



STREETLIGHT

INSIGHT

StreetLight Bicycle Volume Methodology and Validation White Paper

Version 2.0
July 2023

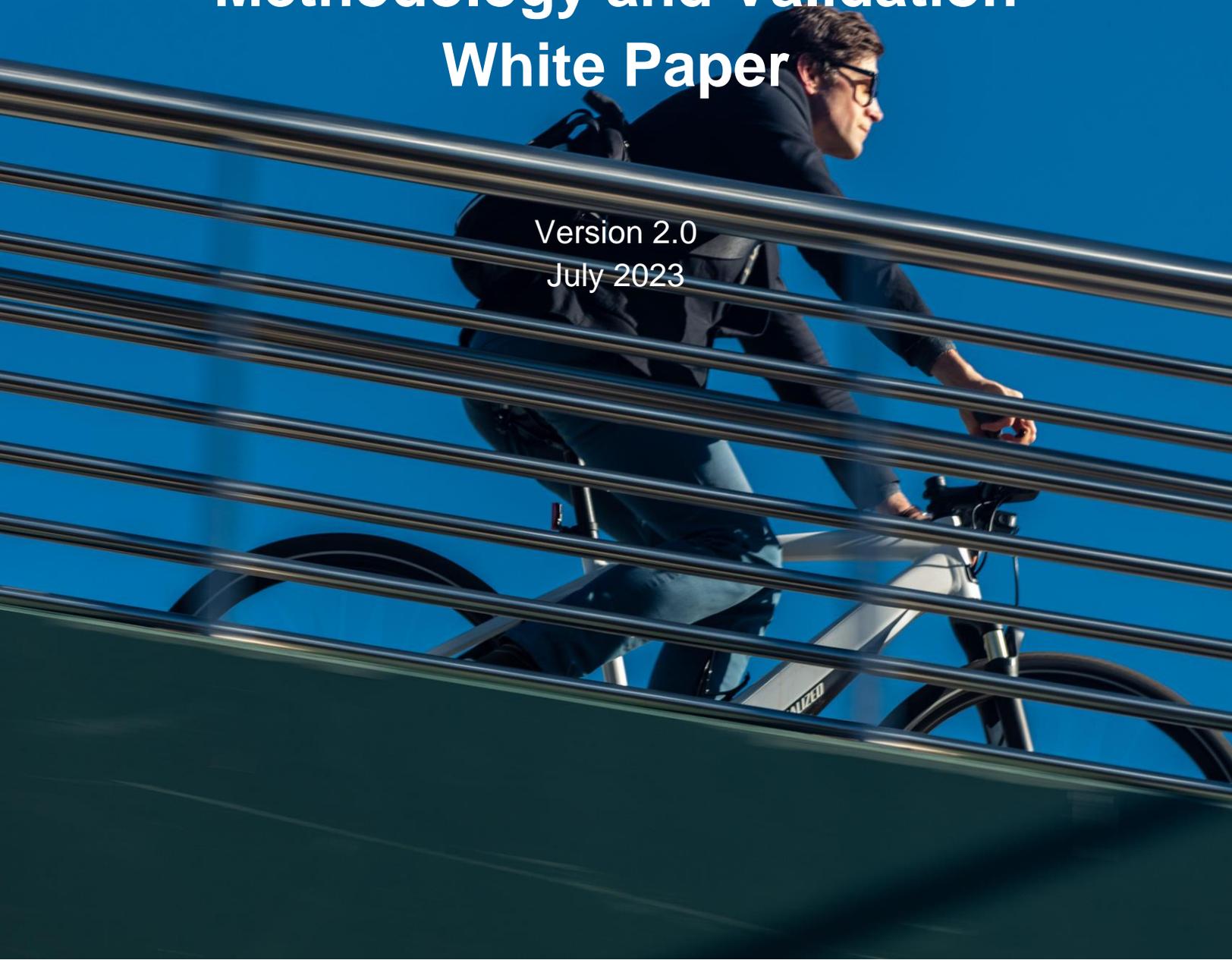


Table of Contents

Table of Contents.....	1
Introduction	2
What’s New in the July 2023 Release?	2
Bicycle Volume Methodology	2
Location Data Sources.....	3
Mode Classification.....	4
Bicycle Volume Model.....	5
Model Summary.....	5
Step 1: Anchoring to Vehicle Penetration Rates	5
Step 2: Scaling to Permanent Bicycle Counters	6
Bicycle Volume Validation	6
Permanent Counter Review	6
Counter Geographic Distribution	7
Counters Temporal Distribution	10
Volume Model Performance.....	10
Accuracy Targets	10
Absolute Error.....	12
Predictive Power	14
Seasonal Error	16
Volume Model Performance in Canada	19
Trip Attributes.....	20
Conclusions	21
About StreetLight	21

Introduction

The Bipartisan Infrastructure Law (BIL) sets aside funds for bikeways, new sidewalks, safe routes to school, and more, but cities, towns, and states struggle with bicycle data gaps for all road types. Good quality bicycle trip data is essential to plan for enhanced and safe bicycle use. Traditional count technologies are challenging because they have limited scope: permanent counters are available on few roads, temporary counts only capture limited hours and days, and surveys typically have small samples and focus on specific use cases like commute trips. Expanding any of these programs is expensive and time-consuming. In contrast, StreetLight Data, Inc. (“StreetLight”) Bicycle Volume Metrics are available for all types of roads, from urban commuting bike lanes to trails. This technical document describes the data sources and methodology employed by StreetLight Data to develop our Bicycle Volume Metrics. It also includes various validation exercises that demonstrate the comparability of our metrics versus ground-truth data from permanent bike counters. This document is updated regularly, so please check our website or reach out to your StreetLight contact to ensure you have the most up-to-date version.

What’s New in the July 2023 Release?

- 1. Improved Bicycle Trips Sample:** We made changes to our mode classification algorithm that filter out slow-moving vehicles from the bicycle trips sample. As a result, we have reduced mode-confusion, particularly in low density areas, airports and shopping malls.
- 2. Larger US Bicycle Counter Network:** Our new volume model is validated against data from 308 permanent bicycle counters from across the US, up from 228 counters in the previous version of the model. We have counts from at least 12 agencies and 75 counters every month, which is a significant improvement in temporal coverage over the previous model version.
- 3. Better Estimates on Low-Volume Roads:** Our effort to filter out slow-moving vehicles greatly reduces overestimation of bicycle volume on low-MADT roads. 95th percentile absolute error is now 48% lower than in the previous version.

Bicycle Volume Methodology

This section outlines the development process and details of our Bicycle Volume Metrics, which estimate the number of bike trips that occur. There are three main sections that correspond to steps in the Bicycle Volume development process.

- 1. Location Data Sources:** StreetLight has several location-based services (LBS) data suppliers, whose data we aggregate into trips and use to validate models.

2. **Mode Classification:** Bicycle trips are identified in our LBS data and differentiated from trips that use other modes (car, bus, etc.).
3. **Bicycle Volume Model Development:** Bicycle Volume is estimated using our bicycle trip sample and data from permanent bicycle and vehicle counters.

The first two sections of this paper solely focus on data and processes specifically relevant to bicycle trips. (For a more thorough description of these steps for all modes, [see the StreetLight Methodology white paper.](#)) The third section of this paper outlines the development of the Bicycle Volume model.

Location Data Sources

The primary data source for our bicycle trips sample is LBS. LBS data has a high spatial precision and regular ping rate. A 2020 literature review paper found that LBS data is well-suited for capturing trips as well as activity patterns.¹

In addition to LBS data, we also rely on several other data sources. These sources are specific to our bicycle algorithms and are used to complement existing LBS data and validate our models.

Data Type	Primary Use	Description
Location-based services (LBS)	Trip sample and metrics	Anonymized location data collected from mobile apps
Geographic features	Trip sample, mode classification	Features from OpenStreetMap (OSM) and the Census are used in bike classifier model

¹ Leea, Kyuhyun and Sener, Ipek. “Emerging data for pedestrian and bicycle monitoring: Sources and applications.” *Transportation Research Interdisciplinary Perspectives*. March 2020. <https://www.sciencedirect.com/science/article/pii/S2590198220300063?via%3Dihub>

Permanent bicycle counts	Volume modeling, validation	Trip counts from permanent bike counters along roads and trails
User-tagged location data	Mode classification	Data from GPS-enabled travel diaries and a cohort of contract data trainers that cover different urban forms and weather regimes who are hired to collect GPS information while biking and send their logs and user-entered mode tags to StreetLight.
Traditional surveys about mode behavior	Validation	Data collected by surveyors about different transportation modes (e.g., average bike trip length)

Table 1. Location and other data sources for bike algorithms

Mode Classification

Mode classification is an essential step in creating our bicycle trip sample. Since LBS data is derived from smartphones, we have to further process the data to identify the mode of transportation. To do this, we have built a multi-pass algorithm to identify the mode of travel associated with each LBS ping using a combination of heuristic and probabilistic algorithms.

Before isolating bicycle trips, we must first identify other modes of travel. Pings attributed to walk, rail, air, and ferry trips are identified using heuristic approaches, as these trips have unique movement patterns that are easily explained. Rail and ferry trips do not use the road network and travel along predefined paths, while air and walk trips travel at extremely high and low speeds, respectively. These pings are removed from consideration before determining whether a ping is associated with a bicycle trip.

Pings associated with bicycle trips are further differentiated from those associated with vehicle and bus trips using a classification algorithm. These modes all share the same road network and follow similar movement profiles, so it can be difficult to distinguish them. To classify bicycle trips, first we use machine learning techniques, specifically a Random Forest model, to assign mode probabilities to every received ping. The model incorporates several features, including speed, distance, time of day, and presence of bike and bus lanes, to assign the mode probabilities.

Mode-tagged pings are then sent to trip creation algorithms, where individual pings are grouped into trips according to movement patterns and expected mode. Finally, mode-tagged trips are filtered with additional quality checks based on overall trip characteristics and geographic context before being included in our trip sample. (For more detail on mode classification, see our [Methodology and Data Sources white paper](#).)

Bicycle Volume Model

We use our Bicycle Volume model to calculate Bicycle Volume — an estimate of the total number of bicycle trips. The model creates a volume estimate using our bike trip sample and other inputs.

Penetration rate, or the proportion of real-world trips captured in our sample, is a critical factor in estimating Bicycle Volume. Our sample varies significantly across space and time, so it's important to normalize across these dimensions.

- The penetration rate fluctuates with changes in data supply. When our suppliers add and remove apps from their LBS data sets, or when popularity and usage of the included apps change, it is reflected in our sample.
- The penetration rate also changes across geography. Regional differences in app usage can affect the penetration rates in different cities or states.

The goal of our Bicycle Volume model is to be able to account for these changes in penetration to accurately estimate bicycle travel across months, years, and geographic regions.

MODEL SUMMARY

The Bicycle Volume estimate is calculated through two main steps:

1. First, we use StreetLight's vehicle penetration rate near the target zone to estimate bike penetration rate.
2. Then, we increase the Bicycle Volume estimate by a constant factor to index more closely to permanent bicycle counters.

STEP 1: ANCHORING TO VEHICLE PENETRATION RATES

The first step in estimating Bicycle Volume uses Vehicle Volume along nearby roads to approximate the bicycle penetration rate. Since both our bicycle and vehicle trip samples are derived from same LBS data, we observe that the two modes have similar penetration rate distributions across space and time. For more details on Vehicle Volume and penetration rate, see our [StreetLight All Vehicles Volume Methodology and Validation white paper](#).

There are several advantages to using Vehicle Volume and penetration rate to estimate Bicycle Volume:

- There is a much more robust network of permanent vehicle counters than bicycle counters. Our Vehicle Volume model is trained on data from over 5,000 permanent vehicle counters, and they are far more geographically diverse than our bike counters.
- The vehicle trip sample is less sensitive to mode confusion because the vast majority of trips on most roads are vehicle trips. Consider a road with 90% vehicles and 10% bikes.

If the mode classifier incorrectly classifies 10% of bike trips as vehicle trips, the total vehicle trip sample increases by only 1.1%, while if the classifier incorrectly classifies 10% of vehicle trips as bike trips, the bike trip sample would increase by 90%.

- The vehicle trip sample is significantly larger than the bike sample. Given the larger sample size, we expect more stability in the sample, allowing for comparison over time.

STEP 2: SCALING TO PERMANENT BICYCLE COUNTERS

The second step is to adjust our Bicycle Volume estimates according to the national difference in bicycle and vehicle penetration rates. Vehicle penetration rates help us estimate Bicycle Volume, but vehicle penetration will never exactly match bicycle penetration. Demographic and behavioral differences across users of the two modes can lead to different sample sizes, and imperfect mode classification can lead us to over- or under-express one mode in our trip sample relative to the next.

We can measure our bicycle penetration rate in several locations by comparing our trip sample to data from permanent bicycle counters provided by various transportation agencies. We chose to train the model only on permanent counters that report counts every month or more frequently because this ensures that we can test the model at the monthly level.

Bicycle Volume Validation

This section includes a review of our ground-truth data and permanent bike counters, and it outlines the performance of our Bicycle Volume model. We focus on model performance in the United States and include ways that the Canada model performance differs.

Permanent Counter Review

In this section, we discuss the geographic and temporal coverage of the 308 permanent bike counters used in model assessment.

Bike counter data was retrieved from transportation agencies and active transportation data clearinghouses wherever available. It is important to characterize the quality of our bike counter data because, though we believe permanent bike counters are the most reliable source of training data for our model, they still face significant challenges in reliably capturing real-world

behavior.² Bike counters vary in quality and sophistication and are known to miss trips due to technical malfunctions or weather conditions, or if riders simply go around the counter. Additionally, they are subject to downtime and typically require both calibration and post-processing for best results. These factors, combined with limited geographic coverage compared to vehicle counters, need to be taken into consideration when evaluating model performance results.

We're constantly looking for new data sources to improve our Bicycle Volume estimation and validation efforts. If you're interested in contributing bike counter data, please reach out.

COUNTER GEOGRAPHIC DISTRIBUTION

While we have data from permanent bike counters all over the country, the counters are not evenly distributed. Bike counters are most common in areas with better bike infrastructure and where biking is popular. The counters we rely on cover regions including the northeast corridor, California, Atlanta, Texas, Minnesota, and Colorado.

² Ozan, Erol et al. "State-of-the-Art Approaches to Bicycle and Pedestrian Counters." NCDOT and Eastern Carolina University. Mar 2021.
<https://connect.ncdot.gov/projects/research/RNAProjDocs/RP2020-39%20Final%20Report.pdf>

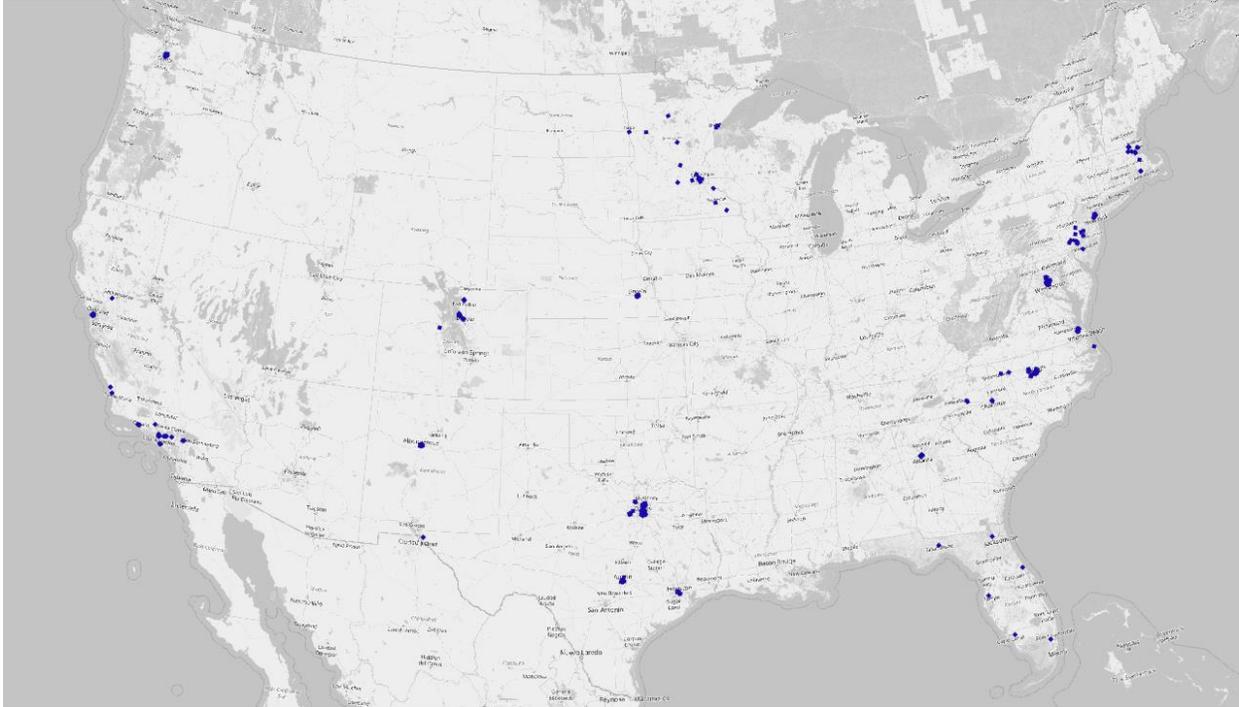


Figure 1. U.S. permanent bike counter locations

We mitigate this geographic bias by anchoring estimates to vehicle penetration rate, which is developed using a much more robust network of over 5,000 vehicle counters. These counters are much more reliable, and they are well distributed across the U.S.

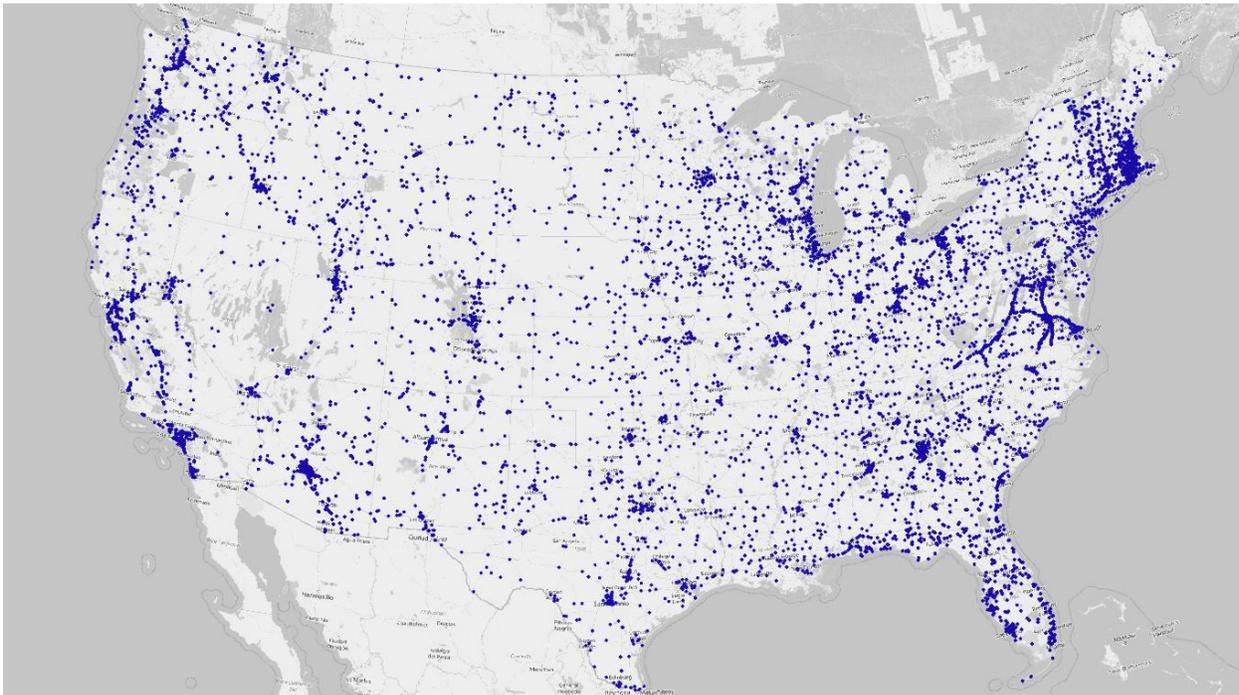


Figure 2. U.S. permanent vehicle counter locations

Permanent bike counters tend to be located in and around cities, while only a small number of bike counters are located in rural areas. See Table 2 below for the distribution of counters by urban classification.

Classification	Number of Counters	Proportion of Counters
Urban	146	47%
Suburban	107	35%
Rural	55	18%

Table 2. Bike counter distribution by urban density

Finally, bike counters are disproportionately distributed across various road volume categories and facility types:

- Bike counters are located disproportionately along bike trails. Most bike trails only have cyclists and some pedestrians, making mode classification much easier. It is more difficult to estimate volume along roads where mode confusion with personal vehicles, buses, and other modes of transportation is more likely.
- Bike counters are more common on low-volume roads. Nearly half of our counters are on roads with under 150 daily bicycles. This is expected, as most U.S. roads have relatively few bicycles.

Bicycle Volume (MADT)	Number of Counters	Percent of Counters	Road Type	Number of Counters	Percent of Counters
Low (25–150)	139	45%	Bike Trail	158	51%
Medium (150–500)	91	30%	Other Roads	150	49%
High (500-1000)	41	13%			

Very High (1000+)	37	12%
----------------------	----	-----

Table 3. Bike counters by bicycle traffic size (left) and type (right) of road.

COUNTERS TEMPORAL DISTRIBUTION

Counter data was not available for every counter every month. Counters sometimes malfunction and need to go offline, and some counters are only active for part of the measurement period. We have counts from at least 12 agencies and 75 counters every month.

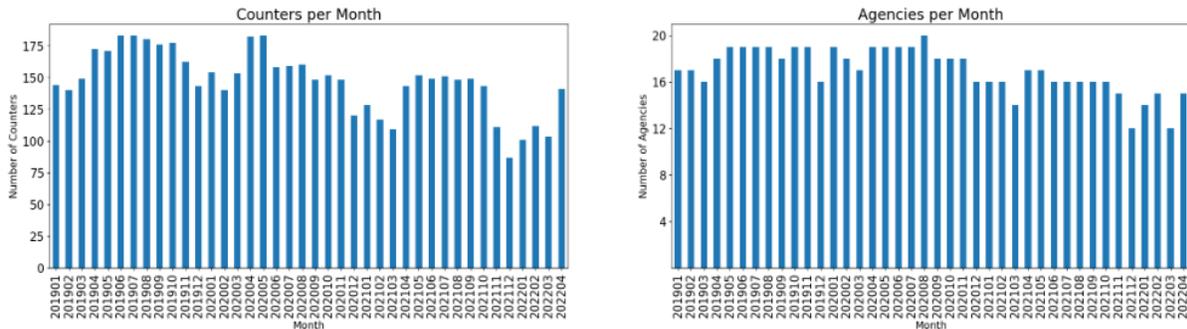


Figure 3. Counters (left) and agencies reporting counts (right) per month

Volume Model Performance

In this section, we compared the model’s predicted Bicycle Volume to counts from the 308 permanent counters described in the previous section.

ACCURACY TARGETS

To define meaningful Bicycle Volume model accuracy targets, we appealed to three primary criteria:

1. Predictions can be less accurate on a percentage basis if the absolute difference in Bicycle Volume is small. This is especially relevant along roads with few bike trips. Consider a road with 50 bike MADT. If we predict a volume of 100 bike trips, our prediction is off by 100%, but it is a difference of only 50 daily bike trips (about 2 bike trips per hour).
2. Accuracy should be consistent or better than results from traditional, short-duration counts or surveys. Most bike counts today rely on short-duration count programs and may use seasonal/annual adjustment factors to estimate Average Annualized Daily Bike Traffic (AADBT). A 2018 TRB paper by WSDOT/Toole Design/UNC Safety Highway Research Center documented Mean Absolute Percent Error (MAPE) as

compared to permanent counters for expanded short-duration counts to range between 30–50% depending on the duration of the count.³ (See Table 4.)

3. Accuracy should be consistent or better than leading research for LBS-derived bike count estimates done by academic institutions and transportation agencies. A 2021 Pooled Fund Study by Portland State University’s TREC used the most thorough and sophisticated methods we’ve found to date. By combining LBS data with static models, they achieved between 42% and 271% MAPE depending on road volume classification.⁴

Based on these guides, we’ve established targets for low-, medium-, and high-activity roads (see Table 4 below). It’s important to note that while the MAPE targets for low-activity roads may seem high, higher targets on small roads are typical for LBS data due to the sample-size limitations. We’ve also included reporting metrics for median percentage error, which may more closely represent expected error in an individual zone than MAPE.

In addition to absolute error, the sections that follow include analysis of predictive power (correlation) and temporal analysis (time trends) to more comprehensively illustrate how the volume model compares to ground-truth data.

Bicycle Volume (MADT)	Target MAPE	Typical MAPE for Short-Duration Counts
Low (25–150)	200%	<i>Data Not Available</i>
Medium (150– 500)	50%	31%
High or Very High (500+)	40%	29%

Table 4. MAPE error targets by monthly average daily bike traffic (bike MADT)

³ Johnstone, Dylan et al. “Annual Average Nonmotorized Traffic Estimates from Manual Counts: Quantifying Error.” Journal of the Transportation Research Board. Aug 2018.

<https://journals.sagepub.com/doi/10.1177/0361198118792338>

⁴ Kothuri, Sirisha et al. “Exploring Data Fusion Techniques to Estimate Network-Wide Bicycle Volumes.” Transportation Research and Education Center. Sep 2021.

https://trec.pdx.edu/research/project/1269/Exploring_Data_Fusion_Techniques_to_Derive_Bicycle_Volumes_on_a_Network

ABSOLUTE ERROR

Bicycle Volume model predictions are shown in the table below. Main takeaways include:

- Overall, our estimates are within the target MAPE on low- and medium-volume roads. Estimates are slightly above targets on high-volume roads.
- While our volume model performs better on higher-volume roads, it performs worse relative to targets because our targets on high-volume roads are more ambitious.
- Median absolute percent error performs significantly better than mean absolute percent error, especially on low-volume roads. This means that the error distribution is skewed such that a few estimates with very high error are disproportionately affecting MAPE on low-volume roads, which is expected.
- The volume model performs better on bike trails than on other roads. Cars and buses can make identifying bike trips on other roads more difficult, and we do not have this problem along trails that only allow bicycles and pedestrians.

Table 5 shows mean and median absolute percent error by road size, which are the typical levels of error observed.

- While mean and median show expected error levels, error necessarily exceeds the mean and median along some roads. To illustrate this, we also include a distribution of error above the median. You can interpret the *n*th percentile as “the absolute percent error is at or below this value on n% of roads.”

Bicycle Volume (MADT)	Target Absolute Percent Error	Mean	Absolute Error Distribution		
			50 th Percentile (Median)	68 th Percentile	95 th Percentile
Low (25–150)	200%	101%	57%	83%	289%
Medium (150–500)	50%	49%	40%	55%	126%
High (500-1000)	40%	47%	35%	55%	116%
Very High (1000+)	40%	44%	45%	58%	84%

Table 5. Absolute error by monthly average daily bicycle traffic (bicycle MADT)

In Table 6, we show the median absolute percent error along roads and trails. Our high-level targets are based on mean absolute percent error because external research on measurement alternatives focuses on the mean, but median error is a better representation of the expected error level in a given zone since outliers with very high errors (see 95th percentile above) affect the median less than the mean.

Bicycle Volume (MADT)	Road Type	Median Absolute Percent Error
Low (25–150)	Road	70%
	Trail	46%
Medium (150–500)	Road	40%
	Trail	39%
High (500-1000)	Road	39%
	Trail	31%
Very High (1000+)	Road	51%
	Trail	26%

Table 6. Median absolute percent error on roads and bike trails

Finally, model performance varies depending on the agency reporting bike counts. This is partially driven by the proportion of counters along trails — excluding agencies with few total counters, agencies with more counters on trails tend to see lower error. Again, we show median absolute percentage error to convey the typical error level.

Agency	Counters	Points (Counter-Months)	% Counters on Trails	Median Absolute Percentage Error
DVRPC	13	366	100%	31%
TXDOT	55	753	92%	37%

Bike Arlington	44	1235	52%	37%
NCDOT	24	71	55%	38%
Austin	12	310	80%	40%
Norfolk	13	131	23%	41%
MnDOT	19	428	57%	44%
Boulder	15	186	39%	45%
NYC	10	258	16%	53%
SFMTA	19	739	0%	53%
SCAG	21	154	46%	55%
MassDOT	12	58	16%	69%
Other (10 agencies)	51	1192	36%	53%

Table 7. Median absolute percentage error by agency

PREDICTIVE POWER

Predictive power describes the model's ability to consistently estimate higher volume where counts are high, regardless of directional bias. The model's overall is $R^2 = 0.57$, so over half of the variation in StreetLight Volume estimates are explained by the actual variation in bike counts.

For comparison, the correlation between StreetLight All Vehicles Volume and permanent vehicle counts is much stronger (0.98). Correlation generally improves as a model is trained on more and better training data – in this case, more bicycle counters and a larger bicycle trip sample. We do not expect to create a Bicycle Volume model that predicts bicycle counts as well as our All Vehicles Volume model predicts vehicle counts because there are far fewer permanent bicycle counters and bicycle trips than vehicle counters and vehicle trips.

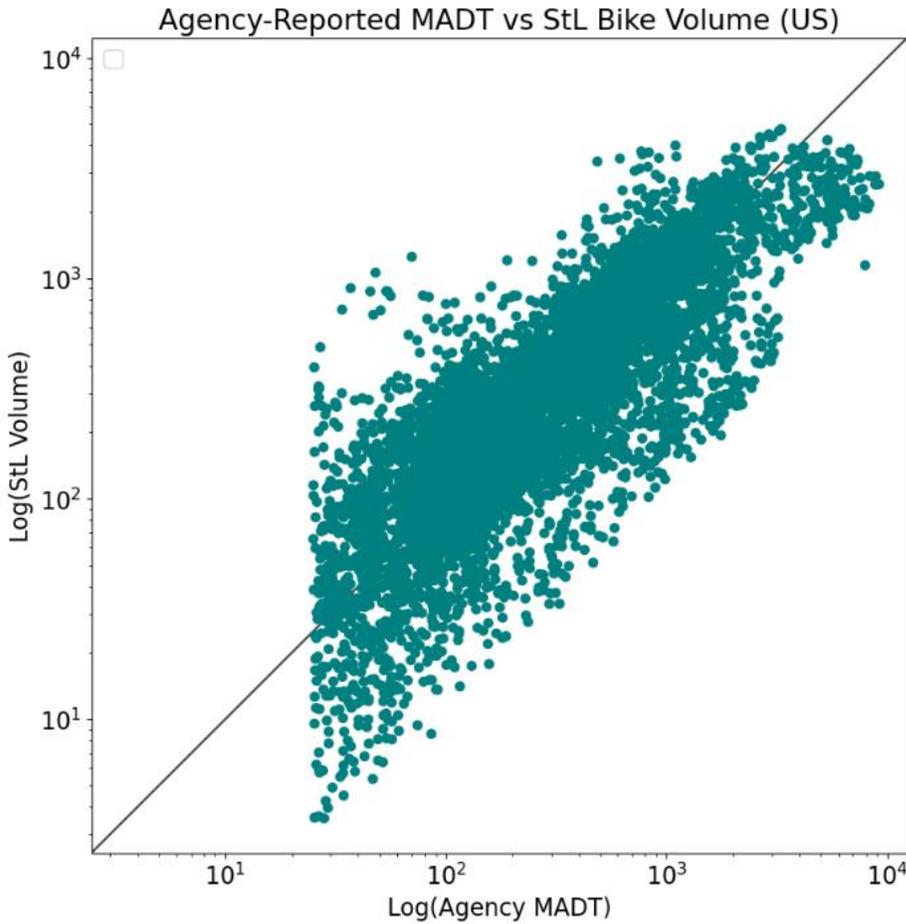


Figure 4. Relationship between bike counts and StreetLight Volume estimates

It's important to keep in mind that there can be strong correlation between bike counters and StreetLight Volume even if volume is consistently over or underestimated. For example, our volume estimates in Atlanta are almost always below bike counts, but R^2 in Atlanta (0.83) is still better than in Texas (0.67) where there is less directional bias. In locations with strong correlation and consistent bias, customers should achieve the best results when using local calibration data to perform their analyses, which can be done selecting the "Single Factor Calibrated Index using user counts (Bicycle Trips)" output option when creating an analysis.

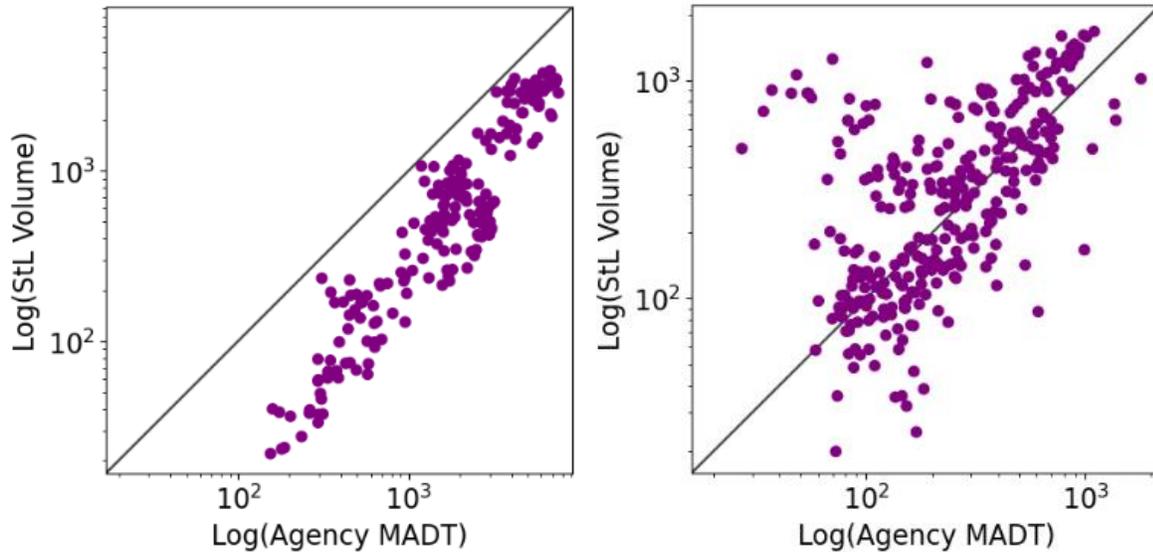


Figure 5. Atlanta (left) and TXDOT (right) counters vs. StreetLight Volume estimates

SEASONAL ERROR

In some regions, bike activity varies dramatically depending on the season, and seasonal fluctuations may not be accurately reflected in short-duration counts, even after adjustments. This is one of the big advantages of LBS data, because even when there are high levels of absolute error, we see very tightly correlated seasonal trends when compared to ground-truth data. This section describes methods and observations from analyzing temporal trends in Bicycle Volume model performance.

To evaluate seasonal trends, we focused on our model’s ability to predict an individual month’s change relative to a three-month rolling average. Results are segmented based on the same four volume categories and are plotted alongside sample counts to illustrate the impact of the Bicycle Volume model.

LBS seasonality typically mirrors counter seasonality. Still, there are clear points in time where seasonal change in the StreetLight bicycle trips sample, shown in orange, varies from the seasonal pattern of permanent counters (blue). The Bicycle Volume model (green) adjusts trip counts to follow a seasonal pattern closer to counters (e.g., when there were dramatic biking increases during spring 2020 COVID lockdowns, when data suppliers have large fluctuations, etc.).

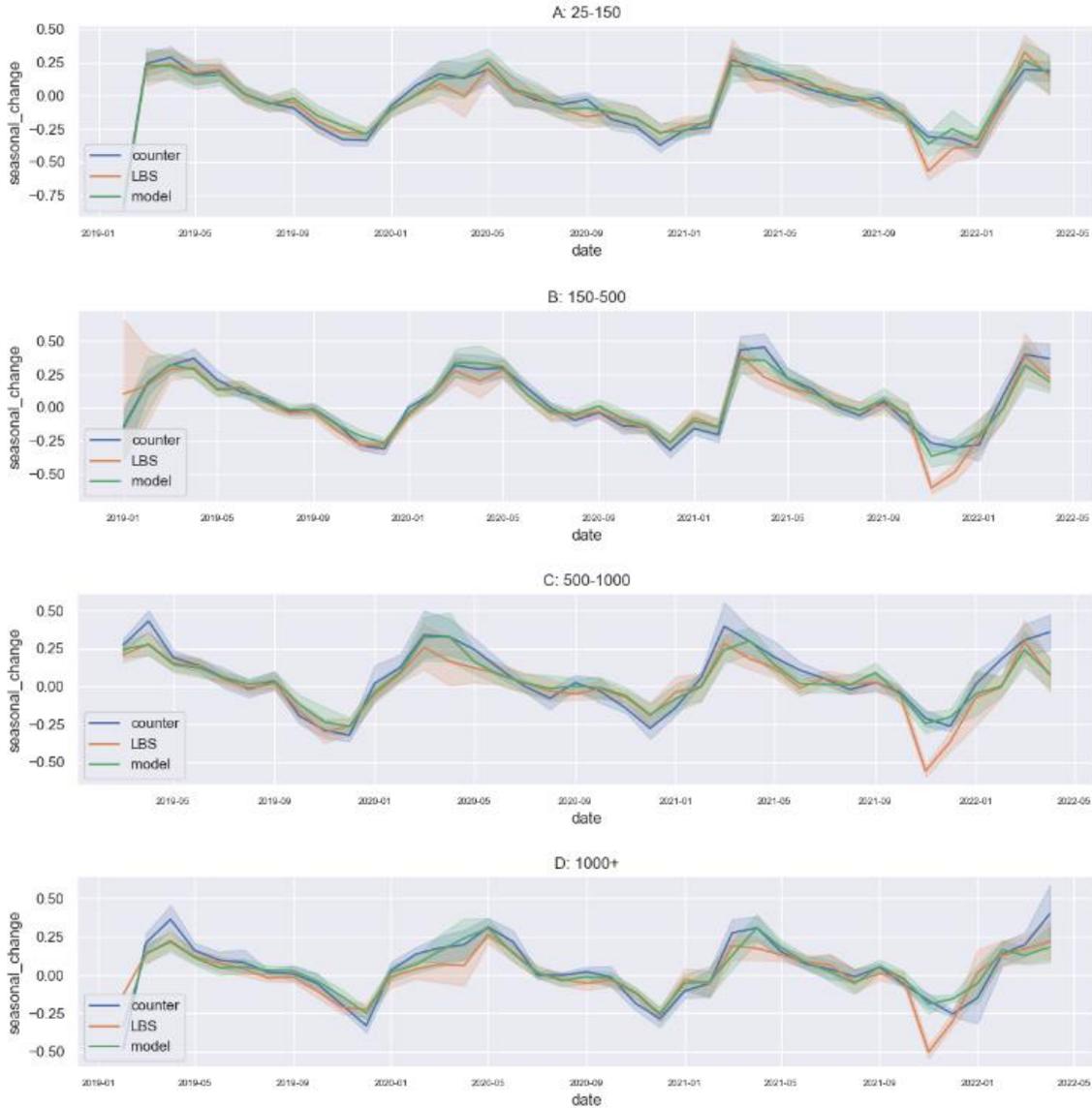


Figure 6. Seasonal (month/rolling 3-month) change across all zones for bike counter counts (counter), LBS total trip sample (LBS) and modeled bike trips (model) by road size

We applied benchmarks of “within 20 percentage points” and “within 10 percentage points” to illustrate how accurately the Bicycle Volume model predicts the month/rolling 3-month trend at individual counter locations. The model successfully predicts within 20 points in most cases and within 10 points almost half the time. It performs better on higher-volume roads.

Bicycle Volume (MADT)	% Within 20 PPTS month/rolling 3-month Error	% Within 10 PPTS month/rolling 3-month Error
Low (25–150)	54%	30%

Medium (150–500)	68%	43%
High (500-1000)	75%	46%
Very High (1000+)	79%	53%

Table 8. Frequency of routes within seasonal error targets by road size

The examples below demonstrate how model performance, despite high absolute error, can still closely follow temporal trends:

- In the first example — Chester Valley Trail outside Philadelphia — the volume model predicts both bike counts and seasonal changes well.
- In the second — White Rock Creek Trail in Dallas — the volume model consistently underestimates bike counts by about 50%, but MoM change in volume is still within 20% of MoM change in counts during most months.

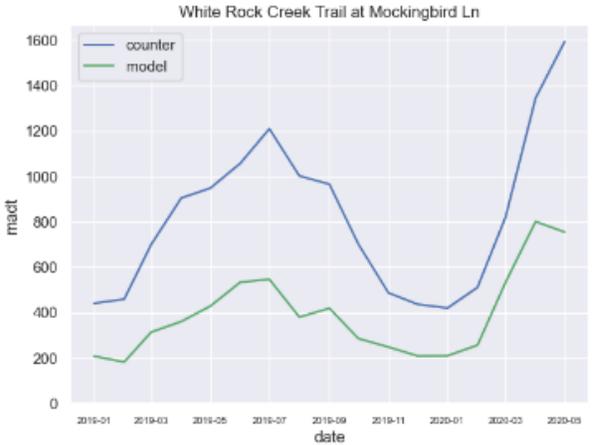
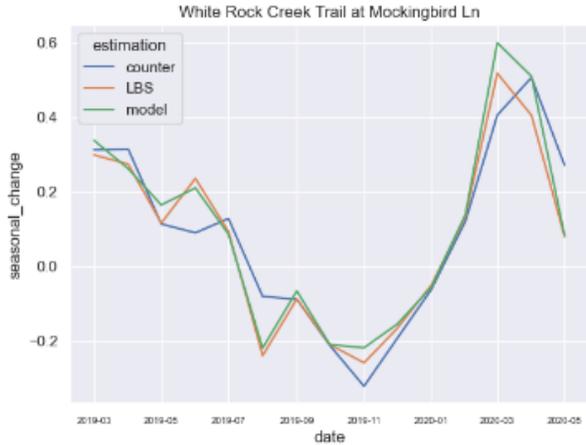


Figure 7. Seasonal change (left) and volume (right) compared on sample routes

VOLUME MODEL PERFORMANCE IN CANADA

In Canada, we have data from 175 permanent bike counters, which is fewer than in the United States. We also have fewer permanent vehicle counters. Given the smaller training set, error levels are unsurprisingly higher in Canada for small and mid-sized roads. Error levels are similar on large roads. Strong performance on high-volume roads may be driven by a larger presence in the counter set — 46% of Canadian counters are on high-volume roads (compared to 25% of U.S. counters).

Bicycle Volume (MADT)	MAPE Target	Mean Absolute Percent Error	Median Absolute Percent Error
Low (25–150)	200%	415%	251%
Medium (150–500)	50%	130%	58%
High (500+)	40%	52%	44%

Table 9. Absolute error by monthly average daily bike traffic (bike MADT)

Overall error levels in Canada are significantly higher due to the poor model performance on counters located in Winnipeg. See agency-level performance in table 10.

Agency	Counters	Points (Counter-Months)	% Counters on Trails	Median Absolute Percentage Error
Calgary	64	567	5%	49%
Ottawa	37	358	12%	49%
Edmonton	25	200	2%	59%
Montreal	41	277	12%	63%
Winnipeg	8	93	20%	182%

Table 10. Median absolute percentage error by agency

Before relying on Bicycle Volume estimates in Canada along low- and mid-volume roads, we recommend you reach out to your StreetLight point of contact to ensure that we have sufficient sample trips in that location.

Trip Attributes

In this section, we compare bicycle Trip Attributes, such as distance and speed, to two surveys — the National Household Travel Survey 2017 (NHTS)⁵ and the California Household Travel Survey 2012-2013 (CHTS)⁶.

There are several limitations in comparisons to travel surveys. Surveys have much smaller sample sizes and tend to yield biased responses. Additionally, the CHTS is older and may not reflect contemporary biking behavior with E-bikes and improved bicycle infrastructure. Small differences in attributes between surveys and StreetLight trips should be expected, but the comparison can still help validate that the attributes from our sample are reasonable.

On average, StreetLight bike sample attributes skew longer than those reported in the surveys, shown in Table 11 below. This may be due to the increased use of E-bikes and the existence of long, athletic bike rides in our trip sample. Median trip attributes are much closer to survey results, as they are less affected by a small proportion of very long trips.

Attribute	NHTS	CHTS	StreetLight - National	
			Average	Median
Trip Length	2.38	1.50	3.26	2.21
Duration	21.92	18.20	32.28	24.28
Speed	6.51*	4.95*	6.65	6.39

**Imputed from trip length and duration*

⁵ Federal Highway Administration. “2017 National Household Travel Survey,” U.S. Department of Transportation. 2017. <https://nhts.ornl.gov>

⁶ National Renewable Energy Laboratory. “2010-2012 California Household Travel Survey,” U.S. Department of Energy. 2013. <https://www.nrel.gov/index.html>

Table 11. Average trip attributes from surveys and StreetLight trip sample

Conclusions

In this white paper, we describe the methodology used to derive StreetLight bicycle trips and how those trips are then expanded to represent real-world Bicycle Volumes. Our results are evaluated against industry research. Model validation results showed that our estimates fall within absolute error ranges expected from both traditional short-duration counters and best-available emerging modeling techniques. Furthermore, we demonstrate how our sample's predictive power and temporal responsiveness can be useful for estimating Bicycle Volume, even for zones with high absolute error.

Because of the significantly lower absolute volumes of bicycle trips as compared to vehicle trips and the much lower number of permanent counters available, it should be noted that customers should expect higher error (MAPE > 40% for all road sizes) as compared to what they may be used to when reviewing vehicle volumes (MAPE < 10% for most road sizes) when comparing StreetLight results to permanent counters. For best results, we recommend customers use calibration data when local, high-quality counts are available for the time period being analyzed in order to best take advantage of StreetLight's predictive power and temporal responsiveness.

We are constantly working to improve our bike classification, volume estimation, and validation methods. To do so, it's critical to assemble comprehensive and representative ground-truth data. If you have access to bike count data that you'd like to be included in future iterations of this work, please contact your StreetLight representative.

About StreetLight

StreetLight Data, Inc. ("StreetLight") pioneered the use of Big Data analytics to shed light on how people, goods, and services move, empowering smarter, data-driven transportation decisions. StreetLight's proprietary data processing engine, Route Science® algorithmically transforms its vast data resources to measure travel patterns of vehicles, bicycles and pedestrians, accessible as analytics on the StreetLight InSight® SaaS platform. Acquired by Jacobs as a subsidiary in February 2022, StreetLight provides innovative digital solutions to help communities reduce congestion, improve safe and equitable transportation, and maximize the positive impact of infrastructure investment. StreetLight powers more than 10,000 global projects every month. For more information, please visit: www.streetlightdata.com.



STREETLIGHT DATA

© StreetLight Data 2023. All rights reserved.