PRELIMINARY EVALUATION OF BIG DATA FOR MODEL DEVELOPMENT

Traffic Volumes by Screenline and Vehicle Miles Traveled (VMT)

Ray Ngo
TPB Transportation Engineer

Travel Forecasting Subcommittee
September 20, 2024

National Capital Region
Transportation Planning Board

Data Focus

- Big Data vendors:

|		
OpenPaths Patterns (Formerly Teralytics)	**REPLICA**	**STREETLIGHT**

- Validation focus:

  - Average Annual Daily Traffic (AADT) by screenline

  - VMT by jurisdiction

# Regional Screenlines for the Gen2 and Gen3 Models

- TPB staff conducted validations of traffic volumes by screenline for both the Gen2 and Gen3 travel models to assess model accuracy and identify model calibration needs.

- Screenlines for the TPB Modeled Area are numbered from 1 to 38.

National Capital Region
Transportation Planning Board
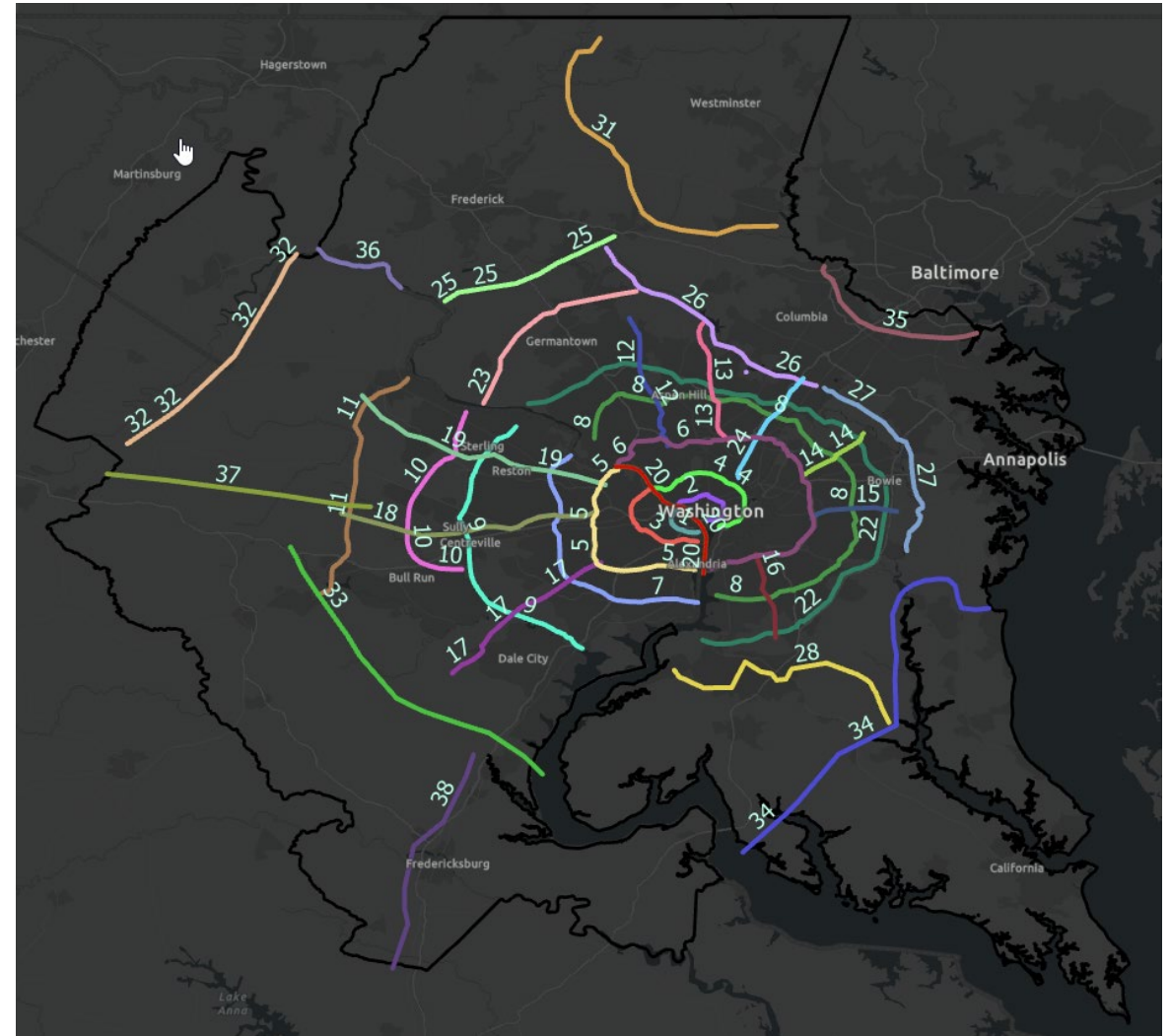
Big Data of Interest

Big Data	Geography	Traffic Volume Data
Teralytics	TPB Planning Area	2019 AADT (12/2023) 2019 AADT (08/2024)*
Replica	U.S. (including TPB Modeled Area)	Fall 2019 Thursdays "Network Link Volumes" data AADT 2019
StreetLight	TPB Planning Area	2018 AADT

* It is not clear to TPB staff what changes were made by Teralytics in the two data set.
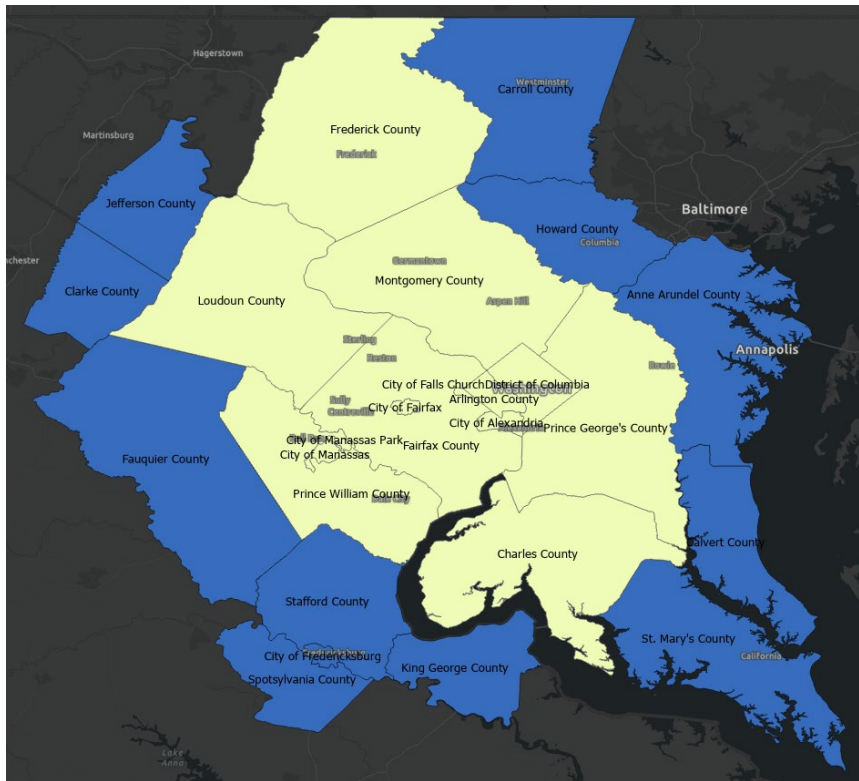
Yellow area is COG and TPB Planning Area. Blue and yellow area combined is the TPB Modeled Area

National Capital Region
Transportation Planning Board

Methods

- Querying the data:

  - Teralytics: Shapefiles from Map module

  - Replica:

    - Network Link Volume shapefiles for Fall 2018 Thur. from "Places Studies"

    - AADT shapefiles from the Annual Average Daily Traffic (AADT) module

  - StreetLight Data:

    - AADT shapefiles from AADT Analysis module (required TPB staff to input screenline shapefiles)

- Processing data:

  - AADT: spatial intersection, duplicate records removed

  - VMT = AADT x distance

Observed data: 2018 AADT collected by state & local government

- 2018 traffic counts (AADT), collected by state & local governments, summarized by screenline was available off the shelf, while the 2019 data was not available by screenline.
- Changes between 2018 AADT and 2019 AADT are relatively small.

Scatter Plot of AADT2018 vs AADT2019 Counts at Stations
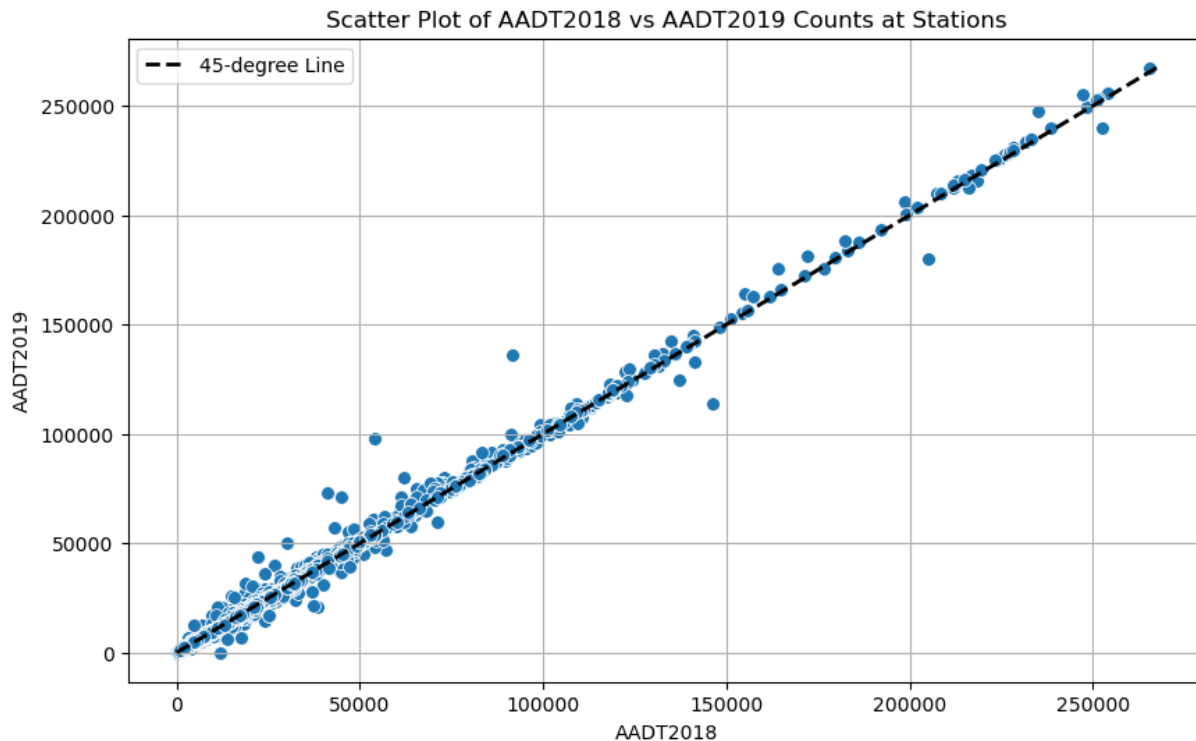
~90%

of links with traffic counts (AADT) show changes of less than 5%

-0.1%

difference in total counted vehicles

National Capital Region
Transportation Planning Board
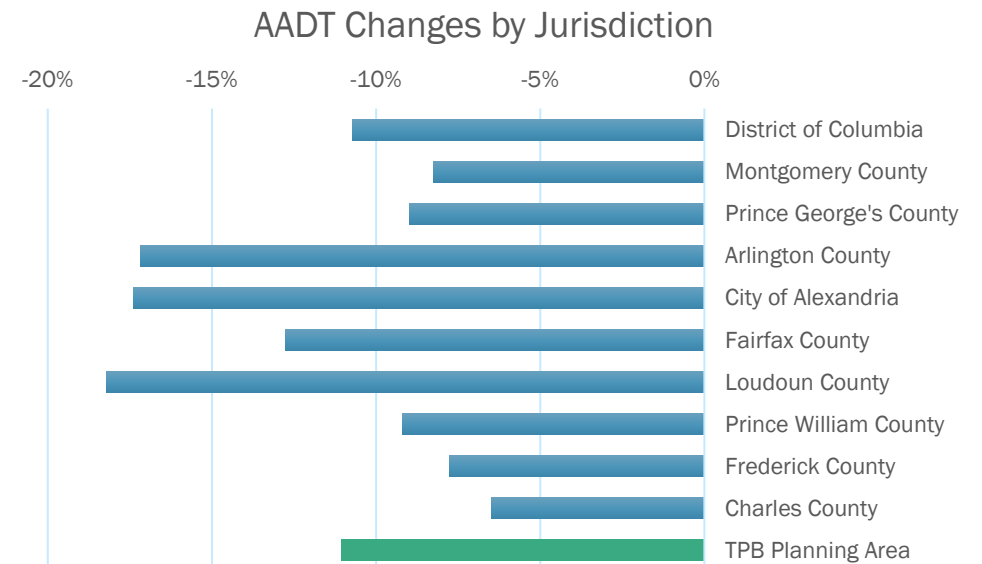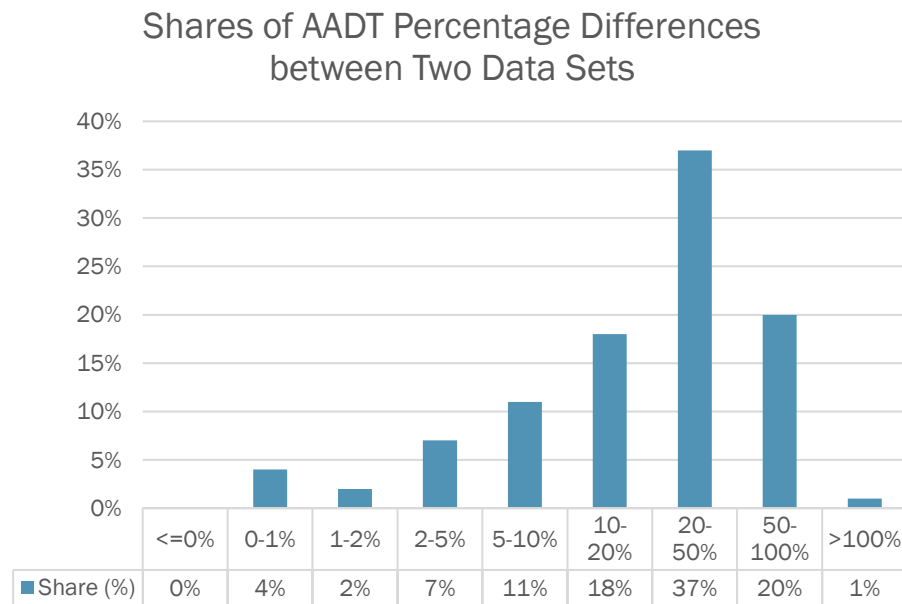
Two 2019 AADT Data Sets from Teralytics

- The 2019 AADT sets downloaded in Dec. 2023 and Aug. 2024 were not identical

- Staff is waiting for a response from Teralytics for clarification, but used the Aug. 2024 data set for this analysis

Shares of AADT Percentage Differences between Two Data Sets

	<=0%	0-1%	1-2%	2-5%	5-10%	10-20%	20-50%	50-100%	>100%
■ Share (%)	0%	4%	2%	7%	11%	18%	37%	20%	1%

AADT Changes by Jurisdiction

District of Columbia
Montgomery County
Prince George's County
Arlington County
City of Alexandria
Fairfax County
Loudoun County
Prince William County
Frederick County
Charles County
TPB Planning Area

Two Data Sets of Replica

- First data set: Fall 2019 Thursday data retrieved from "Network Link Volumes":
    - Private auto: Includes trips made by the drivers of private autos
    - Taxi/TNC: Includes trips made by a passenger of a taxi/transportation network company (TNC) trip
    - Commercial vehicle: Includes trips made by commercial vehicles (freight trips)
    - Public transit: persons traveling by bus
    - ⟹ One should not sum up these four items to get the traffic volume data
    - ⟹ Per request, Replica sent COG staff a link network data file that contains bus volumes [vehicle trips] on September 16, 2024
- Second data set: 2019 AADT data includes private auto, taxi/TNC, commercial vehicle (freight), and bus[vehicle] trips

Primary Mode Filter
The primary travel mode of the activity.

Categories to include in this filter:

☐ Auto passenger

☐ Biking

☑ Commercial vehicle (freight)

☑ Private auto

☐ Public transit

☑ Taxi/TNC

☐ Walking

☐ Other

National Capital Region
Transportation Planning Board

Replica Data Sets Comparison

- We compared two data sets from Replica at the regional, jurisdictional, and link level and found some inconsistencies

- Finding: Replica's "Network Link Volumes" download option limits the number of links to the top 100,000 links in the selected TPB Modeled Area, sorted by highest volumes

- We used Replica's 2019 AADT data set for the validations

National Capital Region
Transportation Planning Board

Comparison Coverage

	Traffic Volume by Screenline	VMT by Jurisdiction
Teralytics	Excluded screenlines: #27, 31-35, 37-38 *	TPB member jurisdictions
Replica	All screenlines	All TPB Modeled Area jurisdictions
StreetLight Data	Excluded screenlines: #27, 31-35, 37-38	N/A

* Due to coverage limits with the data set.

National Capital Region
Transportation Planning Board
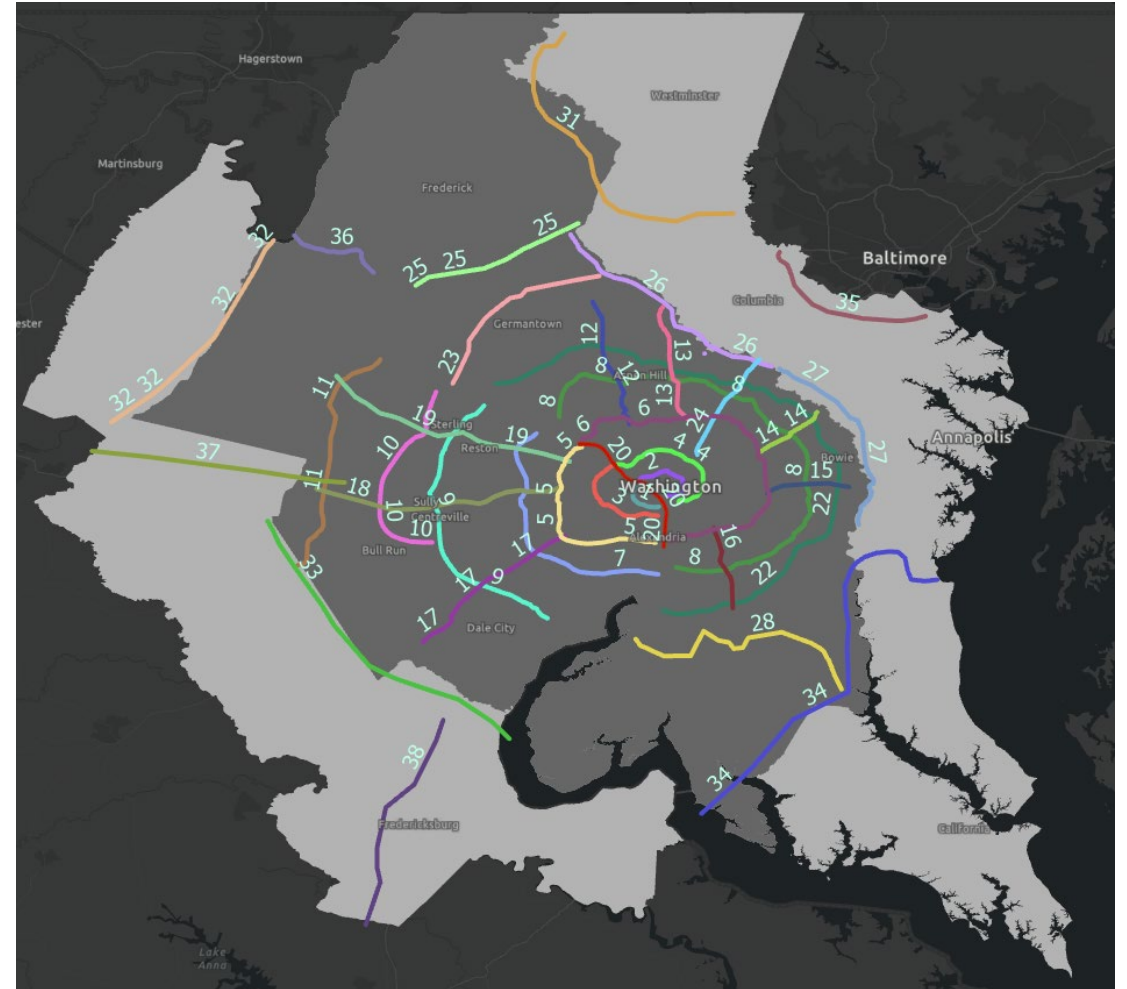
Screenline #20: Potomac River and a key screenline
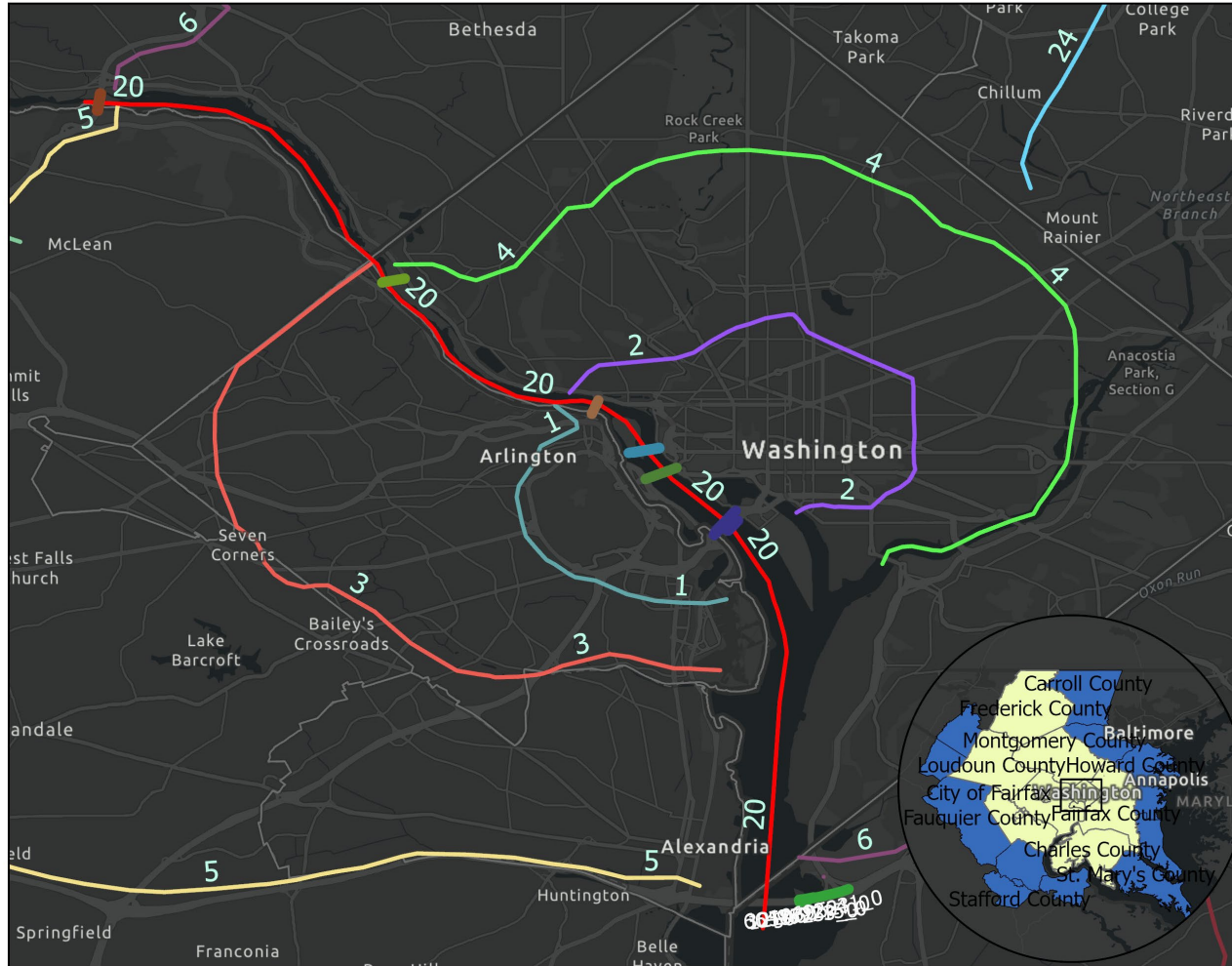
Legend

- Francis Scott Key Memorial Bridge
- 14th St Bridge
- Woodrow Wilson Memorial Bridge
- American Legion Memorial Bridge
- Arlington Memorial Bridge
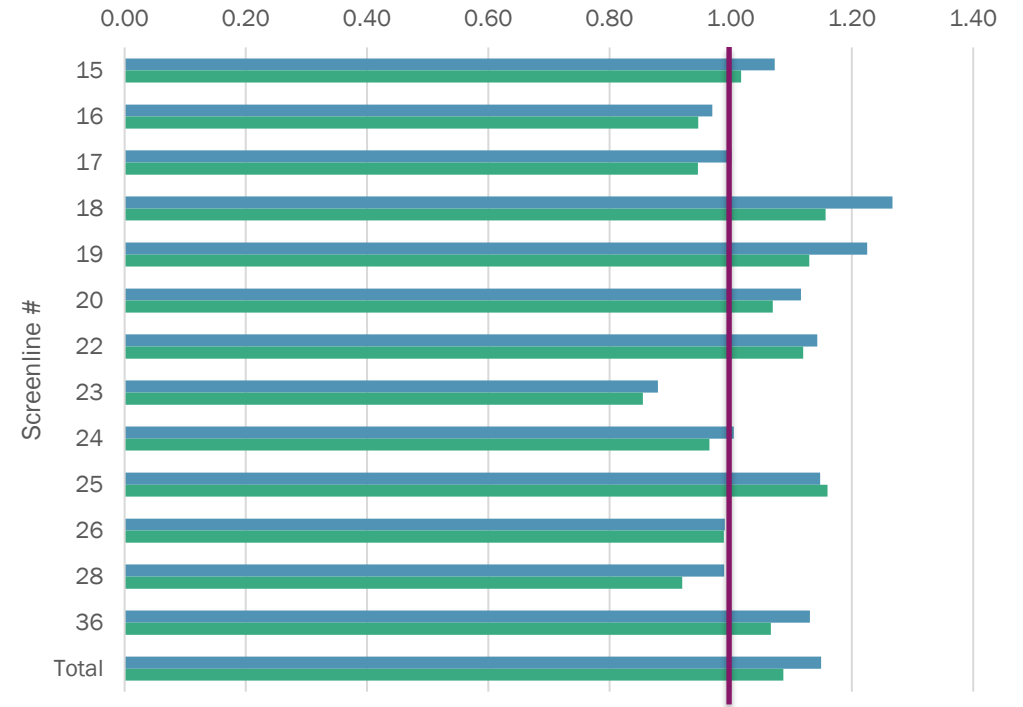- Theodore Roosevelt Bridge
- Chain Bridge

Screenline Validation to Teralytics' AADT Data

Ratios of Teralytics' Traffic Volumes by Screenline to Observed Counts

Ratios of Teralytics' Traffic Volumes by Screenline to Observed Counts (continued)
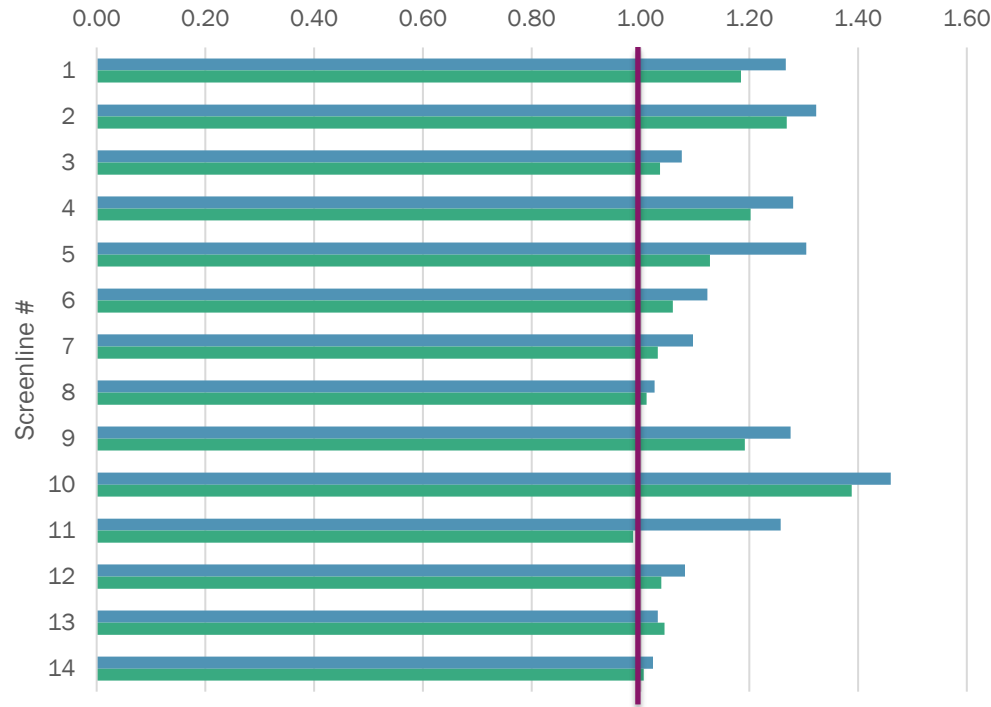
■ AADT (Data downloaded 12/04/2023) / Obs
■ AADT (Data downloaded on 08/23/2024) / Obs

National Capital Region
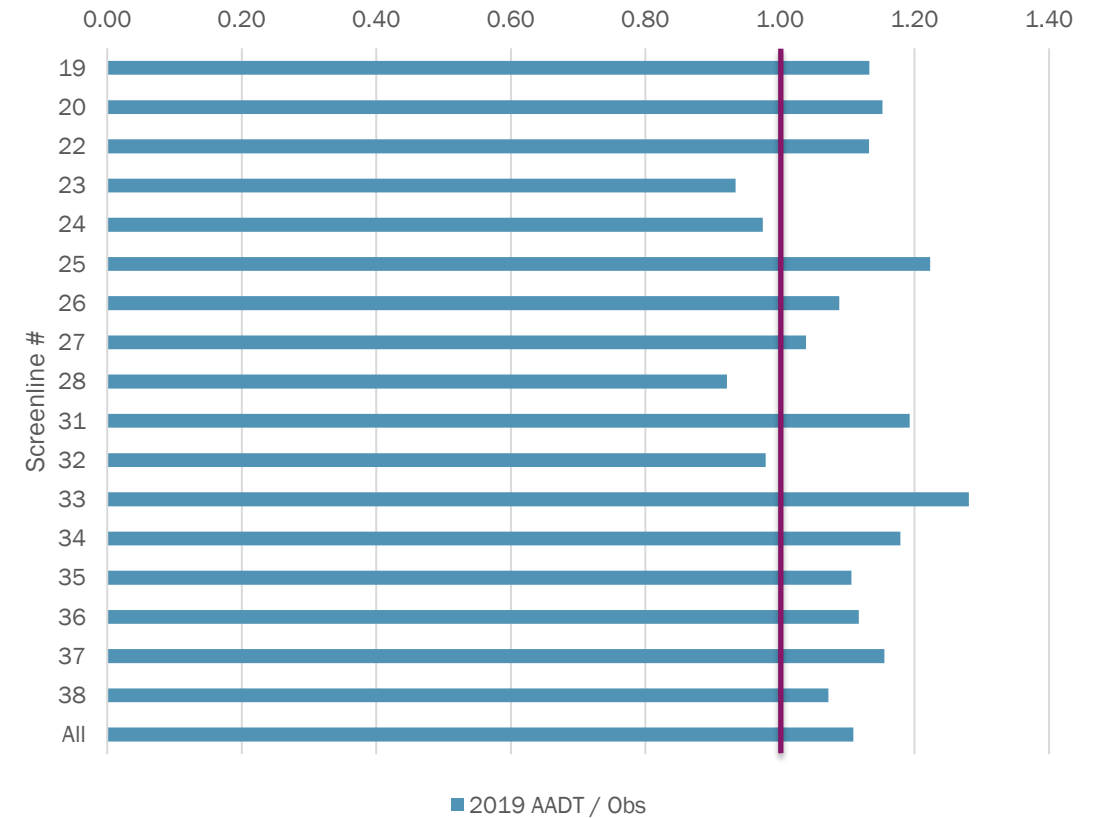**Transportation Planning Board**

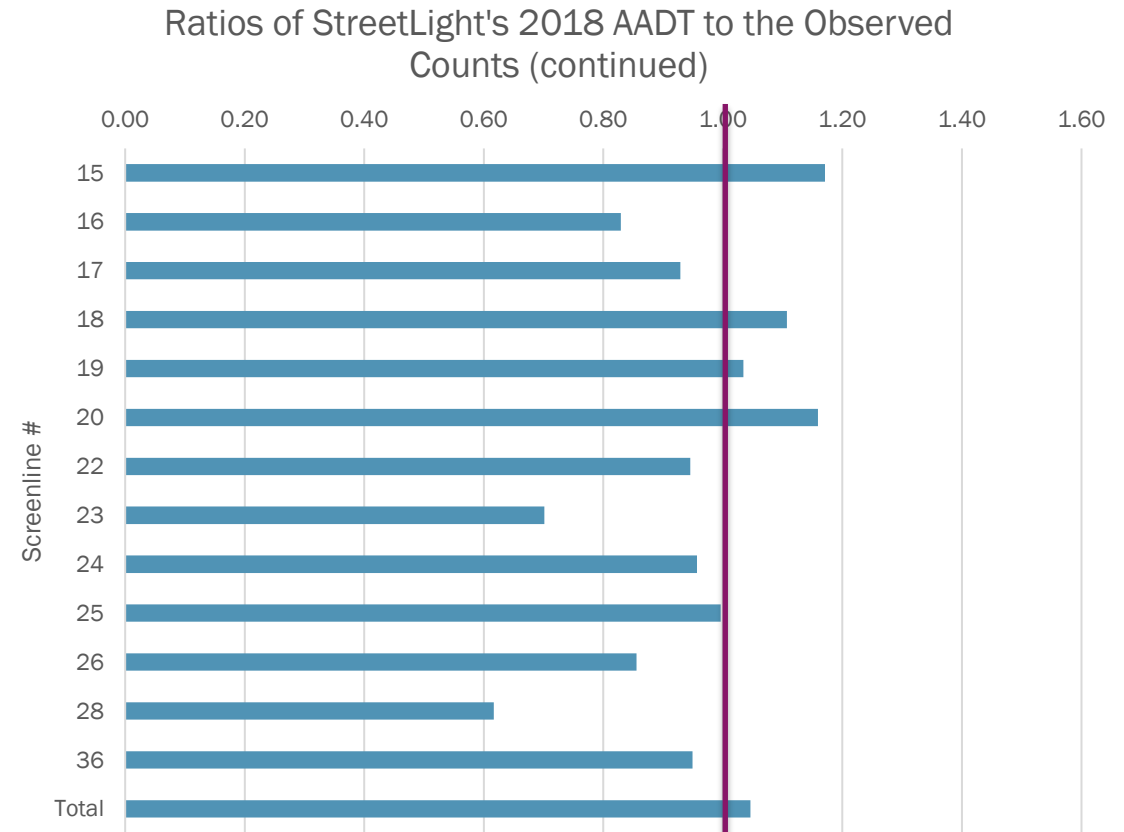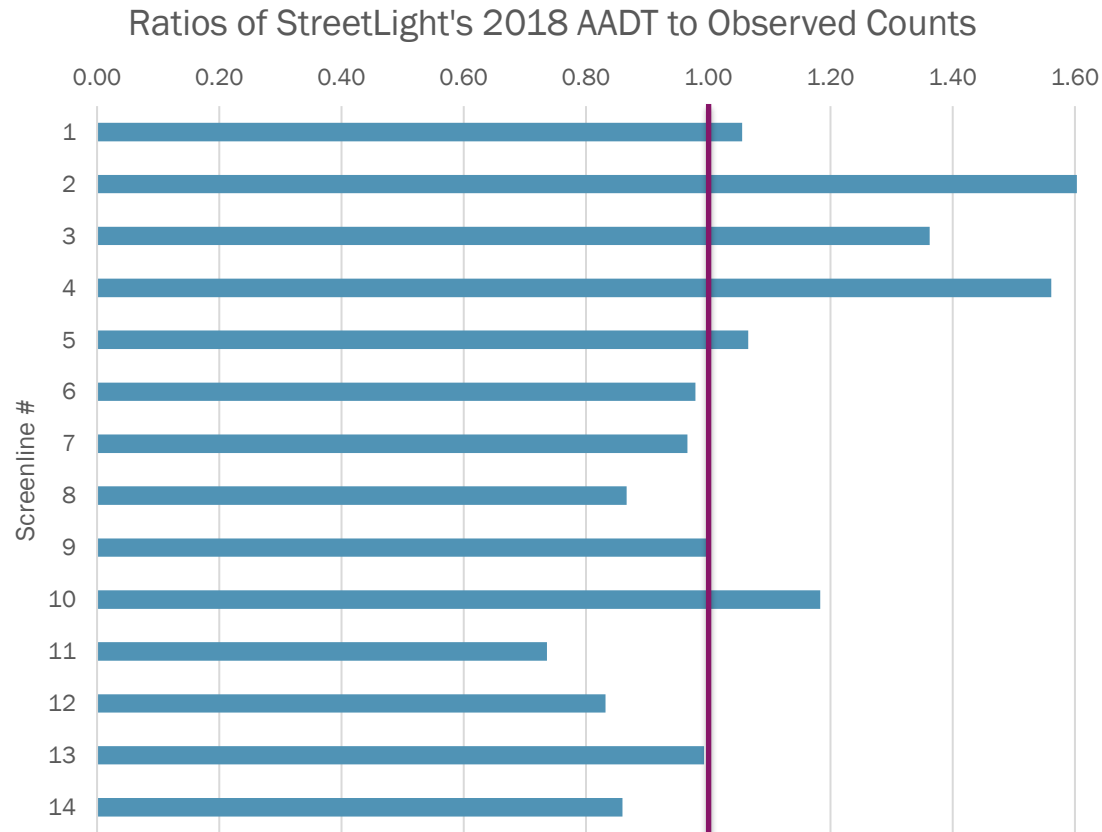# Screenline Validation to Replica's AADT Data

Ratios of Replica's Traffic Volumes by Screenline to Observed Counts
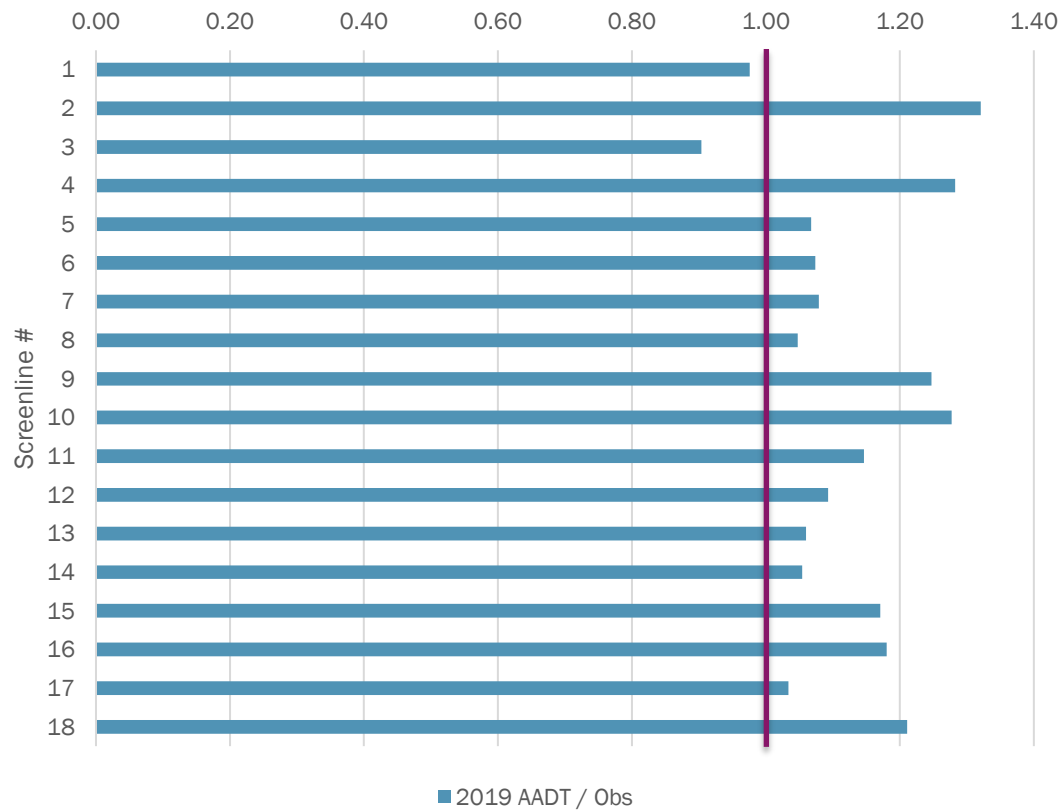
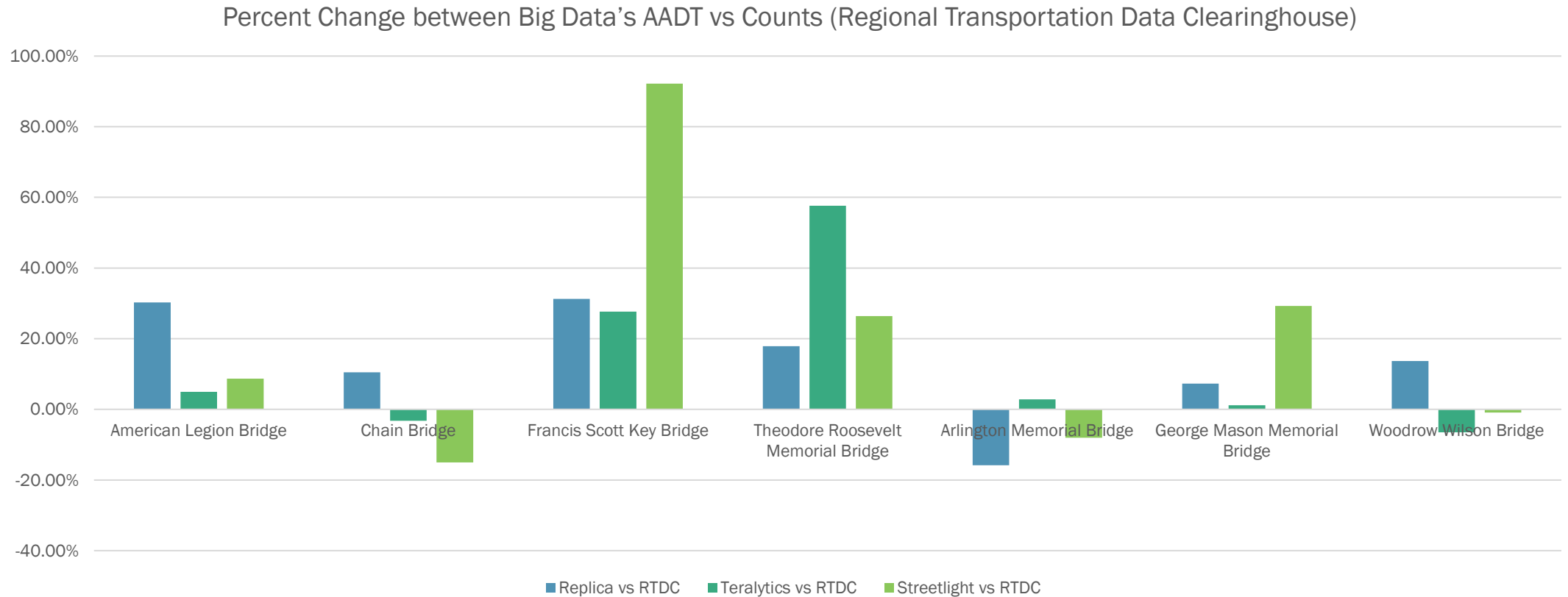Ratios of Replica's Traffic Volumes by Screenline to Observed Counts (continuted)

■ 2019 AADT / Obs

Screenline Validation to StreetLight's AADT Data

Ratios of StreetLight's 2018 AADT to Observed Counts

Ratios of StreetLight's 2018 AADT to the Observed Counts (continued)

Validation for Potomac River Bridges (Screenline #20)

Percent Change between Big Data's AADT vs Counts (Regional Transportation Data Clearinghouse)

Legend: Replica vs RTDC, Teralytics vs RTDC, Streetlight vs RTDC

VMT-by-Jurisdiction Validation to Teralytics Data

Legend: ■ VMT (Data downloaded 12/04/2023) / Observed   ■ VMT (Data downloaded on 08/23/2024) / Observed

National Capital Region
Transportation Planning Board
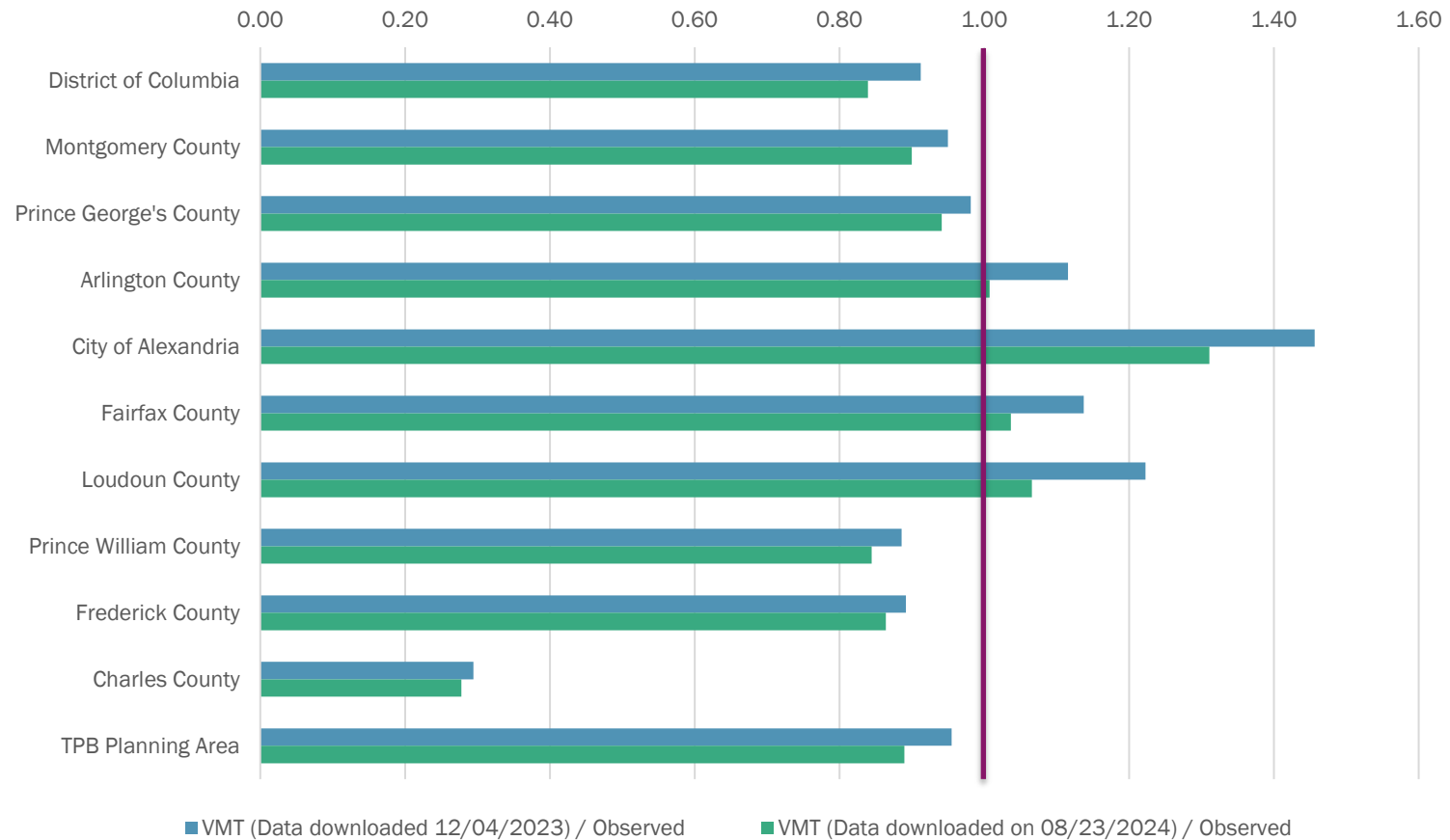
VMT-by-Jurisdiction Validation to Replica Data

Validation of Replica's VMTs by Jurisdiction

■ Replica VMT (based on 2019 AADT) / Observed

National Capital Region
Transportation Planning Board
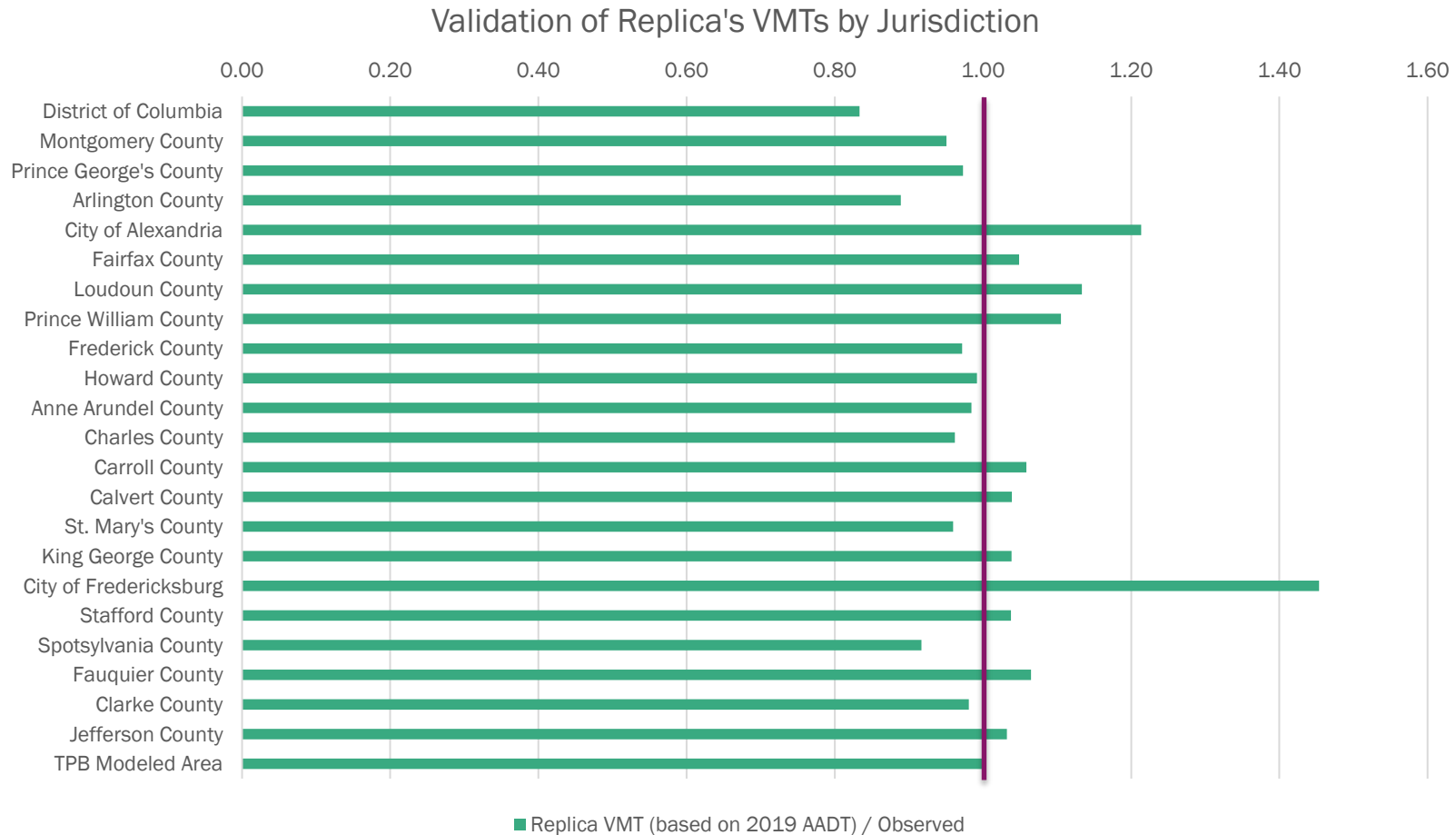
Conclusion

- Understanding Big Data sets is crucial for accurately querying and interpreting relevant data. Communication with Big Data vendors helps!

- Different vendors may have different methodologies and restrictions to query data

- Data may vary depending on when one downloads the data. Important to note timestamp for reproducibility

- The AADT screenline validation:

  - Demonstrates variable accuracy across different screenlines

  - The AADT link validation varies significantly

- The VMT validation:

  - Regional total VMT aligns well with observed data

  - Some jurisdictional VMTs show discrepancies

National Capital Region
Transportation Planning Board

Acknowledgement

- Anant Choudhary, Meseret Seifu, Feng Xie, Mark Moran (COG/TPB)

- Lauren Massey, Albab Noor (Replica)

Ray Ngo

TPB Transportation Engineer

(202) 962-3231

rngo@mwcog.org

Anant Choudhary

TPB Transportation Engineer

(202) 962-3354

achoudhary@mwcog.org

mwcog.org/tpb

Metropolitan Washington Council of Governments

777 North Capitol Street NE, Suite 300

Washington, DC 20002

National Capital Region
Transportation Planning Board