StreetLight All Vehicles Volume Methodology and Validation White Paper

United States

Updated: November 2023

Table of Contents

Introduction

StreetLight Data, Inc. ("StreetLight") All Vehicles Volume provides a quick, easy, and cost-effective way to measure average hourly or daily traffic at the yearly, monthly, or user-defined time period. StreetLight's Volume estimates are available for any location, such as a road, park, Traffic Analysis Zone (TAZ), or a user-defined area. It can also be used to estimate zone-to-zone traffic, providing comprehensive estimates for projects such as intersection studies and travel demand modeling. StreetLight All Vehicles Volume metrics are available for both the U.S. and Canada.

Over the years, StreetLight has provided Volume metrics estimated from both Location-Based Services (LBS) data and Connected Vehicle Data (CVD) sources. In July 2023, StreetLight updated its All Vehicles CVD+ Volume metric, leveraging Connected Vehicle Data (CVD), to take advantage of expanded validation data. All Vehicles CVD+ Volume metrics are available for data months starting with January 2022. For months prior to January 2022, we continue to offer All Vehicles LBS+ Volume. This most recent version of this white paper describes the methodology and validation for both CVD+ and LBS+ All Vehicles Volume metrics.

This white paper provides technical detail about the methodology, algorithm development, validation, and data sources used in StreetLight's All Vehicles Volume output for the U.S. This white paper was first published in October 2019 and is updated periodically to reflect the most recent methodology and validation.

What's New in All Vehicles Volume

The improved All Vehicles CVD+ Volume algorithm includes the following enhancements.

- **Increased validation data coverage:** The updated Volume models are trained and validated using data from more than 10,000 permanent counters in 48 states. The previous version used about 5,000 counters in 22 states. Nearly doubling the number of training sites enables expanded visibility into model performance and improved accuracy in locations with new counters.
- **New activity-based weighting factor as input feature:** A new activity-based weighting factor better accounts for spatial variation in trip sample. Including this feature helps improve overall accuracy.
- **Improved handling of OSM edge cases:** The previous CVD+ Volume algorithm did not robustly handle the rare cases in which OpenStreetMap data has changed significantly. The updated algorithm is more consistent and performs well in these cases.
- **Increased consistency in month-over-month time trends:** In the previous version, a new Volume model was trained each month as we received new CVD data, which resulted in occasional discontinuities in monthly time trends. The updated Volume metric uses just one model for the entire time period, producing more reliable time trends.

Methodology

To generate StreetLight Volume metrics, we start with sample trips and then build models to combine these trip counts with contextual information, such as demographics, road characteristics, weather, and seasonality. Estimated Volume is validated against traffic data collected by state and regional agencies.

Model Overview

We employ multiple machine-learning models, pulling from several data sources, to estimate StreetLight Volume. These models have gone through several iterations over the years as we have gained access to additional data sources and learned from experience, as well as validations conducted by our customers.

In the following sections, we use the terms Volume, MADT, and AADT. Volume refers to traffic volume generally and can apply to any time period. MADT (Monthly Average Daily Traffic) is a more specific metric that refers to the average daily traffic volume in a given month. The term AADT (Annual Average Daily Traffic) refers to the average daily traffic volume over a specific year.

All Vehicles Volume Versions

Since we first released StreetLight All Vehicles Volume in 2019, we have continued to improve the algorithm, incorporating the most up-to-date data and introducing more sophisticated modeling techniques. This white paper discusses methodology and validation for the most recent algorithm versions for the U.S., as summarized in Table 1.

Algorithm Version	Date of Most Recent Release	Key Improvements
All Vehicles LBS+ Volume	June 2022	Improved accuracy and historical trends
All Vehicles CVD+ Volume updated	November 2023	Higher accuracy, better edge case handling, smoother time trends

Table 1: Summary of the most recent All Vehicles Volume algorithm versions for the U.S. See the recent data period availability here.

All Vehicles LBS+ Volume Algorithm

The LBS+ Volume models support StreetLight's All Vehicles LBS+ Volume metric for the data periods seen here. These models predict vehicle volumes by combining Location-Based Services (LBS) trips with contextual features that represent the street network, demographics, climate, and other geographic characteristics. We refer to the models in the plural because the algorithm uses a pair of models, one for most roads and one specialized for low-volume roads. The following sections detail key aspects of these models.

Input Features

The All Vehicles LBS+ Volume models use the following sources of input data:

- **LBS trips:** Our processing algorithms link pings from LBS devices into trips, then classify these trips into certain travel modes. The Volume models use trips that we have identified to be from vehicles.
- **U.S. Census data:** We use demographic data, such as number of housing units and whether or not a location is in an urban area.
- **Population factor:** Using Census data, this feature scales the number of LBS trips to our best guess of the total number of trips based on the population.
- **OpenStreetMap (OSM) data:** OSM provides information on road classification, number of lanes, speed limit, and other figures.
- **Weather data:** Temperature, precipitation, and visibility affect travel and help us better predict seasonal changes.

Ground-Truth Data

To train and validate the models, we rely on ground-truth data, mainly traffic volume collected by state agencies. For the LBS+ Volume models, we acquired and processed permanent traffic counter data from over 6,000 unique counter locations across the U.S., for the years 2019 through 2022.

Table 2 and Figures 1–2 show the locations of counters used in U.S. All Vehicles LBS+ Volume model development. Counter locations span 49 states, plus the District of Columbia, and represent a mix of small and large roads in urban and rural locations. At the time of model development, not all counter data was available for 2021 or 2022, so those months' models use data from 22 states. Even so, the 2021–2022 data represents a range of regions and location types, as shown in the map in Figure 1.

Year of Data	Volume Model	Number of States (Plus D.C.)	Number of Locations
2019	LBS+	49	5,103
2020	LBS+	48	5,095
2021	LBS+	22	3,080
2022 (Jan–Apr)	LBS+	22	3,058

Table 2: Number of permanent counter locations used in training the U.S. All Vehicles Volume model, by

data year. The table shows the number of unique counter locations; each location could have counts for more than one travel direction.

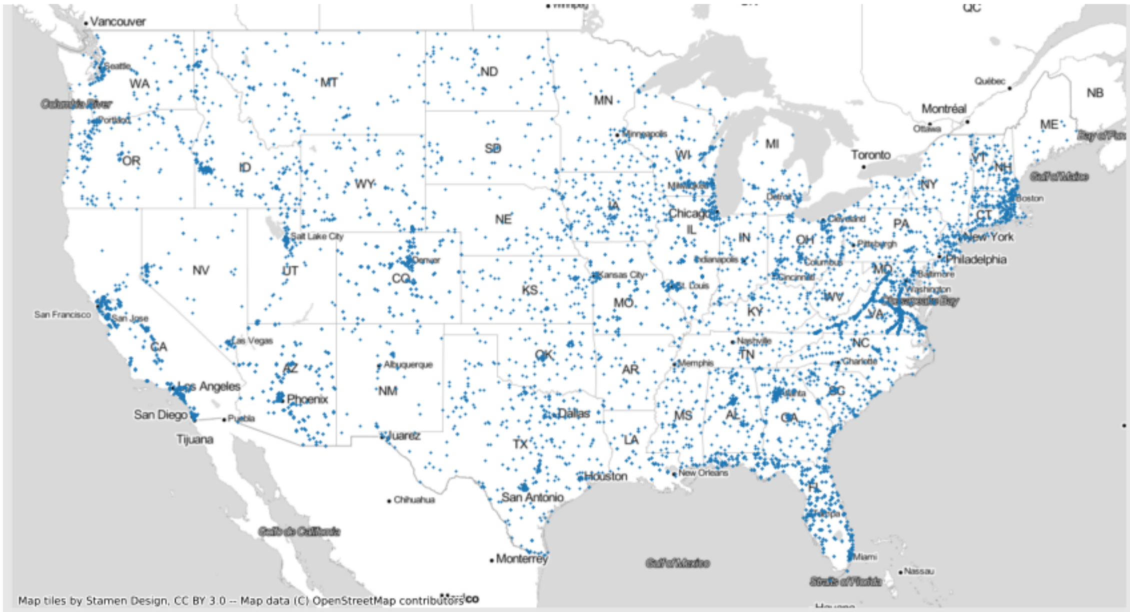

Figure 1: Locations of permanent counters used in the All Vehicles LBS+ 2019–2020 MADT models for the U.S

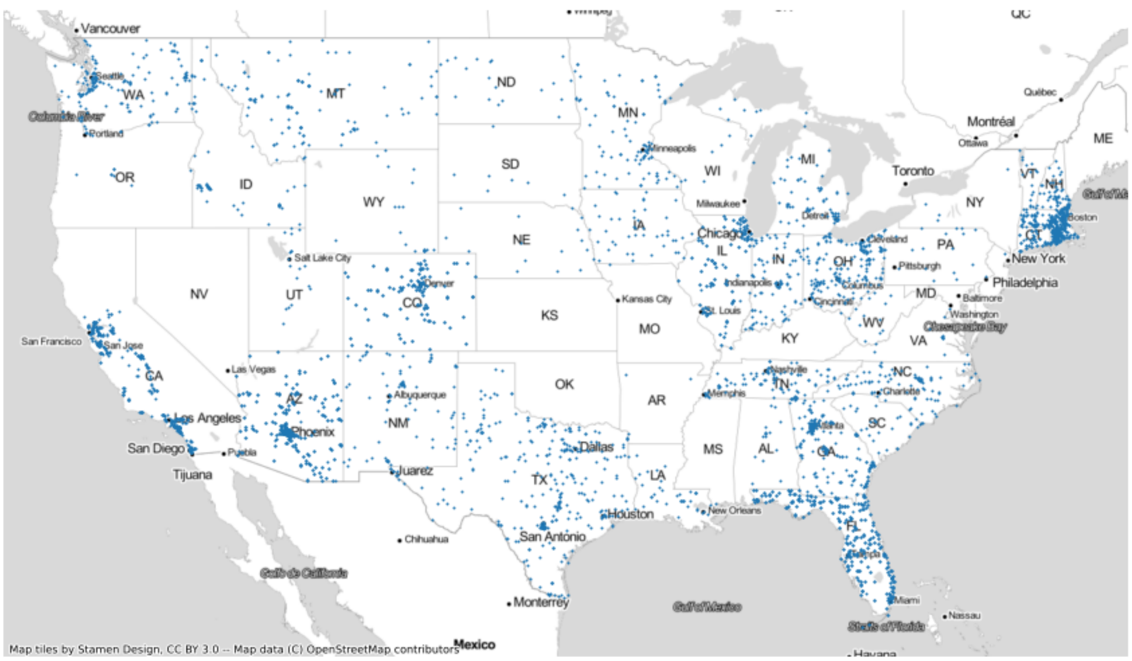

Figure 2: Locations of permanent counters used in the All Vehicles LBS+ 2021 and 2022 MADT models for the U.S. At the time of model development, fewer locations had data available for these years, but coverage is still good.

Algorithm Details

The U.S. All Vehicles LBS+ Volume algorithm consists of a pair of MADT models—one for low-volume roads and one for all others—each of which is trained on all data from January 2019 to the current month. Training a separate model for low-volume roads allows the model to be better tailored to those roads and results in higher accuracy. Both models use Extreme Gradient Boosting algorithms and 23 input features, including our sample of vehicle trips.

All Vehicles CVD+ Volume Algorithm

The All Vehicles CVD+ Volume model, which applies to data months beginning in January 2022, incorporates Connected Vehicle Data (CVD) as a key input. While CVD trips replace LBS trips as the most important input, this model continues to use historical LBS trips as features, allowing the model to gain signal from both data sources.

Aside from trip data source, there are some minor differences between the LBS+ and CVD+ model methodologies. First, CVD+ Volume is estimated from a single MADT model for all roads, instead of from two separate models for low-volume and higher-volume roads. Second, slightly different from a direct population factor, we use a trip weighting factor to scale CVD trips to the estimated total number of trips in the population. Third, in addition to weather inputs, we include a seasonality feature that uses historical travel patterns to quantify the seasonality of a locale.

In other respects, the CVD model methodology remains the same as for LBS. The model continues to use features that represent demographics, road attributes, and weather conditions. It also employs the same Extreme Gradient Boosting technique, and prediction intervals are calculated in the same way. Methodology for estimating Volume for area zones, day parts and day types, and Origin-Destination analyses remain the same.

Input Features

The following features are used in the updated CVD+ Volume:

- **LBS trips:** Our processing algorithms link pings from LBS devices into trips, then classify these trips into certain travel modes. The Volume models use trips that we have identified to be from vehicles. While LBS trips are the key feature in the LBS+ Volume models, historical LBS trips also figure into CVD+ Volume.
- **CVD trips:** Our algorithms create trips using pings from personal vehicles embedded with location technology. The CVD+ model is based mainly on CVD trips. CVD trips do not need mode classification, because they come from vehicles.
- **U.S. Census data:** We use demographic data, such as number of housing units and whether or not a location is in an urban area.

- **Trip weighting factor:** Using Census population data, we introduce a trip weighting factor that scales the number of CVD trips to our best guess of the number of trips in the population. This is analogous to the population factor used in the LBS+ algorithm.
- **OpenStreetMap (OSM) data:** OSM provides information on road classification, number of lanes, speed limit, and other figures.
- **Weather data:** Temperature, precipitation, and visibility affect travel and help us better predict seasonal changes.
- **Seasonality:** Historical travel patterns are used to quantify the seasonality of a given area.

Ground-Truth Data

The CVD+ Volume models are trained and validated using data from over 10,000 permanent counters, representing more than 6,000 unique locations across the U.S. For the most recent update, we obtained 2022 training data from nearly 5,000 additional counters, almost doubling the number of counter sites from the original v1, as shown in Table 3. (Note that not every counter reports data in every month.) Significantly, the updated version uses 2022 data from 48 states, including the District of Columbia, whereas the original version's training covered only 22 states. Therefore, we expect improved model accuracy in states that were originally not represented in the training data. The map in Figure 3 below shows the locations of permanent counters used in the updated CVD+ Volume model. For the updated version, we still only had access to 2023 data in the subset of 22 states. Even so, we expect the inclusion of the new counters to improve accuracy in 2023.

Year of Data	Volume Model Version	Number of States	Number of Locations	Number of Counters
2022–2023 (Jan–Apr)	CVD+ v1(old)	22	2,836	5,514
2022	CVD+ updated	48	6,086	10,195
2023 (Jan–May)	CVD+ updated	22	3,215	5,227

Table 3: Number of permanent counter locations used in training the U.S. All Vehicles CVD+ Volume models, by data year. The table shows the number of unique counter locations; each location could have counts for more than one travel direction.

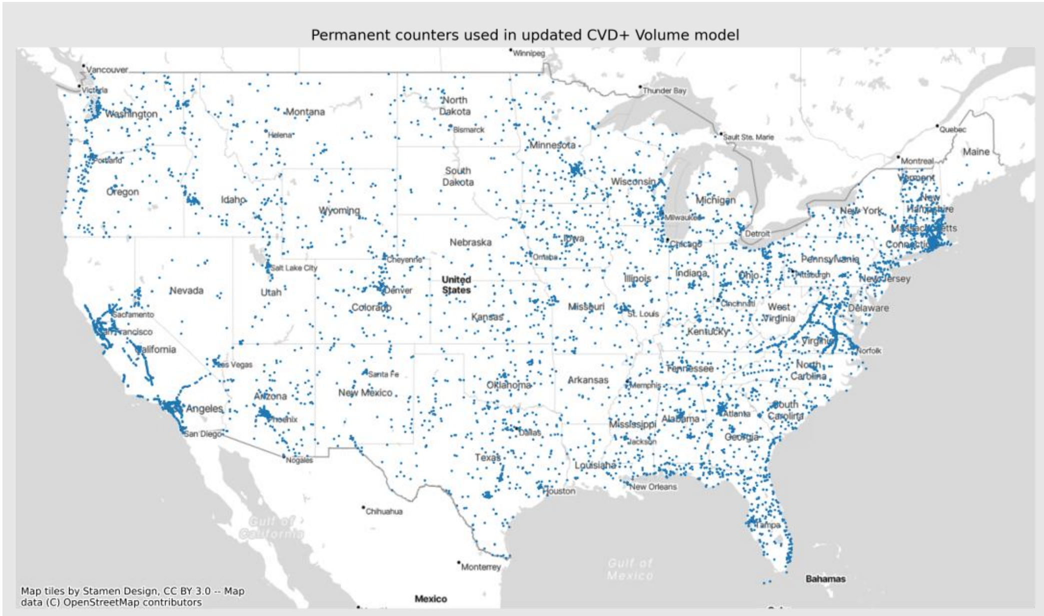

Figure 3: Locations of permanent counters used in the All Vehicles CVD+ Volume model.

ALGORITHM DETAILS

The CVD+ Volume metric is estimated using two gradient-boosting algorithms. The main model predicts MADT on most roads. A secondary model is used when it's determined that the historical LBS trip count is unreliable. This could happen if OSM information has changed since the LBS data was recorded, or if some other anomaly or error affected LBS data. The introduction of the secondary model allows better handling of edge cases and results in overall more reliable estimates.

Prediction Intervals

Both LBS+ and CVD+ All Vehicles Volume estimates include prediction intervals that represent the likely upper and lower bounds of the true value. The 95% prediction interval is an estimate of the range within which the true value is expected to lie 95% of the time. To construct these intervals, we fit a quantile regression model to the MADT errors. The prediction intervals included with StreetLight's Volume metrics are only available for all day, average day estimates (not for specific day types or day parts) and only when there are at least 25 days in the input period. Prediction intervals are only reported when analyzing pass-through zones and are not available when analyzing area zones or conducting Origin-Destination analyses.

Estimating Volume from the MADT Model

VOLUME FOR SPECIFIC TIME INTERVALS

The Volume models described in the previous section estimate monthly traffic volumes (MADT). To generate Volume estimates for day parts or day types, we use MADT as the base, then apply a time-specific scaling factor. The scaling factor is equal to the ratio of LBS or CVD trips for that time interval to LBS or CVD trips for that month. Our assumption here is that the ratio of sample trips to actual trips for a location is stable throughout the month. Possible trip sample bias by day part or day type is not presently accounted for in our Volume metric.

VOLUME FOR AREAS

The MADT models are designed to estimate Volume for roads, because ground-truth data from permanent counters is available for road segments. We use an additional algorithm to develop Volume estimates for areas. To estimate Volume for a specified area, the algorithm selects a subset of roads with trips that start, end, or pass through the zone area, depending on if the Volume requested is for trips starting, ending, or passing through the zone. See Figure 4 for an example where a specified area (shaded blue) is accompanied by a subset of randomly sampled roads (yellow polygons) in the surrounding area. The number of sampled roads will depend on the size and location of the area zone.

Figure 4: Example area zone with selected gates (yellow polygons) used to calculate MADT for trip starts and stops to the blue area.

For each sampled road, our algorithm does the following:

1. Generates an estimate of MADT from each sampled road.
2. Obtains an estimated trip sample penetration at each sampled road. This assumes that the sampled roads represent the StreetLight trip sample penetration for the nearby area.
3. Calculates a weighted average trip penetration from the sampled roads and applies that to the trip count of the area zone to derive a final estimated StreetLight Volume.

This results in an estimated Volume for the defined area based on trip starts and ends.

Volume for Origin-Destination Analysis

After estimating Volume outputs for individual zones (both road segment and area zones), we apply them to Origin-Destination (OD) analysis, which allows for evaluating the number of trips that span locations. The goal is to generate an OD Volume that provides a number representing a reasonable estimate of the real-world number of trips. This approach includes the following steps:

1. Calculate the total Volume at each origin and destination zone (described in the previous sections).
2. Return the sample trip counts between each OD zone pair.
3. Scale the OD sample trip counts to Volume based on the estimated total Volume at each origin and destination.

Validation

Validation Methodology

To validate the MADT model, we used stratified k-fold cross-validation with k equaling 10.1. K-fold cross-validation was a good choice in this case, where we have a large amount of training data, because it reduces bias and minimizes over-fitting without being too computationally intensive. In performing the cross-validation, we stratified by road volume category to ensure minimal bias by road size.

We use the following set of error metrics to evaluate estimation accuracy:

- **Mean absolute percentage error (MAPE)** is a standard measure of error for predictive machine-learning models. It describes errors well on small roads and treats errors of all sizes equally.

- **Normalized root mean square error (NRMSE),** another standard error metric in machine learning, penalizes large errors more, making it more sensitive to the accuracy of AADT estimation on high-volume roads.
- **68th percentile absolute error** represents one standard deviation from the mean within a standard bell curve and thus describes the typical error across the road segments. The 68th percentile absolute error can be interpreted as the expected typical error.
- **95th percentile absolute error** represents the upper limit of the expected absolute error.
- **Median percentage error** is a measure of bias in our model, which also indicates how accurate an aggregate value (such as the sum of Volume on all roads in this bin) will be. Values close to zero suggest that our model has low bias, while positive values would indicate more overestimation and negative values indicate more underestimation.

VALIDATION RESULTS FOR ALL VEHICLES LBS+ VOLUME

This section presents validation results for the All Vehicles LBS+ Volume model for months from January 2019 through December 2021. Streetlight MADT estimates are closely correlated with actual MADT values for each year covered by the model, as demonstrated by an R^2 of 0.98 in Figure 5 below.

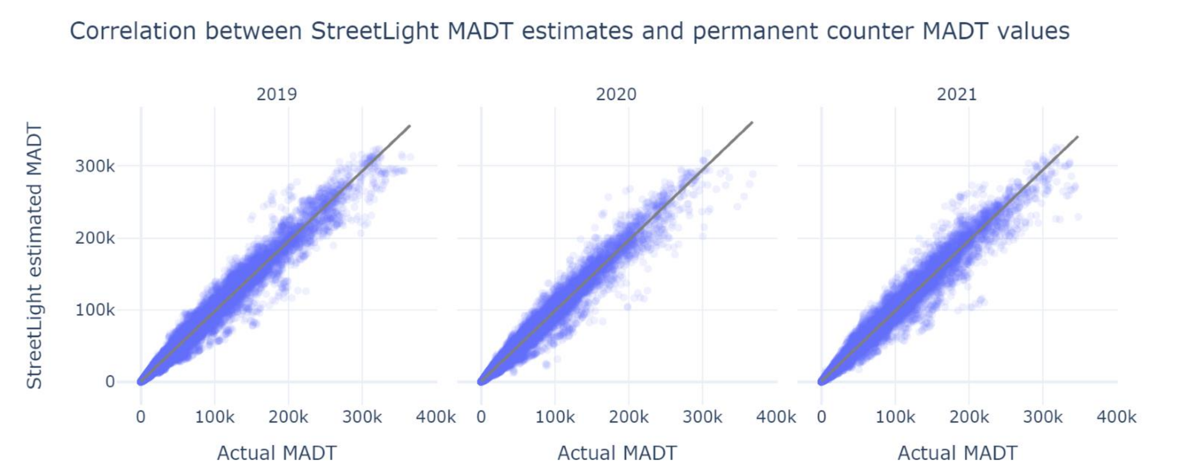

Figure 5: Correlation between StreetLight LBS+ MADT and permanent counter MADT by year. The best fit lines all have R^2 values of 0.98.

Table 4 below shows error metrics for StreetLight MADT for the U.S., by road size. The results show minimal bias and low error across road classes; error metrics improve as road size increases.

Road Size (AADT)	# of Counters	Median Percent Error (Bias) (%)	MAPE	68^{th} Percentile Abs Error (%)	95^{th} Percentile Abs Error (%)	NRMSE

A: 0–499	404	4.3	20.6	24.9	51.7	27.0
B: 500–1,999	1,629	2.1	14.9	18.0	39.0	19.6
C: 2,000–4,999	2,544	1.0	12.0	14.1	32.3	15.9
D: 5,000–9,999	3,027	1.1	10.7	12.8	28.3	13.9
E: 10,000–19,999	3,419	0.6	9.4	11.2	25.1	12.4
F: 20,000–34,999	2,961	-0.2	8.7	10.1	23.7	11.8
G: 35,000–54,999	2,278	0.2	8.5	9.5	24.5	11.8
H: 55,000–84,999	1,721	-0.5	7.8	8.5	23.1	11.0
I: 85,000–124,999	1,044	-0.8	7.7	8.7	22.3	10.9
J: 125,000±	642	-1.7	7.4	8.3	21.6	10.8

Table 4: Cross-validation results for All Vehicles LBS+ Volume by road size. The metrics cover all months from 2019–2021. Road size category is based on the measured AADT of the counter. Number of counters refers to unique counter locations; each counter can have multiple data points for different months/years and travel directions.

Accuracy and Comparability Across Time

The plots in Figure 6 and Figure 7 below show how MAPE and median percent error (bias) vary by month for different road sizes. MAPE varies little over time, especially for roads over 500 AADT. For roads under 500 AADT, MAPE fluctuates more in 2021, potentially due to the model using fewer training sites that year. Median percent error shows a slight trend toward overestimation in 2021 for roads 500–4,999, but overall, the magnitude of bias and the change over time are very small. Note that bias rarely exceeds 5% and consistently stays within a few percentage points for higher volume roads.

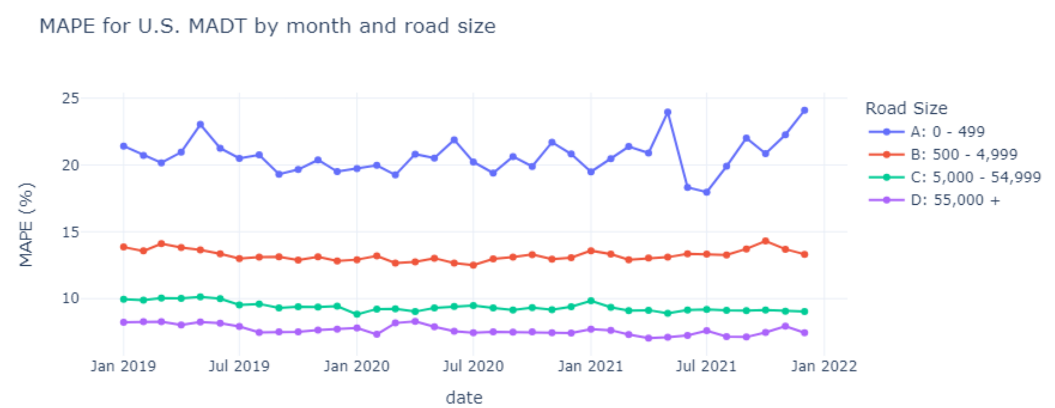

Figure 6: MAPE by month for All Vehicles LBS+ Volume, stratified by road size.

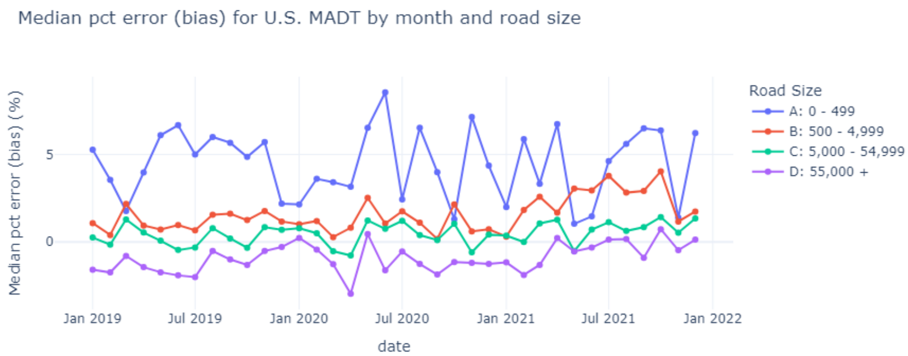

Figure 7: Median percent error (bias) by month for All Vehicles LBS+ Volume, stratified by road size.

To evaluate how well the model captures time trends, the plot in Figure 8 shows estimated and actual MADT for a random selection of counters, stratified by road size. The results suggest that although there is some variability in the low-volume road categories, estimated MADT tracks actual values across time well. Specifically, estimates appear to generally capture the changes in traffic volume during the COVID-19 pandemic, which had greatly varied effects across locations. Overall, the results by month should provide confidence in the metric's time trends.

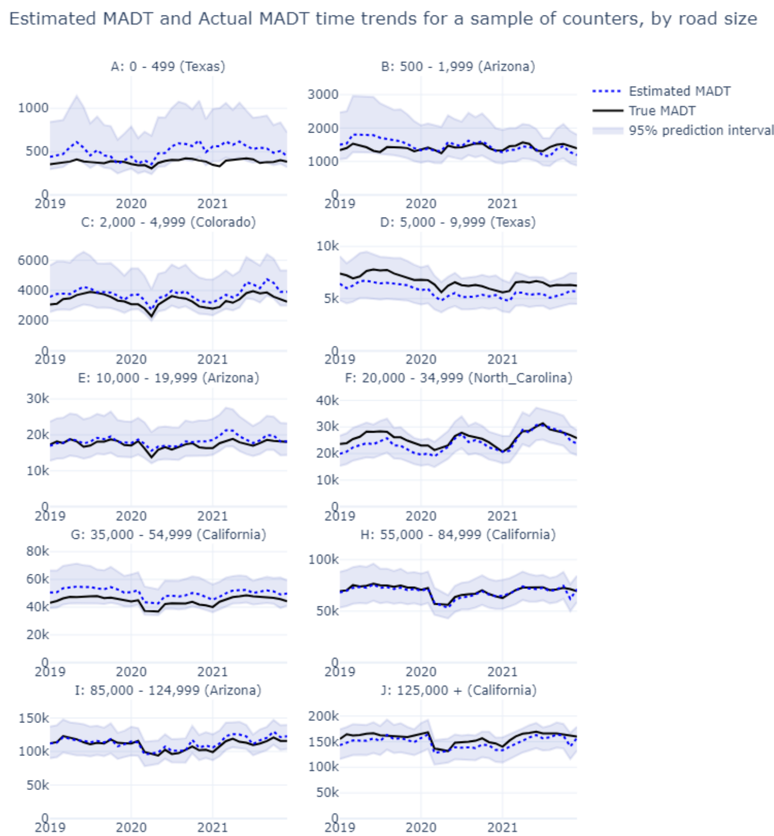

Figure 8: Estimated LBS+ MADT and counter MADT by month for a random sample of counters in U.S., stratified by road size. These plots indicate a close correlation between estimated and actual time trends.

VALIDATION RESULTS FOR ALL VEHICLES CVD+ VOLUME

In this section, we report cross-validation results for the latest All Vehicles CVD+ Volume algorithm, including comparisons with the previous CVD+ version. As shown in Figure 9, the correlation between estimated MADT and permanent counter MADT remains strong, with an R^2 of 0.97 for both 2022 and 2023. The correlation plot for 2023 contains fewer points because, at the time of model development, permanent counter data was available only for the months of January through May, and from a smaller number of counter sites, as discussed in the Ground-Truth Data section.

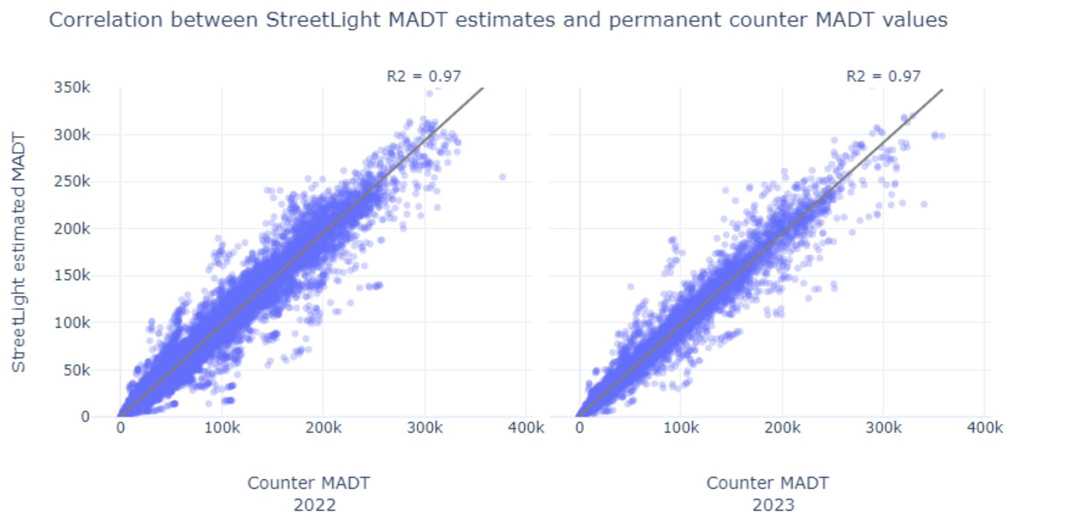

Figure 9: Correlation between StreetLight CVD+ MADT and counter MADT by year. For 2023, the plot shows results for the months of January–May. The best fit lines have R^2 values of 0.97.

Table 5 below, shows the cross-validation results for updated CVD+ Volume, for all months from January 2022–May 2023, broken down by road size category. The results show minimal bias, with absolute values of less than 1% on road sizes over 2,000 in average daily traffic. On all road sizes, the absolute bias is less than 8%. On other metrics like MAPE, errors are relatively low and decrease with road size, as expected.

Road Size Category (Based on MADT)	# of Counters	Median Percent Error (Bias) (%)	MAPE (%)	68th Absolute Percentile Error (%)	95th Absolute Percentile Error (%)	NRMSE (%)
A: 0– - 499	408	7.6	28.3	73.8	26.4	29.9
B: 500–1,999	1,606	2.0	17.2	43.6	15.1	21.0
C: 2,000–4,999	2,230	0.4	13.5	32.6	11.8	16.6
D: 5,000–9,999	2,282	0.4	11.5	27.8	10.3	15.8
E: 10,000–19,999	2,562	0.1	10.2	26.3	9.5	15.5
F: 20,000–34,999	2,070	0.1	8.8	23.9	8.4	13.8
G: 35,000–54,999	1,477	0.1	8.9	25.9	8.6	14.1
H: 55,000–84,999	1,165	-0.6	8.1	24.2	7.8	12.3
I: 85,000–124,999	788	-0.6	9.1	25.7	8.9	13.8
J: 125,000±	706	-1.6	8.2	24.2	7.6	11.7

Table 5: Cross-validation results for updated All Vehicles CVD+ Volume by road size. The metrics cover all months from January 2022–May 2023. Road size category is based on the measured AADT of the counter. Number of counters refers to unique counter locations; each counter can have multiple data points for different months/years and travel directions.

For a view of accuracy over time, Figures 10 and 11 show error metrics by month and road size category. MAPE (Figure 10) and bias (Figure 11) changed little over time throughout 2022, indicating stable model accuracy during this time period. For the lowest volume roads, those under 500 MADT, MAPE and bias increased in 2023. This is because additional training data from the new permanent counters was available for 2022, but only a subset of counters had data for 2023. This more limited set of counters affects very low-volume roads the most. In fact, MAPE for roads over 500 MADT is consistent even in 2023. Bias for larger roads exhibits more variation in 2023, but the overall bias values are small. The stability in error metrics over time also suggests the model produces reliable time trends. Time trends are discussed in more detail in the following section.

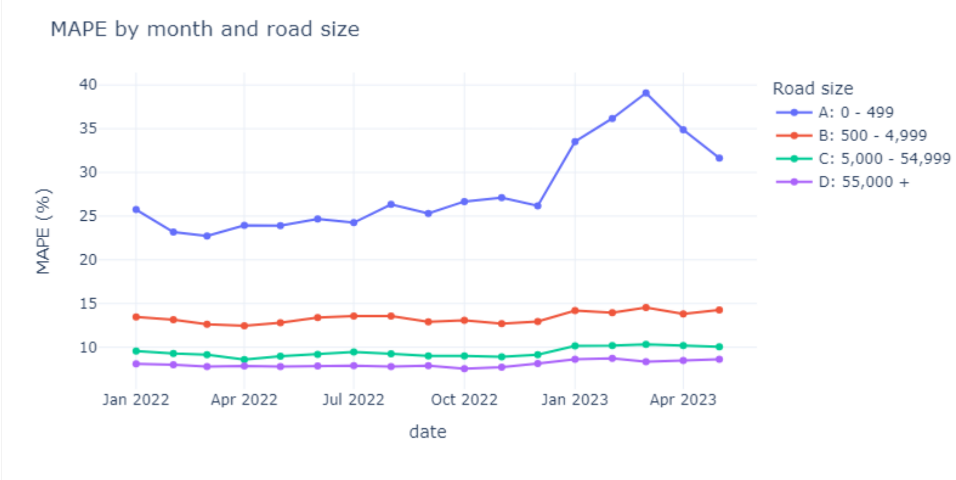

Figure 10: MAPE for CVD+ Volume v2 by month and road size, for all counters.

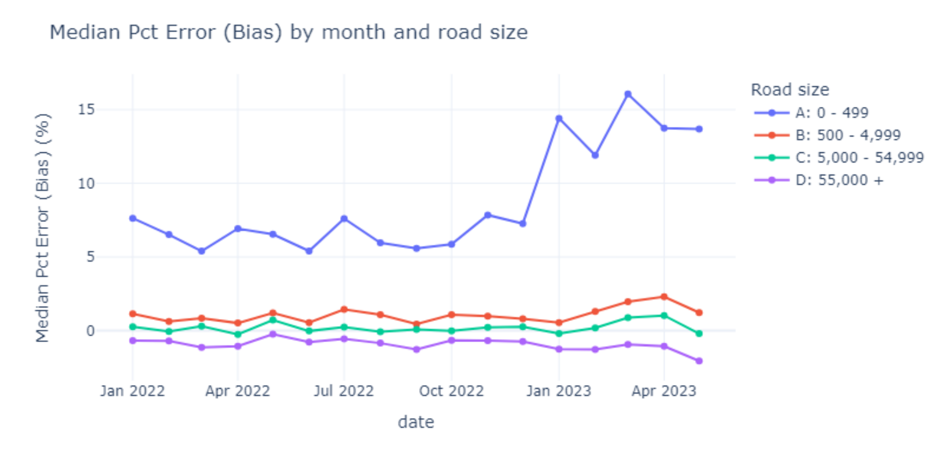

Figure 11: Median percent error (bias) for CVD+ Volume v2 by month and road size, for all counters.

Comparing the Updated CVD+ Model with the Previous Version

The plots below compare error metrics for the updated CVD+ Volume to the previous v1 version, including the more than 10,000 permanent counters used to train the updated models—that is, both old and new counters. We can see that the updated version performs much better overall at these counter locations in terms of MAPE (Figure 12) and bias (Figure 13). We see this sizeable improvement because nearly half of the counters in the updated model were not originally available for v1 training. In particular, the high bias values in Figure 13 indicate that v1 tended to heavily overestimate on roads of less than 20,000 in MADT.

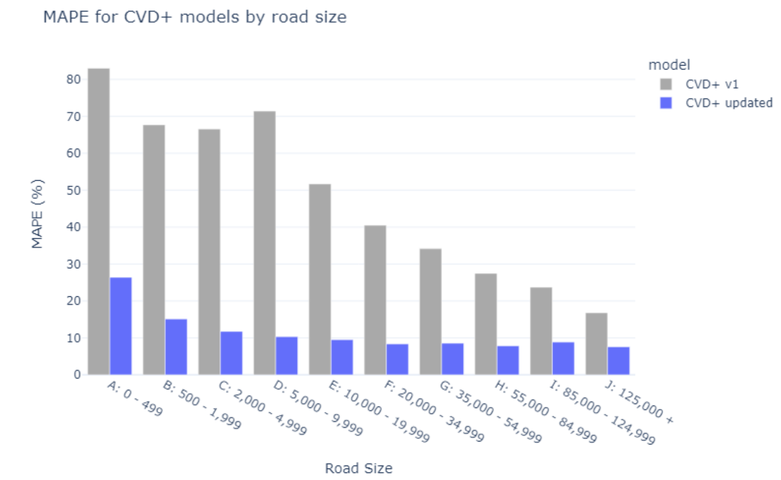

Figure 12: MAPE for the updated CVD+ compared to v1, for all months. Errors are shown for all 9,800 counters used in training the updated models; only about half of these were available for training the v1 models.

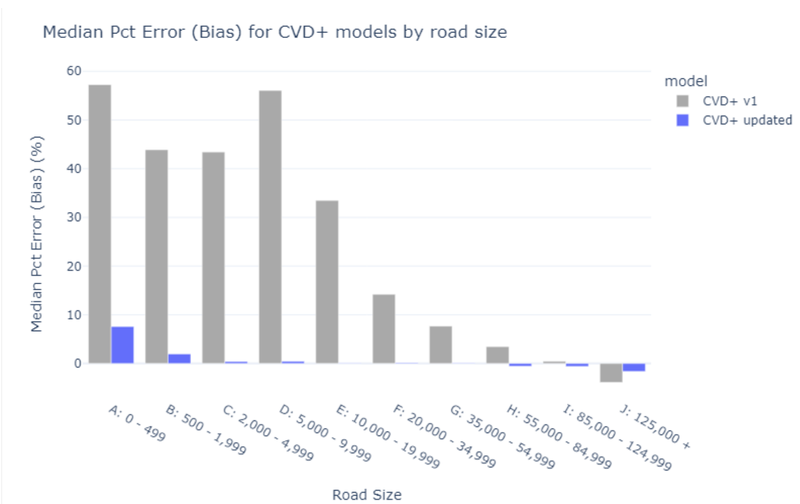

Figure 13: Median percent error (bias) for the updated CVD+ compared to v1, for all months. Errors are shown for all counters used in training the updated models; only about half of these were available for training the v1 models.

We want to make sure the improvements in accuracy obtained by adding new counters did not come at the expense of the accuracy at old counter locations. To check, we can restrict the results to only the 5,514 counters used in training CVD+ v1. A small increase in errors is expected because, when we add new counters, a model that was specialized to a smaller set of counters must now fit a wider range of conditions. As shown in Figure 14, when we look only at these counters, we see accuracy at the pre-existing counters has largely been maintained. Roads of less than 5,000, as well as road size I, had a slight increase in MAPE. Thus, with the updated model, overall accuracy improves dramatically, with only a slight decrease in accuracy on some pre-existing counters, mainly on low-volume roads.

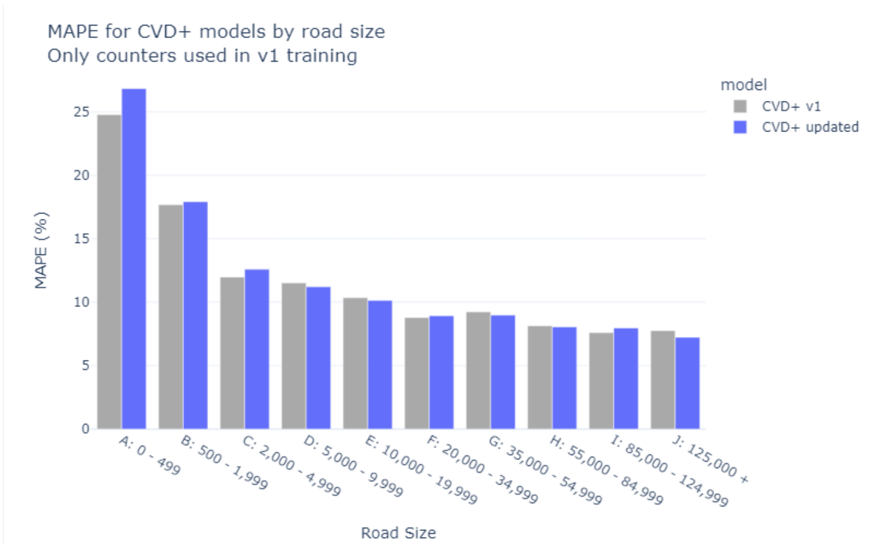

Figure 14: MAPE for updated CVD+ compared to v1 for all months. Errors are shown only for the 5,514 counters used in training the v1 models.

As shown in Figure 15, considering only the original set of counters, bias has increased on some road sizes, particularly those with less than 2,000 MADT. On higher volume roads, bias remains low with little change. To reiterate, some increase in error at pre-existing training sites is expected. Overall, though, the updated Volume model largely maintains accuracy at the pre-existing training counters.

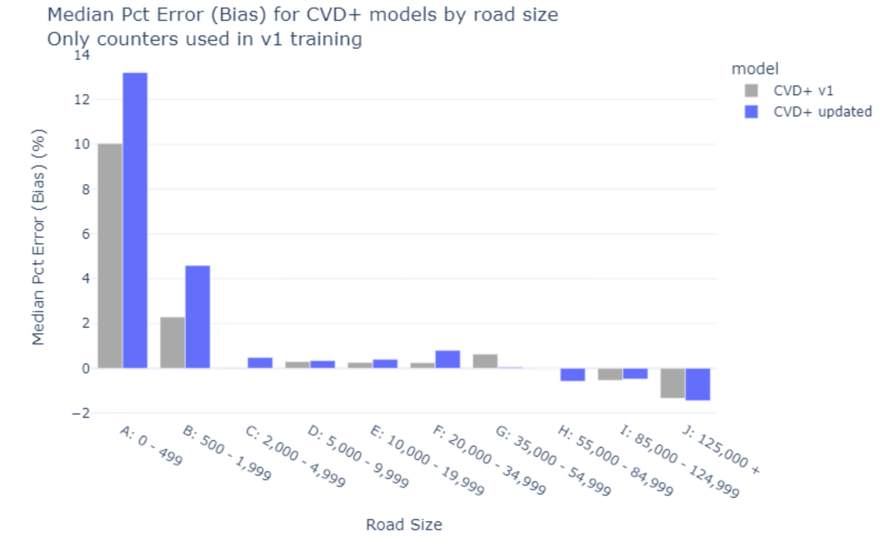

Figure 15: Median percent error (bias) for updated CVD+ compared to v1 for all months. Errors are shown only for the 5,514 counters used in training the v1 models.

COMPARABILITY ACROSS TIME

To better understand how well our models reflect time trends, we examine MADT by month for a random selection of counters. Figure 16 below shows time trend plots for random counters that were available in training the original v1 CVD+ Volume model, stratified by road size. The updated CVD+ trends match counter data closely, except for the road in the upper right (road size B, Montana), where CVD+ underestimates traffic, and in the third row on the left (road size E, Texas), where both versions overestimate. Note, in the road size B example, the updated version picks up the seasonal trend a little better than does v1.

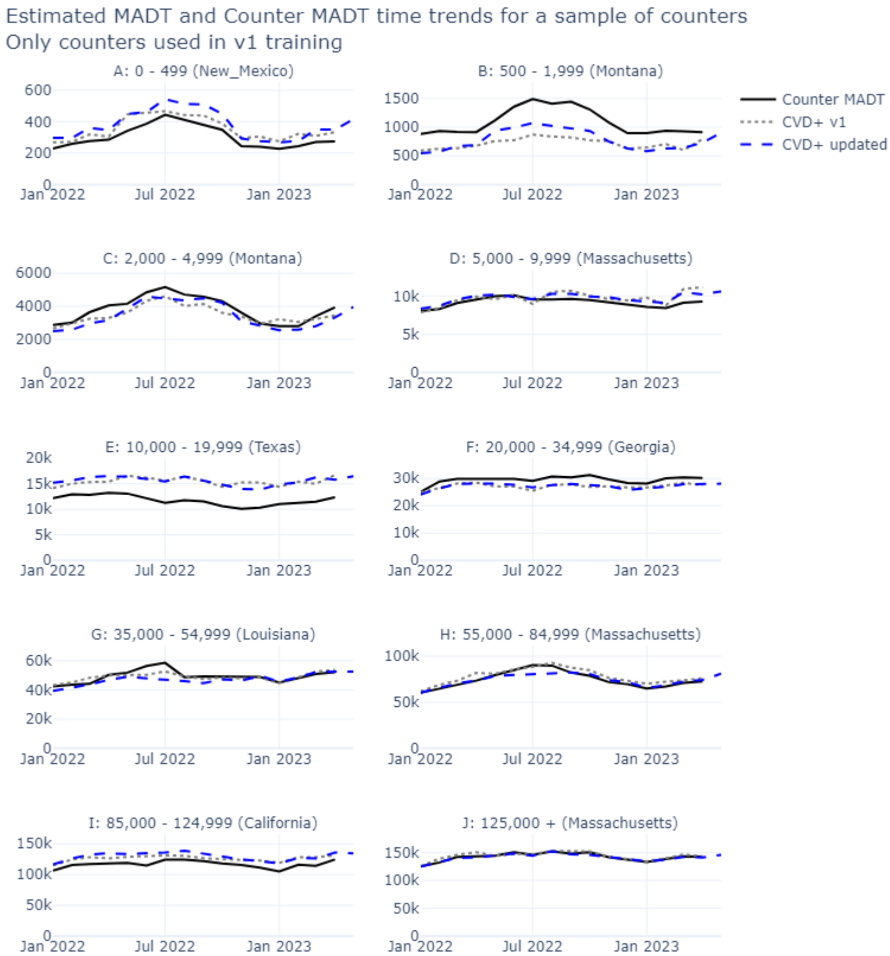

Figure 16: Estimated CVD+ MADT compared to permanent counter MADT by month for a random selection of counters, stratified by road size. This figure shows only counters used in training the original v1 CVD+ model.

Looking at time trends only for counters that were new to the updated CVD+ model, we see a much bigger divergence between model versions. As shown in Figure 17 with a random selection of these counters, updated estimates tend to match counter data more closely than the v1 estimates. Moreover, v1 predicted rises and falls in Volume not seen in counter data. This is because the v1 estimates were not just generated by a single model for all months; for each new month (or batch of months) of data, a new model was fit. New fitting of a model is what causes the discontinuities observed on some roads, like those in the second row (Virginia and Nevada). In these cases, one model was fit for the months of May–July 2022 and another for August–October 2022, resulting in a jump in Volume at the transition point. In contrast, updated CVD+ Volume time trends are smoother across the time range.

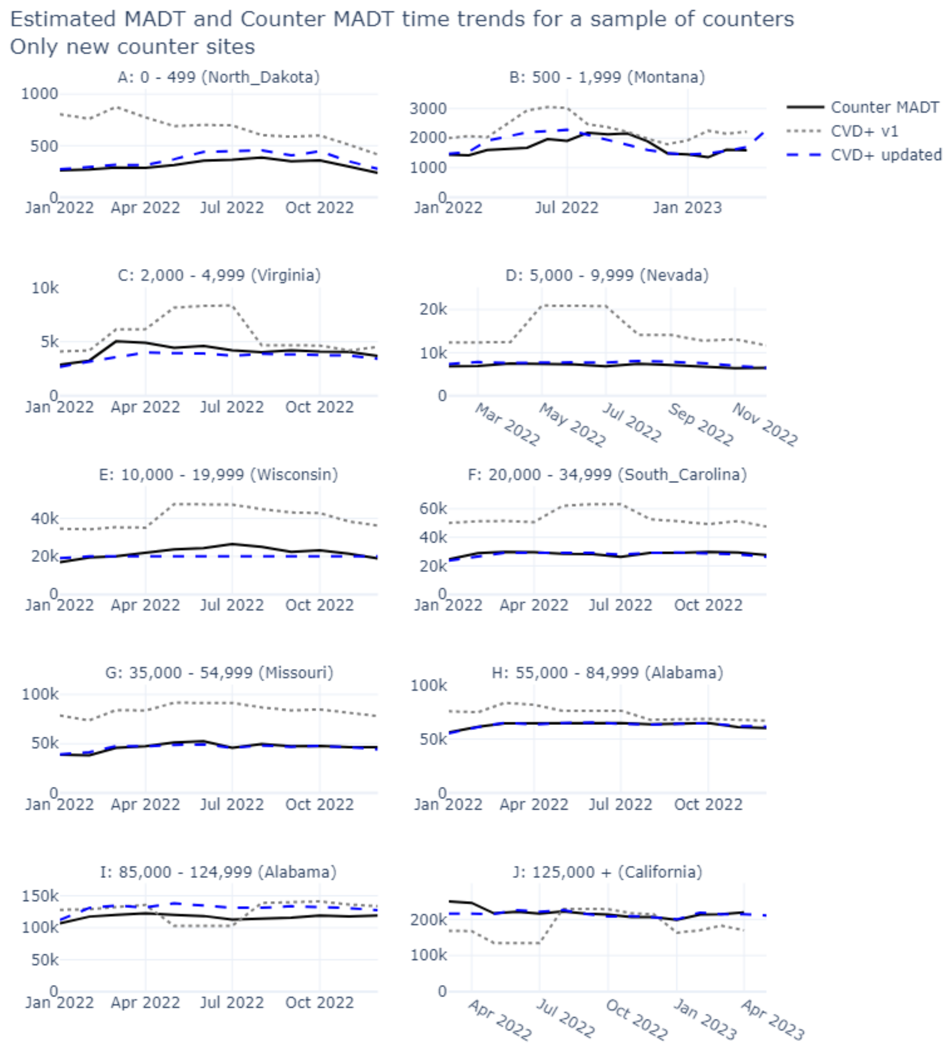

Figure 17: Estimated CVD+ MADT compared to permanent counter MADT by month for a random selection of counters, stratified by road size. This figure shows only new counters; these counters were not used in training v1.

Conclusions

StreetLight's U.S. All Vehicles Volume metric has continually maintained accuracy across regions and road types. The models have a few limitations—most importantly, little visibility into accuracy at locations without permanent counters. We continue to address those limitations by increasing coverage of training and validation data and iterating on the algorithms. Future versions of All Vehicles Volume will focus on integrating new sources of trip data and expanding historical comparability.

About StreetLight Data

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 DATA