function was able to find an alternative route that utilized existing cycling infrastructure. However, this is not always the case. In areas where there is little existing cycling infrastructure, the route with the lowest impedance may not actually be a "low impedance" cycling route. This will be reflected in the calculated route impedance which can be used to tell if the route is high in impedance relative to route against routes of similar travel time that do utilize cycling infrastructure.

It is important to keep in mind that link impedance does not consider all the mode choice factors that go into the decision to cycle or consider the likelihood of a cyclist cycling on a specific link. Impedance is a measure of the relative cost associated with cycling on a link, and the reduction of impedance can be used as a leading indicator for future increases in cycling activity.

In addition, the nature of shortest path algorithms and the method used to calibrate the impedance functions makes these impedance functions suitable for only destination-oriented trips. This excludes trips (usually, exercise and recreational trips) where the origin and destination are the same or there is not a particular destination in mind.

PROPOSED LINK AND TURN IMPEDANCE FUNCTIONS

The proposed link impedance function is shown in Equation 9. Where $aa_{a\,a}$ is the link impedance of link gg, $ll_{a\,a}$ is the length of link gg, $S\,S$ is a cyclist's assumed average speed, ii. $\boldsymbol{\beta\beta}_{uuuuu}$ is the link attribute coefficient vector, and $\boldsymbol{xx}_{aa}$ is the attribute value vector for link

Commented [BD11]: Formatting: missing equations

This equation is modified from Schweizer et al. (2016) in that all link attributes are considered to lead to a percent increase or decrease in the link's travel time. This is similar in concept to the percent change in distance metric (Equation 1) used in the

cycling route choice literature. Percent change in travel time implies that the attribute is experienced along the entire link.

$$aa(\boldsymbol{\beta\beta} \quad , xx) = \frac{ll_{aa}}{SS} * (1 + \beta\beta' \quad xx)$$

Equation 9: Proposed link impedance function

The proposed turn function is shown in Equation 10. Where $ss_{aa,aa}$ is the turn cost from link gg to link bb, $\boldsymbol{\beta\beta}_{ttttttll}$ is the turn attribute coefficient vector, and $\boldsymbol{xx}_{\boldsymbol{aa,bb}}$ is the attribute vector for the turn from link gg to link bb. While Lowry et al. (2016) assigned a base impedance to turns, the proposed function assumes that there is no base turn cost. Instead, the turn cost itself is calibrated which means that the $\beta\beta'_{ttttttll}xx_{aa,bb}$ term represents the travel time increase in minutes. The travel time increase implies that turn attributes are an event impedance.

$$ss_{aa,aa}(\beta\beta_{ttttttll}, xx_{aa,bb}) = \beta\beta'_{ttttttll}xx_{aa,bb}$$

Equation 10: Proposed turn impedance function

The objective function, shown in Equation 11, has also been slightly modified from Schweizer et al. (2016). Where $LL_{aa_{ii} \cap qq_{ii}}$ is the length of all the links shared by the chosen route (gg_{ii}) and the modeled route (qq_{ii}) and $LL_{aa_{ii} \cup qq_{ii}}$ is the length of all the unique links in the chosen route (gg_{ii}) and the modeled route (qq_{ii}). The ss term represent the total number of trips used in the calibration process, so the objective function represents the average overlap between the chosen and modeled routes. Unlike Equation 6, Equation 11 does not place more emphasis on longer trips.

$$zz(\boldsymbol{\beta\beta}\quad,\boldsymbol{\beta\beta}\quad) = -\frac{1}{\_}\boxed{?}\ \frac{LL_{aa_{ii}\cap qq_{ii}}}{\underline{\qquad\qquad}}$$

$$llllllll\quad tttttttll\qquad SS\qquad \forall_{ll} LL_{aa\ \cup qq}$$
$$ii\quad ii$$

Equation 11: Objective function considering the overlap of the modeled routes and chosen routes

IMPEDANCE CALIBRATION

The impedance calibration process follows the particle swarm optimization (PSO) approach demonstrated in Schweizer et al. 2016. The "stochopy" package was used to perform PSO (Luu, 2023). The population size (i.e., the number of particles searching for a solution) was set to five. The network used for map matching the CycleAtlanta data was also used for impedance calibration except this time it also contained the turns in the pseudo dual graph structure outlined in Chapter 3. The assumed average cycling speed, SS, in Equation 9 was nine miles per hour. All 1,792 map-matched trips across the 577 users were used to calibrate the link and turn impedance functions.

RESULTS

Several combinations of link and turn attributes were tested for the impedance calibration process. At the time this report was published, the best-performing combination of attributes was the presence of a multi-use path, the presence of a bike lane, the number of lanes, and whether the link's average grade was above 4% for the link impedance function and the presence of an unsignalized turn across a major road for the turn impedance function.

The calibration process took about three hours on a machine with a 10-core processor and 32 Gb of memory. Because new impedance factors are generated for every round of the optimization process, it is difficult to use heuristics to speed up the routing process.

The average overlap using the calibrated impedance functions was 30.8%, meaning that, on average, the modeled and matched route shared 30.8% of the link by length. Meanwhile, the average overlap using the shortest travel time impedance was 30.2% (a 2% increase). The values of the parameters used in the impedance functions are shown in Table 20.

Table 20: Calibrated Beta Parameters

Type	Parameter	Search Bounds	Value	Units
Link	Multi-Use Path	[-1,0]	-0.184	Percent Time
Link	Bike Lane	[-1,0]	-0.398	Percent Time
Link	Lanes	[0,4]	0.126	Percent Time
Link	> 4% Mean Grade	[0,4]	0.325	Percent Time
Turn	Unsignalized Crossing of Major Road	[0,4]	0.324	Minutes

Commented [BCD12]: How did vehicular speed and volume perform as parameters compared to number of lanes? I'm surprised they weren't high performing.

Commented [TP13R12]: At the time of the report they were left out because they were highly correlated with lanes. After testing them, speed and volume did perform worse

Commented [TP14R12]: Add section on additional variables tested

Commented [BD15R12]: I don't see the new section for additional variables, am I missing it?

DISCUSSION

The signs of the parameters in Table 20 are as expected, and it doesn't appear that the calibrated values approached the search bounds. These search bounds were selected based on the feasible minimum and maximum increase or decrease in impedance from an attribute. For instance, the lanes variable was expected to have an impedance coefficient above zero because this variable had been shown to decrease route choice probability in the cycling route choice literature. The maximum search bound of four (400% time increase) was selected because any coefficient above that would likely be unrealistic.

The increase in overlap from the shortest travel time path was minimal. For this report, the values of the multi-use path and bike lane parameters will allow us to demonstrate the BikewaySim framework for assessing new cycling infrastructure. To increase the overlap

further, the research team plans on testing more impedance functions as well as segmenting the data to see if there are multiple impedance functions required.

The distribution of overlap values amongst the trips is shown using the shortest travel time impedance (grey) and the calibrated impedance (purple) in Figure 34. While the calibrated impedance factors have shifted the distribution slightly, there are still many trips that are not being predicted using the impedance functions. Additionally, some trips are well explained by the shortest travel time impedance but are not well explained by the calibrated impedance.

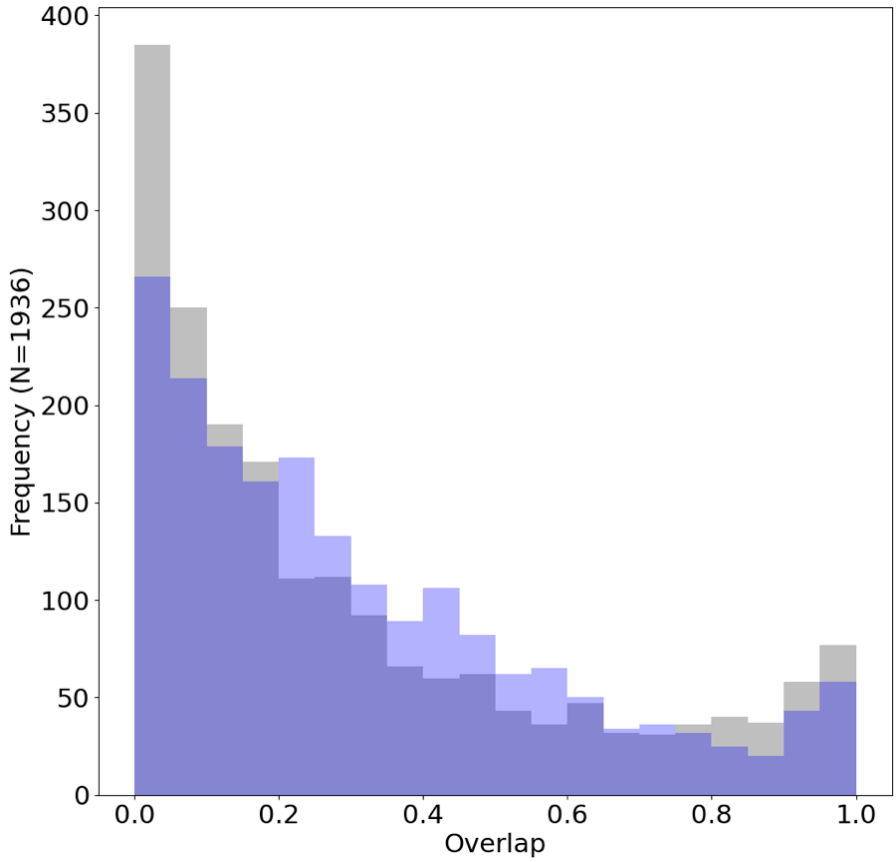

Figure 34: Distribution of overlap values

Another way to examine these parameters is to look at the possible combinations of link attributes to see how they affect the total percent time change as in Table 21. The range of total percent time change values reflects the flexibility of the link impedance functions in altering impedance across the links. In this case, that range was from -27% to 58%. While the bike lane coefficient was more negative than the multi-use path coefficient, the lane variable made them more comparable (-0.185 and -0.272).

Table 21: Possible combinations of link attributes and their impact on the link percent time change

	Multi-Use Path	Bike Lane	Lanes	> 4% Mean Grade	Total Percent Time Change
Pedestrian Path (base)	0	0	0	0	0
				1	0.325
Multi-Use Path	1	0	0	1	0.141
				0	-0.184
Road	0	1	1	1	0.053
			2		0.179
			3		0.305
			1	0	-0.272
			2		-0.146
			3		-0.020
		0	1	1	0.451
			2		0.577
			3		0.703
			1	0	0.126
			2		0.252
			3		0.378

Commented [BCD16]: I'm surprised that bike lanes on a four-lane road were comparable to a multi-use path and outperformed a path on a two-lane road. What's driving this preference? While more confident and experienced riders may prefer the bike lanes, most users would strongly prefer the path.

One advantage of using impedance functions is that attribute coefficients can be applied to the network to visualize their impact on link impedance. Figure 35 depicts the percent change in travel time due to the link attributes, calculated using the calibrated impedance factors. Figure 35 also shows the locations of unsignalized crossings of major roads. These crossings are along the roads with higher impedance.

Most road links had a percent increase in travel time, as few roads had cycling infrastructure. The links with the highest increase in impedance are multi-lane arterials, and cyclists will avoid them if possible. Road links with cycling infrastructure still increased in impedance, but the presence of cycling infrastructure mitigated the impacts of stressful road attributes. The multi-use path links all decreased in impedance because their number of lanes attribute equaled zero. This reflects the attractiveness of these facilities to cyclists as they offer separation from motor vehicles. There were also a few links that saw no change in impedance because they reflected the base case.

In this report, the base case was a pedestrian path that was not explicitly designated as a multi-use trail for cycling. These pedestrian paths included the paths within Piedmont Park and the college campuses in the study area. These pedestrian paths tended to provide internal connectivity (i.e. within the park or college campus); multi-use trails such as the BeltLine and Stone Mountain PATH provided cross study area connectivity (i.e., the trails spanned multiple blocks and connected to many streets). In the future, the effect of pedestrian volume (e.g., pedestrians, strollers, toddlers, dogs-on-leash, etc.) on a pedestrian path or multi-use trail should be considered, particularly on the BeltLine.

Commented [BCD17]: Why was this used as the base case?

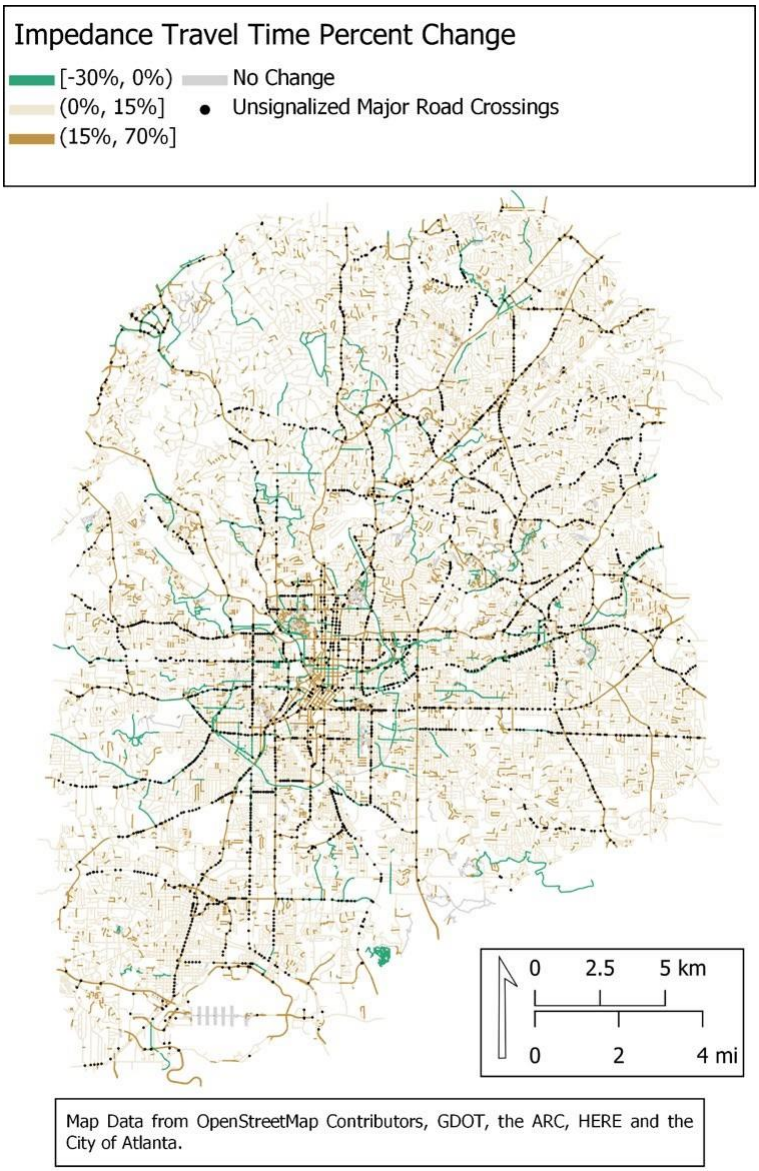

Figure 35: Map showing the impedance travel time percent change and unsignalized major road crossings

IMPEDANCE CALIBRATION SUMMARY

Although there was only a slight improvement in overlap from the shortest travel time path, the calibrated impedance factors can be used to demonstrate BikewaySim's ability to assess new bicycle facilities.

The next step for the researchers is to investigate the trips that have not been well explained using the calibrated impedance function. Just as Fitch and Handy (2020) demonstrated the wide preferences that users had for route attributes, it is unlikely that one impedance function will be able to explain the routing of every trip. As such, the trips will likely need to be segmented before calibration. Also, the researchers will continue to experiment with more configurations of the current impedance functions or test different impedance functions entirely.

CHAPTER 6. BIKEWAYSIM DEMONSTRATION FOR THE STUDY AREA

This chapter demonstrates BikewaySim's ability to assess the impacts of new bicycle facilities. Using the initial impedance functions developed in the previous chapter, 38 planned bicycle facilities in the City of Atlanta are assessed using 3.6 million trips across 127,000 origin-destination pairs from the Atlanta Regional Commission's TIP Amendment Six Activity Based Model run (Atlanta Regional Commission, n.d.). Metrics on trip impedance reduction and the additional minutes of travel due to impedance are visualized at the traffic analysis zone (TAZ) level to communicate the broad impacts of the planned cycling facilities. Link betweenness centrality is used to visualize network flow changes as a result of impedance and the planned bicycle facilities. The impedance reduction impact of each planned bicycle facility is calculated, and the planned bicycle facilities are then ranked in order of increasing impedance reduction per mile. An additional visualization, a bikeshed, is also presented to show the localized impact of cycling impedance and new bicycle facilities.

The results are discussed, and recommendations are provided as to how planners and engineers can best use the results from BikewaySim in the bicycle facility planning process. Content from this chapter has been expanded on from the material originally published in Passmore et al. (2024).

The workflow for assessing new bicycle facilities using BikewaySim is shown in Figure 36. Starting from "OD Pairs," the first step is to provide origin-destination (OD) pairs and match each origin and destination to the nearest network node. Each origin and destination should include a unique identifier that relates to coordinates, traffic analysis

zones (TAZs), or census blocks. The OD pairs should also indicate the number of trips between the pairs for weighting the impedance reductions, but when those data are not available, a scenario considering all potential origin-destination pairs can be run (Passmore et al., 2024).

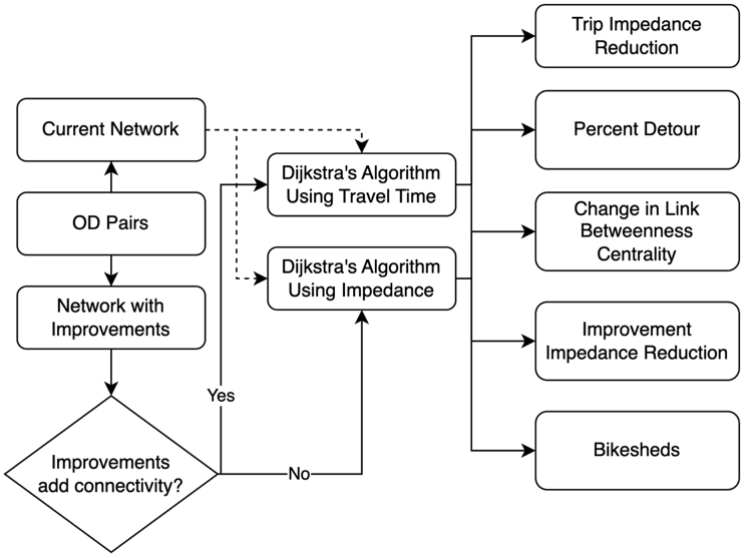

Figure 36: BikewaySim assessment framework workflow

Origins and destinations are matched to the nearest network node via Euclidean (i.e., as the crow flies) distance. The Euclidean distance is also used to calculate the access time from the origin to the network and the egress time from the network to the destination. Access and egress travel times are calculated assuming a walking speed of 2.5 miles per hour. However, access/egress times were not incorporated into the overall impedance calculation for these analyses as the focus was on quantifying impedance reduction from bicycle facilities.

If the input origins and destinations represent polygons or zones (e.g., census blocks, TAZs), the centroid is used to find the nearest network node and calculate access and egress travel times. In sparsely populated areas with geographically large TAZs or census blocks, the centroid may not be located in a plausible location (e.g., centroid located in a body of water) or near a network node. In these cases, the trips into or out of a TAZ could be assigned to parcels or building footprints within the TAZ. This would increase the number of origins and destinations to evaluate in the shortest path routing step, which would in turn increase computation time.

For this project, the OD pairs were from the Atlanta Regional Commission's TIP Amendment Six Activity-Based Model run for the year 2030, where each origin and destination corresponded to a traffic analysis zone (TAZ), where the model includes more than 5,900 TAZs (Atlanta Regional Commission, n.d.). The OD data represent individual complete tours (i.e., all trips taken throughout the day), so there are multiple trips per person. All trip purposes and travel modes were considered.

These data were filtered to only include trips that were within the study area, less than or equal to three miles in length and non-intrazonal (i.e., trips went from one TAZ to a different TAZ). TAZ centroids that had a Euclidean distance to the nearest network node greater than 500 feet (approximately 2.5 minutes of walking at 2.5 mph) were also removed. Table 22 shows the results of each filtering step. In the end, the trimmed data set included 1,135 TAZs, 127,682 unique OD pairs and 3,579,964 trips taken by 653,522 people across 483,576 households.

Table 22: ARC Trip Data Filtering Process

Step	Number of TAZs	Number of OD Pairs	Number of Trips	Number of Households	Number of Persons
Initial count	5,922	4,372,623	59,381,744	2,481,224	5,427,425
Within study area	1,327	743,987	9,476,763	784,992	1,187,904
Matching distance less than or equal to 500 feet	1,139	565,395	7,342,741	672,054	988,552
Trip less than or equal to 3 miles	1,135	127,682	4,138,944	517,835	708,466
Non-intrazonal	1,135	126,641	3,579,964	483,576	653,522

The distribution of trip origins (and by extension destinations) is important, as many of the output metrics are weighted by the number of trips. This means that bicycle facilities that connect OD pairs with a higher number of trips will have a greater contribution to reducing trip impedance. As shown in Figure 37, the average number of trips originating from a TAZ was 3,154, with a range of seven to 48,090 trip origins. The skewed distribution makes sense, because many TAZs in the study area represent low-density, single-family zoned residential areas that would have relatively few origins.

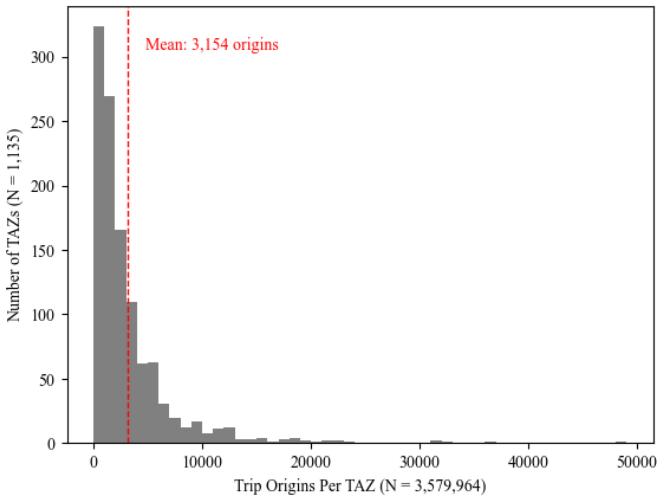

Figure 37: Number of trips origins per TAZ

Figure 38 shows the spatial distribution of the trip origins per TAZ within the study area. Trip origins per TAZ were binned into five levels for symbolization using natural breaks. TAZs within the study that were filtered out are shown in grey. Labels were added to contextualize the study area. As suspected, many of the TAZs with a low number of trip origins were the outlying low-density, single-family zoned TAZs. The few trips with more than 10,000 origins mostly corresponded to Downtown Atlanta, Midtown, Buckhead, Georgia Tech, Emory University, and Hartsfield Jackson International Airport. Bicycle facilities that connect to or from these TAZs have a greater potential to reduce impedance.

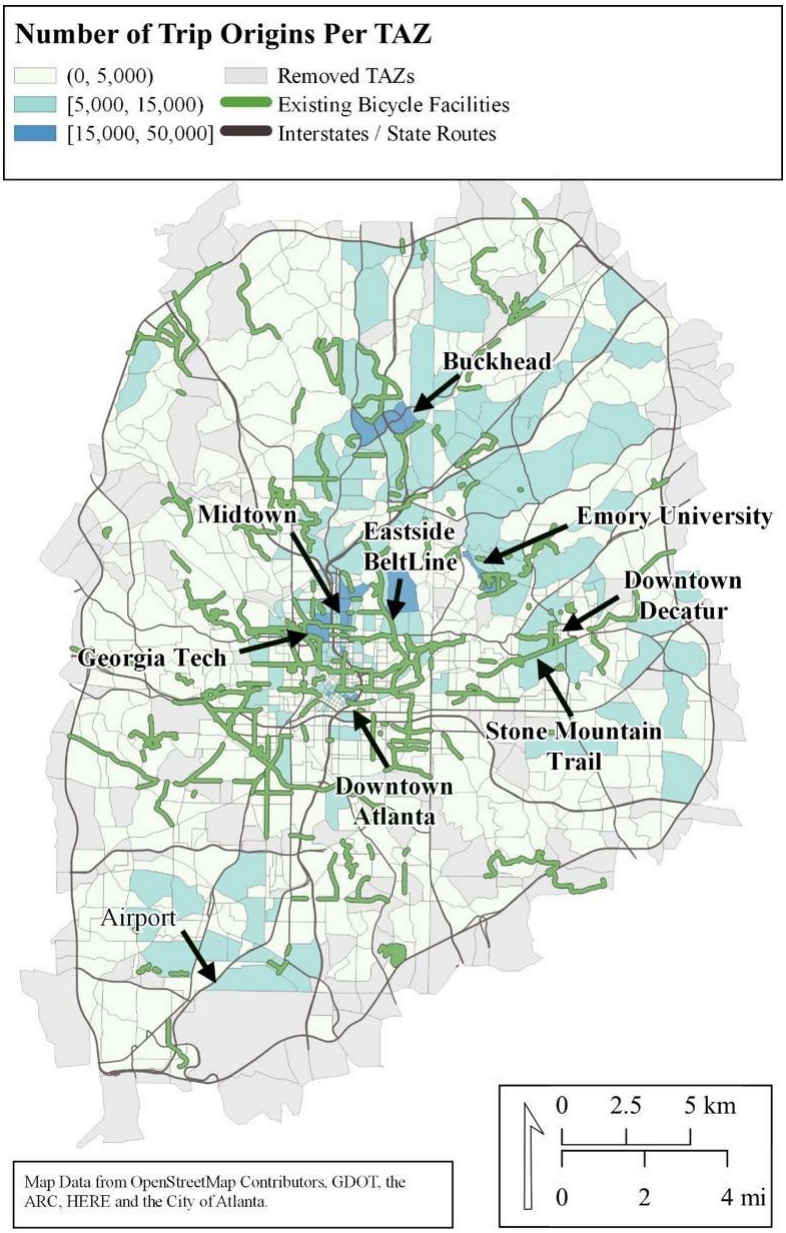

Figure 38: Number of trip origins per TAZ

In Figure 36, the next step is to prepare the “Current Network” and the “Future Network.” The current network represents the street network as it is now, while the future network contains proposed improvements (i.e., bicycle facilities in this case). Improvements must either add new connectivity (e.g., adding links and/or new nodes to the network) or relate to a potential change in one of the attributes in the link or turn impedance functions used for BikewaySim impedance. At the time of this report, the impedance attributes of interest for testing include the presence of a multi-use path, the presence of a bike lane, the number of vehicle lanes per direction, the presence of a road grade above 4%, and the presence of an unsignalized crossing across a major road.

The network data used for the study area were the same reconciled OSM network used for map matching in Chapter 4 and impedance calibration in Chapter 5 with 76,912 links (144,458 directed links), 66,151 nodes, and 222,764 link-to-link turn opportunities in the pseudo dual graph.

Improvements can be provided and evaluated all at once, or one at a time, depending on the intended application. Evaluating the improvements all at once demonstrates how the interactions between the improvements yield a combined reduction in impedance. It also demonstrates the relative impact of each improvement within that combined set of improvements. This aids in evaluating cycling network master plans where the projects within a set of investments need to be ranked against each other, but not as if they were being implemented alone (without any other connecting projects, which is important from a systems perspective).

Evaluating the improvements one at a time, on the other hand, still produces useful metrics and visuals that demonstrate the independent utility of each project. This can be

useful in evaluating options within bicycle facility project alternatives, such as deciding what level of separation from traffic should be provided.

Currently, improvements are added by editing the network link or turn attributes, or by manually drawing in new links or nodes using GIS software. For this project, only alternatives that improved existing links were considered, and these improvements were modeled by editing the network link attributes. No new links were added to the future network for these model runs.

The network improvements for this project are funded bicycle lanes, multi-use paths, and cycletracks from the City of Atlanta's bicycle facility inventory. Cycletracks were treated as multi-use paths to correspond with the calibrated "presence of a multi-use path" impedance. This also meant that the "lanes" attribute was set to zero for links that were converted to cycletracks. Once the improvements were added to the links in the network, link and turn impedance factors were re-calculated. Figure 39 shows where these improvements were located and to what extent the improvements reduced link impedance. For clarity, impedance reductions are shown in terms of the decrease in percent change in travel time rather than in minutes. Also, when the impact of the improvement was different between the forward and reverse links, the value with the highest impedance reduction was shown.

In Figure 39, the link impedance reductions from the proposed bicycle lanes (orange) and multi-use paths/cycletracks (brown) are categorized using natural breaks. The range of impedance reduction was 4% to 56%. For the most part, the multi-use paths/cycletracks near the center of the map that connects Midtown and Downtown Atlanta as well as the one that is adjacent to State Route 13 (Lee Street Southwest) had the highest link

impedance reduction. This makes sense as these segments have higher values for the "lanes" attribute.

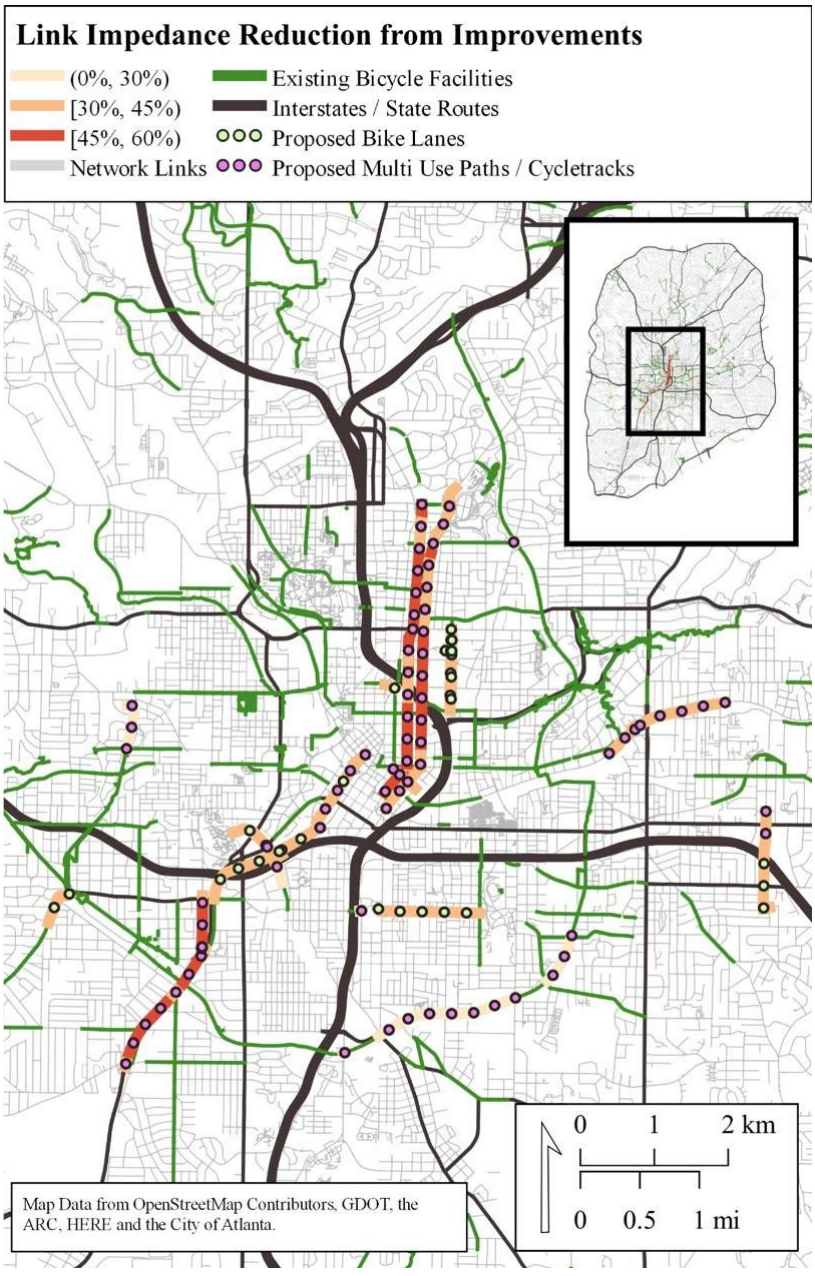

Figure 39: Link impedance reduction from improvements

The BikewaySim Dijkstra's algorithm approach was used to find the shortest paths between all the unique origin-destination pairs in the current and future networks (Figure 36). For the current network, the shortest paths were calculated twice. Once using the link travel time, and once using the calibrated link and turn impedance functions. If there had been improvements that added new connectivity, then the shortest paths would also be calculated twice for the future network. Without new connectivity in the future network, the travel time shortest paths are identical between the current and future networks.

The shortest path calculations are the most computationally intensive step in the workflow, and for 126,641 unique OD pairs, each set of shortest path calculations (one for the current network using travel time impedance, one for the current network using the calibrated impedance factors, etc.) took about three hours to run on a computer with a ten-core processor and 32 Gb of RAM. The computation time will vary with network size and the number of unique origin-destination pairs.

The shortest path results are processed to calculate the trip impedance for each OD pair, the travel time for each OD pair, and the link betweenness centrality for each link in both the current and future networks. From these results, five different assessments (shown on the rightmost side of Figure 36) were used to assess the impact of the improvements.

TRIP IMPEDANCE REDUCTION

Trip impedance measures the "perceived travel time" of a trip, as it factors in cyclists' preferences for attributes. Trip impedance (CC_{aa}) is calculated for trip (gg) by summing all the link impedance values (aa_{aa}) and turn impedance values ($ss_{aa,aa}$) along a trip's route as in Equation 12.

$$CC_{aa} = \sum_{cc_{aa} \in aa} aa_{aa} + \sum_{aa_{aa,bb} \in aa} ss_{aa,aa}$$

Equation 12: Trip Impedance Function

This metric is most effectively used when calculating the difference in trip impedance after adding network improvements. Impedance reduction is calculated for each OD pair by subtracting the improved network trip impedance from the current network trip impedance. This reduction is weighted by the number of trips for an OD pair (if provided), so the impact of reducing the impedance by a small amount for an OD pair with many trips can, depending on the circumstances, be comparable to reducing the impedance by a large amount for an OD pair with much fewer trips.

There were 576 OD pairs that saw no reduction in impedance. Of the OD pairs that did experience a reduction in impedance, the weighted average impedance reduction from the current network to the improved one was 0.75 minutes and the maximum impedance reduction was 23.7 minutes.

Figure 40 shows the average weighted impedance reduction per origin TAZ symbolized using natural breaks. The map has been zoomed in to focus on the TAZs where the weighted average impedance was reduced. The proposed bike lanes (orange), proposed multi-use paths/cycletracks (brown), existing cycling facilities (green), and Interstates and state routes (black) are shown for reference. The darker grey TAZs saw no reduction in impedance and the light grey TAZs had been filtered out.

As shown in Figure 40, the TAZs with the highest weighted average impedance reduction (between 0.5 and 2.0 minutes) were close to multiple proposed bicycle facilities that also

tended to be longer. In particular, the TAZs near the proposed bicycling facilities connecting Downtown Atlanta to Midtown Atlanta and along State Route 13 (Lee Street Southwest) saw the largest reductions in weighted average impedance.

The weighted average impedance reduction begins to drop for improvements made further away from the proposed bicycle facilities. But this decline in impedance reduction is not always symmetric around the proposed bicycle facilities, due to either the underlying network connectivity or the OD flows. An OD pair will see a reduction in impedance only when the shortest path based on calibrated impedance utilizes the improvements. If an origin or destination TAZ isn't located near an improvement, and/or there are only a few trips originating from or going to that TAZ, it is unlikely that trips will use the facility improvements, and unlikely for the TAZ to experience any significant reduction in weighted average impedance.

Overall, visualizing the weighted average impedance reduction by origin TAZ helps convey the extent to which proposed bicycle facilities reduce the impedance of cycling, even when the bicycle facility is not necessarily near the origin TAZ.

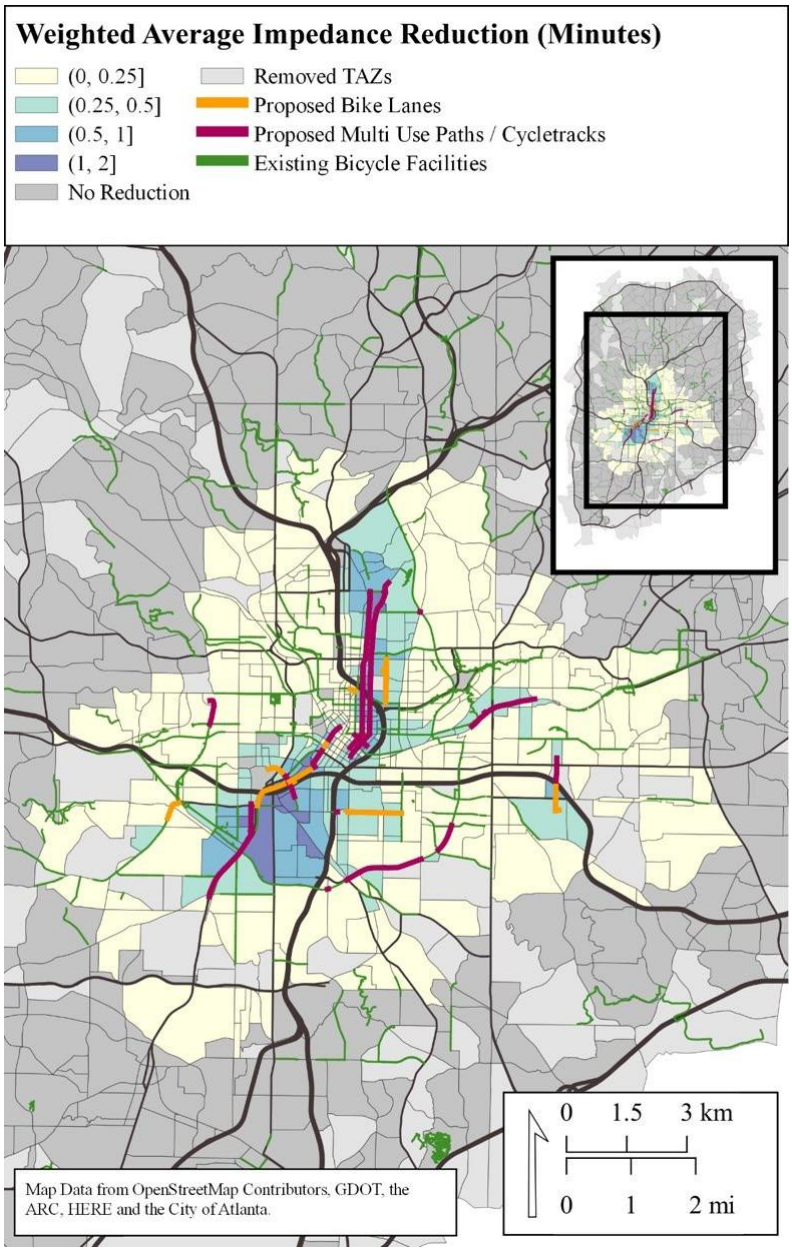

Figure 40: Average impedance reduction per TAZ

PERCENT DETOUR

While BikewaySim primarily relies on cycling impedance to calculate the shortest path between ODs, it is also important to monitor the travel time increase that was traded-off against other impedance reductions. This tradeoff can be assessed using a percent detour metric. When bicycle facilities are sparse, cyclists may have to go substantially out of their way to access them. As more bicycle facilities are provided, cyclists don't have to detour as much.

The travel time change between the shortest travel time route and the impedance route is an important metric to monitor for cycling network planning as it measures the directness of trips in the context of the current and/or future network conditions. To account for the variation in trip distance, travel time change is converted to percent detour, which measures the percent increase in travel time from the shortest travel time route to the least impedance route. Percent detour for a trip is calculated using Equation 13, where $TTTT_{aammppaaddaappccaa}$ represents the travel time for the route calculated using the calibrated impedance factors and $TTTT_{ddhppaaaaaaddaa\ aaaaaattaam\ aaummaa}$ represents the travel time for the route calculated using the travel time impedance.

$$\%_{ggeessttddgg} = \frac{TTTT_{aammppaaddaappccaa} - TTTT_{ddhppaaaaaaddaa\ aaaaaattaam\ aaummaa}}{TTTT_{ddhppaaaaaaddaa\ aaaaaattaam\ aaummaa}} * 100$$

Equation 13: Percent detour

The closer the travel time of the route calculated using the calibrated impedance functions is to the travel time of the route calculated using the travel time impedance (i.e., closer to 0% detour), the better. At 0% detour, the preferred route, calculated using the

calibrated impedance functions, is also the most direct route, which is a desirable property for effective cycling networks (CROW, 2009).

The trip percent detour ranged from 0% to 43% with an average of 1.8%. The average is percent detour is much smaller than found in typical literature values (8-18%). The distribution of trip percent detour values is shown in Figure 41. From this, it seems that a relatively small number of trips had to detour. In past research, this study area has been shown to have average detour rates as high as 30% (Misra, 2016). The low average detour rate is likely, in part, due to the calibrated impedance functions not capturing the detour present in the GPS trace data used for calibration. The research team is currently working on improving the calibration results.

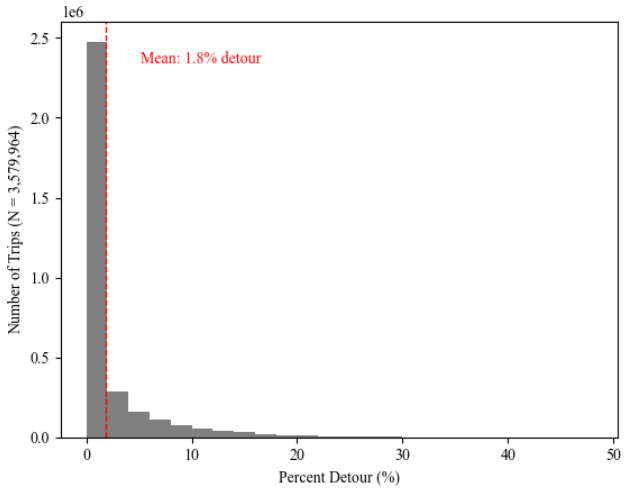

Figure 41: Histogram of trip percent detour

Figure 42 shows the weighted average percent detour per origin TAZ symbolized using natural breaks for the current network. From this figure, it appears that the weighted

average percent detour was highest for areas near Midtown and Downtown Atlanta. Part of this could be due to trips from these TAZs being shorter, meaning that relatively small detours could result in a comparatively higher percent detour. The weighted average percent detour is lowest in the outlying TAZs, in part because the trips starting from these TAZs are longer.

Another observation is that the TAZs near bicycle facilities generally appear to have a higher weighted average percent detour. It should be noted that in areas with homogenous link attributes, the calibrated impedance route is similar to that of the least travel time route. This results in a lower percent detour but a comparatively higher impedance if no bicycle facilities were available.

Another metric to monitor is the difference in trip percent detour between the current and future networks. While not calculated for this report, the difference in percent detour measures whether the added bicycle facilities in the future network increased or reduced the amount of impedance-based detour onto longer paths. In cases where the percent detour decreased, trips were able to become more direct due to the improvements that were made to the most direct links. However, in other cases the percent detour increased, meaning that improvements along another route created a more pleasant route that users were willing to further divert from the shortest path to enjoy. An increase in percent detour isn't necessarily bad, but it shows that adding network improvements that provide circuitous connections may provide tradeoffs. To both reduce the impedance and the percent detour, network improvements need to be placed on links that are direct. The next metric, link betweenness centrality, provides a way to measure to identify these "direct links."

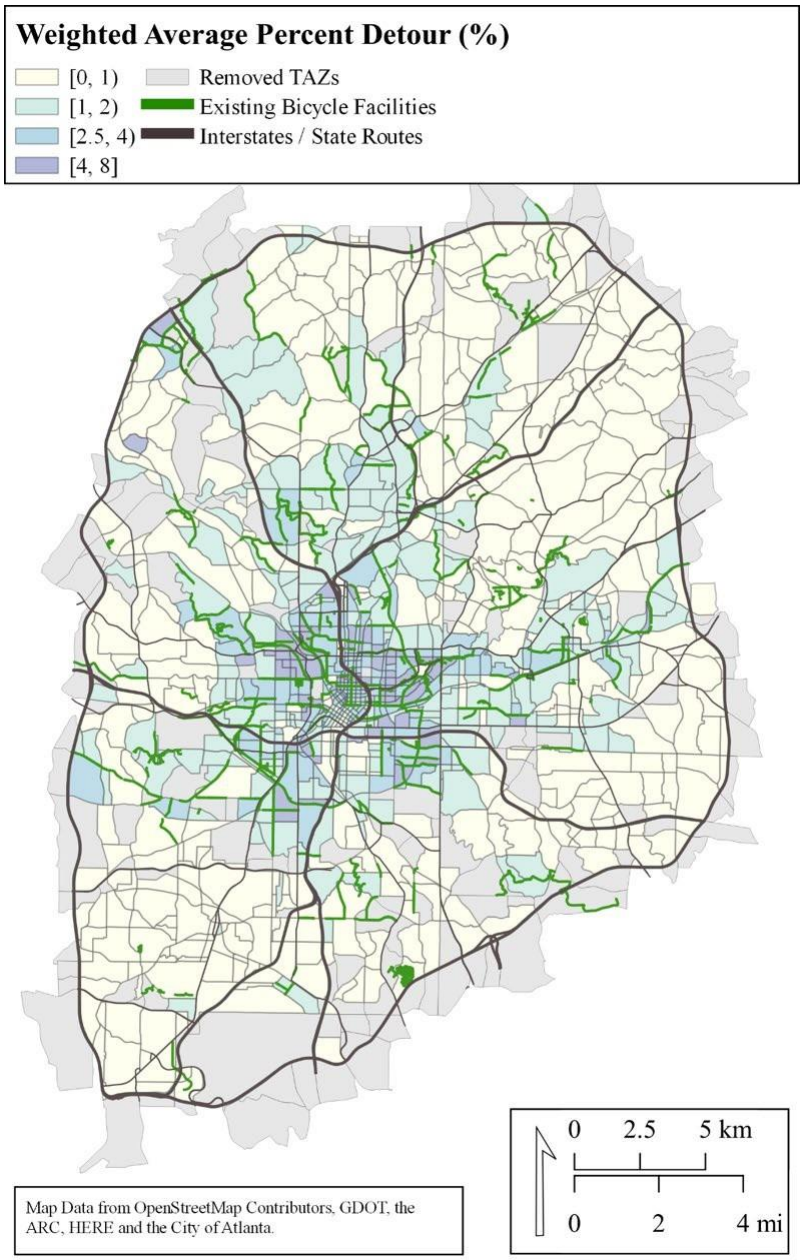

Figure 42: Weighted average percent detour per origin TAZ

CHANGE IN LINK BETWEENNESS CENTRALITY

Next, change in link betweenness centrality (LBC) was calculated. A link's LBC represents the number of trips that utilized that link as part of the trips' calculated shortest route. A high LBC signifies that a link is a crucial connection.

The map on the left in Figure 43 shows the travel time LBC on the current network (i.e. the LBC when using travel time as the impedance in Dijkstra's algorithm). Here, the links with the highest LBC are the most direct links for the most trips. The streets in Midtown and Downtown Atlanta appear to carry the most trips, but there are other clusters of links with high LBCs in Buckhead and near Emory. However, some of the links with a high LBC may possess undesirable attributes for cycling. The map on the right in Figure 43 shows the calibrated impedance LBC on the current network. The links with the highest LBC are generally the least impedance links for the most trips.

To highlight where LBC changed, the travel time LBC is subtracted from the calibrated impedance LBC (i.e. the LBC when using calibrated impedance as the impedance in Dijkstra's algorithm) for the current network. This difference is visualized in Figure 44.

The change from travel time LBC to calibrated impedance LBC can be used as a measure of the directness of the current or future cycling network. A decrease in LBC (orange links) indicates that fewer trips utilized that link because it and other links around it had a higher impedance than before. An increase in LBC (purple links) indicates that the link's attributes and the attributes of surrounding links make it worth the detour. In other words, an increase in LBC indicates increased importance while a decrease indicates reduced importance. Like percent detour, when the difference in link betweenness centrality is close to zero, the most direct links also possess preferable attributes. The existing bicycle

facilities are shown for reference as many of the links that saw an increase in LBC were bicycle facilities or roads connecting to bicycle facilities.

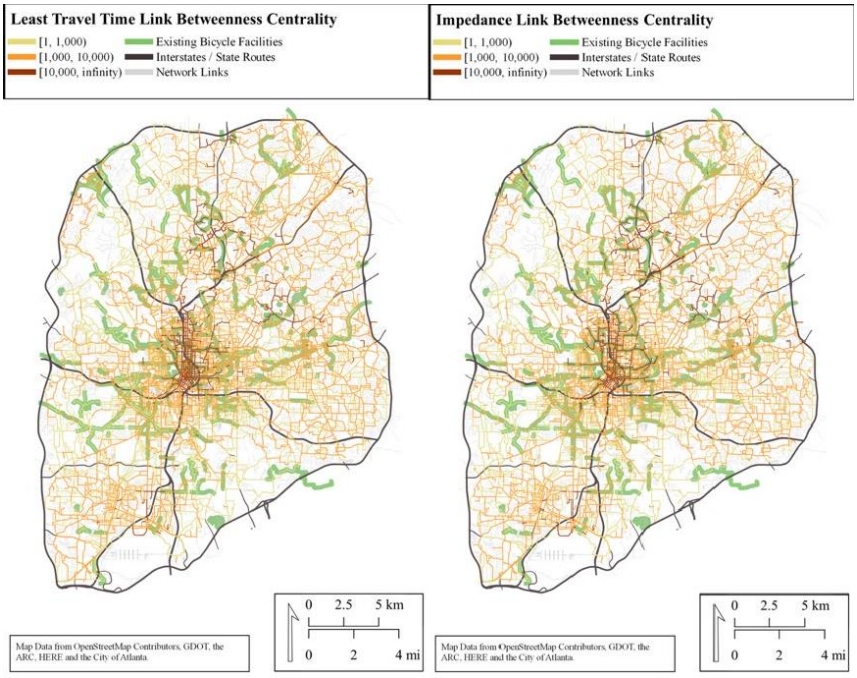

Figure 43: (left) link betweenness centrality calculated from the least travel time impedance (right) link betweenness centrality calculated from the calibrated impedance

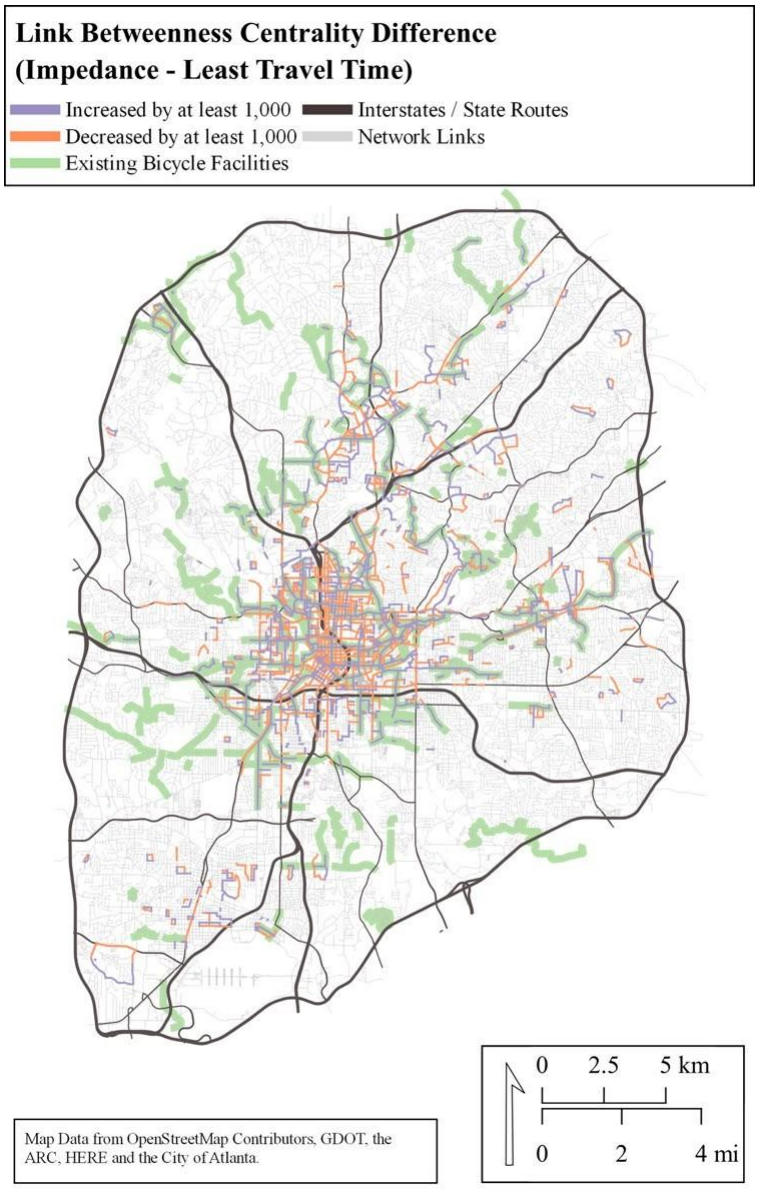

Figure 44: Impedance LBC minus the travel time LBC for the current network

On the other hand, the change in calibrated impedance LBC from the current to the future network highlights the changes in trip flows that result from the added improvements. Increases in LBC indicate that the respective links or nearby links decreased in impedance due to the improvements, while decreases in LBC indicate that the respective links lost trips due to the improvements because a more pleasant route had become available.

Figure 45 visualizes the difference in LBC from the current to the future network. As with Figure 44, the links that increased in LBC are purple and the links that decreased in LBC are orange. The proposed bicycle facilities are shown for reference. Several of the LBC increases were along the links with the proposed bicycle facilities. This indicates that the corresponding bicycle facilities both reduced impedance and offered a more pleasant route. Some of the proposed bicycle facilities on the outskirts yielded no changes in LBC, indicating that these facilities were too far out of the way to detour to.

Planners can use LBC to sketch out the most important corridors for cycling improvements. Transportation plans can prioritize cycling infrastructure on routes with a high travel time impedance LBC but a calibrated impedance LBC to increase the directness of the cycling network. The calibrated impedance LBC could also be cross-referenced against observed cycling data to see if these expectations match reality (Cooper, 2017). However, another application of LBC is to calculate the relative impedance reduction across improvements, for use in project selection.

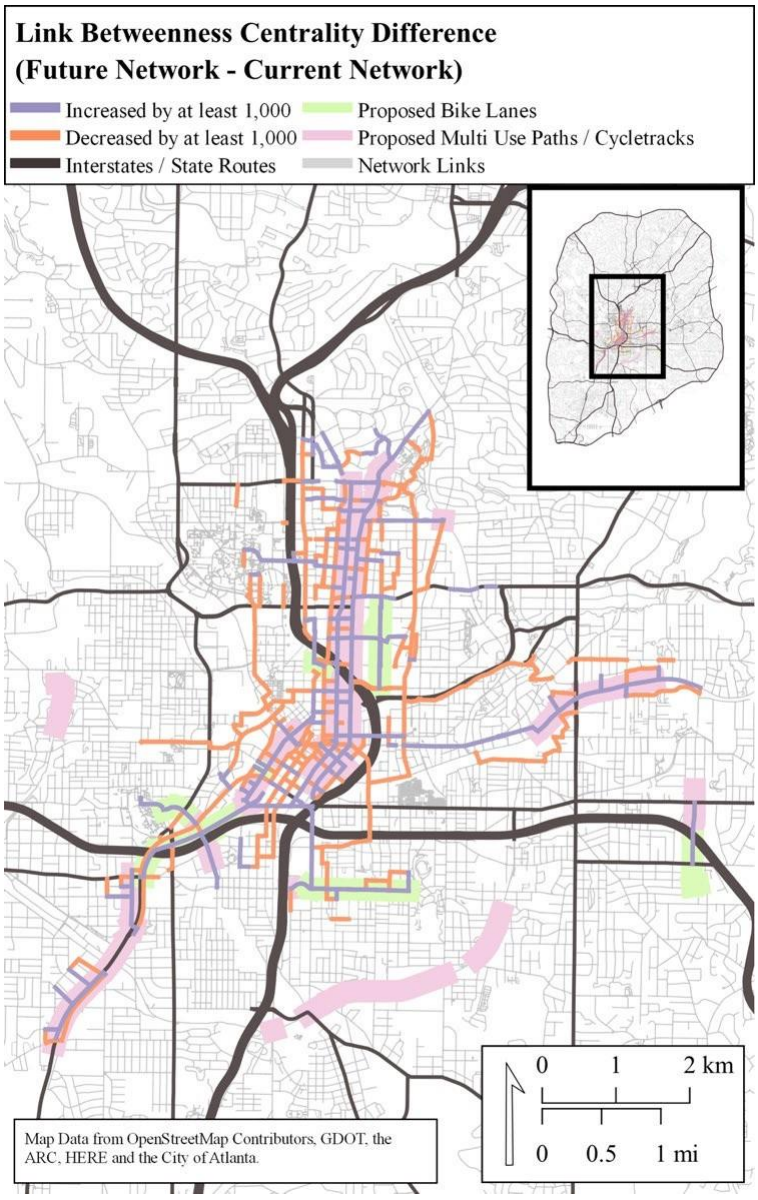

Figure 45: Impedance LBC difference between future and current network

IMPROVEMENT IMPEDANCE REDUCTION CONTRIBUTION

A bicycle facility's impedance reduction contribution represents its effectiveness at reducing trip impedance. The more trips that utilize a bicycle facility, the higher the bicycle facility's contribution to impedance reduction. Planned bicycle facilities often span across several links (where each link represents a section of a longer road), so the impedance reduction contribution is averaged across all of the links. This prevents longer bicycle facilities from dominating shorter bicycle facilities that may actually provide a greater reduction in total trip impedance.

Using Equation 14, a bicycle facility's (AA) impedance reduction contribution (ZZ_{AA}) is calculated by multiplying the link impedance reduction ($CC_{aa,ccaaaaaaaappaa} - CC_{aa,ooaaaaaaaaaa}$) and future LBC ($LLBBCC_{aa,ooaaaaaaaaaa}$) for each link ($gg$) that is part of the bicycle facility ($gg \in AA$). The link impedance reduction is calculated by subtracting the current network link impedance ($CC_{aa,ccaaaaaaaappaa}$) by the future network link impedance ($CC_{aa,ooaaaaaaaaaa}$). These values are summed and then divided by the bicycle facility's overall length in feet (LL_{AA}) to yield the bicycle facility's impedance reduction contribution.

$$ZZ_{AA} = \frac{\sum_{aa \in AA}(CC_{aa,ccaaaaaaaappaa} - CC_{aa,ooaaaaaaaaaa})LLBBCC_{aa,ooaaaaaaaaaa}}{LL_{AA}}$$

Equation 14: Improvement Impedance Reduction

Figure 46 shows a map of each bicycle facility's impedance reduction contribution. The bicycle facilities with the highest impedance reduction contribution are represented with thicker green lines and the grey circles demark the extents of the bicycle facilities.

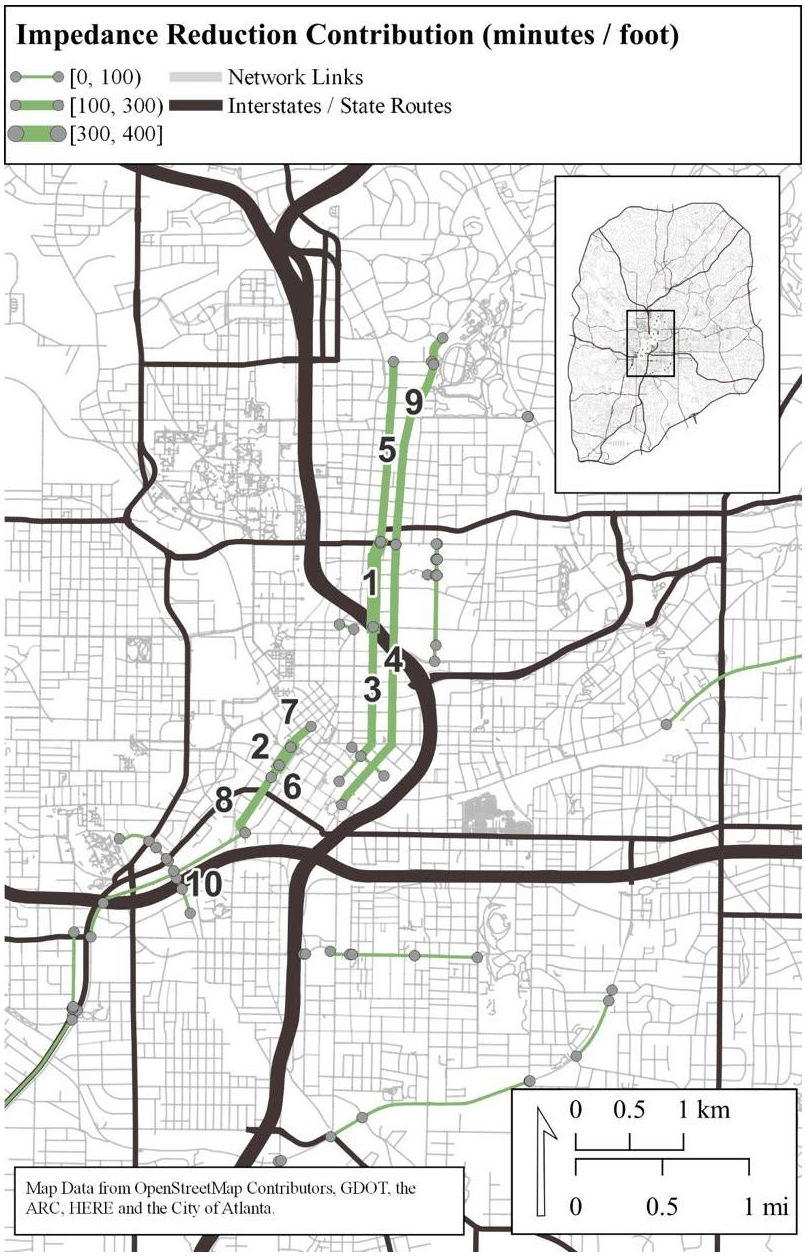

Figure 46: Impedance reduction from bicycle facilities

The top 10 bicycle facilities by impedance reduction contribution are labelled in Figure 46 from 1 to 10 (where 1 is the highest impedance reduction contribution). These bicycle facilities almost exclusively facilitated northbound and southbound bicycle travel through Midtown and Downtown Atlanta. Bicycle facilities located outside of these areas had lower impedance reduction contributions. This makes sense in the context of Figure 43, because the LBC tended to be highest for links within Midtown and Downtown Atlanta given the larger number of trips that originated or ended in that area. It also makes sense because the current link impedance values within Midtown and Downtown Atlanta were high, due to the presence of one-way roads with many travel lanes.

The detailed metrics for the top 10 bicycle facilities are listed in Table 23. The leftmost column indicates the bicycle facility's ranking which can be cross-referenced with the labels in Figure 46. The total impedance reduction contribution and length of the bicycle facility is given in the two rightmost columns. The bicycle facilities are all of different lengths, and some, such as #7 and #10, are less than 1,000 feet. It should be noted that Table 23 represents the top ten facilities that are the most effective at reducing trip impedance given the set of inputted trips. The rankings would change for a different set of trips.

Table 23: Impedance reduction by bicycle facility

Rank	Name	Segment	Type	Impedance Reduction Contribution Per Foot (minutes / foot)	Total Impedance Reduction Contribution (minutes)	Length (feet)
1	Courtland St NE	Ralph McGill Blvd to North Ave NE	Two-Way Cycle Track	390	1,008,901	2,590
2	Forsyth St SW	MLK Jr Dr SW to Alabama St SW	Protected Bike Lane	361	470,215	1,304
3	Courtland St NE	Gilmer St SE to Ralph McGill Blvd	Protected Bike Lane	260	1,054,868	4,059
4	Piedmont Ave NE	MLK Jr Dr SE to North Ave NE	Protected Bike Lane	221	1,875,179	8,486
5	Juniper St NE	North Ave NE to 14th St NE	Protected Bike Lane	214	1,176,355	5,508
6	Forsyth St SW	Mitchell St SW to MLK Jr Drive SW	Bike Lane	198	172,874	875
7	Forsyth St SW	Alabama St SW to Walton St NW	Protected Bike Lane	181	323,836	1,788
8	Forsyth St SW	Memorial Dr SW to Mitchell St SW	Protected Bike Lane	171	687,543	4,021
9	Piedmont Ave NE	North Ave NE to 15th St NE	Protected Bike Lane	137	1,023,880	7,493
10	McDaniel St SW	Whitehall Terr to RDA Fwy	Separated Bike Lane	83	60,907	738

BIKESHEDS

The last visualizations presented are bikesheds. Unlike the previous visualizations, bikesheds are localized to a specific origin. The bikesheds shown in this report are reproduced from Passmore et al. (2024). These bikesheds were generated using a separate set of time + attribute impedance factors that were not yet calibrated with GPS traces. The network presented in the figures that follow are only for a subset of the full network used in this report. Nevertheless, monitoring changes in bikesheds that result from

network improvements that reduce link impedance is a good way of comparing the benefits of various improvement strategies.

Bikesheds represent the extent of the network (or destinations) one can access by bike within specified constraints. For this report, the bikesheds were limited to every node on the network that one could reach in ten minutes at a speed of 8.0 miles per hour. In this report, bikesheds are measured in linear network miles (i.e., the total length of all unique traversable links within a bikeshed), but bikesheds can also be measured by area by buffering the links within a bikeshed by a designated access volume and then taking the area of this buffered region while accounting for overlaps.

Figure 47 depicts the bikeshed using the travel time impedance (Figure 47a) and the time + attributes impedance (Figure 47b). The travel time impedance bikeshed size was 22 network miles, and the time + attributes impedance bikeshed size was 15 network miles. This overall decrease in bikeshed reflects the lack of cycling infrastructure and low-stress roads in this area.

Figure 47: 10-min bikesheds for (a) travel time impedance and (b) time + attributes impedance, and (c) links added and removed when changing from the travel time impedance to the time + attributes impedance.

Figure 47c shows how the bikeshed expands in some areas and contracts in other areas, due to the changes in impedance across various facility types. Some links are added to the bikeshed, and some are removed from the bikeshed, when changing from the travel time impedance to time + attributes impedance are shown in blue and red, respectively. When

links are removed, either these links or the links leading up to the links experienced an increase in time + attributes impedance, compared to travel time only impedance. For example, the discomfort of multi-lane travel may penalize some routes. When links are added to the bikeshed, the new links that appear, or the links leading up to these new links, experienced a lower time + attributes impedance compared to the original travel time only impedance. For example, the presence of a dedicated bike lane or multi-use paths reduces the impedance along some routes (i.e., some people are willing to travel further on some links compared to high-stress links). The expansion of the bikeshed to the North and South is due to the presence of the Atlanta BeltLine and Freedom Parkway multi-use paths, which provided for a more enjoyable route.

The bikesheds were then re-calculated to include the two network improvements. The orange lines in Figure 48 represent links added to the bikeshed because of the network improvements (cross-hatched). The size of the resulting bikeshed increases to 16 miles (6% increase). This example highlights how bikesheds can be used to assess the impacts of infrastructure improvements on bikeability. Bikesheds can also demonstrate the potential effects of reduced speed limits and various traffic calming measures. In recommending bicycle infrastructure, one could further penalize distance traveled on 40 miles per hour roads to address safety concerns, which would restrict the size of the bikeshed (Lowry et al. 2016). In addition to these transportation planning contexts, these bikesheds can be used as a tool on an individual level, similar to WalkScore (Walk Score, n.d.). A person could look up their residence and see all the locations within this time + attributes 10-min bikeshed.

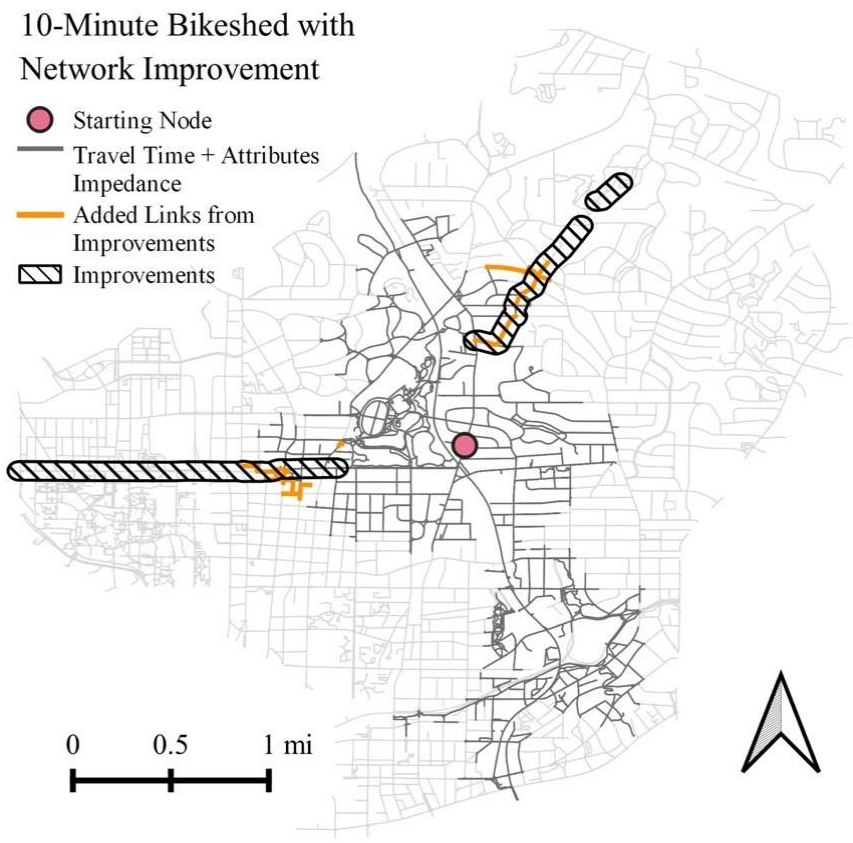

Figure 48: Change in the time + attributes impedance 10-minute bikeshed after the proposed improvements were added to the network

BIKEWAYSIM DEMONSTRATION SUMMARY

This chapter presented how BikewaySim can be used to systematically assess the impacts of infrastructure improvements using a set of specific metrics (changes in bikeshed mileage, percent detour, change in impedance, change in LBC, etc.). These metrics can all be visualized to help communicate the impacts to planners and the public.

BikewaySim cannot be used alone to create an optimal cycling network; well-designed and thought-out cycling improvements and network plans must still be developed as the inputs to BikewaySim. But BikewaySim can serve as a tool for planners and engineers to compare and evaluate cycling facility alternatives or cycling network buildouts, so that all plans and projects are compared using consistent methods, and for communicating the predicted impacts of these potential infrastructure improvements to decision-makers and the public.

CHAPTER 7. DEMONSTRATION OF BIKEWAYSIM IN SAVANNAH, GA

The ability to utilize BikewaySim in other cities within Georgia, as well as across the country, was a fundamental goal of this research. As such, this chapter showcases the transferability of BikewaySim in the Savannah Metro Area by assessing 57 proposed bicycle facilities from the Savannah Metropolitan Planning Commission's non-motorized transportation plan (MPC, n.d.).

Using BikewaySim's network construction scripts, OSM data were downloaded and processed into an all-paths network using the Savannah Metro Area boundaries. The all-paths network was then reconciled with the available GDOT and USGS data to provide additional road attributes for impedance calculation. After the network was pruned to remove links that cyclists weren't allowed to traverse, the network contained 107,397 links (214,794 directed links), 91,820 nodes, and 350,746 link-to-link turn opportunities in the pseudo dual graph.

The 57 proposed bicycle facilities include bike lanes, cycletracks, and multi-use paths. Only on-street features (i.e., bike lanes, cycletracks) were assessed, so these facilities were added to the network be editing link attributes. No new connectivity (i.e., new links) to the network was added. The study area network, existing bicycle facilities, and proposed bicycle facilities are shown in Figure 49. The few existing bicycle facilities in the study area are shown in green, the proposed bike lanes are shown in orange, and the proposed multi-use paths/cycletracks are shown in brown. The Interstates and state routes are shown for reference in dark grey, and the network links are shown in light grey. The discontinuities in the street network are a result of the wetlands surrounding Savannah.

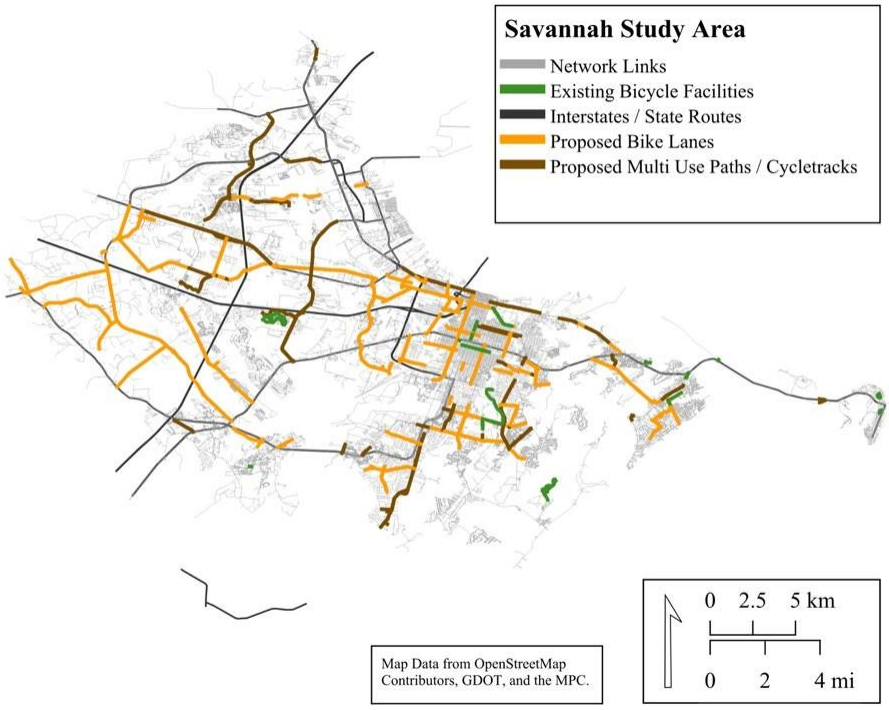

Figure 49: Savannah study area

Longitudinal Employer-Household Dynamics Origin-Destination Employment Statistics (LODES7) data from the US Census Bureau was used in place of a travel demand model OD matrix (U.S. Census Bureau, 2010). These data corresponded to census blocks and represented home-to-employment ODs for 2010. The LODES7 data only represents work commutes whereas the ODs used in the previous chapter represented all types of trips. People for Bikes also uses LODES7 data in their bicycle network analysis tool (People for Bikes, n.d.).

Commented [BCD18]: How does the LODES7 data for Atlanta compare to the ODs from Cycle Atlanta? Are they comparable enough for LODES7 to serve as a proxy for all ODs?

Commented [TP19R18]: LODES7 is meant to be comparable to the OD data provided by the ARC ABM model. A major limitation of LODES7 compared to both CycleAtlanta ODs and the ARC ABM ODs is the lack of any non-work trips. LODES7 should only serve as a placeholder if no other OD data are available.

After using the same procedure from the previous chapter to filter the OD data, there were 17,856 trips between 15,275 unique origin-destination pairs representing 2,921 census blocks. The shortest path routing was run on both the current network and the network that included the proposed bicycling facilities, using the impedance functions calibrated in Chapter 5.

After running the shortest path calculations, trip impedance reductions were calculated. The weighted average trip impedance reduction was 1.82 minutes, and the maximum trip impedance reduction was 14.6 minutes. Figure 50 shows the weighted average impedance reduction by census blocks. The map zoom level in Figure 49 highlights downtown Savannah.

The weighted average impedance reduction was greatest for census blocks towards the outskirts of the gridded area. One possible reason for this could be that the streets in the gridded area generally had few vehicle travel lanes and unsignalized crossings of major roads, so the cycling route impedance was already low. On the other hand, the census blocks on the periphery also may have fewer low impedance routes for traveling to other census blocks, so the introduction of new bicycle infrastructure had a comparatively larger reduction in impedance.

The Savannah application demonstrates the transferability of BikewaySim to another study area in the State of Georgia, but cycling GPS trace data should be collected for each study area to ensure that the impedance factors represent the local cycling population. In study areas outside of Georgia, but still within the United States, alternative data sources on the number of lanes and other desired attributes will need to be identified.

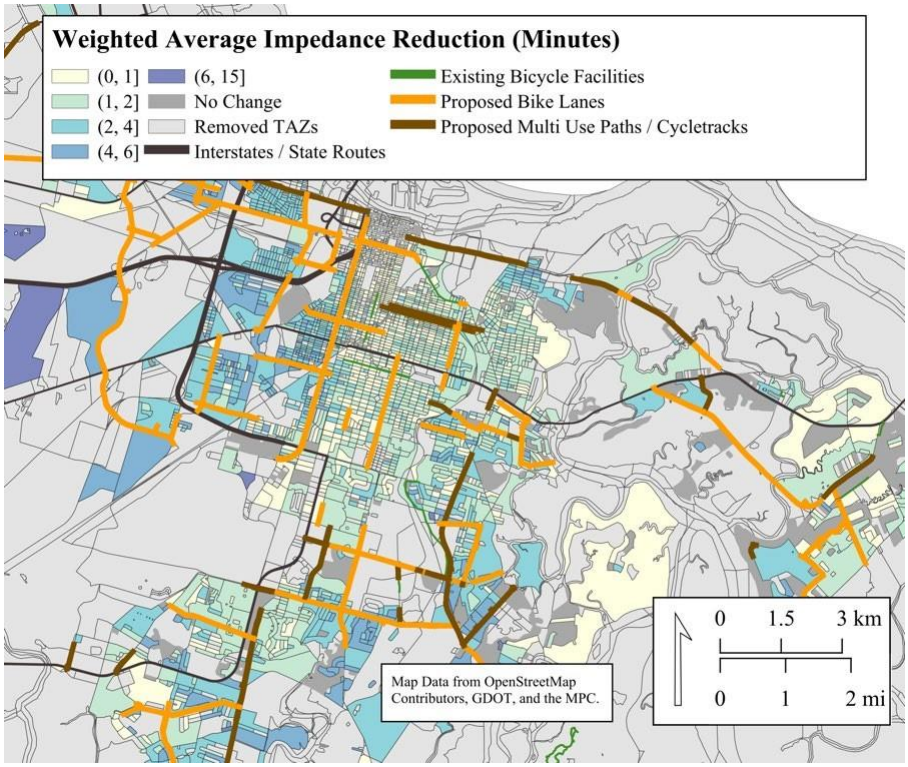

Figure 50: Weighted average impedance reduction by TAZ

CHAPTER 8. CONCLUSIONS

The bicycle infrastructure funding process needs political and public buy-in. Cycling facilities can be met with strong opposition, especially in cases that require removing on-street parking or vehicle lanes. Such opposition can result in cycling infrastructure being built in places that are politically feasible, but not as effective in encouraging ridership, or, in worst-case scenarios, can lead to cycling infrastructure being placed in dangerous locations or not being built at all.

BikewaySim is a tool that planners and engineers can utilize to visualize and calculate the mobility and connectivity impacts of planned bicycle facilities. Planners and engineers can present BikewaySim's results to decision-makers and the public as supporting evidence for building bicycle facilities. BikewaySim relies on cycling impedance which measures the relative cost of cycling taking into consideration travel time, exposure to automobiles, hills, and the provision of cycling infrastructure. The reduction of cycling impedance can be used as a leading indicator for future increases of cycling activity in areas with little existing cycling infrastructure and few cyclists.

The cycling route choice literature informed the development of BikewaySim's cycling impedance functions. In total, 19 cycling route choice studies helped the researchers develop an intuition for how cyclists perceive route attributes (e.g., trip distance, percent of trip on bicycle facilities, etc.) in the context of a change in distance or travel time. The measures representing these perceptions, the marginal rates of substitution, were consistent in sign but not magnitude within and across the cycling route choice studies. This fact, coupled with the overhead required to implement these models in the travel

demand planning process led to the researchers pursuing an alternative method for developing BikewaySim's cycling impedance functions. In addition, existing studies rarely applied the models that they estimated to new situations, or applied individual or collective findings from the cycling route choice literature to model impacts of proposed bicycle facilities.

Critical to calibrating cycling impedance functions and assessing bicycle facilities was the development of an all-paths network. The all-paths network, constructed from OpenStreetMap, was reconciled with additional data from GDOT, ARC, HERE, USGS, and the City of Atlanta to add attributes on the number of lanes, motor vehicle speed, elevation, and bicycle facility installation dates. With the exception of elevation data, most data were reconciled through a process that considered the Hausdorff distance and whether the street names matched. Elevation was added through DEM sampling and processing based on Liu et al. (2018). Classified LiDAR point cloud data were used to resample elevation for bridges. Lastly, the network links where cycling travel was not viable (highways, sidewalks, etc.) were removed to yield a final network with 76,912 links, 66,151 nodes, and 222,764 link-to-link turn opportunities.

One limitation of this work was the lack of more detailed road attribute data which may have limited the research team's ability to calibrate impedance factors. Information on the attributes of cycling infrastructure such as the width of the facility and the presence of street parking would have been informative for impedance calibration. Field data collection forms were created to collect this data for future impedance calibration and asset management.

Once the all-paths network was constructed, cycling GPS traces from the CycleAtlanta app were processed and map-matched to the all-paths network for impedance calibration. There were 1,936 trips representing 586 users that were successfully map matched to the network. For the trips that were not successfully map matched, in future research, the researchers intend to investigate if there are missing links (i.e., informal shortcuts) in the all-paths network that need to be created.

The map matched traces were then used to calibrate link and turn impedance factors using stochastic optimization techniques. The calibrated link and turn impedance functions considered the number of lanes, the presence of a multi-use path, the presence of a bike lane, the presence of a road grade above 4%, and unsignalized crossings of major roads. The presence of bicycle facilities reduced impedance while the other attributes increased impedance. The link impedance coefficients correspond to a percent change in travel time while the turn coefficient (unsignalized crossing of major roads) corresponds to an increase in travel time.

The increase in overlap between the impedance modeled routes and the map matched routes from the shortest travel time modeled routes was minimal. In future research, the researchers will increase this overlap through testing new permutations of the impedance functions and by segmenting the map matched data. There may be certain users that warrant modeling separately due to their routing behavior.

Future work should also focus on collecting cycling GPS traces on a regular basis so that cycling impedance factors can be updated over time as bicycle infrastructure becomes available and new cyclists emerge. Since the CycleAtlanta study, e-bikes have become more widely available, and additional impedance factors likely need to be created for e-

bike riders. Work by Meister et al. (2022) suggests that cyclists on e-bikes are less sensitive to hills which means that BikewaySim could also likely be used to evaluate the anticipated impedance reductions from incentive programs such as Atlanta's E-Bike Rebate.

If current cycling volumes in an area are not high enough to warrant the collection of GPS traces, researchers could ask local cyclists and bicycle advocacy organizations to draw out a series of safety-oriented routes between a variety of places to provide an a priori assumption on an area's impedance factors. These impedance factors can then be adjusted once GPS traces become available.

The calibrated coefficients in this report were used to demonstrate the BikewaySim network assessment framework in the Metro Atlanta study area. Using trip data from the ARC's activity based model and planned bicycle facilities from the City of Atlanta, the framework first performs shortest path routing calculations using a travel time impedance and the calibrated impedance on the current network and future network (i.e., current network with planned bicycle facilities). The calculated routes were processed to create four different perspectives on the impacts of the planned bicycle facilities.

Trip impedance reduction (visualized as weighted average impedance reduction by origin TAZ) showed which areas in the study area benefited from the planned bicycle facilities. Percent detour (visualized as weighted average percent detour by origin TAZ) showed which areas required detouring from the minimum travel time route to access lower impedance routes with more cycling infrastructure or streets with fewer vehicle travel lanes. Link betweenness centrality showed which links were the most critical in connecting the origin destination pairs using both the travel time impedance and

calibrated impedance. The impedance reduction contribution for each bicycle facility was calculated to rank which facilities did the most to reduce impedance. Finally, an example bikeshed was presented to demonstrate the local impacts of the impedance factors and the planned bicycle facilities.

Planners and engineers should take a holistic view of these visualizations and metrics in deciding which bicycle facilities to prioritize. In addition, they should consider testing other forms of origin-destination pairs, such as access to grocery stores, restaurants, shopping, and other activities as adding new cycling infrastructure might generate new types of trips that aren't considered in travel demand models.

In addition to the Metro Atlanta study area, BikewaySim was successfully demonstrated for another study area, the Savannah Metro Area. BikewaySim's scripts were used to develop an all-paths network, and the calibrated impedance factors from the CycleAtlanta data were used to assess planned bicycle facilities using US Census Bureau home-to-employment origin destination data. This indicates that BikewaySim is currently applicable to any study area within Georgia and will be applicable in places across the United States that possess the necessary road attributes.

For GDOT, BikewaySim currently offers a way to test the impacts of new bicycle facilities on state routes, but it can also test the impact of road-narrowing or traffic calming once these impedance factors are developed. BikewaySim may be particularly useful in the areas of Georgia where there is little cycling infrastructure and few cyclists. BikewaySim will continue to be developed by the researchers, and because it is open source, BikewaySim can be adopted and modified by other state DOTs, MPOs, or cities.

GDOT can continue to support the development of BikewaySim through the collection of bicycle infrastructure attribute data and cycling GPS traces so that more nuanced cycling impedance factors can be developed. Additionally, the researchers recommend incorporating BikewaySim in the evaluation process for active transportation grant programs. This would help local jurisdictions propose projects that focus on building cycling facilities that build towards a connected cycling network.

APPENDIX A: BICYCLE FACILITY INSPECTION FORM

Bike Facility Inspection Report

Inspector(s): _____ Inspection Date: _____ Inspection Time: _____

[OSM Map Location 3"x3"]	[Photo(s)]

Bike Facility ID:		**Zone/Neighborhood:**	
Latitude:		**Longitude:**	
Parallel Street		**Nearby Landmark**	

ROW Elements	Motor Veh.	Pedestrians	**General Elements**	
Bike Lane Alignment	(L/C/R/M)	(L/C/R/M)	Bike Route Signage	y/n
Type of Horiz. Separation			Signage Visibility	GFP/A
Horiz. Separation Condition			Bike Facility Directionality	(1W/2W)
Buffer Width			Directional Separation Type	
Type of Vert. Sep.			Lane Width (per direction)	
Vert. Sep. Condition	(GFP/A)	(GFP/A)	Shared Bus/Bike Lane?	(y/n)
Shared Lane Markings?	(y/n)	(y/n)	Running Slope	
Marking Condition	(GFP/A)	(GFP/A)	Cross Slope	
			Surface Material	
			Colored Surface	(y/n)
			Potholes Present?	(y/n)
Adjacent Travel Behavior Elements			Cracking Present?	(y/n)
Adjacent Street Parking			Loose Materials Present?	(y/n)
Motor Vehicle Volume			Rutting Present?	(y/n)
Motor Vehicle Speed			Gutter Width	
Pedestrian Volume			Gutter Depth	

Field Notes:

.

See reverse side for key

Version 07/23/24

Key:

Metadata

Inspector(s) – Write the name(s) of the person(s) performing the inspection

Inspection Date – Record the Date of the inspection

Inspection Time – Record the Time of the inspection

OSM Map Location – [for pre-/post-processing] Insert the location of the inspection on a map

Photo(s) – [for pre-/post-processing] Insert photo(s) of the bicycle facility that was inspected

Bike Facility ID – [for pre-/post-processing] Identification number for bicycle facility, for organizational purposes if there is a coordinated / planned effort to inspect several bicycle facilities in an area

Zone/Neighborhood - [for pre-/post-processing] Subarea in which the bicycle facility is located, for organizational purposes if there is a coordinated / planned effort to inspect several bicycle facilities in an area

Latitude – Latitude of the inspection location

Longitude – Longitude of the inspection location

Parallel Street – Street that is concurrent with / parallel to bicycle facility

Nearby Landmark – Identifiable feature near inspection location

Right of Way (ROW) Elements

Bike Lane Alignment – Alignment of the Bicycle Facility with respect to motor vehicle and pedestrian traffic. (Left of traffic, right of traffic, center of traffic, or mixed with traffic)

Type of Horiz. Separation – Type of horizontal separation from traffic (Striping, Striping with buffer, other, absent, N/A)

Horiz. Separation Condition – Condition of the horizontal traffic separation feature (good, fair, poor, absent, N/A)

Buffer Width – Width of the horizontal buffer between bicycle traffic and motor vehicle or pedestrian traffic (numeric value

Type of Vert. Separation – Type of vertical separation from traffic (n/a if there is no vertical separation)

Vert. Separation Condition – Condition of vertical separation from traffic (good, fair, poor, absent)

Shared Lane Markings – Note if there are surface markings indicating that bicycles share the lane with motor vehicle and/or pedestrian traffic (yes, no, N/A)

Marking Condition – Condition of the shared lane markings (good, fair, poor, absent)

Adjacent Travel Behavior Elements

Adjacent Street Parking – Is there street parking immediately adjacent to the bicycle facility?

Motor vehicle volume – What is the observed volume of motor vehicles adjacent to / on the bicycle facility? (veh/time)

Motor vehicle speed – What is the observed speed of motor vehicles adjacent to / on the bicycle facility?

Pedestrian Volume – What is the observed volume of pedestrians adjacent to / on the bicycle facility?

General Elements

Bike Route Signage – note if there are signs indicating that the bicycle facility is a bike route (yes, no, N/A)

Signage Visibility – Note the level of visibility of bike route signage (good, fair, poor, absent, N/A)

Bike Facility Directionality – Does the bicycle facility have one-way or two-way bicycle traffic (1 way, 2 way, N/A)

Directional Separation Type – Type of separation between directions of bicycle traffic (Striping, Striping with buffer, other, absent, N/A)

Lane Width per Direction – Width of the facility usable by bicycles (per direction if directions are separated)

Shared Bike/Bus Lane – Is the bicycle facility intended for exclusive use by bikes AND buses? (yes, no, N/A)

Running Slope – Slope of the surface in the direction of travel (numeric value reported as a %)

Cross Slope – Slope of the surface perpendicular to the direction of travel (numeric value reported as %)

Surface Material – Material of bike facility surface (Asphalt, Concrete, Dirt, Gravel, Rubberized, Other)

Colored Surface – Note if the bicycle facility surface is a different color from adjacent roadway/walkway (yes, no, N/A)

Potholes Present – Note if potholes are present within the bicycle travel way (yes, no, N/A)

Cracking Present – Note if substantial pavement cracking is present within the bicycle travel way (yes, no, N/A)

Loose Materials Present – Note if there is substantial non-permanent debris buildup within the bicycle travel way (

Rutting Present – Note if substantial ruts have formed within the bicycle travel way (yes, no, N/A)

Gutter Width – Note the width of any street gutter within the bicycle travel way (numeric value, N/A)

Gutter Depth - Note the width of any street gutter within the bicycle travel way (numeric value, N/A)

Field Notes – Record any additional pertinent details regarding the bicycle facility, and/or any data entries that warrant additional context / explanation

Version 07/23/24

APPENDIX B: BICYCLE CONFLICT POINT INSPECTION FORM

Bicycle Conflict Point Inspection Report

Inspector(s): _____ Inspection Date: _____ Inspection Time: _____

[OSM Map Location 3"x3"]	[Photo(s)]

Bike Facility ID:		**Zone/Neighborhood:**	
Latitude:		**Longitude:**	
Parallel Street		**Nearby Landmark**	

General Elements		**Intersections**	
Parked cars within 20 feet	(y/n)	Intersection veh. directions	((ex. 3W, 4W)
Conflict point visibility	(y/n)	Adjacent veh. travel lanes	
Vehicles yield to bikes signage?	(y/n)	Cross street veh. travel lanes	
Mid-Block Pedestrian Crossing Conflict		Buffer width (in)	
Bikes yield to peds sign / marking	(y/n)	Traffic control type	
Crosswalk present?	(y/n)	2 stage bike left turn?	(y/n)
Crosswalk condition?	(gfp/a)	Bike box present?	
Raised crosswalk?	(y/n)	Bike box condition	(GFP/A)
Mid-Transit Bus Stop Conflict		Signal cycle length (sec)	(y/n)
Tot. duration of conflicts (min/hr)	(min/hr)	Bike specific signals?	(y/n)
Bus only pavement markings?	(y/n)	Lane marks in intersection?	(y/n)
Bikes yield to bus signage?	(y/n)	Lane marking condition?	(GFP/A)
Loading Zone / Accessible Parking Mid Block Conflict		Veh. cross over bike lane?	(y/n)
Crosswalk present?	(y/n)	Veh. queue blocks bike ln?	(y/n)
Crosswalk condition?	(y/n)	Merge area length?	(ft)
Bike lane width reduction (ft)	(ft)	Merge area dist. to intersect	(ft)

Field Notes:

.

See reverse side for key

Version 07/24/24

Key:

General Elements

Parked cars within 20 feet – Are there cars parked within 20 feet of the conflict point? (yes/no)

Sight triangle issues – Are there any obstructions that would block visibility at the conflict point? (yes/no)

Vehicles yield to bikes sign? - Is there signage present indicating that vehicles should yield to bikes? (yes/no)

Mid-Block Pedestrian Crossing Conflict

Bikes yield to peds sign / marking? – Are there sign(s)/marking(s) indicating that bikes should yield to pedestrians? (yes/no)

Crosswalk Present – Is there a crosswalk present? (yes/no)

Crosswalk Condition – Condition of the crosswalk (good, fair, poor, absent)

Raised Crosswalk – Is the crosswalk raised above the level of the bike lane? (yes/no)

Mid-Block Transit Stop Conflict

Tot. duration of conflicts – The total duration of time in which transit vehicles are blocking the bike lane, in min/hr.

Bus only pavement markings – Are there pavement markings indicating that the space is for buses only? (yes/no)

Bikes yield to bus signage – Are there signs / markings indicating that bikes should yield to buses?

Loading Zone / Accessible Parking Mid Block Conflict

Crosswalk Present – Is there a crosswalk present? (yes/no)

Crosswalk Condition – Condition of the crosswalk (good, fair, poor, absent)

Bike Lane Width Reduction – Reduction in width of the bike lane at the conflict point, in feet.

Intersection Conflict Points

Intersection Veh. directions – Number of directions from which motor vehicles enter the intersection. Ex: 4W

Adjacent veh. travel lanes – Number of vehicle travel lanes adacent to the bike lane

Cross-street veh. travel lanes - Number of vehicle travel lanes on intersecting street

Traffic Control Type – Denote the type of traffic control or signalization used at the intersection (i.e., 4-way stop, signal, etc.)

Buffer width – Width of the buffer zone between the bike lane and adjacent motor vehicle lane, in inches.

2 stage bike left turn – Do bikes making a left turn have to do so in two stages? (yes/no)

Signal cycle length – Length of time required to complete one full signal cycle, in seconds)

Bike specific signals – Are there signals specifically for bicycle traffic?

Lane markings in intersection? – Is the bike lane clearly marked? (yes/no)

Lane marking condition – What is the condition of the bicycle lane markings (good, fair, poor, or absent)

Veh. cross over bike lane? – Do turning motor vehicles cross over the separated bike lane to complete their turn? (yes/no)

Merge area length – length of the area in which motor vehicle traffic merges into / through bicycle traffic, in feet)

Merge area dist. To intersect – Distance between the end of the merge area and the intersection, in feet.

Field Notes – Record any additional pertinent details regarding the bicycle facility, and/or any data entries that warrant additional context / explanation

Figure 1: Two-Stage Bicycle Left Turn Example

Figure 2: Bicycle Box Example

Version 07/24/24

ACKNOWLEDGMENTS

The authors would like to thank Diyi Liu and Dr. Ziyi Dai for their work on the BikewaySim repository. We would also like to thank staff members from the City of Atlanta and Rebecca Serna of Propel for providing data on proposed cycling facilities for the Metro Atlanta study area. Finally, we would like to thank our Georgia Department of Transportation technical advisors Sam Harris, Connor Booth, and Ron Knezevich and research manager Jordan Kamari.

REFERENCES

AASHTO Task Force on Geometric Design. (2012). *AASHTO Guide for the Development of Bicycle Facilities*.

Alattar, M. A., Cottrill, C., & Beecroft, M. (2021). Modelling cyclists' route choice using Strava and OSMnx: A case study of the City of Glasgow. *Transportation Research Interdisciplinary Perspectives*, *9*, 100301. https://doi.org/10.1016/J.TRIP.2021.100301

Alt, H., & Godau, M. (1995). Computing the Fréchet Distance between Two Polygonal Curves. *International Journal of Computational Geometry & Applications*, *05*(01n02), 75–91. https://doi.org/10.1142/S0218195995000064

Arias, D., Ederer, D., Rodgers, M. O., Hunter, M. P., & Watkins, K. E. (2021). Estimating the effect of vehicle speeds on bicycle and pedestrian safety on the Georgia arterial roadway network. *Accident Analysis & Prevention*, *161*, 106351. https://doi.org/10.1016/j.aap.2021.106351

Atlanta Regional Commission. (n.d.). *Activity-Based Modeling*. https://atlantaregional.org/what-we-do/transportation-planning/modeling/

Atlanta Regional Commission. (2022). *2022 Regional Bikeway Inventory*. https://opendata.atlantaregional.com/datasets/regional-bikeway-inventory-2022/explore

Barajas, J. M. (2021). Biking where Black: Connecting transportation planning and infrastructure to disproportionate policing. *Transportation Research Part D: Transport and Environment*, *99*, 103027. https://doi.org/10.1016/j.trd.2021.103027

Ben-Akiva, M., & Bierlaire, M. (2003). Discrete Choice Models with Applications to Departure Time and Route Choice. In *International Series in Operations Research & Management Science* (pp. 7–37). Kluwer Academic Publishers. https://doi.org/10.1007/0-306-48058-1_2

Bernardi, S., Geurs, K., & Puello, L. L. P. (2018). Modelling route choice of Dutch cyclists using smartphone data. *Journal of Transport and Land Use*, *11*(1), 883–900. https://doi.org/10.5198/jtlu.2018.1143

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

Broach, J., & Dill, J. (2016). Using Predicted Bicyclist and Pedestrian Route Choice to Enhance Mode Choice Models. *Transportation Research Record: Journal of the Transportation Research Board*, *2564*, pp 52-59. https://doi.org/10.3141/2564-06

Broach, J., Dill, J., & Gliebe, J. (2012). Where do cyclists ride? A route choice model developed with revealed preference GPS data. *Transportation Research Part A: Policy and Practice*, *46*(10), 1730–1740. https://doi.org/10.1016/j.tra.2012.07.005

Buehler, R., & Pucher, J. (2021). *Cycling for sustainable cities*. The MIT Press.

Cabral, L., Kim, A. M., & Shirgaokar, M. (2019). Low-stress bicycling connectivity: Assessment of the network build-out in Edmonton, Canada. *Case Studies on Transport Policy*, *7*(2), 230–238. https://doi.org/10.1016/j.cstp.2019.04.002

Casello, J. M., & Usyukov, V. (2014). Modeling Cyclists' Route Choice Based on GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2430*(1), 155–161. https://doi.org/10.3141/2430-16

Chaloux, N., & El-Geneidy, A. (2019). Rules of the Road: Compliance and Defiance among the Different Types of Cyclists. *Transportation Research Record: Journal of the Transportation Research Board*, *2673*(9), 34–43. https://doi.org/10.1177/0361198119844965

Chen, P., Shen, Q., & Childress, S. (2017). A GPS data-based analysis of built environment influences on bicyclist route preferences. *International Journal of Sustainable Transportation*, *12*(3), 218–231. https://doi.org/10.1080/15568318.2017.1349222

City of Atlanta Department of Public Works. (2024). *Bike Facilities Public View*. ArcGIS. https://dpwatl.maps.arcgis.com/home/item.html?id=3f8907a069e94fefb33215406f7be228

Cooper, C. H. V. (2017). Using spatial network analysis to model pedal cycle flows, risk and mode choice. *Journal of Transport Geography*, *58*, 157–165. https://doi.org/10.1016/J.JTRANGEO.2016.12.003

CROW. (2009). *Design manual for bicycle traffic*. CROW.

DiGioia, J., Watkins, K. E., Xu, Y., Rodgers, M., & Guensler, R. (2017). Safety impacts of bicycle infrastructure: A critical review. *Journal of Safety Research*, *61*, 105–119. https://doi.org/10.1016/j.jsr.2017.02.015

Dijkstra, E. W. (1959). A note on two problems in connexion with graphs. *Numerische Mathematik*, *1*(1), 269–271. https://doi.org/10.1007/bf01386390

Douglas, D. H., & Peucker, T. K. (1973). Algorithms for the Reduction of the Number of Points Required to Represent a Digitized Line or its Caricature. *Cartographica:*

The International Journal for Geographic Information and Geovisualization, *10*(2), 112–122. https://doi.org/10.3138/fm57-6770-u75u-7727

Dutta, V. (2020). Ride easy with new biking features in Google Maps. *The Keyword*. https://blog.google/products/maps/ride-easy-new-biking-features-google-maps/

Ester, M., Kriegel, H. P., Sander, J., & Xu, X. (1996). A density-based algorithm for discovering clusters in large spatial databases with noise. *Kdd*, *96*(34), 226–231.

Federal Highway Administration. (2017). *National Household Travel Survey*. Federal Highway Administration. https://nhts.ornl.gov

Ferster, C., Fischer, J., Manaugh, K., Nelson, T., & Winters, M. (2020). Using OpenStreetMap to inventory bicycle infrastructure: A comparison with open data from cities. *International Journal of Sustainable Transportation*, *14*(1), 64–73. https://doi.org/10.1080/15568318.2018.1519746

FHWA. (2018). *Guidebook for Measuring Multimodal Network Connectivity*.

Fitch, D. T., & Handy, S. L. (2020). Road environments and bicyclist route choice: The cases of Davis and San Francisco, CA. *Journal of Transport Geography*, *85*. https://doi.org/10.1016/j.jtrangeo.2020.102705

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry*, *40*(1), 35. https://doi.org/10.2307/3033543

Geofabrik. (n.d.). *OpenStreetMap Data Extracts*. http://download.geofabrik.de

Georgia Department of Transportation. (n.d.). *Road & Traffic Data*. https://www.dot.ga.gov/GDOT/Pages/RoadTrafficData.aspx

Georgia Department of Transportation. (2022). *Georgia Crash Portal*. https://gdot.numetric.net/crash-data#/

Ghanayim, M., & Bekhor, S. (2018). Modelling bicycle route choice using data from a GPS-assisted household survey. *European Journal of Transport and Infrastructure Research*, *18*(2), 158–177. https://doi.org/10.18757/ejtir.2018.18.2.3228

Glennon, R. (2015). *Find better bicycle routes with new Valhalla options*. Mapzen. https://www.mapzen.com/blog/valhalla-bicycle-routing-options/

Groeger, L. V. (2016). Unsafe at Many Speeds. *Propublica*. https://www.propublica.org/article/unsafe-at-many-speeds

Hood, J., Sall, E., & Charlton, B. (2011). A GPS-based bicycle route choice model for San Francisco, California. *Transportation Letters*, *3*(1), 63–75. https://doi.org/10.3328/TL.2011.03.01.63-75

Khatri, R., Cherry, C. R., Nambisan, S. S., & Han, L. D. (2016). Modeling Route Choice of Utilitarian Bikeshare Users with GPS Data. *Transportation Research Record: Journal of the Transportation Research Board*, *2587*(1), 141–149. https://doi.org/10.3141/2587-17

Koch, T., & Dugundji, E. R. (2021). Taste variation in environmental features of bicycle routes. *Proceedings of the 14th ACM SIGSPATIAL International Workshop on Computational Transportation Science, IWCTS 2021*. https://doi.org/10.1145/3486629.3490697

Liu, H., Li, H., Rodgers, M. O., & Guensler, R. (2018). Development of road grade data using the United States geological survey digital elevation model. *Transportation Research Part C: Emerging Technologies*, *92*, 243–257. https://doi.org/10.1016/j.trc.2018.05.004

Lowry, M. B., Furth, P., & Hadden-Loh, T. (2016). Prioritizing new bicycle facilities to improve low-stress network connectivity. *Transportation Research Part A: Policy and Practice*, *86*, 124–140. https://doi.org/10.1016/J.TRA.2016.02.003

Luu, K. (2023). *stochopy: Python library for stochastic numerical optimization* (Version v2.3.0) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.4058008

Marshall, W. E., Piatkowski, D., & Johnson, A. (2017). Scofflaw bicycling: Illegal but rational. *Journal of Transport and Land Use*, *10*(1). https://doi.org/10.5198/jtlu.2017.871

Meert, W., & Verbeke, M. (2018). *HMM with non-emitting states for Map Matching*. ECDA, Paderborn, Germany.

Meister, A., Felder, M., Schmid, B., & Axhausen, K. W. (2022). *Route choice modelling for cyclists on dense urban networks*. https://doi.org/10.3929/ETHZ-B-000561710

Meister, A., Gupta, J., & Axhausen, K. W. (2021). *Descriptive route choice analysis of cyclists in Zurich*. https://doi.org/10.3929/ethz-b-000504160

Meister, A., Liang, Z., Felder, M., & Axhausen, K. W. (2024). Comparative study of route choice models for cyclists. *Journal of Cycling and Micromobility Research*, *2*, 100018. https://doi.org/10.1016/j.jcmr.2024.100018

Menghini, G., Carrasco, N., Schüssler, N., & Axhausen, K. W. (2010). Route choice of cyclists in Zurich. *Transportation Research Part A: Policy and Practice*, *44*(9), 754–765. https://doi.org/10.1016/j.tra.2010.07.008

Min, D., Zhilin, L., & Xiaoyong, C. (2007). Extended Hausdorff distance for spatial objects in GIS. *International Journal of Geographical Information Science*, *21*(4), 459–475. https://doi.org/10.1080/13658810601073315

Misra, A. (2016). *Mapping bicyclist route choice using smartphone based crowdsourced data* [PhD Thesis]. Georgia Institute of Technology.

Misra, A., Gooze, A., Watkins, K., Asad, M., & Le Dantec, C. A. (2014). Crowdsourcing and Its Application to Transportation Data Collection and Management. *Transportation Research Record: Journal of the Transportation Research Board*, *2414*(1), 1–8. https://doi.org/10.3141/2414-01

Misra, A., & Watkins, K. (2018). Modeling Cyclist Route Choice using Revealed Preference Data: An Age and Gender Perspective. *Transportation Research Record: Journal of the Transportation Research Board*, *2672*(3), 145–154. https://doi.org/10.1177/0361198118798968

MPC. (n.d.). *Bicycle & Pedestrian Planning*. https://www.thempc.org/Core/Bpp#gsc.tab=0

NACTO. (2014). *Urban Bikeway Design Guide*. Island Press. https://nacto.org/publication/urban-bikeway-design-guide/

Nassir, N., Ziebarth, J., Sall, E., & Zorn, L. (2014). Choice Set Generation Algorithm Suitable for Measuring Route Choice Accessibility. *Transportation Research Record: Journal of the Transportation Research Board*, *2430*(1), 170–181. https://doi.org/10.3141/2430-18

Newman, M. E. (2008). The mathematics of networks. *The New Palgrave Encyclopedia of Economics*, 1–12.

Newson, P., & Krumm, J. (2009). Hidden Markov Map Matching Through Noise and Sparseness. *17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems (ACM SIGSPATIAL GIS 2009), November 4-6,*

Seattle, WA, 336–343. https://www.microsoft.com/en-us/research/publication/hidden-markov-map-matching-noise-sparseness/

Osmium Tool. (n.d.). https://osmcode.org/osmium-tool/

Passmore, R. (2024). *BikewaySim Network Reconciliation* [Computer software]. https://github.com/gti-gatech/BikewaySim/blob/master/network/Step_2_Network_Reconciliation.ipynb

Passmore, R., Watkins, K. E., & Guesnler, R. (2021). BikewaySim Technology Transfer: City of Atlanta. *UC Davis: National Center for Sustainable Transportation*. https://doi.org/10.7922/G2CF9NDV

Passmore, R., Watkins, K., & Guensler, R. (2024). Using shortest path routing to assess cycling networks. *Journal of Transport Geography*, *117*, 103864. https://doi.org/10.1016/j.jtrangeo.2024.103864

People for Bikes. (n.d.). *Bicycle Network Analysis*. https://bna.peopleforbikes.org/#/

Prato, C. G., Halldórsdóttir, K., & Nielsen, O. A. (2018). Evaluation of land-use and transport network effects on cyclists' route choices in the Copenhagen Region in value-of-distance space. *International Journal of Sustainable Transportation*, *12*(10), 770–781. https://doi.org/10.1080/15568318.2018.1437236

Reichard, W. (2023). *Design and Implementation of an ADA Intersection Crossing Inspection Form by Undergraduate Students*. 2023 Transportation Research Board Annual Meeting.

Reichard, W., Allen, L., & Guensler, R. (2024). Design and Implementation of Bicycle Facility Inspection Forms for use in the United States. *Submitted for Presentation at the 2025 Annual Transportation Research Board and Publication in TRR*.

Reichard, W., Hatch, A., Adkins, B., Axelson, V., Bateman, V., Bentley, W., Cedeno, C., Changivy, L., Curtis, G., Goff, J., Hatch, A., Kornegay, C., Schmidt, C., Taylordean, L., Tillman, R., & Guensler, R. (2022). *Sidewalk, Ramp, Curb Cut, Crossing, and Bus Stop Inventory and Condition Assessment: Westside Study Area, Atlanta, GA.* Submitted to Atlanta City Council.

RSG. (2019). *NCHRP 08-36, Task 141 Evaluation of Walk and Bicycle Demand Modeling Practice*. *May*.

Schultheiss, W., Sanders, R. L., & Toole, J. (2018). A Historical Perspective on the AASHTO Guide for the Development of Bicycle Facilities and the Impact of the Vehicular Cycling Movement. *Transportation Research Record: Journal of the Transportation Research Board*, *2672*(13), 38–49. https://doi.org/10.1177/0361198118798482

Schweizer, J., Rupi, F., & Poliziani, C. (2020). *Estimation of link-cost function for cyclists based on stochastic optimisation and GPS traces*. *14*(13), 1810–1814. https://doi.org/10.1049/iet-its.2019.0683

Scikit-learn. (n.d.). *Demo of DBSCAN clustering algorithm*. https://scikit-learn.org/stable/auto_examples/cluster/plot_dbscan.html

Shah, N. R., & Cherry, C. R. (2021). Different Safety Awareness and Route Choice between Frequent and Infrequent Bicyclists: Findings from Revealed Preference Study Using Bikeshare Data. *Transportation Research Record: Journal of the Transportation Research Board*, *01764312*. https://doi.org/10.1177/03611981211017136

Skov-Petersen, H., Barkow, B., Lundhede, T., & Jacobsen, J. B. (2018). How do cyclists make their way? A GPS-based revealed preference study in Copenhagen. *International Journal of Geographical Information Science*, *32*(7), 1469–1484. https://doi.org/10.1080/13658816.2018.1436713

Tags. (n.d.). OpenStreetMap Wiki. https://wiki.openstreetmap.org/wiki/Tags

Ton, D., Cats, O., Duives, D., & Hoogendoorn, S. (2017). How Do People Cycle in Amsterdam, Netherlands?: Estimating Cyclists' Route Choice Determinants with GPS Data from an Urban Area. *Transportation Research Record: Journal of the Transportation Research Board*, *2662*(1), 75–82. https://doi.org/10.3141/2662-09

U.S. Census Bureau. (2010). *LEHD Origin-Destination Employment Statistics* [Dataset]. https://lehd.ces.census.gov/data/

USGS. (n.d.-a). *TNM Access API*. https://tnmaccess.nationalmap.gov/api/v1/docs

USGS. (n.d.-b). *What is a digital elevation model (DEM)?* https://www.usgs.gov/faqs/what-digital-elevation-model-dem

Valhalla. (n.d.). *Overview of how routes are computed*. Valhalla Docs. https://valhalla.github.io/valhalla/route_overview/

Walk score. (n.d.). https://www.walkscore.com

Watkins, K., Mokhtarian, P., Clark, C., & Passmore, R. (2019). BeltLine Bicyclist Facility Preferences and Effects on Increasing Trips. In *DOT National Transportation Integrated Search—ROSA P*. FHWA-GA-19-1638. https://rosap.ntl.bts.gov/view/dot/40257

Winter, S. (2002). Modeling Costs of Turns in Route Planning. *GeoInformatica*, *6*(4), 345–361. https://doi.org/10.1023/a:1020853410145

Zimmermann, M., Mai, T., & Frejinger, E. (2017). Bike route choice modeling using GPS data without choice sets of paths. *Transportation Research Part C: Emerging Technologies*, *75*, 183–196. https://doi.org/10.1016/j.trc.2016.12.009