

Webinar on Phase II: Ubiquitous Traffic Volume from Probe Data

November 13, 2019



Conference call number: xxx and enter xxxx at the prompt



Webinar & Audio Information

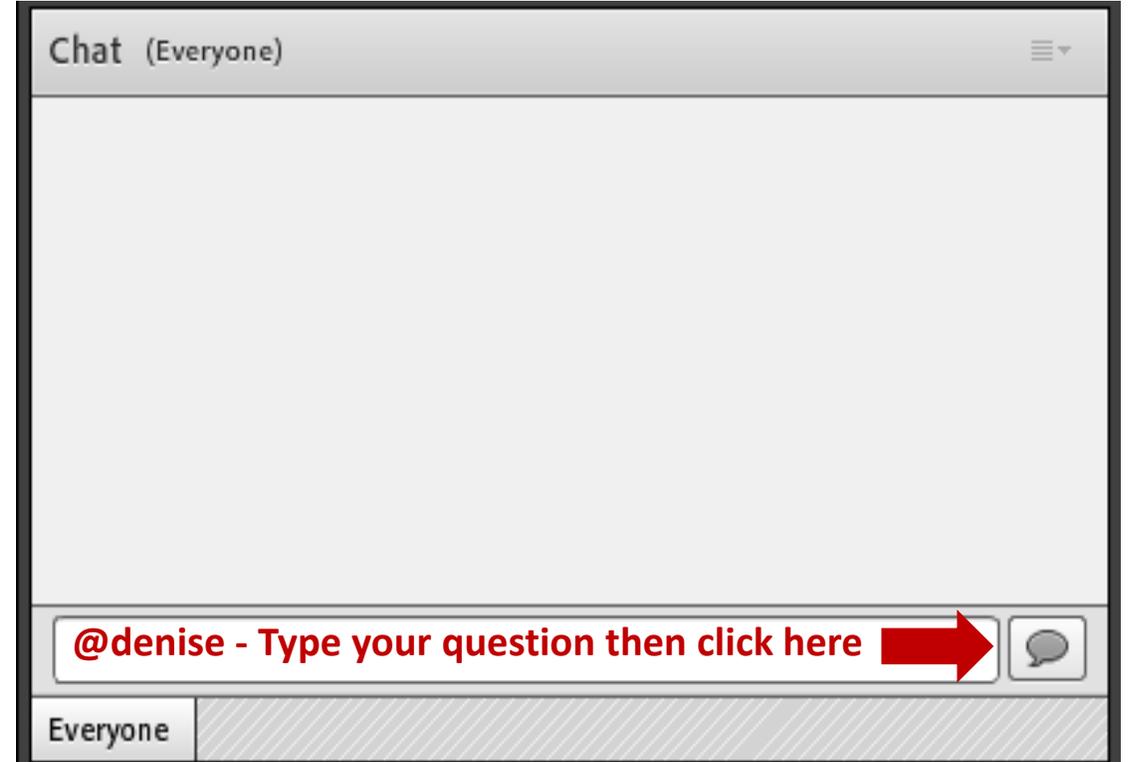
- The call-in phone number is: **xxxx & enter xxxx at the prompt**
- **Participants will be in “Listen Only” mode throughout the webinar**
- Please press *0 to speak to an operator for questions regarding audio
- Please call Justin Ferri at xxxx for difficulties with the web or audio application
- This webinar will be recorded.
- Presentations will be posted to the I-95 Corridor Coalition website. Participants will receive a link to the presentations after they are posted.



Asking Questions



- Please pose your questions using the **chat box**
- Questions will be monitored then answered by the speakers either at the end of their presentation or at the end of the webinar
- Please direct your question to the appropriate speaker



Welcome & Introductions



Denise Markow, PE

I-95 Corridor Coalition

TSMO Director

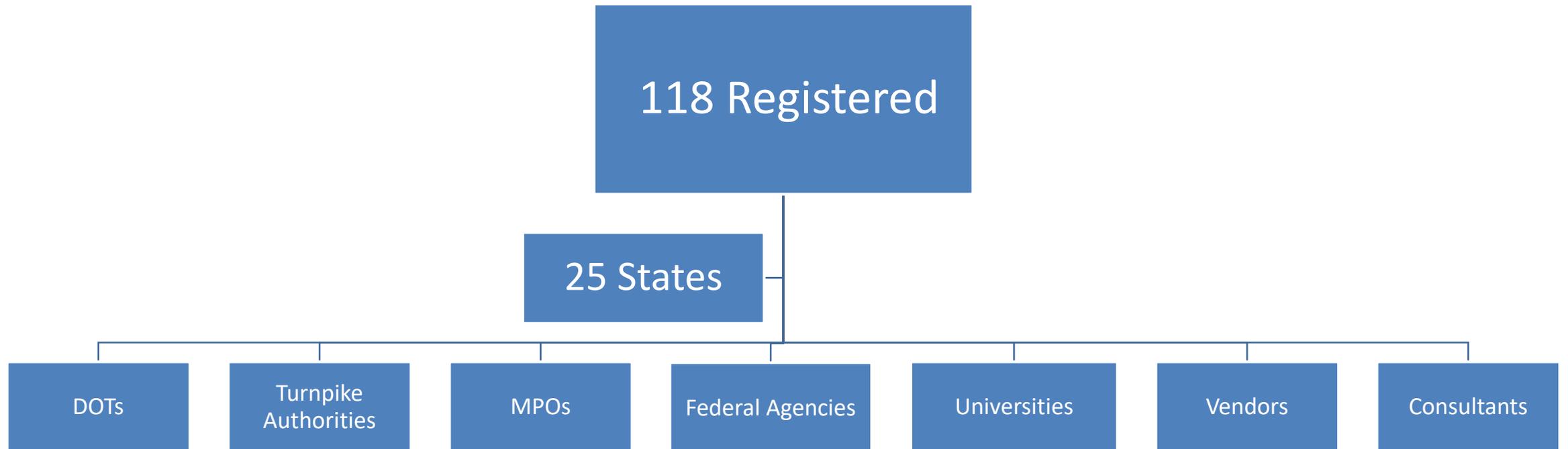


Agenda

	Topic	Speaker
1	Introductions & Welcome	Denise Markow, I-95 Corridor Coalition
2	NREL Work Effort	Stan Young & Venu Garikapati, NREL
3	UMD CATT Work Effort	Kaveh Sadabadi, UMD CATT
4	Questions & Wrap Up	Denise Markow, I-95 Corridor Coalition



I-95 Corridor Coalition Sponsored Event



Project Brief

I-95 Corridor Coalition VTM – Ubiquitous Traffic Volume from Probe Data Phase I RECAP



A new frontier in probe data & analytics – Phase I Summary

Through a Multistate Corridor Operations and Management Program (MCOMP) grant, the I-95 Corridor Coalition sponsored research to achieve accurate volume and turning movement estimates through outsourced probe data for both operations and planning purposes. **Phase I tasks are now complete and findings are available.**

Project Value

For many agencies, network-wide volume and turning movement data remain key missing dimensions for complete and actionable situational awareness, accurately assessing transportation system performance and developing targeted, cost-effective mobility projects and programs. Having the ability to easily access and leverage these data (both in real-time and historic) along with probe speed and travel time data, offers these substantial benefits:

- Improves incident management monitoring and action
- Enhances work zone monitoring, impact analysis, and safety
- Adds additional insight to anticipate and verify “jam” conditions
- Provides more accurate user delay cost reporting for weather, sporting or other events
- Improves traffic signal system timing management, enabling more cost effective, timely, and accurate updates to signal timing plans
- Provides data for more complete after-action reviews
- Advances travel demand modeling accuracy
- Better addresses air quality and emissions requirements and energy analysis inquiries

From Point Data To Ubiquitous Traffic Volume Data



Phase I Objectives Accomplished

- ✓ Created a practical and logical framework for the delivery of probe-based volume estimate.
- ✓ Documented the properties and requirements to support a variety of DOT applications.
- ✓ Developed methods to ensure and measure the accuracy of the volume estimator.
- ✓ Developed the algorithms and methods using machine learning.
- ✓ Demonstrated the process in collaboration with industry, setting expectations for fidelity, form, granularity, and usability.
- ✓ Estimated the cost and resources needed to create, support, and maintain such a system at a statewide, or even national level.

Principal Investigator

Kaveh Sadabadi
Center for
Advanced Transportation Technology
kfarokhi@umd.edu

Co-Principal Investigators

Venu Garikapati
National Renewable
Energy Laboratory
Venu.Garikapati@nrel.gov

Denise Markow
I-95 Corridor Coalition
dmarkow@umd.edu

Stanley Young – Program Manager
National Renewable Energy
Laboratory
Stanley.Young@nrel.gov

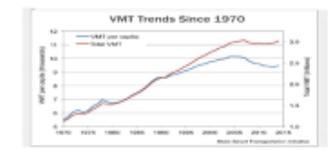
October 2019

I-95 Corridor Coalition VTM – Ubiquitous Traffic Volume from Probe Data Phase II

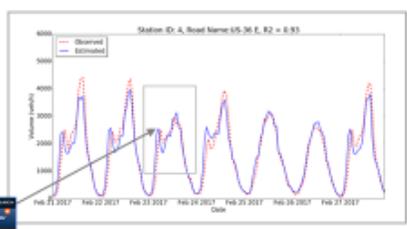


Next Steps: Moving VTM from Theory to Implementation

- Testing implementation of Phase I with agency partners
- Confirming volume estimates can be used for AADT, ADT and real-time operations applications
- Expanding calibration to arterials and turning movements
- Quantifying acceptable error bounds / thresholds for planning uses and for operations
- Exploring if probe data can be used to test accuracy of non-ATR counters
- Summarize lessons learned and tips to address conflation needs
- Maintain neutral third party aspect of research



Measures VMT more efficiently and measuring VMT increasingly matters



Effectively captures volume changes due to February snow storm

What our members are saying

"Real-time volume data would be of great value to NCDOT, especially for incident and work zone management - including timelier detouring or route diversions - better control of evacuations in the event of a hurricane, and improved special event traffic management."



Kelly Wells, PE
State Traveler Information Engineer
North Carolina Department of Transportation

"Having robust estimated volume and turning movement data derived from probe data would be a tremendous asset for DVRPC, complementing the speed and travel time data we're already using from the VPP Project to facilitate analysis of our entire road network, including problem identification, project development, and comprehensive, accurate system performance evaluation."



Jesse Buark
Capital Project Development Manager
Delaware Valley Regional Planning Commission

October 2019

Available on the Coalition website - https://i95coalition.org/wp-content/uploads/2015/02/VTM-Project-Profile-Phase-2_10-1-2019.pdf?x70560



Speakers



Stanley Young, PhD, PE
National Renewable Energy
Laboratory
(NREL)
Stanley.young@nrel.gov



Venu Garikapati, PhD
National Renewable Energy
Laboratory
(NREL)
Venu.Garikapati@nrel.gov



Kaveh Sadabadi, PhD
Center for Advanced
Transportation Technology -
University of Maryland
(UMD CATT)
kfarokhi@umd.edu



Ubiquitous Estimation of Traffic Volumes



Stanley Young, PhD, PE

National Renewable Energy Laboratory (NREL)





Ubiquitous Estimation of Traffic Volumes

Venu Garikapati and Stan Young
National Renewable Energy Laboratory
November 2019

Motivation and Background

Volume
Estimation

- Colorado
- North Carolina
- Chattanooga, Tennessee

Data
Processing

- Harrisburg, Pennsylvania

Closing Remarks

NREL Team Members



Stanley Young



Venu Garikapati



Yi Hou



Chris Hoehne



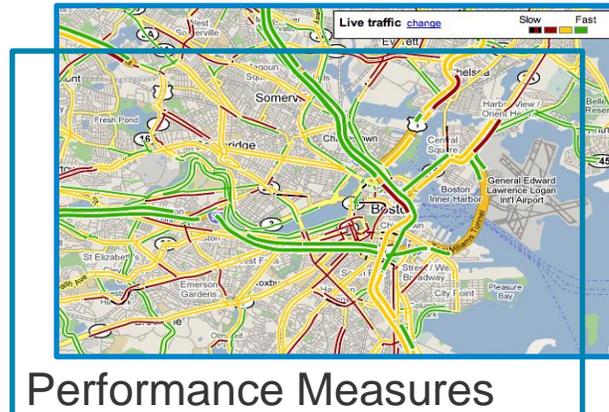
Jinghui Wang



Kevin Kasundra

Why Do We Need More and Better Volume Data?

- **Operation**
 - Detect real-time traffic volume in the network
 - Traffic volume during inclement weather and special events
- **Planning & Performance measure**
 - Assess user costs
 - Utilization of existing capacity
 - AADTs – measured, not modeled
- **Economic and energy assessment**
 - Estimate economic impact of congestion
 - Quantify VMT and energy use



Ubiquitous Traffic Volumes



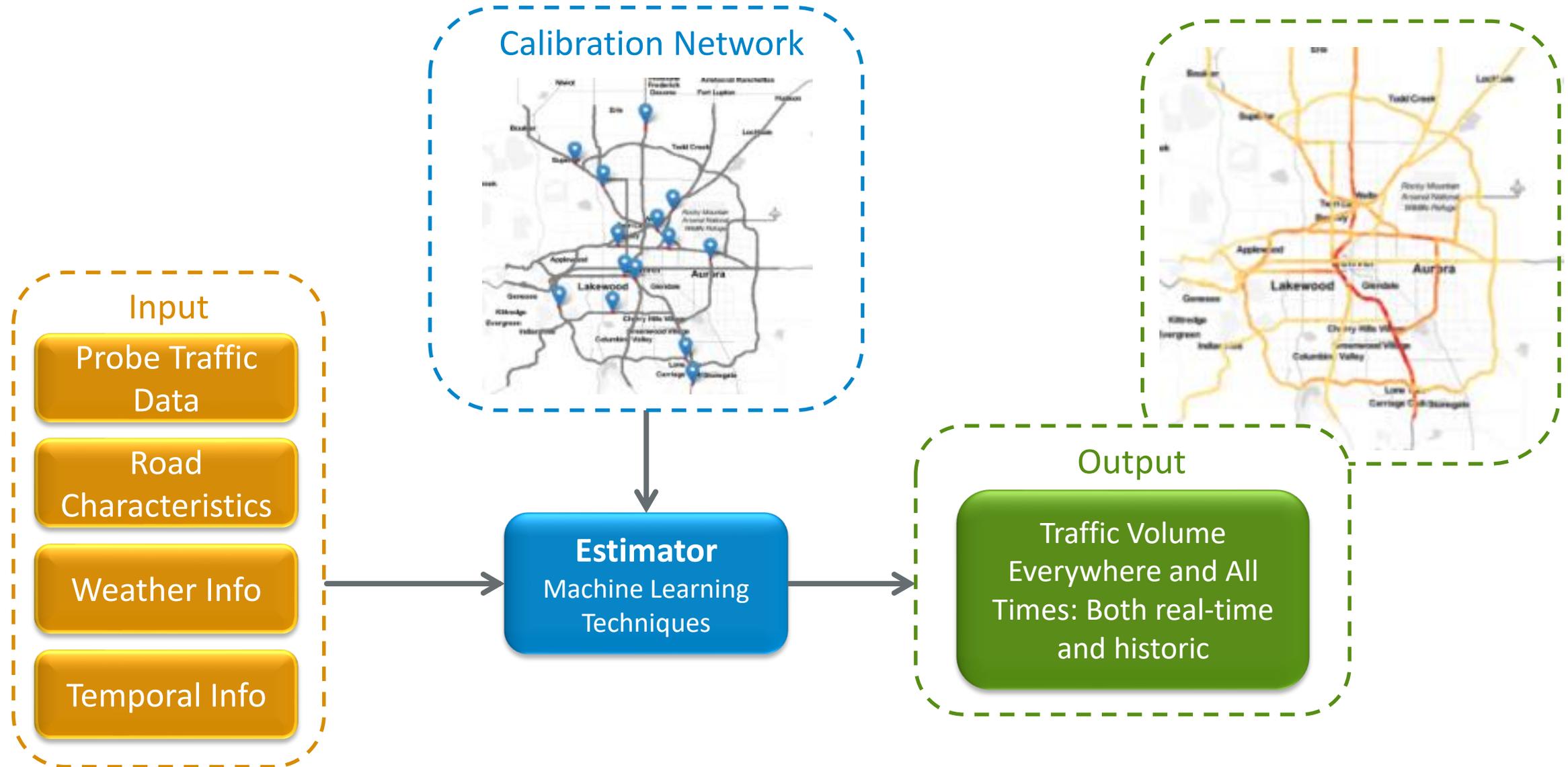
Ubiquitous network observability

- Ideal but expensive to achieve with sensors

Best alternative

- Utilize and fuse existing high-quality yet sparse data with probe data to predict traffic volumes on every road

Proposed Solution



Phases of Research

Phase I

- I-95 Corridor Coalition Volume-Turning Movement Research Initiative

Phase II

- Prototype Phase – ‘Getting from the Lab to the Streets’

Phase III

- Path forward for AADT

Data Sources

- State Department of Transportation (**DOT**) **continuous count** stations or 48-hour **short-term counts**
- Federal Highway Administration (**FHWA**) - Travel Monitoring Analysis System (**TMAS**) **volume data** -
 - Primarily continuous count data from states, but quality checked extensively.
- **Weather data** from The Weather Company
- **Vehicle Probe Data** data from TomTom
 - Number of observed probes and average speed

Input Variables

TomTom
vehicle probe
data

- **Probe vehicle counts, hourly average speed**

Weather

- Temperature, precipitation, wind, snow

Road
characteristics

- Functional road class, speed limit, longitude, latitude, direction, number of lanes

Temporal
information

- Month, day of week, hour of day

Model Estimation Process – Cross Validation

- Split data into training data and test data
- Training locations were randomly and evenly divided into 10 groups
- Repeat this for 10 times
 - 9 groups are used for model training
 - 1 group is used for model validation
- Find model hyperparameters that yield the best estimation results
- Train a model using all training data and test model on test data



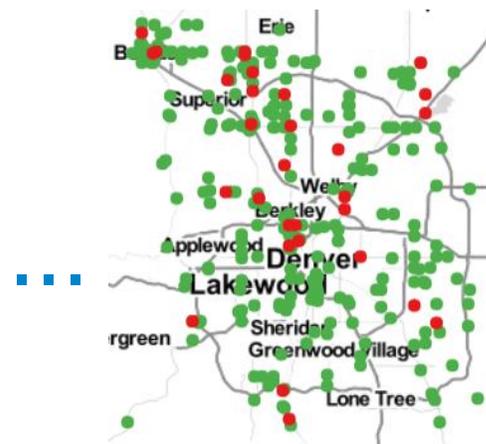
1st iteration



2nd iteration



3rd iteration



10th iteration

Red: Validate
Green: Train

- **Machine Learning**

- Extreme Boost Machine (XGBoost)
- One of the most successful ML algorithm for prediction
- Applied in travel demand and travel time predictions

- **Advantages**

- Does not require detailed mathematical forms and assumptions on variable distributions
- Suitable for capturing the underlying relationships among different variables in an environment of uncertainty
- Fast and scalable to large datasets

- **Disadvantages**

- Only predicts within bounds of training – no extrapolation

How Good is Good Enough?

Traffic Engineer

- **Mean Absolute Percentage Error (MAPE)**
 - Volume dependent - estimate
 - 10-15% High Volume
 - 20-25% Mid Volume
 - 30-50% Low Volume
 - (Mean Absolute Error may be appropriate)

Statistician/
Planner

- **R^2 Coefficient of Determination**
 - >70% good >80% better >90% best

Highway Operations

- **Error to Capacity (ETCR) or Max Flow (EMFR)**
 - < 10% becomes useful < 5% is target
 - {For highway operations, reflective of capacity constraint situations}

MAPE is Volume Dependent!

AADT Range	Decreasing (-)	Increasing (-)
0 - 19	-100%	400%
20 - 49	-40%	50%
50 - 99	-30%	40%
100 - 299	-25%	30%
300 - 999	-20%	25%
1000 - 4,999	-15%	20%
5,000 - 49,999	-10%	15%
50,000+	-10%	10%

MNDOT Example

- **Error to Maximum Flow Ratio (EMFR)**

- Reflects error relative to the max volume observed, lower is better

$$\text{EMFR} = \frac{1}{N} \sum_{i=1}^N \frac{|V_i - \hat{V}_i|}{V_{max}}$$

- **Coefficient of Determination (R^2)**

- Explanatory power of model
- Between 0 and 1, higher is better

$$R^2 = 1 - \frac{(\hat{V}_i - V_i)^2}{(V_i - \bar{V})^2}$$

- **Mean Absolute Percentage Error (MAPE)**

- Reflects error relative to measured volume, lower is better,
- Not ideal error measure for low volumes

$$\text{MAPE} = \frac{1}{N} \sum_{i=1}^N \frac{|V_i - \hat{V}_i|}{V_i}$$

- **Mean Absolute Error (MAE)**

- Reflects simple magnitude of error, independent of the actual volume

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^N |V_i - \hat{V}_i|$$

Cross Walk Between FHWA and TomTom Road Classes

TomTom Functional Classification

Road Class	Description
0	Motorway, freeway or other major road
1	Major road, less important than a motorway
2	Other major road
3	Secondary road
4	Local connecting road
5	Local road of high importance
6	Local road
7	Local other (neighborhood street)

HPMS Functional Classification (2010)

Road Class	Description
1	Interstate
2	Other Freeways and Expressways
3	Principal Arterial
4	Minor Arterial
5	Major Collector
6	Minor Collector
7	Local



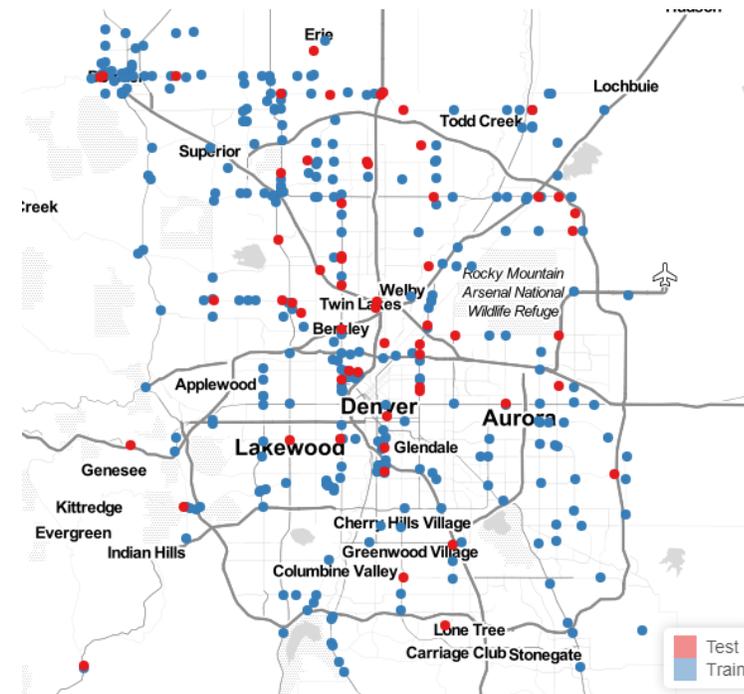
Phase – 1

Colorado Volume Estimation Freeways and **Off-freeways**

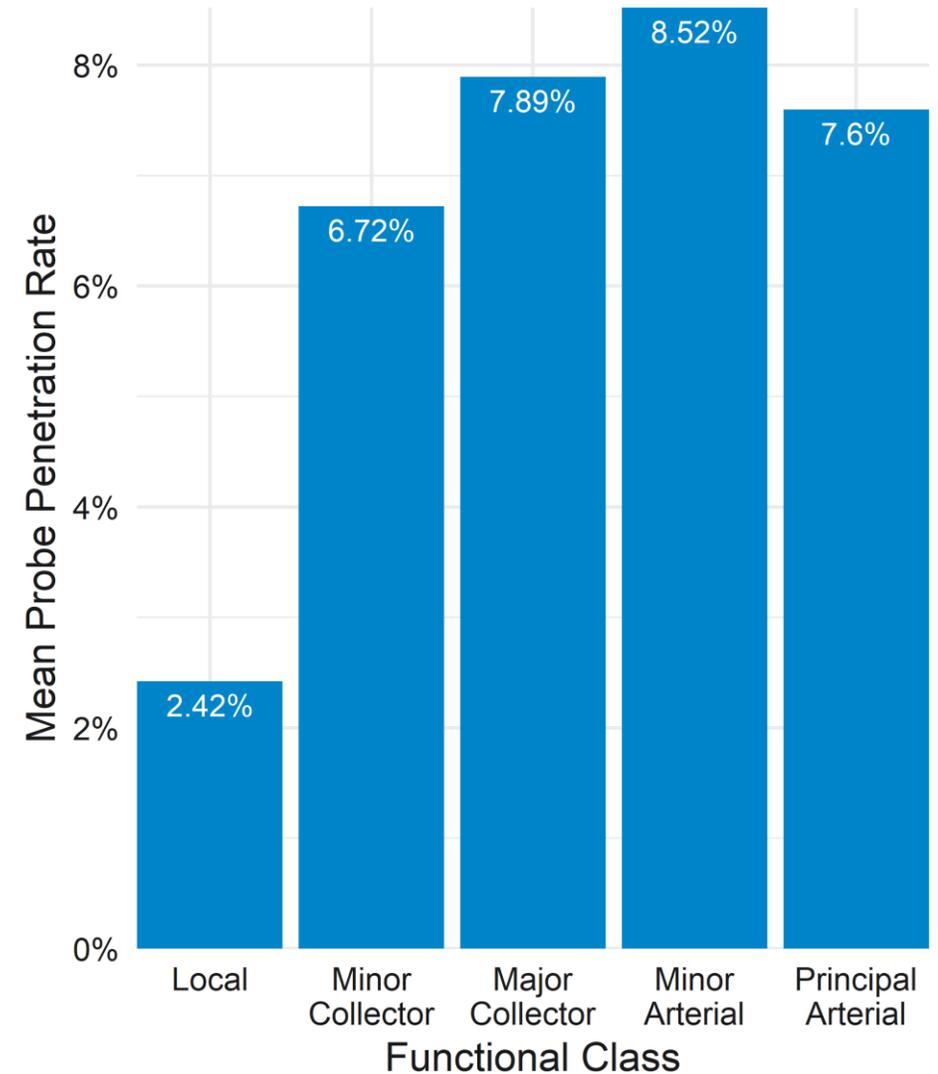
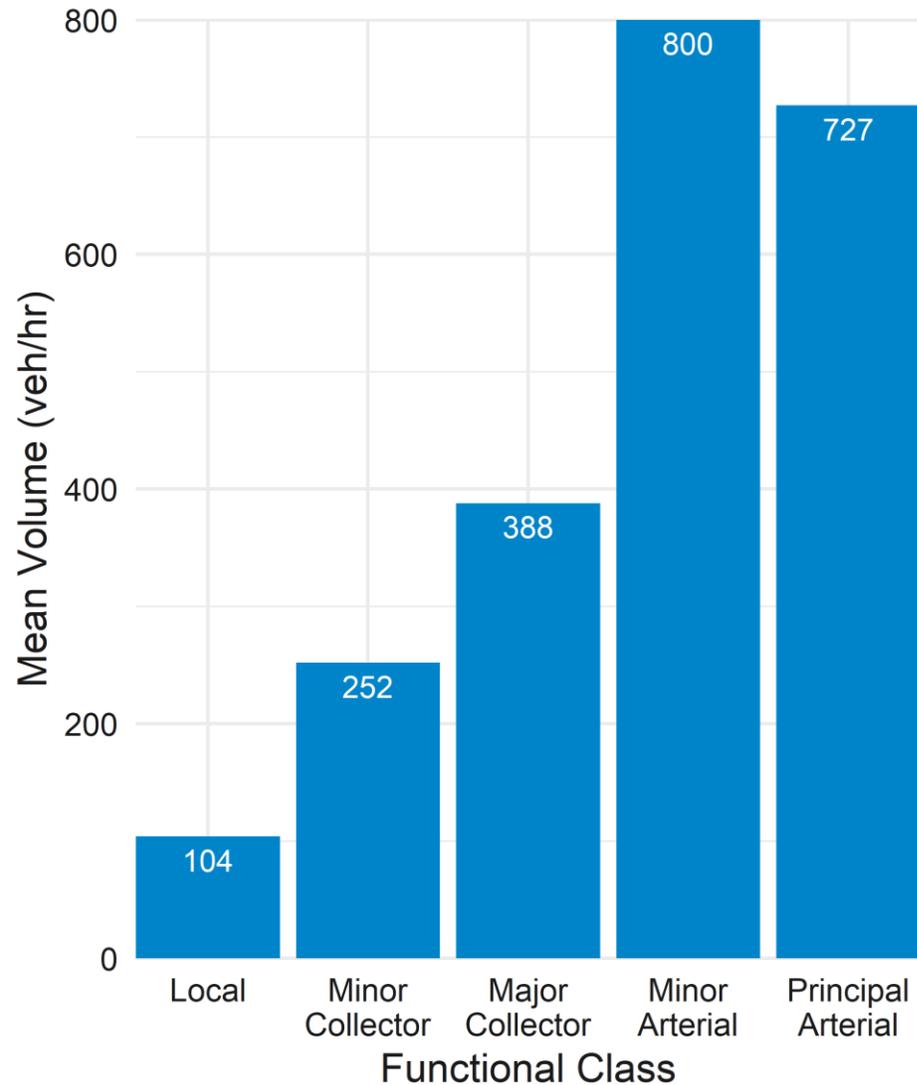
Phase I – Volume Estimation on Lower Functional Class Roads

	Lower Class Roads	Freeways
Volume data source	48-hour short-term count	Continuous count stations
Number of locations	359	14
Data collection period	Jan. – Sep., 2017 (9 months)	Feb. – Apr., 2017 (3 months)

- 300 locations for training/calibrating
 - Total of 30,096 data points
- 59 locations for testing
 - Total of 5,118 data points

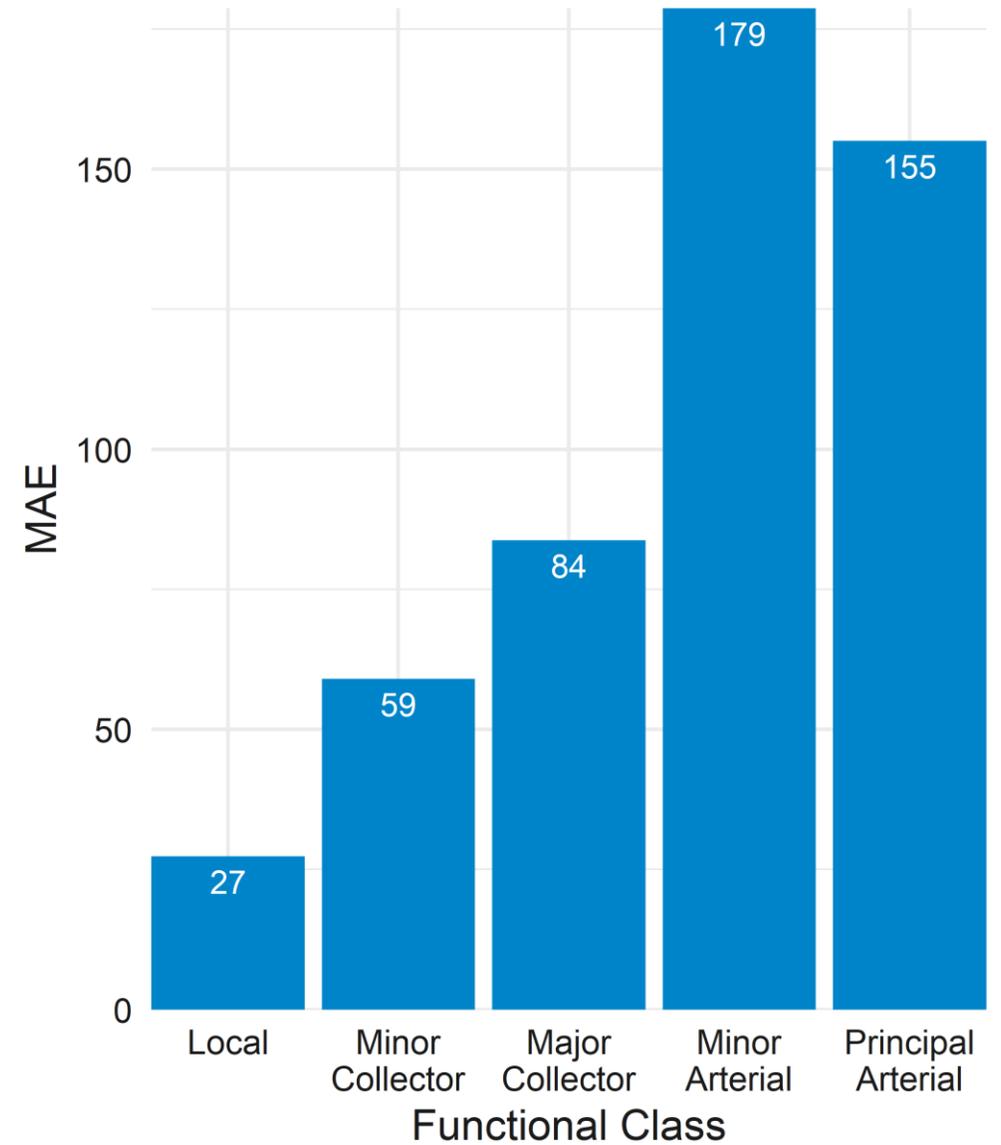


Phase I – Mean Volume and Probe Vehicle Penetration (Lower Functional Class Roads)

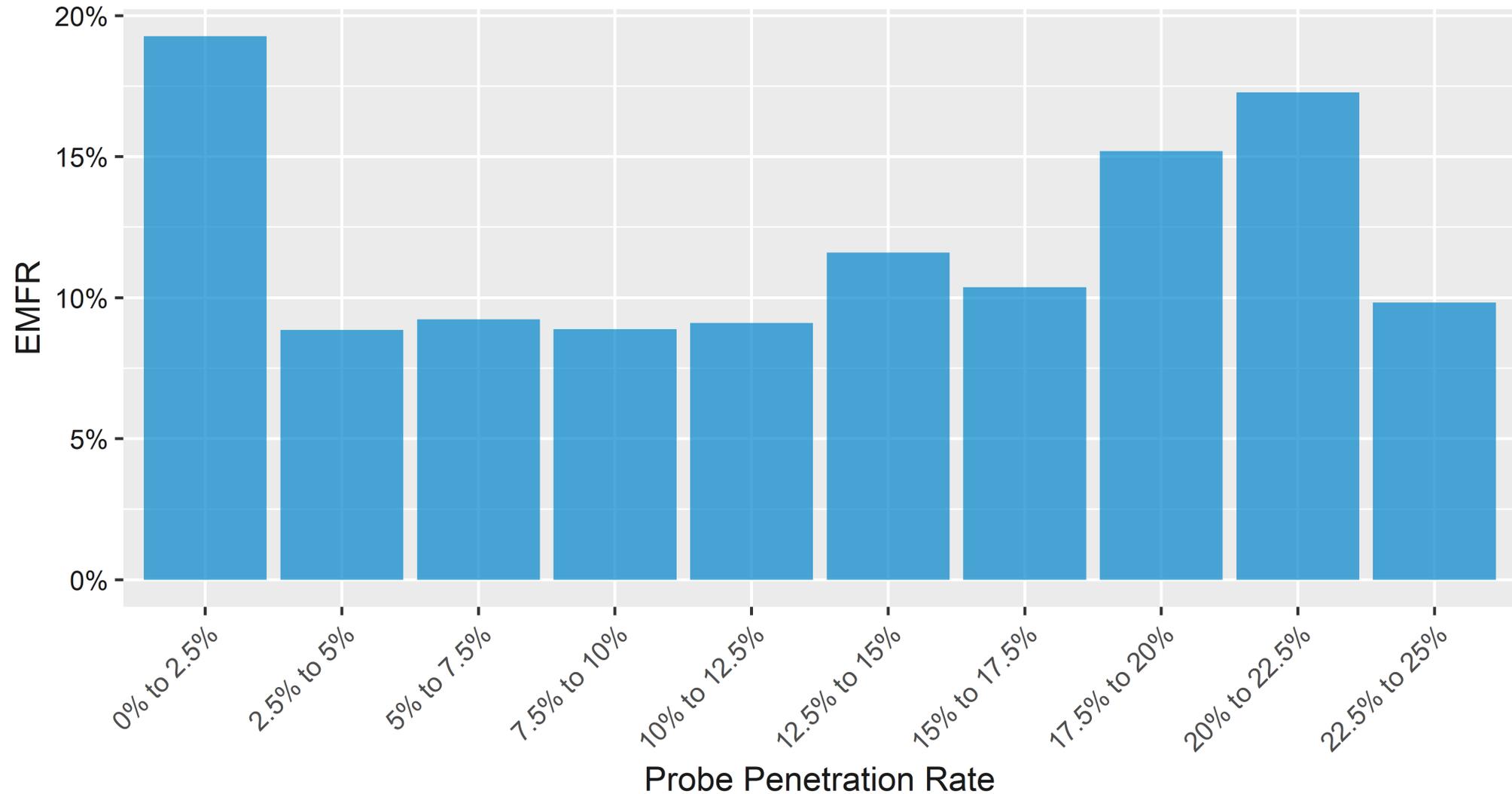


Predicted vs. Actual

Roadway Type	Model	MAE	EMFR	R2
Off-Freeway	XGBoost	89	13.2%	0.88
Freeway	XGBoost	357	5.3%	0.91

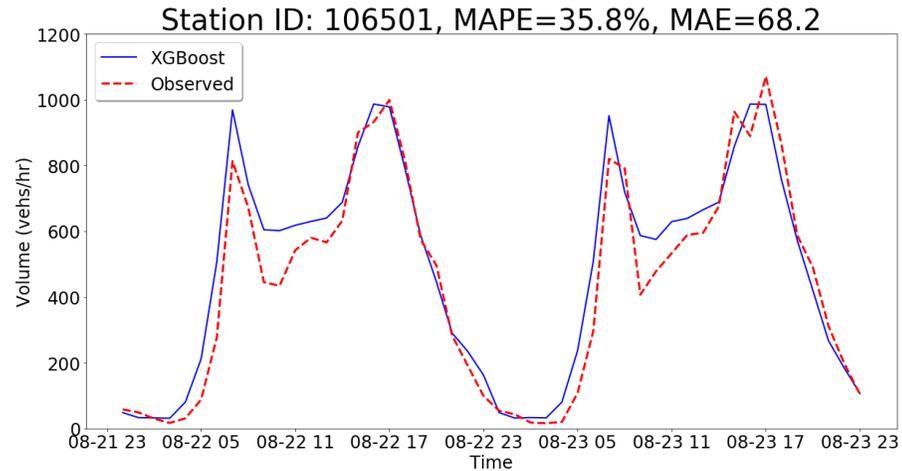


Error by Probe Penetration Rate - Lower Functional Class Roads

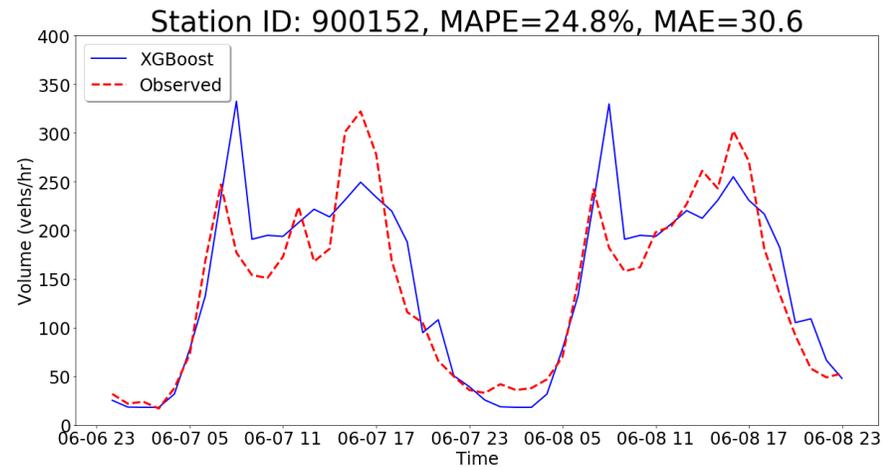


Model Performance

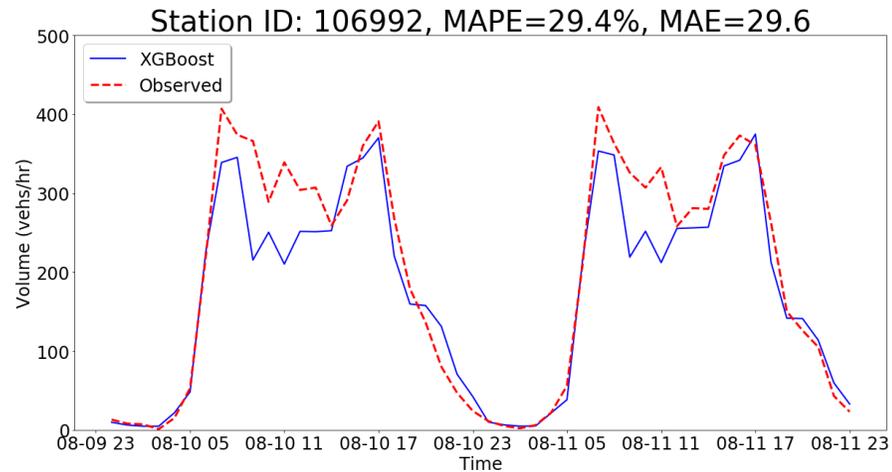
Principal Arterial



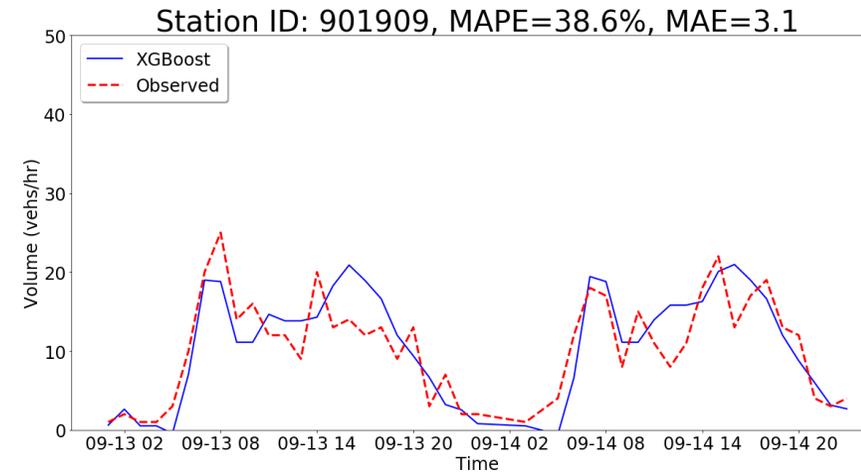
Minor Arterial



Major Collector



Local Street



- **Assessment**

- Both freeway and off-freeway model accuracies are close to expected thresholds
- Off-freeway results are slightly inferior compared to freeway results

- **Next Steps**

- Refine the model further
- Investigate reasons for slightly lower accuracies on off-freeways

Phase – II

North Carolina

Off-Freeways

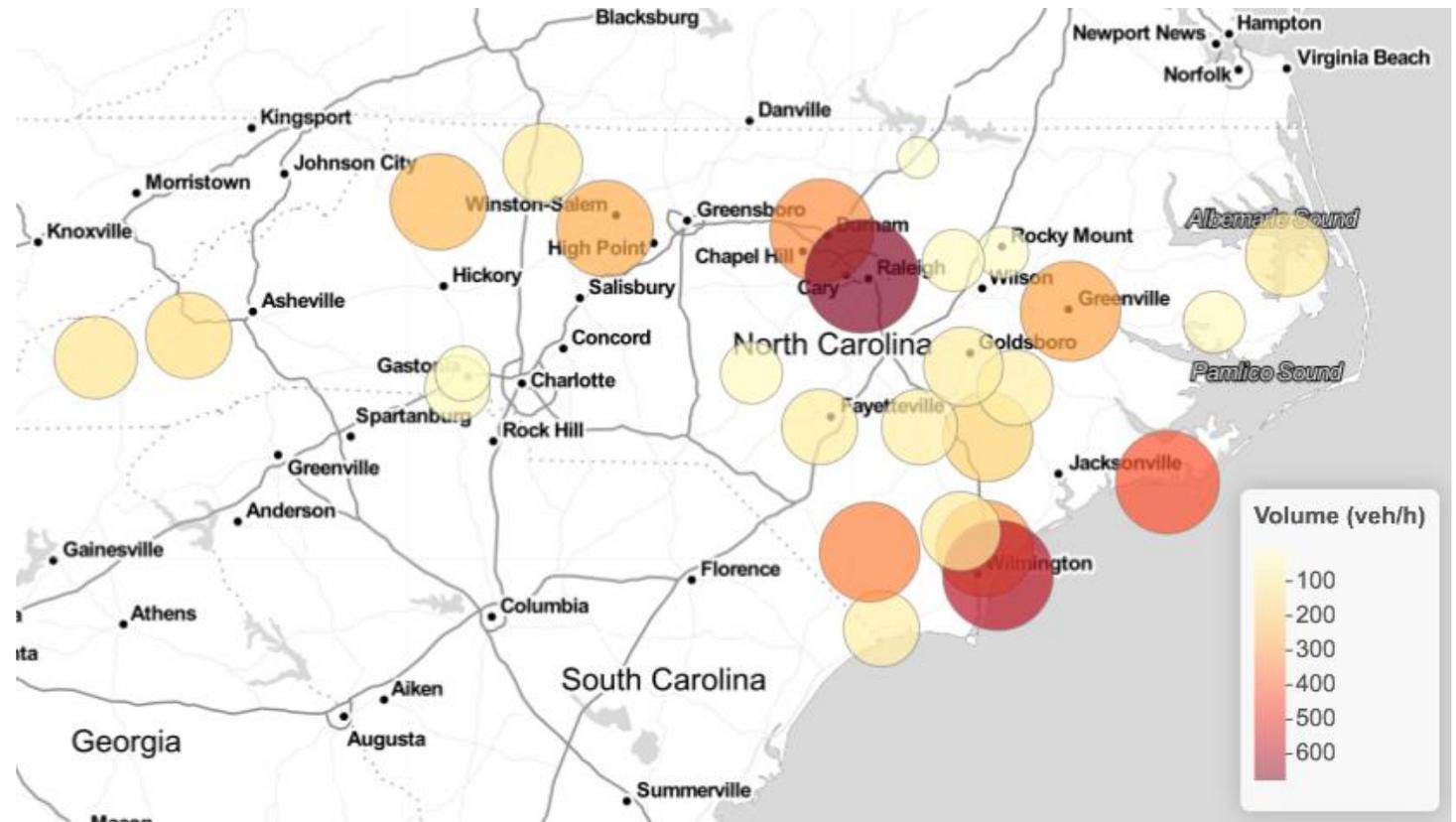
*Preliminary Results

Continuous Count Stations

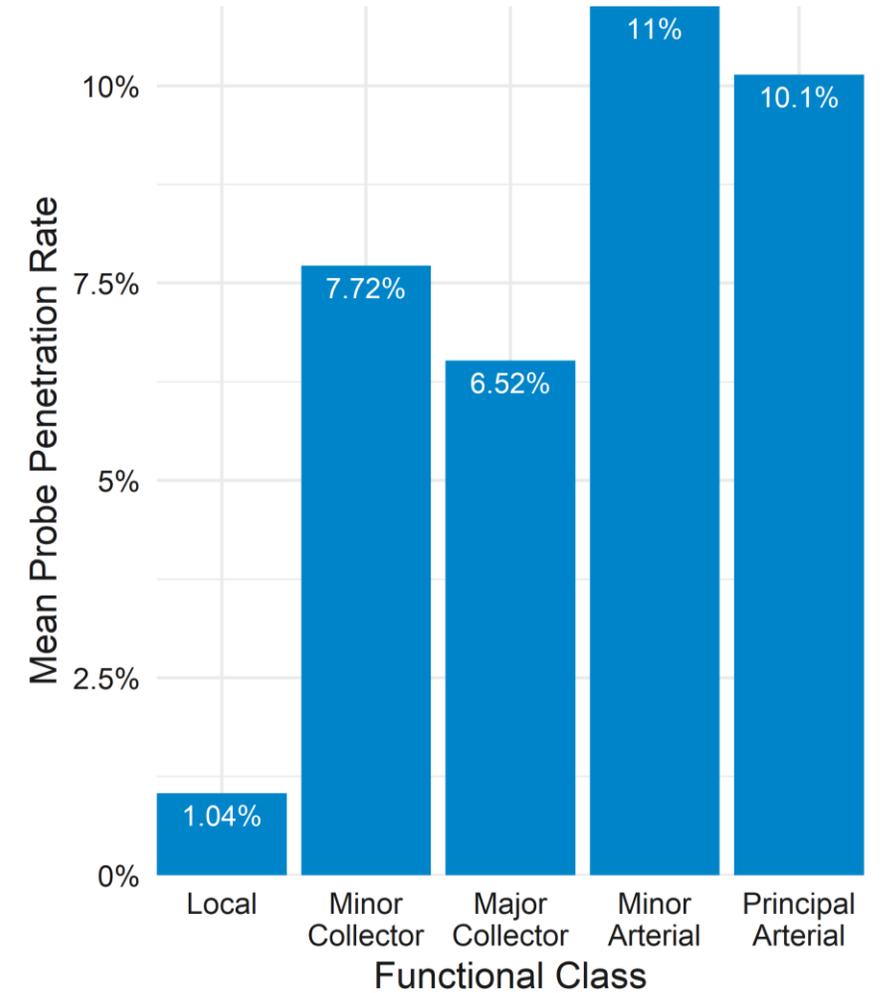
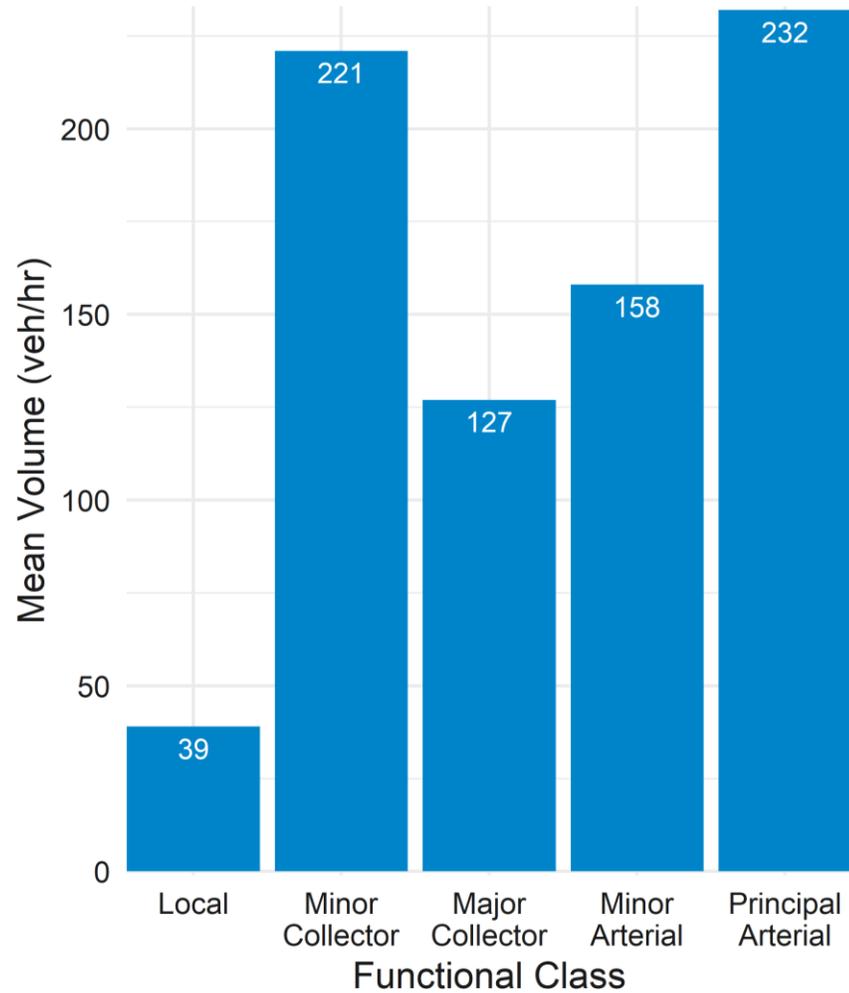
- July 1th to Dec 31th, 2018
- 27 Stations
- 144,923 observations

FHWA Functional Class	# of CCSs
Principal Arterial	5
Minor Arterial	10
Minor and Major Collectors	6
Local	6

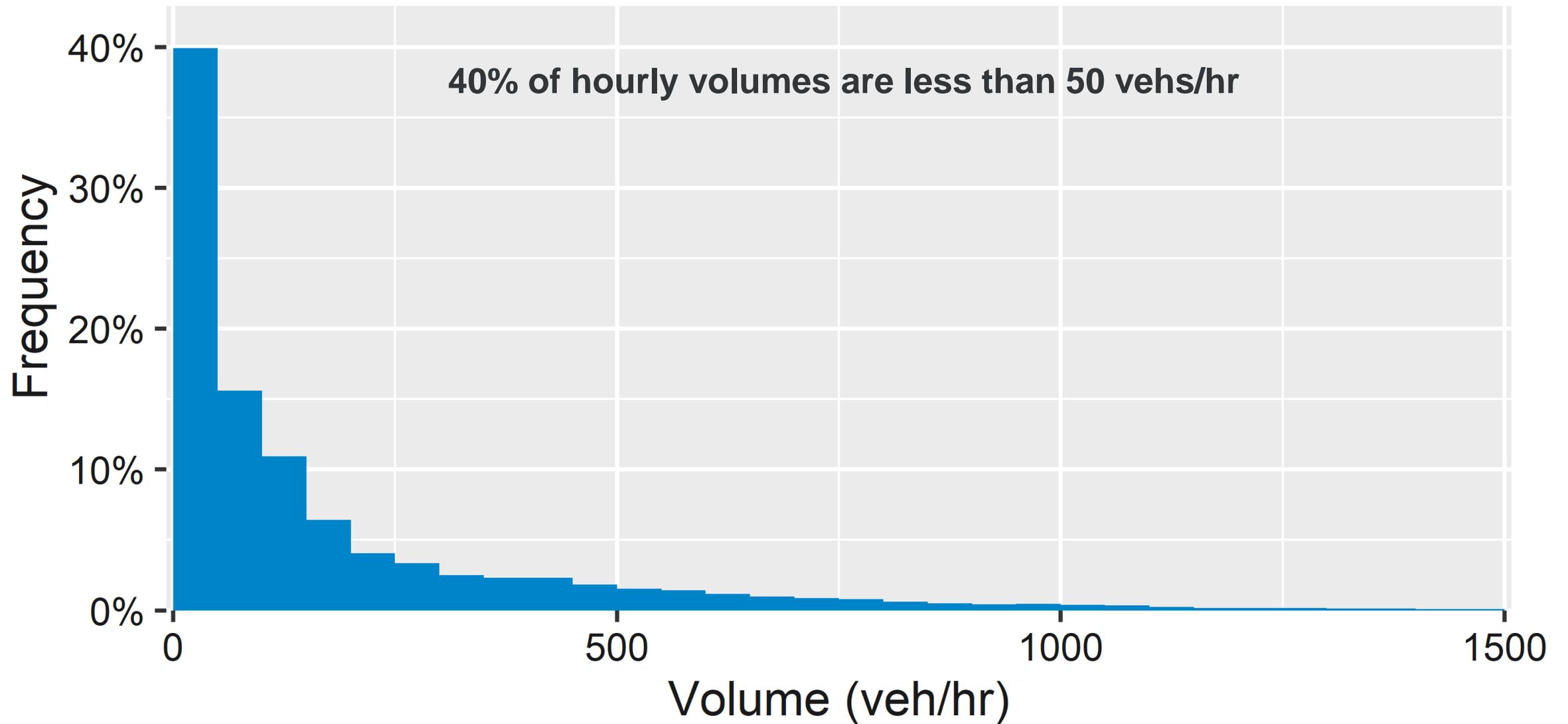
Average Traffic Volume by Station



Mean Volume & Penetration Rates by Road Class

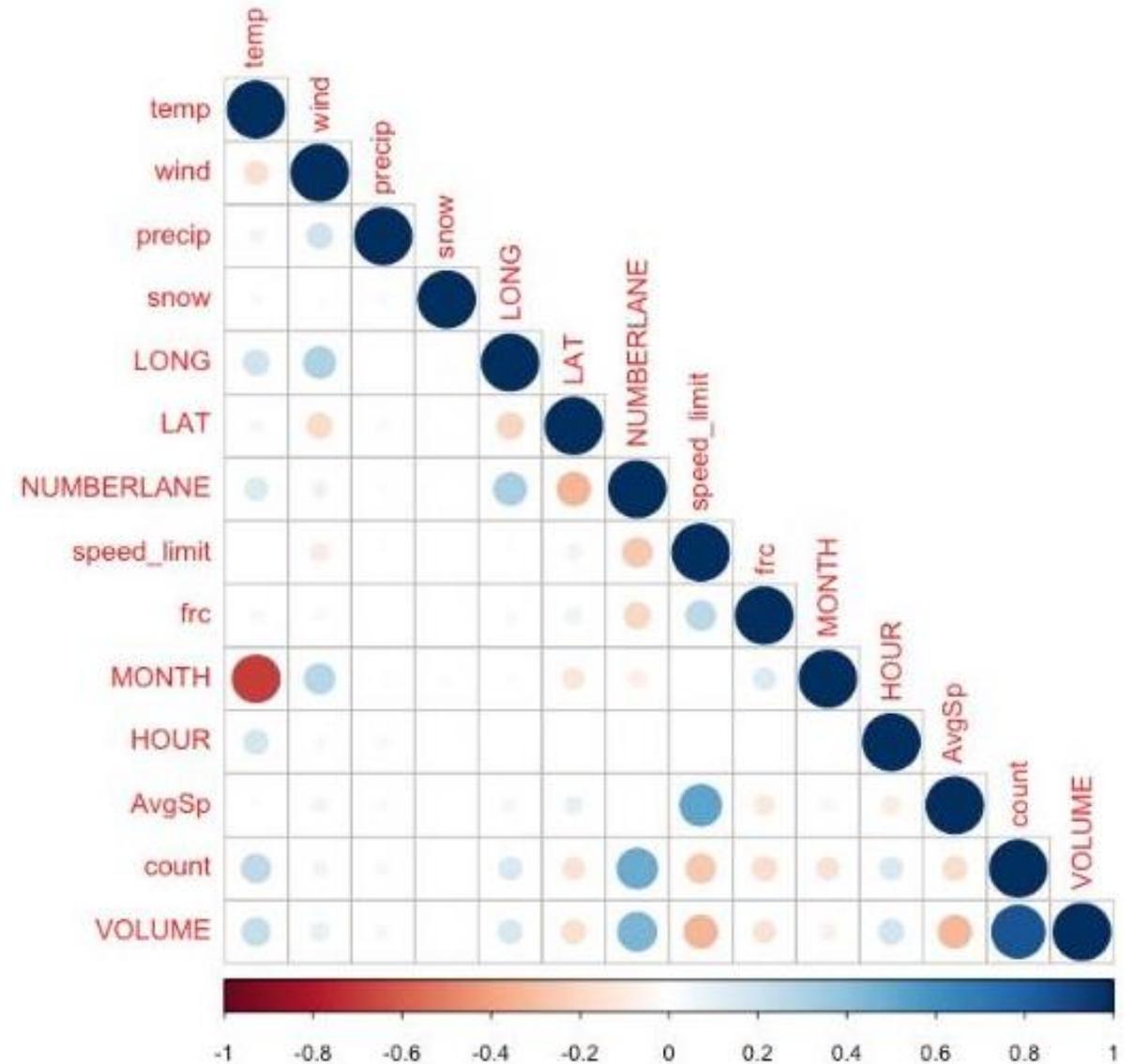


Traffic Volume Distribution



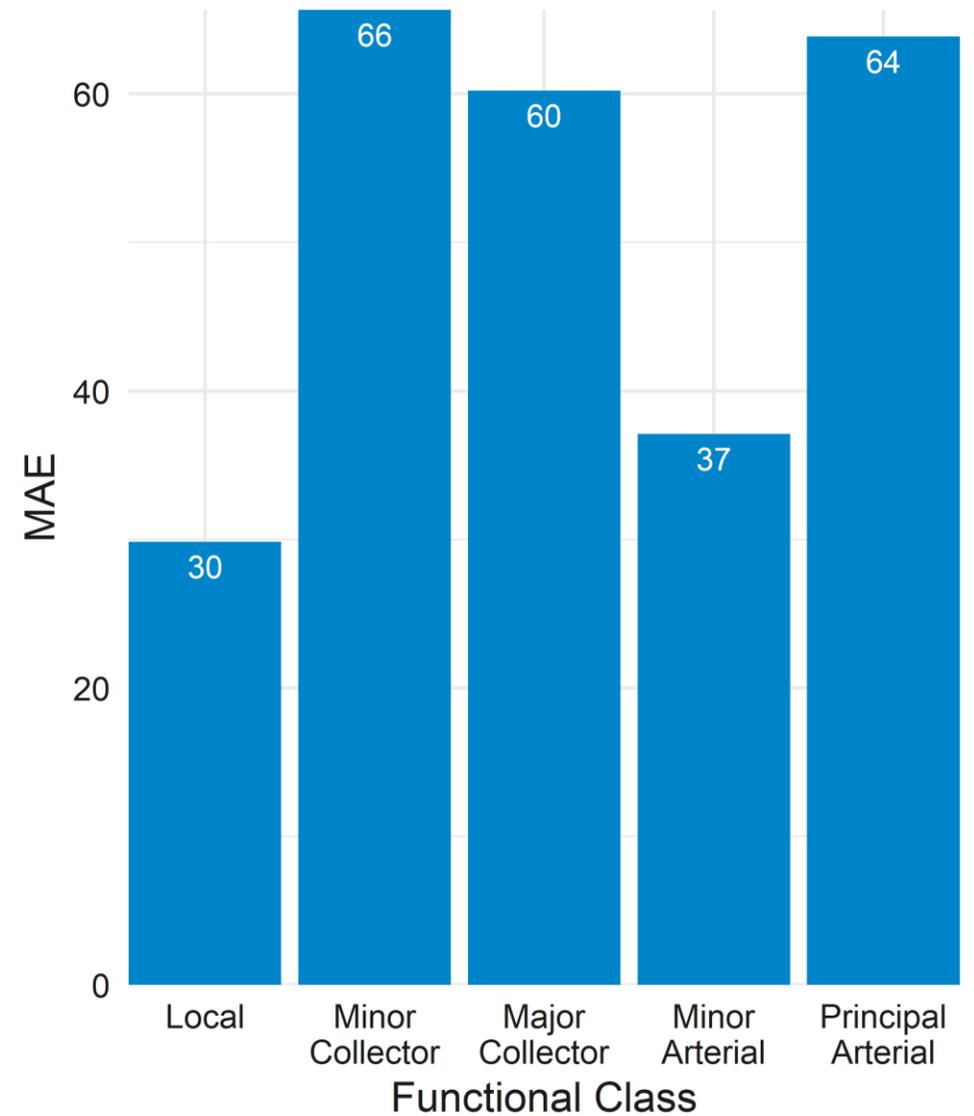
Correlation Between Variables

- Highly positive correlated variables
 - Temperature
 - Number of lanes
 - Probe counts
- Highly negative correlated variables
 - Speed limit
 - Average speed
- Weak or no correlation
 - Snow
 - Wind
 - Precipitation
 - Month



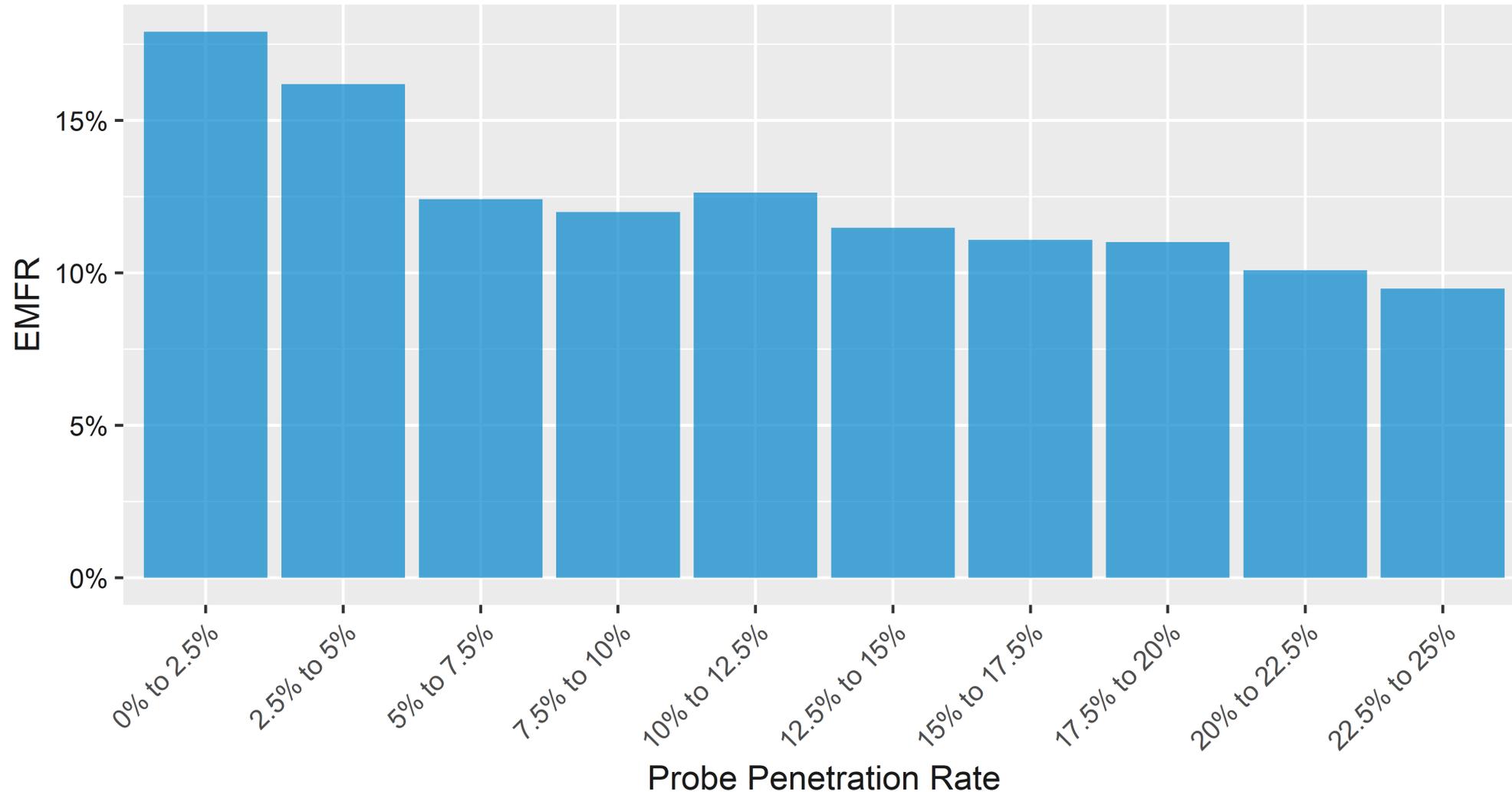
Volume Estimation Results

Roadway Type	Model	MAE	EMFR	R2
Off-Freeway	XGBoost	57	14.2%	0.87



Error by Probe Count

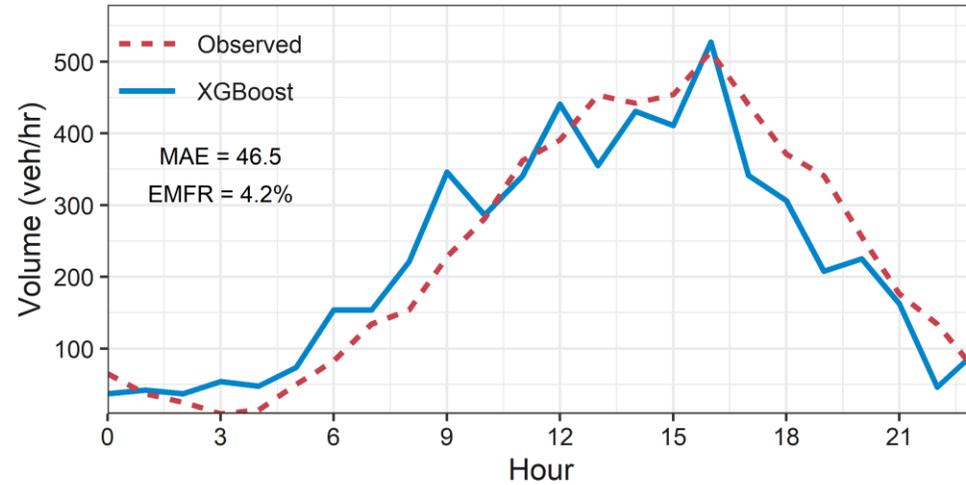
EMFR decreases as probe penetration increases



Model Performance

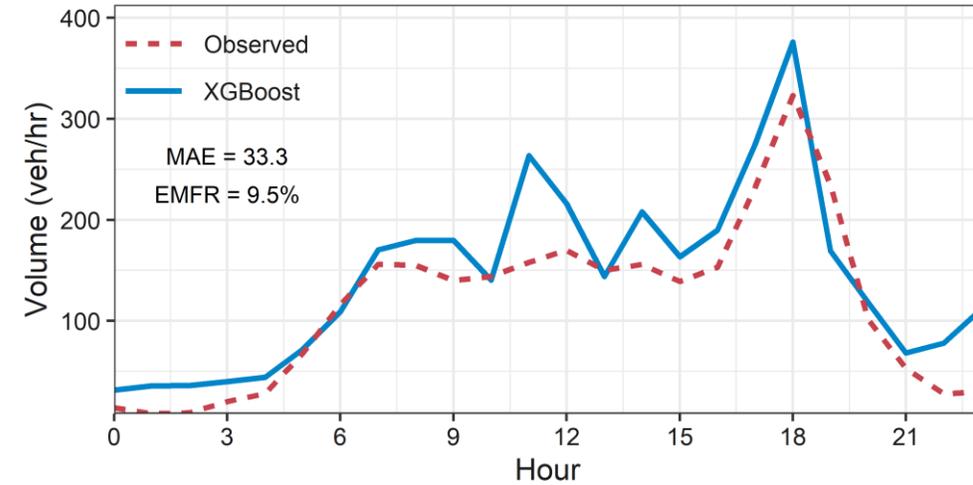
Principal Arterial

Station 0A6401 on 2018-09-03



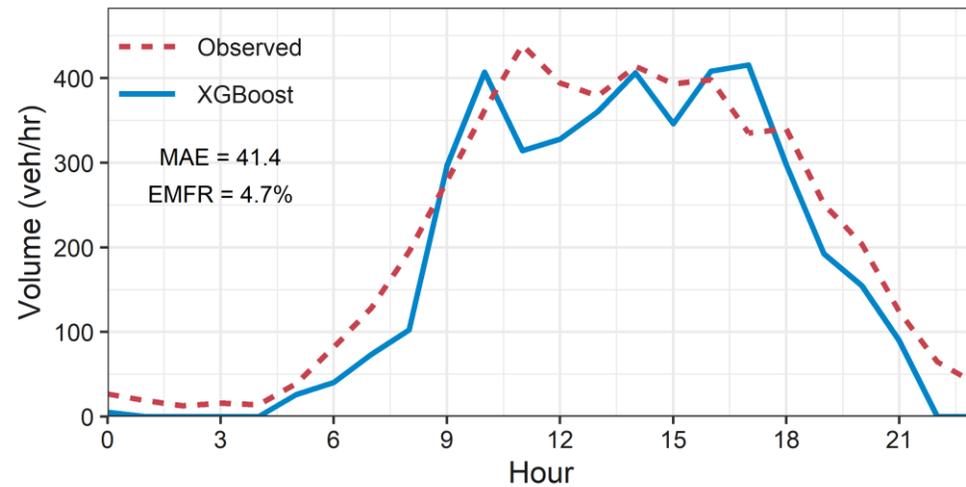
Minor Arterial

Station 0A9501 on 2018-08-22



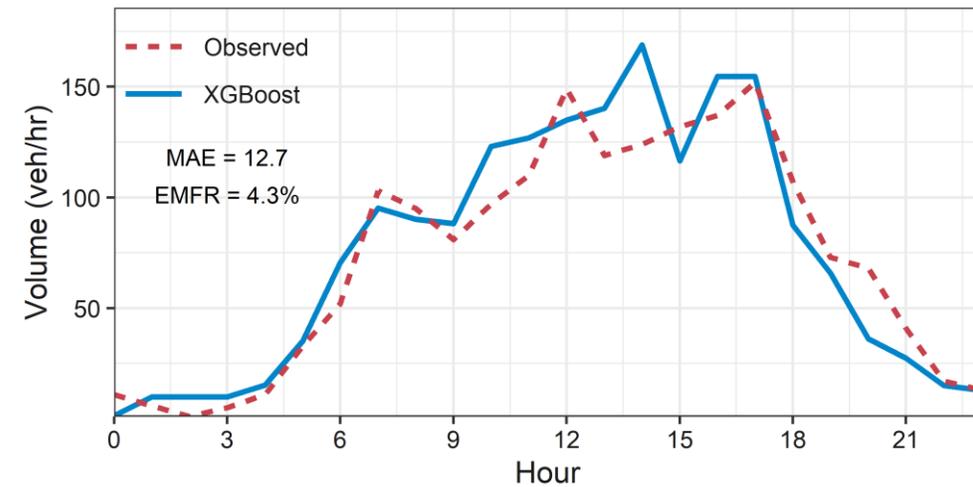
Major Collector

Station 0A3303 on 2018-09-03



Local

Station 0A8101 on 2018-08-29



- **Assessment**

- Model producing reasonable accuracies across all functional classes
- Predictions on major collectors inferior to those of other states
 - Possible reasons: Lower probe penetration; lower sample size

- **Next Steps**

- Refine the model further
- Analyze any anomalies in input data for major collectors

Phase - II

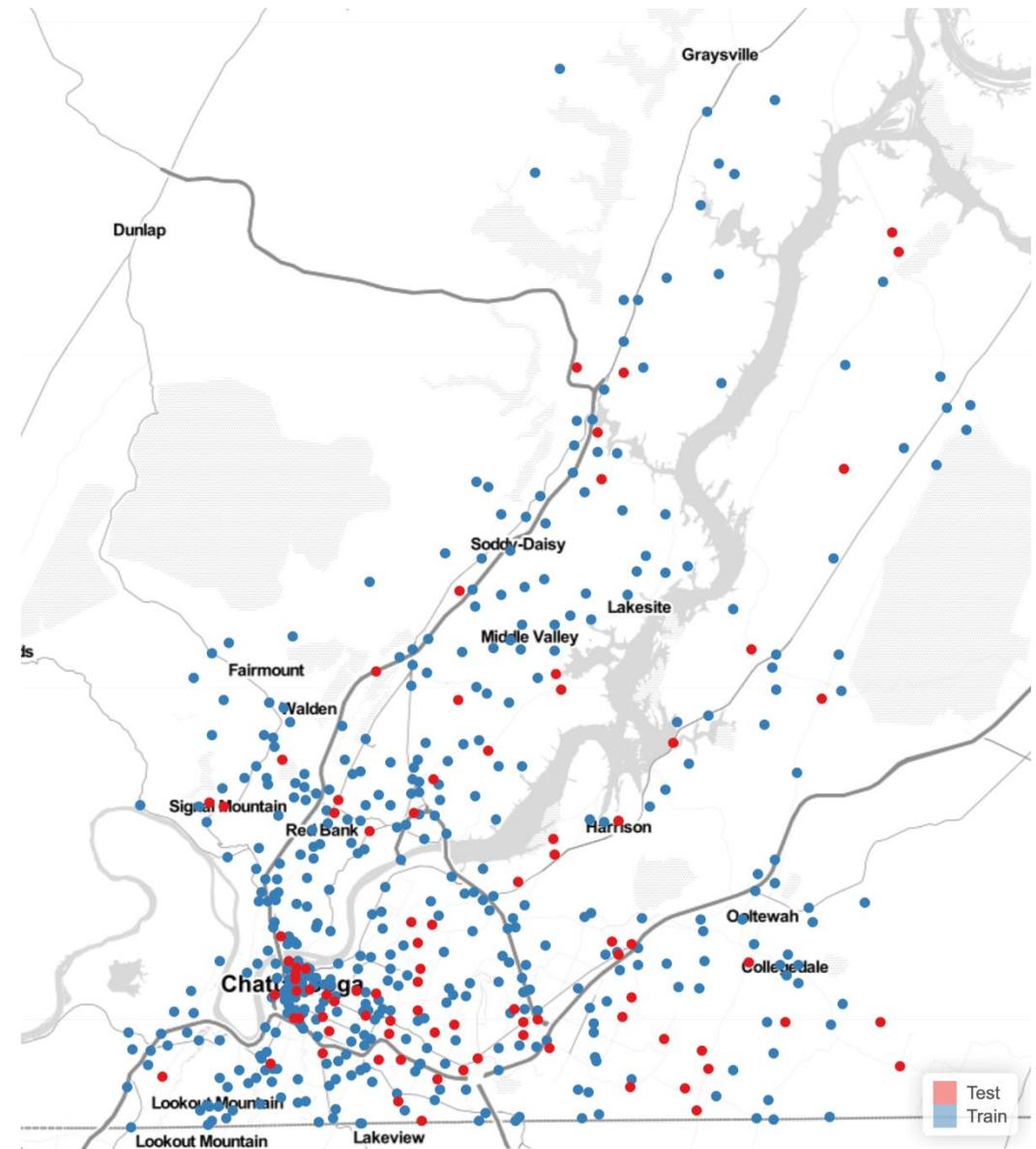
Chattanooga, Tennessee

Freeway and **Off-Freeway**

*Preliminary Results

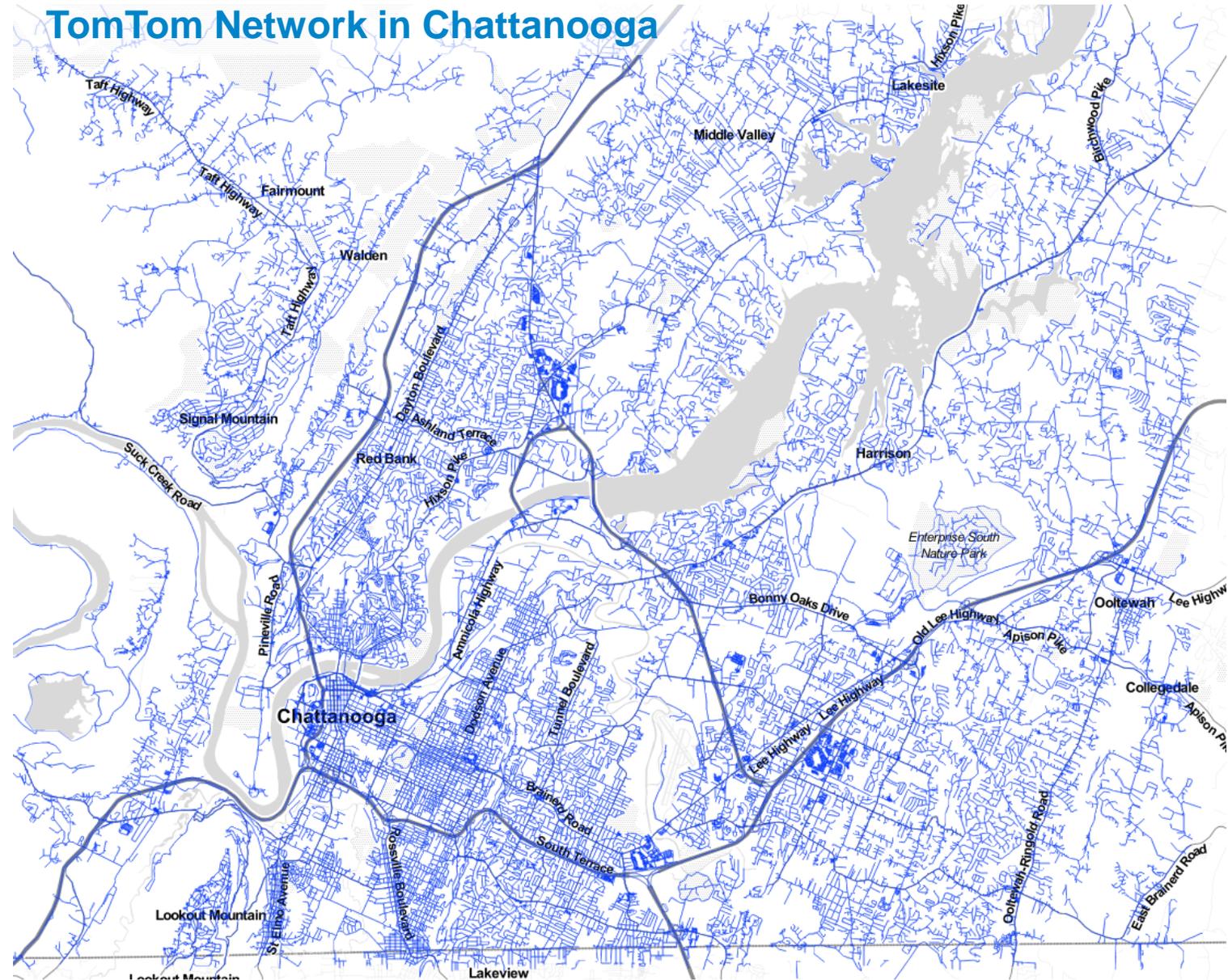
Locations of Data Collection in Chattanooga

- Jan 3th – May 24th, 2018
- A total of 516 locations in Chattanooga
- 430 locations for model training
 - Total of 13,324 data points
- 86 locations for model testing
 - Total of 2,742 data points
- **Volume data is non-directional**

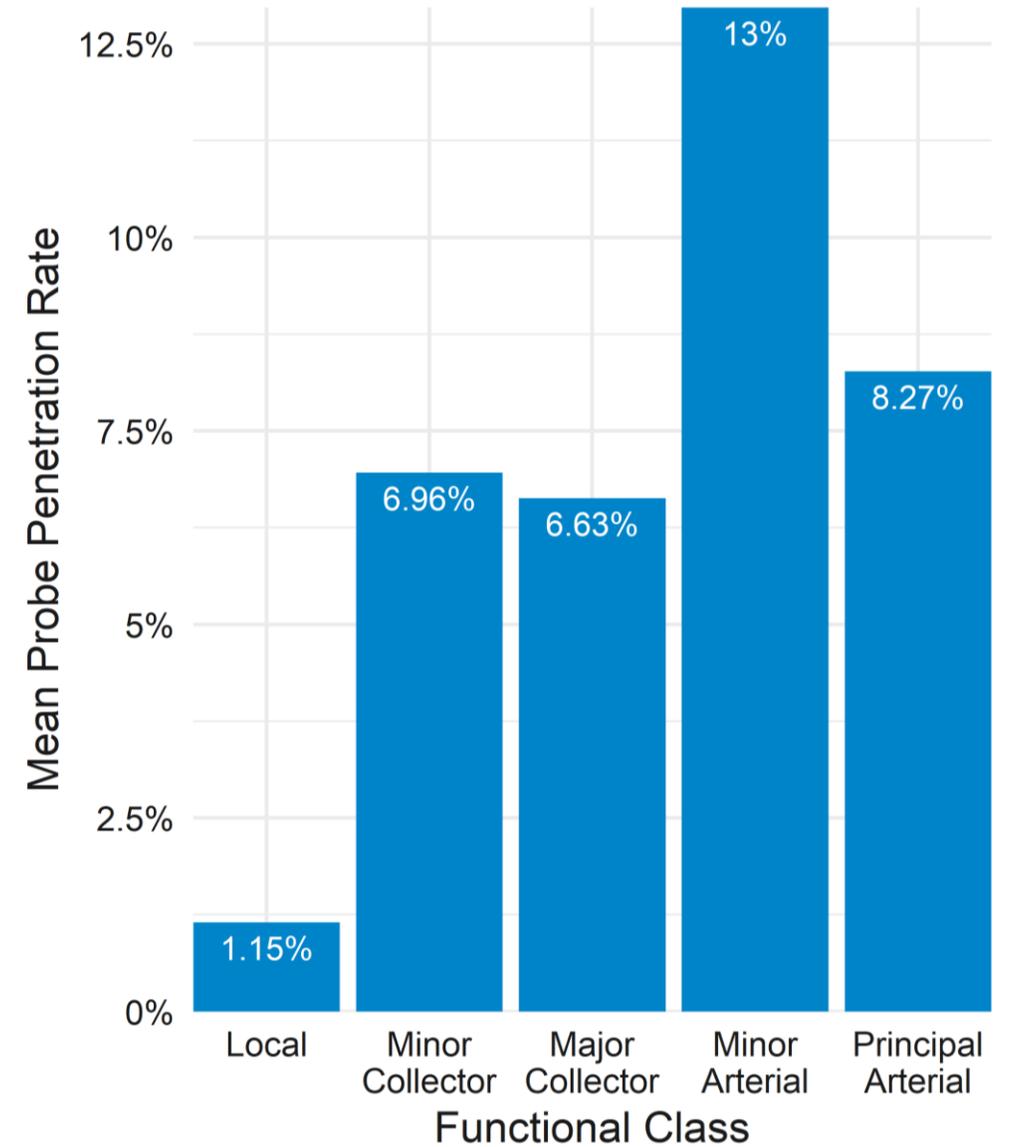
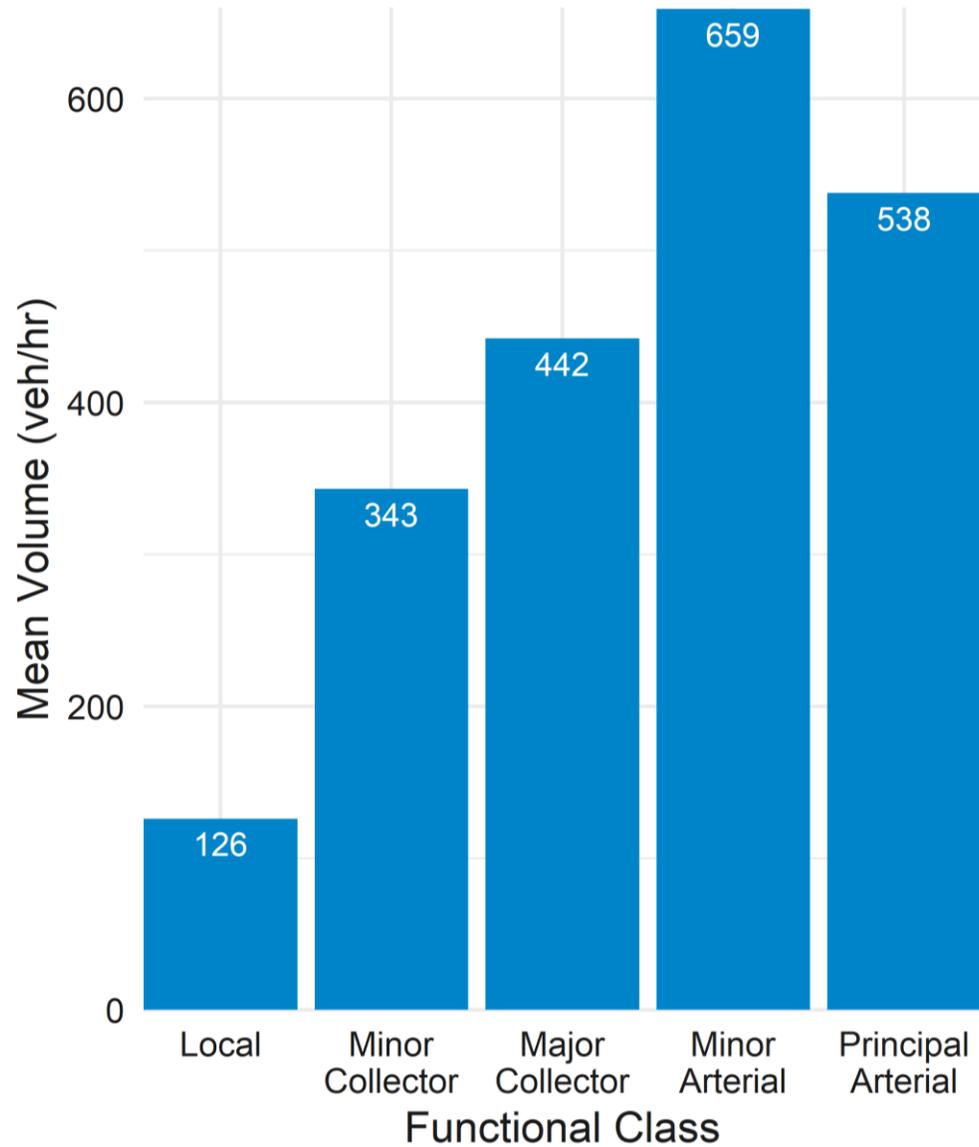


Input Variables

- TomTom GPS data
 - Hourly Average speed and probe counts
- Weather
 - Temperature, precipitation, wind, snow
- Road characteristics
 - Road functional class, speed limit, AADT, longitude, latitude
- Temporal information
 - Day of week, hour of day

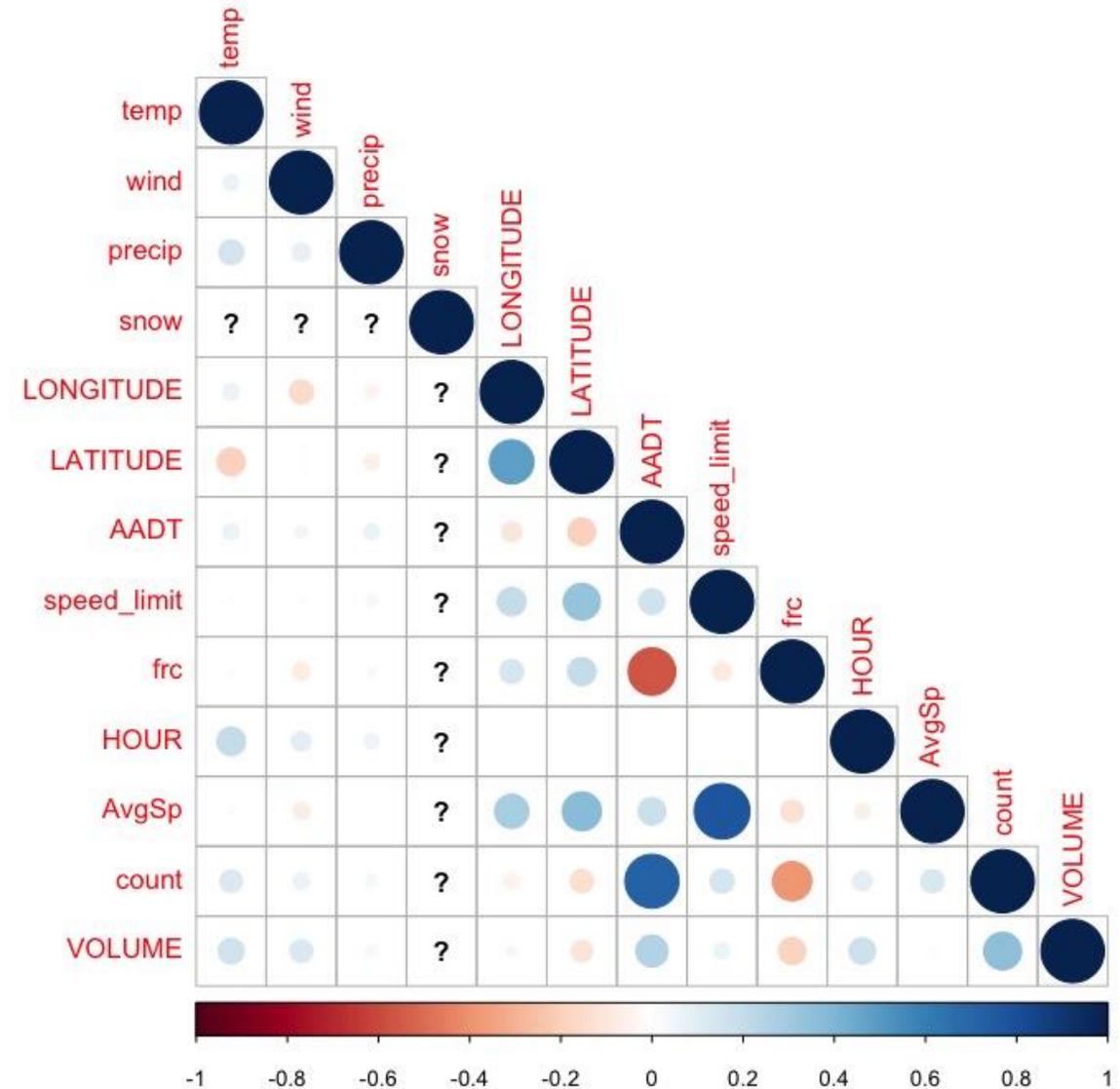


Mean Volume & Penetration Rates

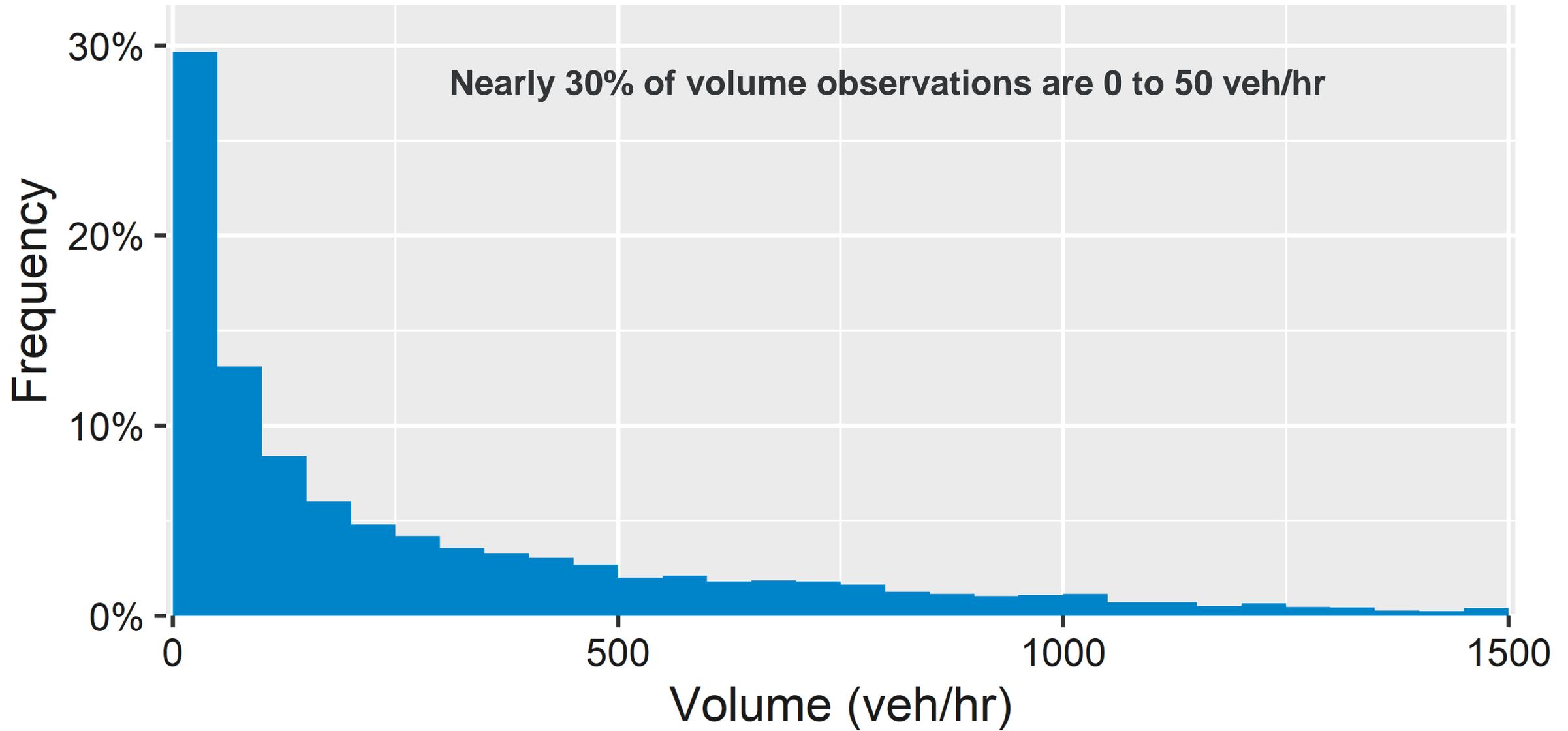


Correlation Between Variables

- Positive correlated variables
 - AADT
 - Probe counts
- Variables with weak or no correlation
 - Weather variables
 - Speed limits

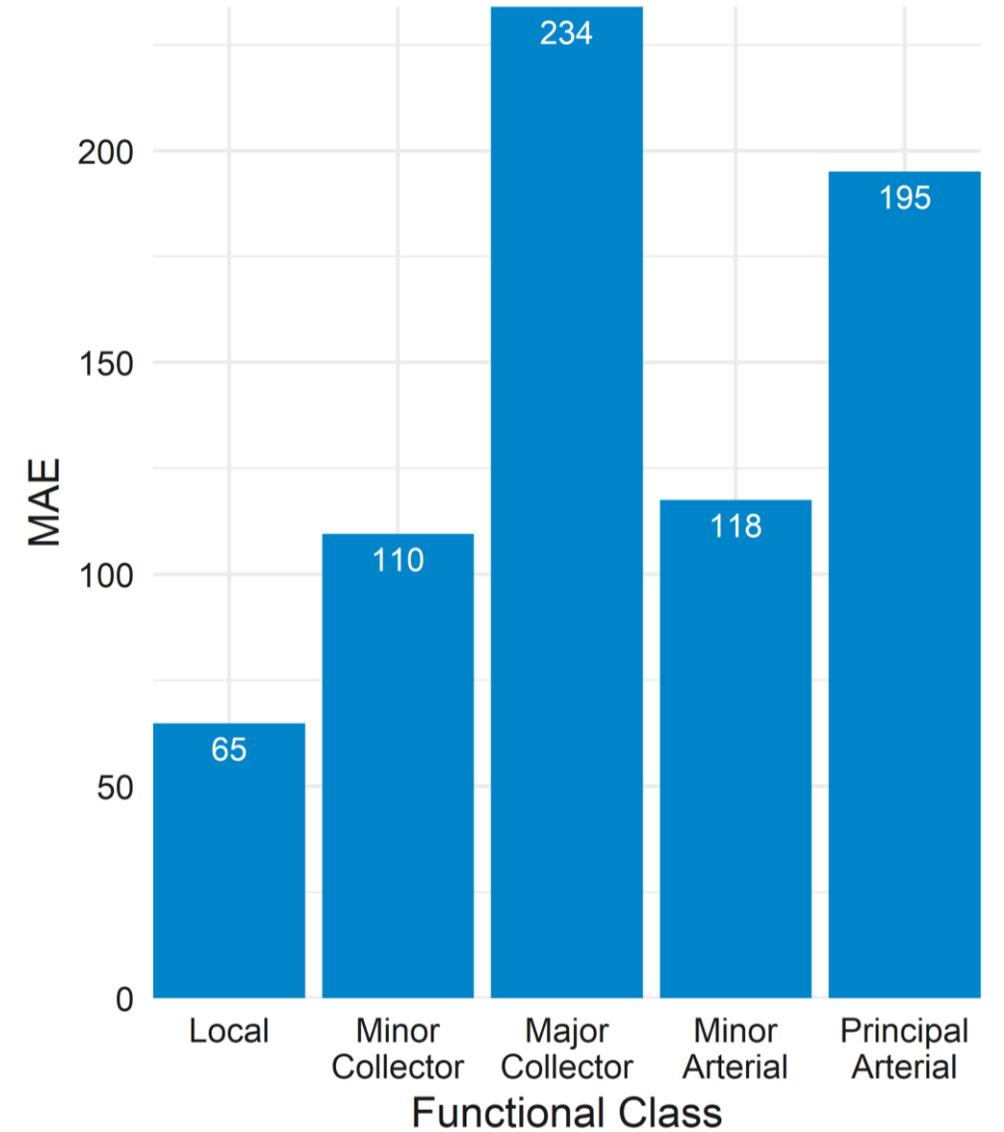


Traffic Volume Distribution



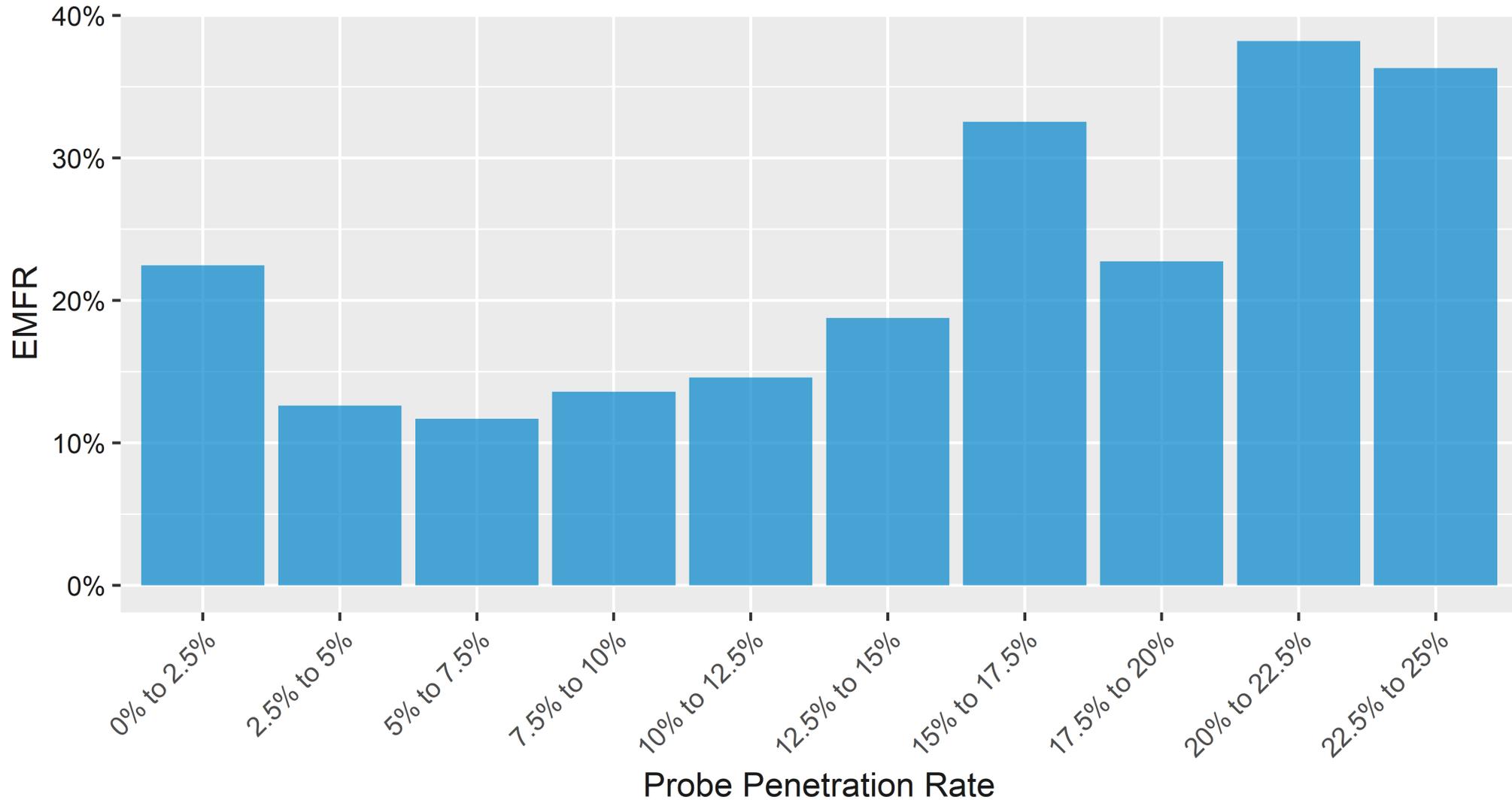
Model Results

Roadway Type	Model	MAE	R ²	EMFR
Off-Freeway	XGBoost	139	0.66	18.6



Error by Probe Count

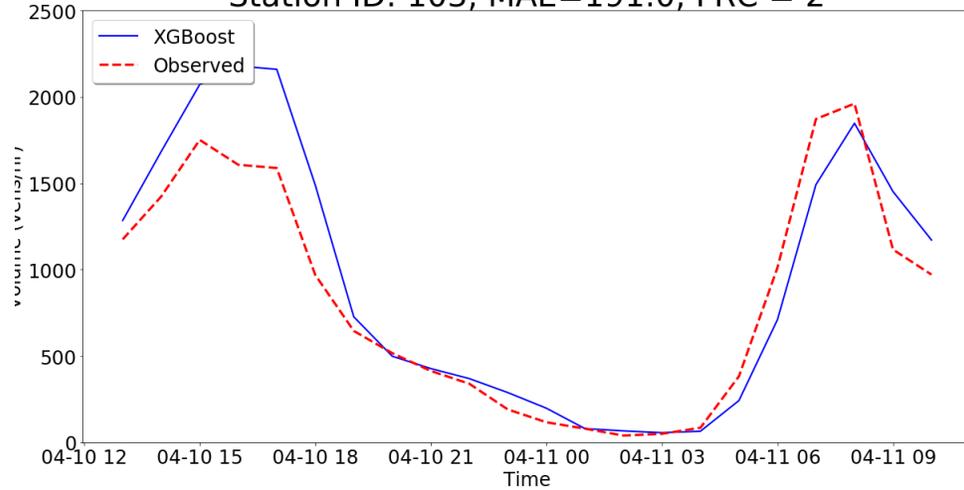
Not as clear a trend as we see in North Carolina



Model Performance

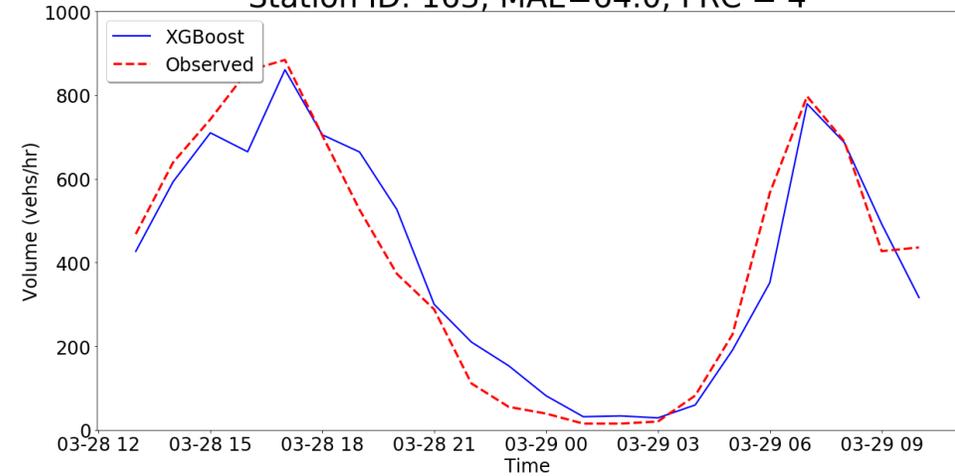
Principal Arterial

Station ID: 103, MAE=191.0, FRC = 2



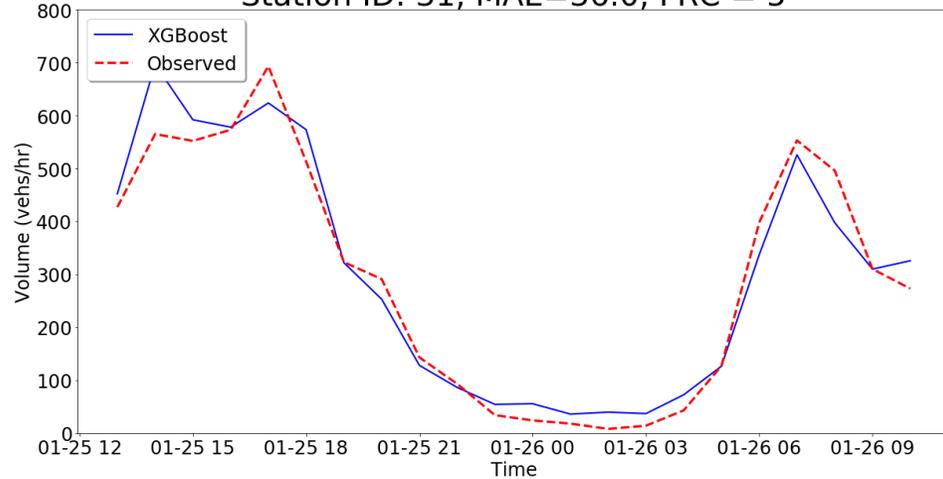
Major Collector

Station ID: 163, MAE=64.0, FRC = 4



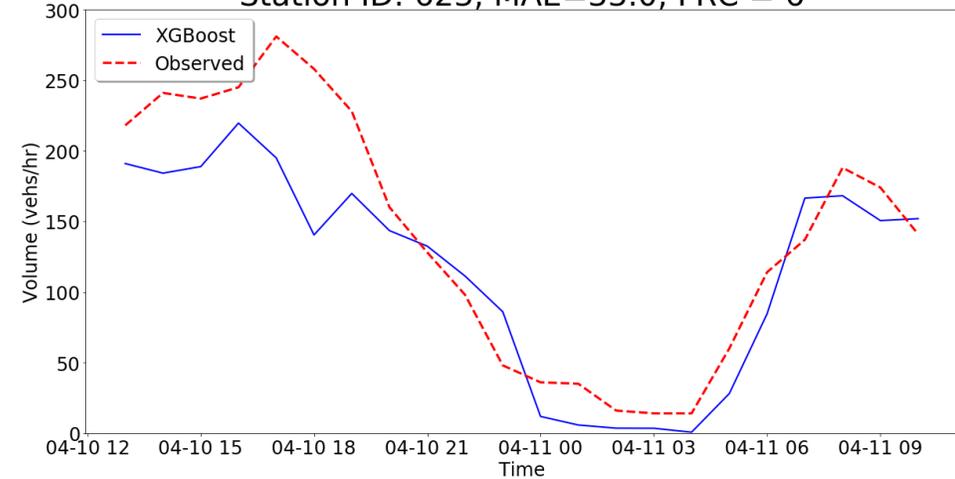
Minor Collector

Station ID: 31, MAE=36.0, FRC = 5



Local Road

Station ID: 623, MAE=33.0, FRC = 6



Chattanooga Volume Estimation – Observations and Next Steps

- **Assessment**

- Initial results from based on XGBoost estimation are not as accurate comparison to Colorado (CO), and North Carolina (NC) State results
- Difference is that CO, and NC counts are directional, but Chattanooga counts are not

- **Possible Reasons**

- Sensor volume data issue : Count data not as accurate, or a referencing issue
- TomTom probe data issue : Is something different about TN?
- Methodology problem : Aggregation of volume data in both directions introduced

- **Next Steps**

- Exploring Long Short-term Memory (LSTM) method
- Rigorous checks of input data

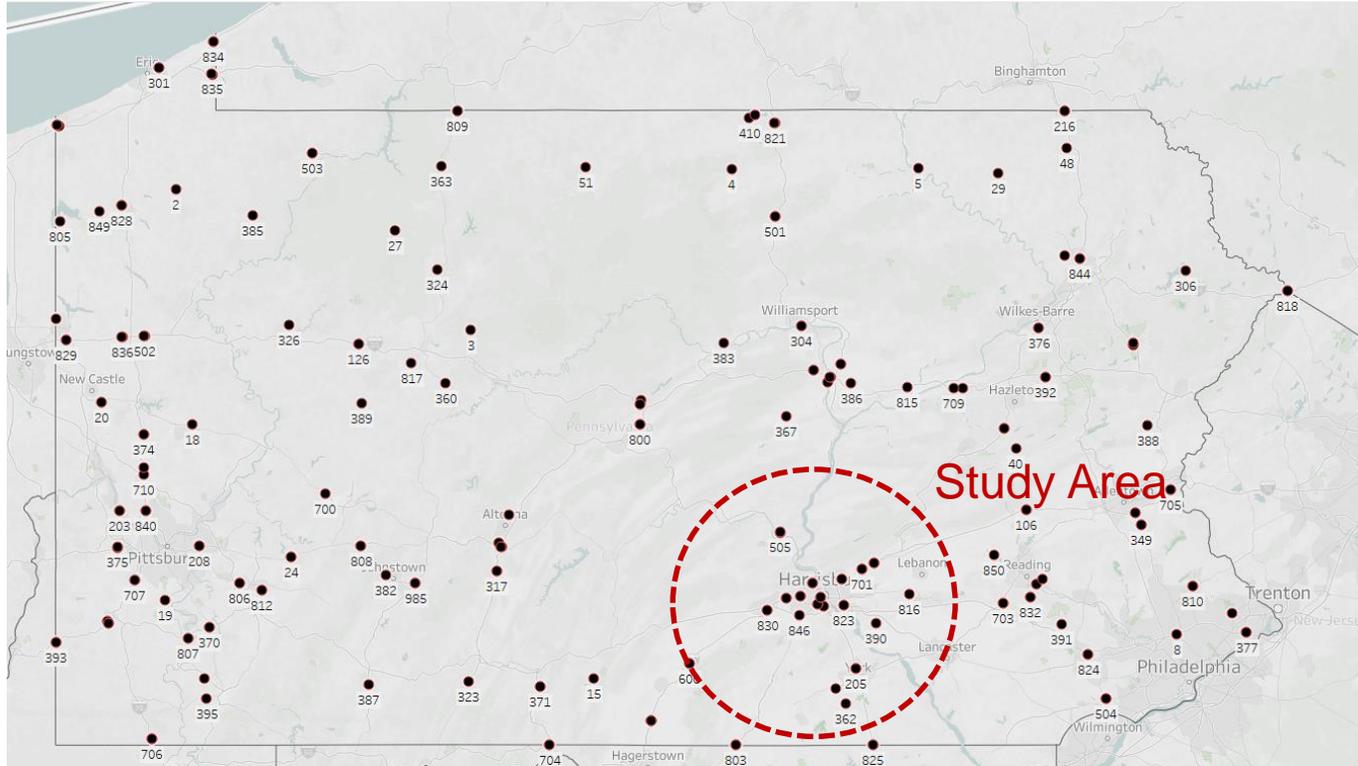
Phase – II

Pennsylvania (Harrisburg)

Volume Estimation

Data Description

Continuous Count Station (CCS) Locations in Pennsylvania



- Study area: Harrisburg MPO
 - Cumberland, Dauphin, and Perry Counties
- Continuous count station data available as a separate file per station
- Short term counts not tagged with lat/lon information
 - Station identification currently underway

Characteristic	Pennsylvania	Harrisburg MPO
Model Year	2018-19	2018-19
No. of CCS	214	25
Total number of observations (short term counts)	125,904 (~5,250 days)	--

PennDOT Open Data Portal

RMSTRAFFIC (Traffic Volumes)

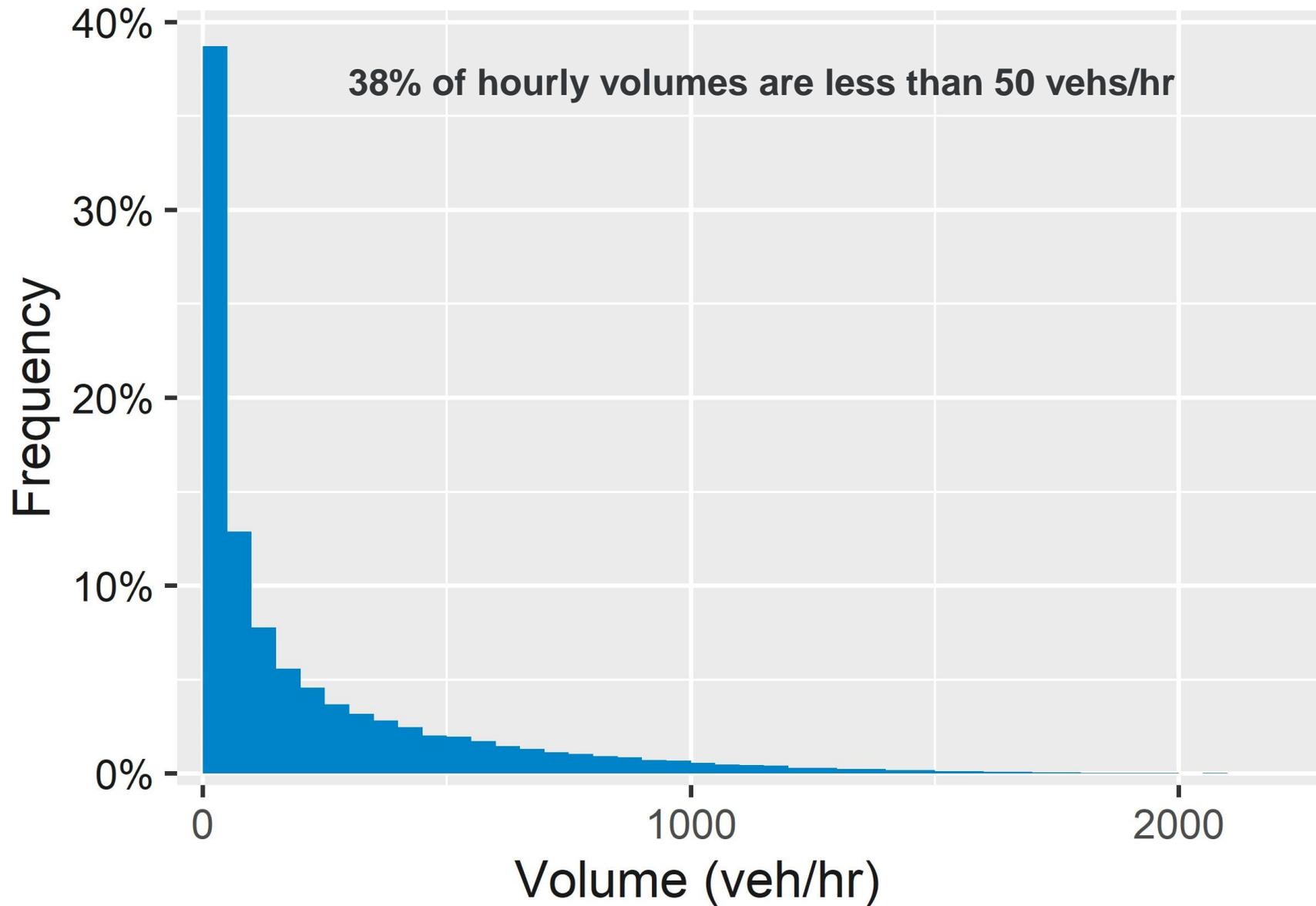
Last updated 4 days ago | 50,339 Records



11/3/2019 Feature Layer No license specified

Download APis

Traffic Volume distribution (Short term counts) – Full State



- **Incorporate Average Annual Daily Traffic (AADT) estimation, in-addition to hourly volumes**
 - Current research focused on operations support
 - Enables both operations & planning by having an archive of estimated volumes
- **AADT estimation efforts**
 - In Colorado, AADT accuracy significantly better than hourly volume estimates
 - North Carolina AADT estimates currently in development
 - In Harrisburg, PA – AADTs will be developed in conjunction with hourly volume estimates
 - Efforts will also be made to quantify acceptable AADT accuracy ranges
- **Methodological Enhancements**
 - Anomaly detection algorithms to identify erroneous input data
 - Spatial transferability of the prediction algorithm (Colorado ↔ North Carolina)

Thank you

www.nrel.gov



Traffic Volume Estimation using GPS Traces



Kaveh Sadabadi, PhD

Center for Advanced Transportation Technology - University of Maryland
(UMD CATT)





Traffic Volume Estimation using GPS Traces

Analysis Performed by:
Przemyslaw Sekula, Zachary Vander Laan

Presented by:
Kaveh Sadabadi

VTM Steering Committee Meeting
November 13, 2019

Today's Presentation

- Research objectives & background
- Data-Driven approach
- Phase 1: Brief summary of accomplishments
 - Maryland 2015
 - Florida
 - New Hampshire
- Phase 2: Ongoing efforts
 - Maryland 2018
 - Harrisburg, PA
- Summary & next steps

Case for “Ubiquitous” Count Data

- Traffic data (**speed** and **count**) needed in ...
 - Operations (incident response, work zone, and event management)
 - Planning (road construction, maintenance decisions)
 - Performance measurement and reporting (HPMS, MAP-21)
- Ubiquitous **speed** data (1-5 minute @ all TMC segments) is available through vehicle probes
 - I-95CC Vehicle Probe Project (VPP)
 - FHWA National Performance Management Research Dataset (NPMRDS)
- Ubiquitous **count** data is not widely available yet, but it is getting there!
 - I-95CC Volume & Turning Movement Project (VTM)
 - <https://i95coalition.org/projects/vpp-marketplace/>

ubiquitous adjective

ubiq·ui·tous | \ yü- 'bi-kwə-təs

Definition of ubiquitous

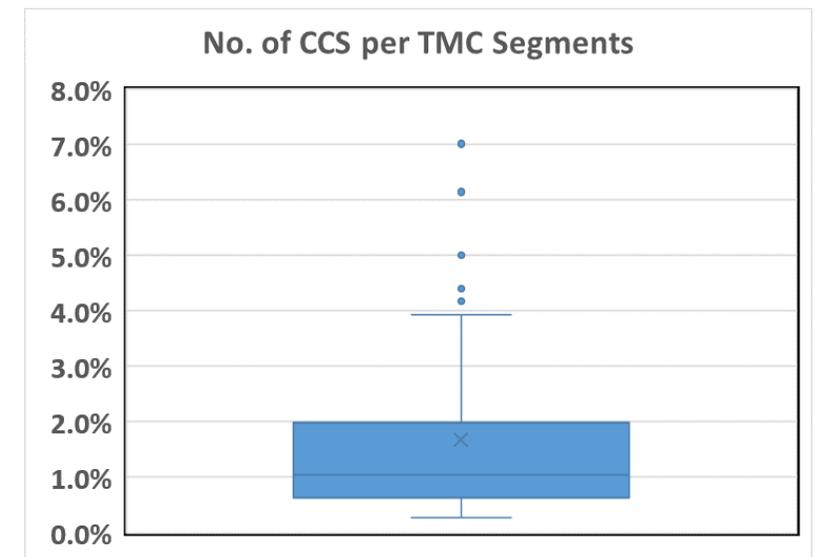
: existing or being everywhere at the same time : constantly encountered : WIDESPREAD



Status Quo of Count Data

- How **Ubiquitous** is the current count data measurements?
 - Continuous Count Stations (CCS)
 - CCS numbers vary
 - 12-803, average: 146
 - CCS coverage varies
 - Directional mile/CCS: range (14-457), mean: 163
 - TMC segments/CCS: range (14-360), mean: 119
 - CCS/TMC Segment: range (0.3%-7.0%), mean: 1.7%
 - Short Duration Counts
 - Periodic coverage counts (mainly for HPMS reporting)
 - Special needs count (projects, data collection, etc.)
- State agencies have limited access to traffic count measurements
 - Maryland has 85 Continuous Count Stations (CCS)
 - This covers only 0.6% of TMC road segments in the state

FHWA, Traffic Monitoring Guide, October 2016. (https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_17_003.pdf)

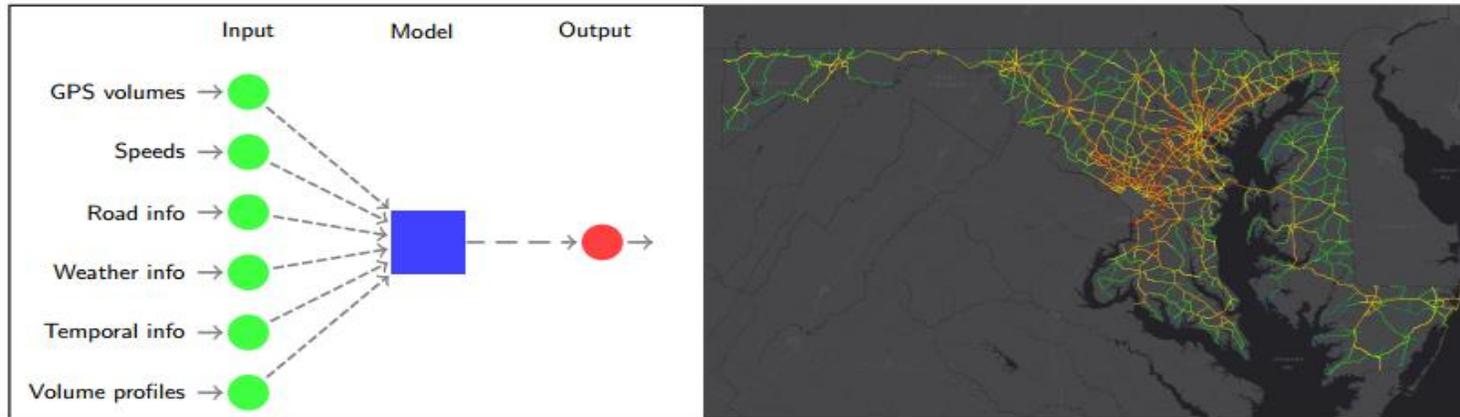
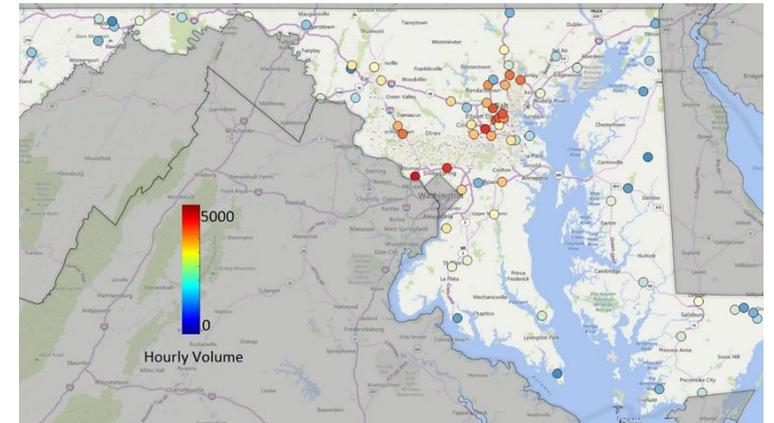


Research Objectives

- How to fill the gap b/w status quo and desired state?
 - Introduction of probe (trajectory) data helps fill the gap
 - Data-driven approaches (ML & AI) as useful tools for this purpose
- How to make ubiquitous count data available at scale?
 - Design pipelines to digest very large datasets
 - Develop routines to pre-process spatial data at scale
- How to measure “success”?
 - Define metrics to evaluate performance
 - Identify and explore limits in data and methods

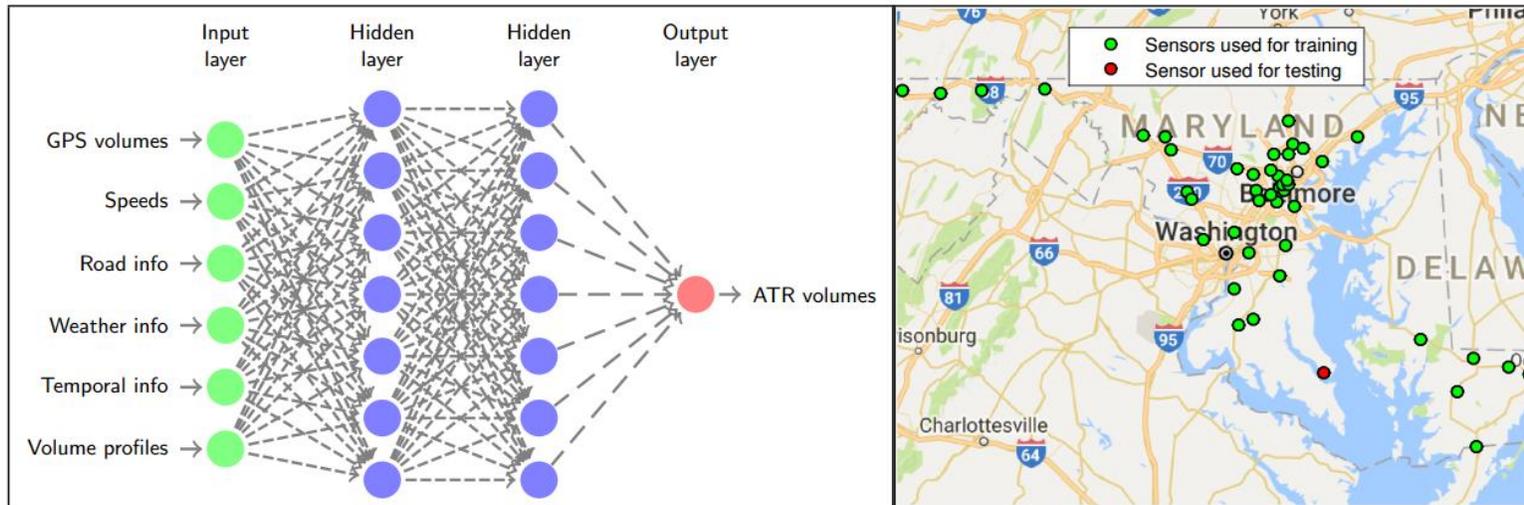
Phase 1 Recap: Research Problem Statement

- Given the following:
 - Probe volumes (processed from GPS traces of a subset of vehicles),
 - Other archived data (speeds, road geometry, weather, etc.), and
 - CCS counts
- Can we build a model to accurately estimate statewide volumes?



Phase 1 Recap: Data-Driven Approach

- Model: “Dense” Artificial Neural Network (ANN)
- Cross validation (repeat 47 times)
 - Train model using data from all but one continuous count stations
 - Generate model predictions using data from remaining station



- Evaluation: Compare estimates with actual volumes & generate metrics

Phase 1 Recap: Quantifying Model Accuracy

y_i = observed volume, \bar{y}_i = average observed volume, \hat{y}_i = model volume estimate, y_{\max} = max observed volume

- Error to Capacity (EMFR)

- Captures accuracy relative to capacity (max observed flow)
- < 10% becomes useful, < 5% target

$$EMFR = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_{\max}} \right| \right) \times 100$$

- Mean Absolute Percentage Error (MAPE)

- Reflects absolute volume accuracy
- *Good: 10-15% (high volume),
20-25% (mid volume)
30-50% (low volume)*

$$MAPE = \left(\frac{1}{n} \sum_{i=1}^n \left| \frac{\hat{y}_i - y_i}{y_i} \right| \right) \times 100$$

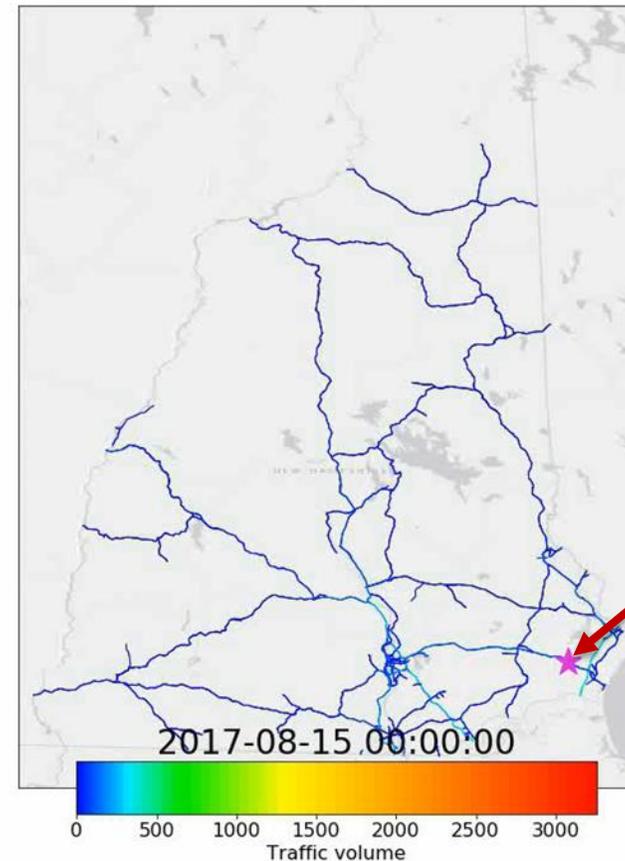
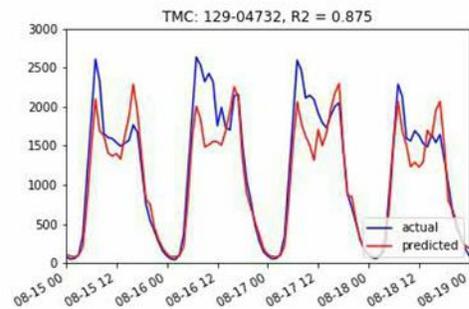
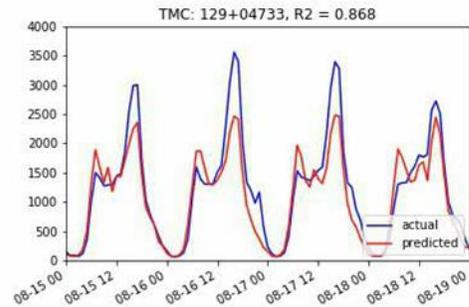
- Coefficient of Determination (R²)

- Shows explanatory power of model
- > 0.70 good, > 0.80 better, > 0.90 best

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

Phase 1 Recap: Statewide Hourly Volumes

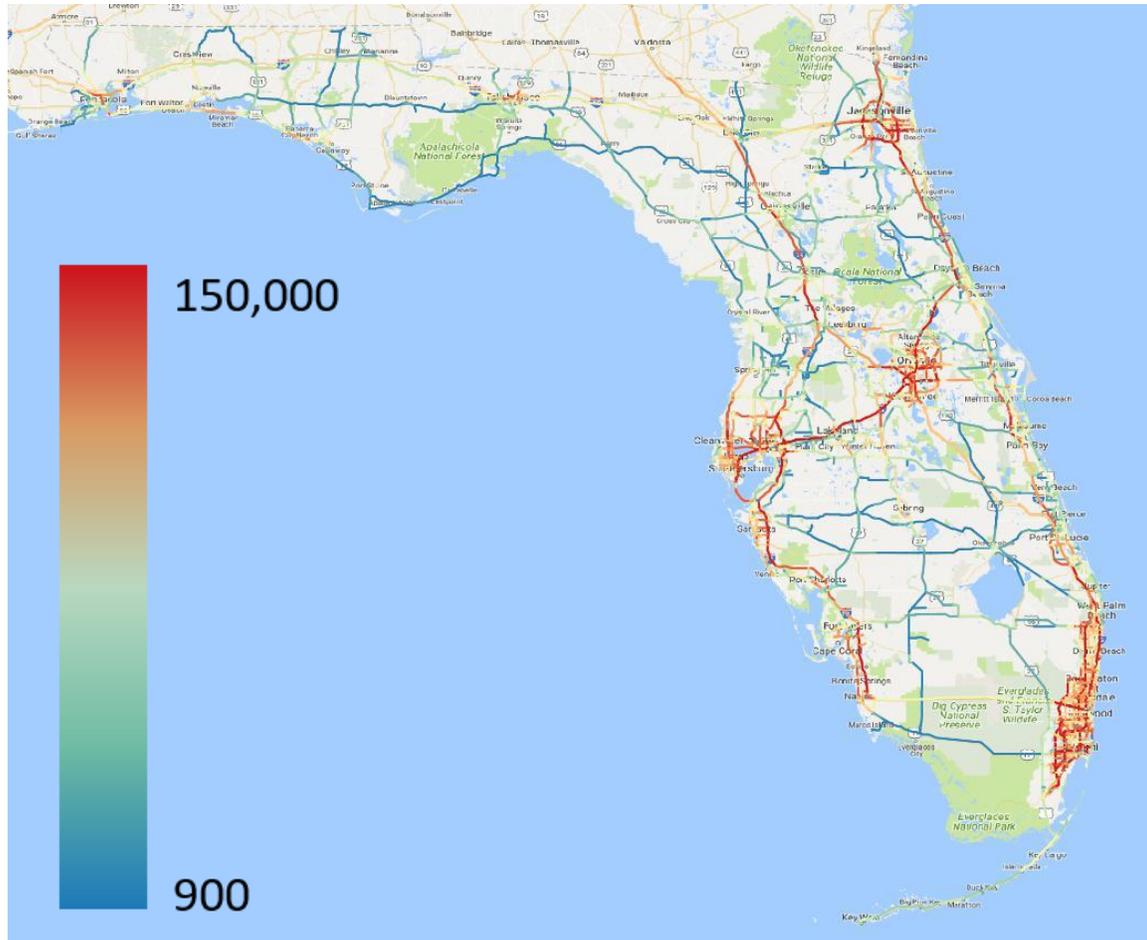
- Statewide models have been prepared for Florida, New Hampshire and Maryland
- Example: New Hampshire statewide model



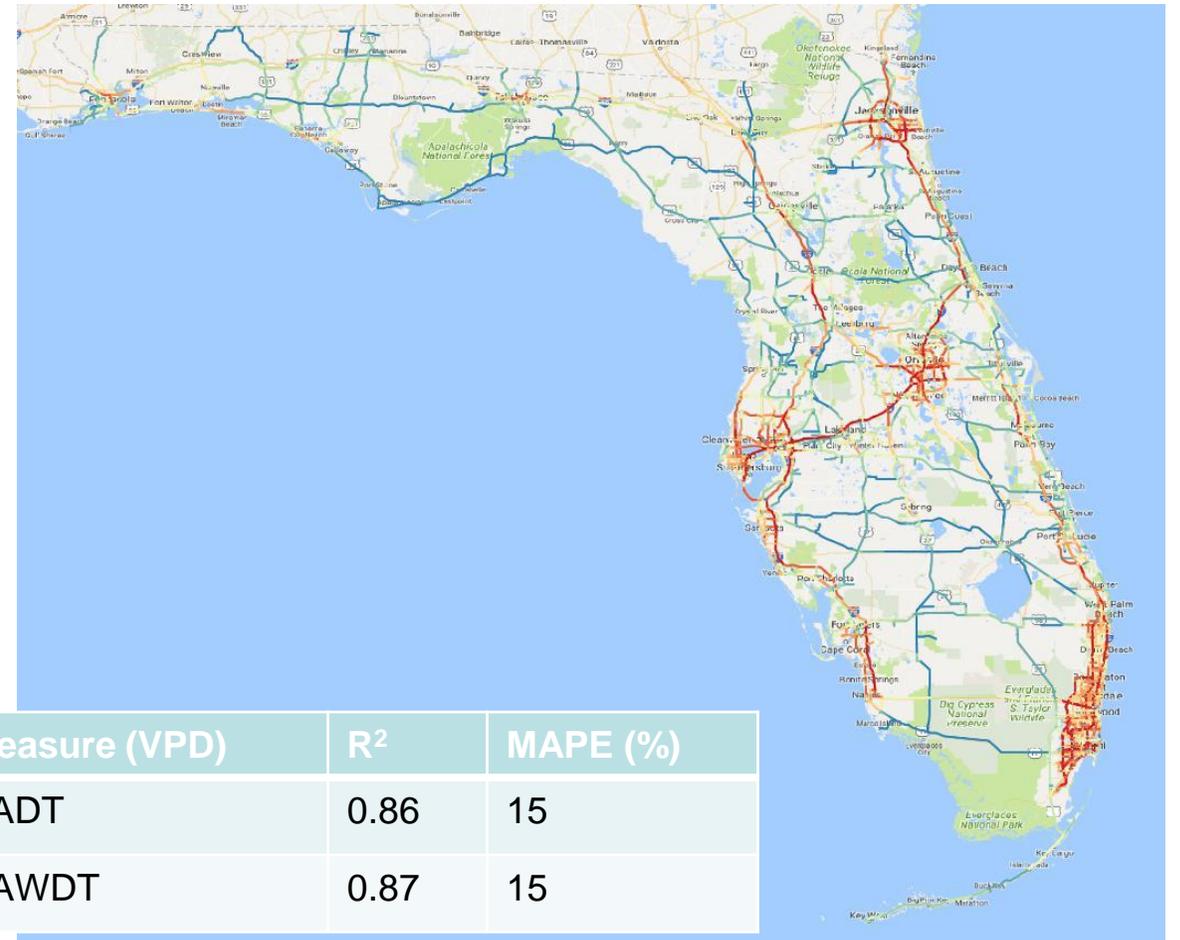
Continuous count station selected that exhibits typical (median) model performance

Phase 1 Recap: AADT & AAWDT Estimation

AADT



AAWDT



Measure (VPD)	R ²	MAPE (%)
AADT	0.86	15
AAWDT	0.87	15

Phase 1 Recap: Statewide Hourly Freight Volumes

- Florida case study
- Apply model to estimate hourly freight volumes
- Leverage highly-granular FDOT continuous count data
- Initial freight volume results look promising, particularly on higher functional road classes



shutterstock.com · 418741906

All Trucks

FHWA Class 5-13	R ²	MAPE (%)
Overall	0.77	38
FRC 1	0.83	24
FRC 2	0.76	42
FRC 3 & 4	0.65	49

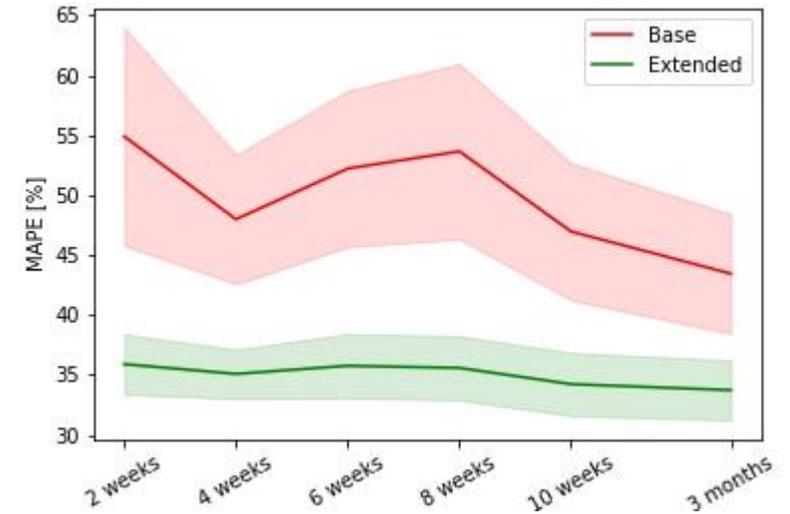
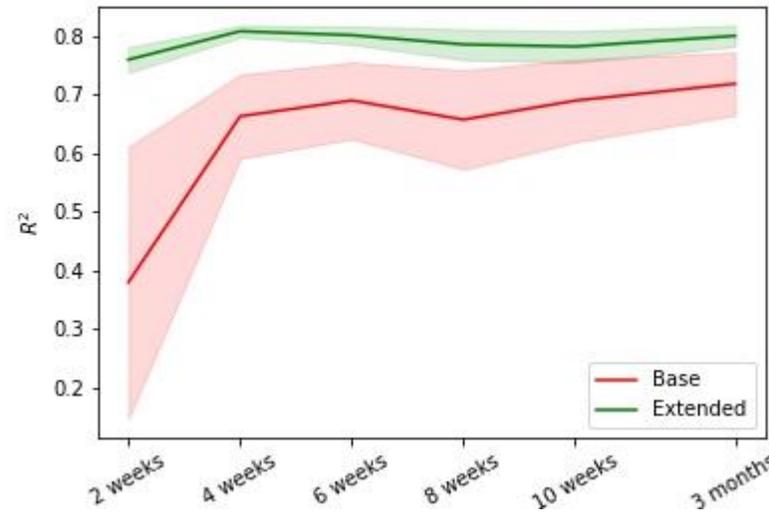
Heavy Trucks

FHWA Class 7-13	R ²	MAPE (%)
Overall	0.66	44
FRC 1	0.80	26
FRC 2	0.62	49
FRC 3 & 4	0.38	54

* Median error metrics

Phase 1 Recap: Power of Data! Leveraging Large Datasets

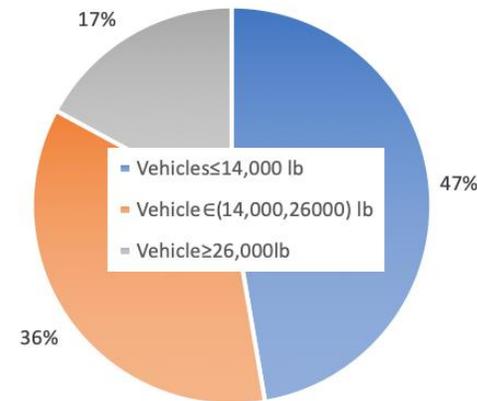
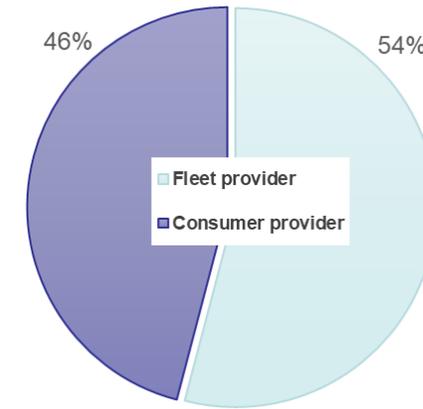
- When Florida dataset is combined with only 2 weeks of New Hampshire data for training, model predictions are reasonably good!
- Over time, only small amount of data for each new geography is needed to create powerful models!
- Potentially, the statewide traffic count data collection (and its associated cost) can be optimized!



Model Name		R ²		MAPE	
		2 w.	3 mo.	2 w.	3 mo.
Base (NH Only)	mean	0.38	0.72	54.9	43.4
	median	0.77	0.83	34.3	27.5
Extended (NH + FL)	mean	0.76	0.80	35.9	33.7
	median	0.81	0.84	27.8	27.3

Phase 2: Maryland Dataset (CY 2018-2019)

- 12 months of INRIX TRIPS data (entire 2018)
 - 130 million trips, 7.1 billion waypoints
 - Mean length of a trip is 22 miles
 - Mean trip duration is 35 minutes
 - 92% of the trips last shorter than 90 minutes
 - Penetration rates
 - Up to 8.5%
 - Mean: 3.2%
 - Median: 3.0%
- Very promising improvement in coverage and consistency over earlier 2015 dataset
- We are targeting **15 minute** flow rate estimation!



Phase 2: Initial Results

Maryland - 15 Min Flow Rates

→ Overall median error metrics:

- R2 = 0.78
- MAPE = 30%
- EMFR = 7.6%

Summary

Promising model performance, even over a variety of scenarios

Observations

- ↑ Road class = ↑ Accuracy
- ↑ Avg. hourly volume = ↑ Accuracy
- ↑ Avg. hourly GPS counts = ↑ Accuracy

Median Error Metrics by Scenario

Road Classification	R2	MAPE (%)	EMFR (%)	Obs
FRC 1 (Interstates)	0.87	19	6.6	937,862
FRC 2 (Other Freeways & Expressways)	0.71	34	8.4	1,061,675
FRC 3 & 4 (Other principal & minor arterials)	0.69	42	9.2	231,540

Hourly Volume (vph)	R2	MAPE (%)	EMFR (%)	Obs
0-1k	0.68	42	8.3	1,171,100
1k-2k	0.75	28	8.3	260,352
2k-3k	0.84	27	7.2	310,300
3k+	0.92	15	5.9	489,325

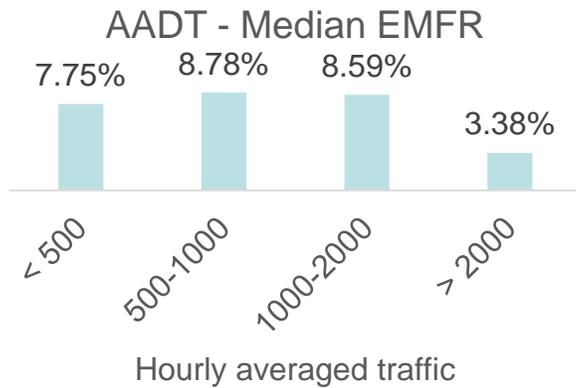
Avg probe counts / hr	R2	MAPE (%)	EMFR (%)	Obs
“Low” [0-19]	0.59	42	9.2	723,047
“Medium” [19-48]	0.78	32	7.3	733,576
“High” [48-177]	0.90	17	6.1	723,047

AADT & AAWDT

AADT

Median EMFR: 5.4%

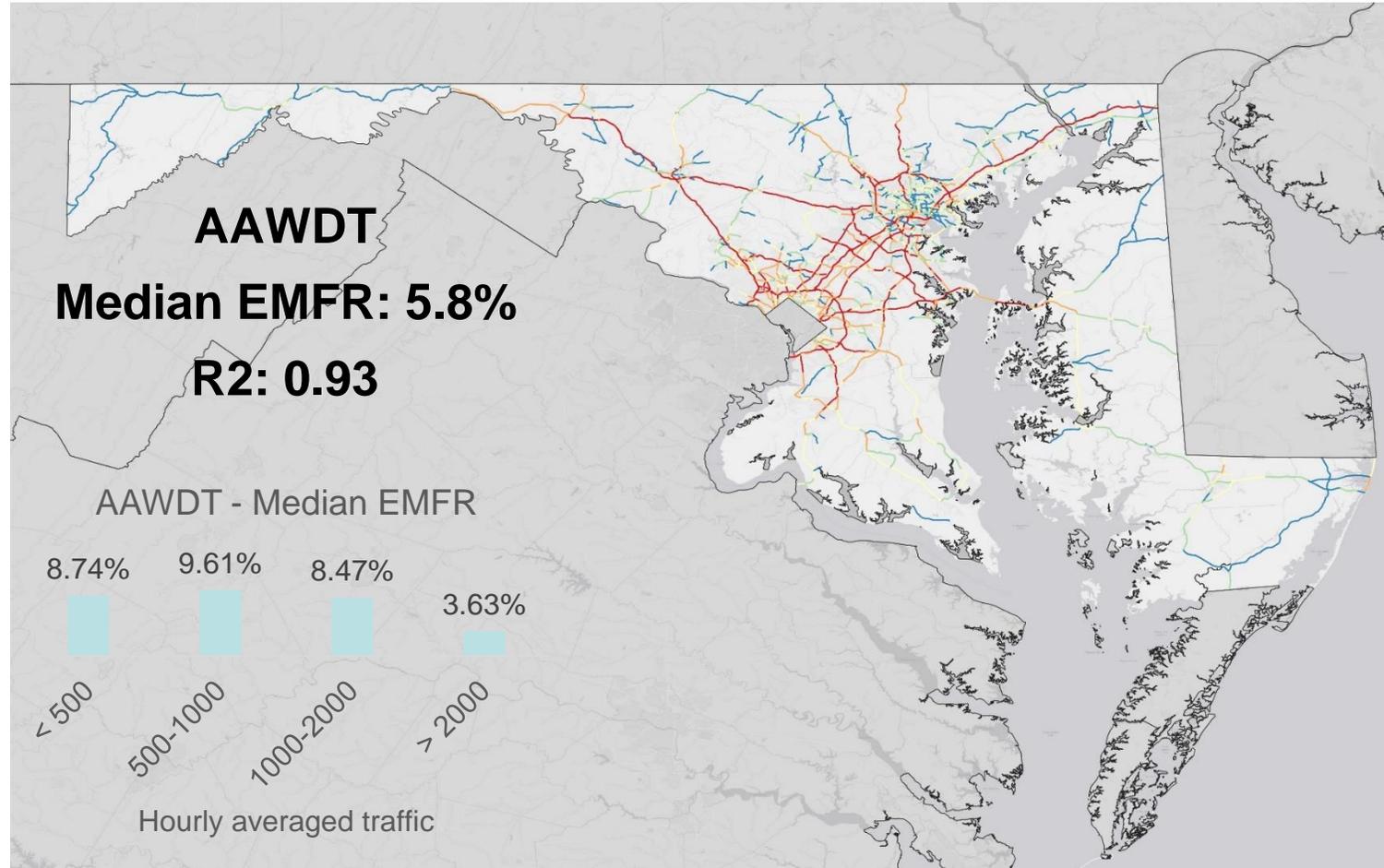
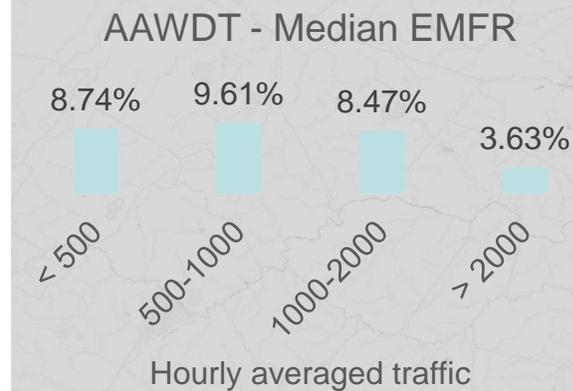
R2: 0.93



AAWDT

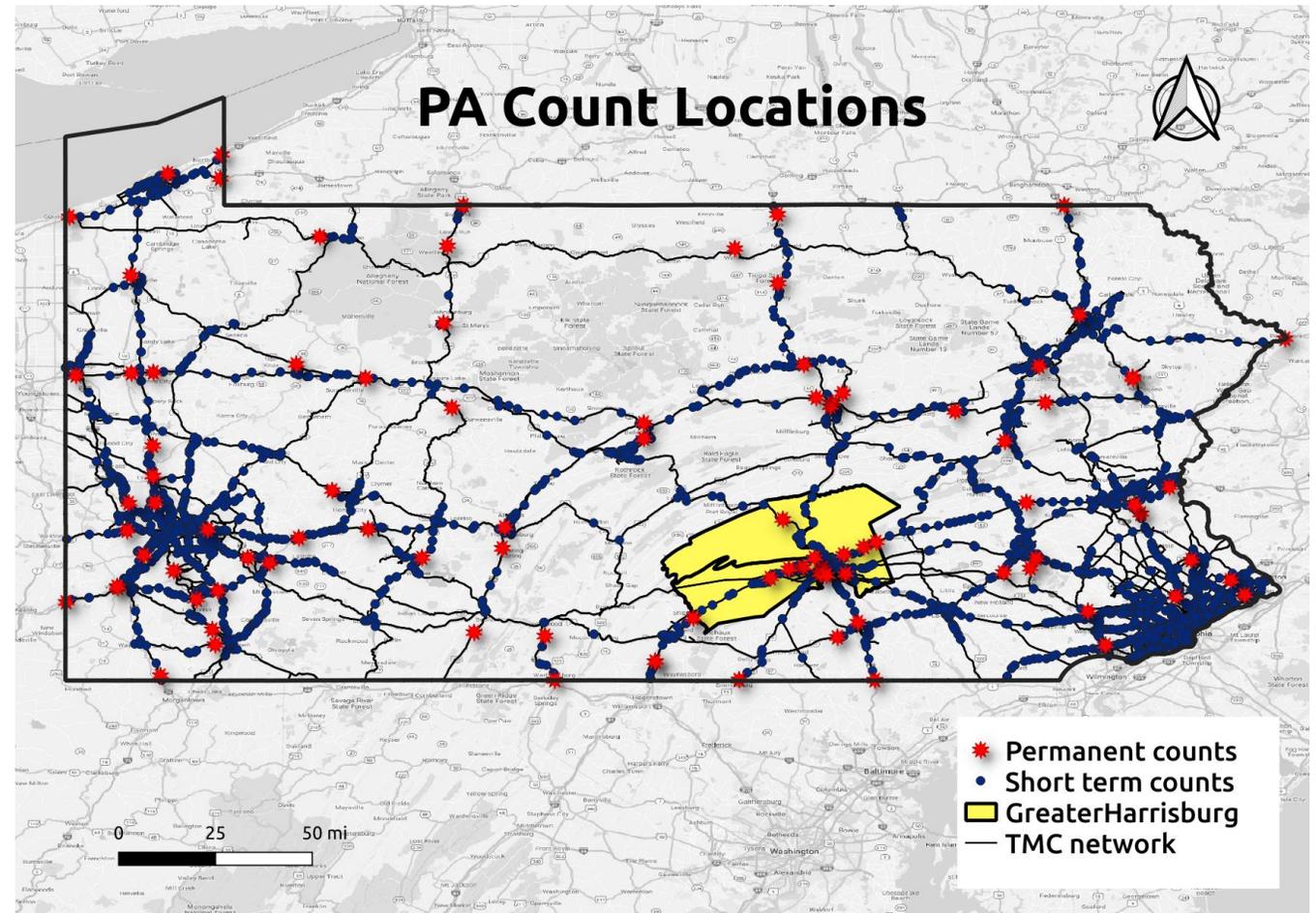
Median EMFR: 5.8%

R2: 0.93



Phase 2: Harrisburg, PA

- Jun-Aug 2018
- Tri-County Region
 - Cumberland
 - Dauphin
 - Perry
- 11 CCS (118)
- 73 Short-term/Directional (1612)
- 137 Short-term/Bi-directional (4757)
- Processing trajectory and short-term counts is underway
- Hourly and AADT volumes will be generated



Summary & Next Steps

- **Phase 1** confirmed probe data as key ingredient in filling the gap between status quo and desired ubiquitous volume data at scale
- **Phase 2** started with Maryland DOT purchase of 2 years of INRIX TRIPS dataset (2018-2019):
 - 12 months of data delivered (CY2018) – The biggest dataset in VTM so far!
- And, it continues with Harrisburg, PA study of 3 months of INRIX TRIPS dataset – (summer 2018):
 - 3 months of data delivered – Data preparation is underway
- Existing data pipelines have reduced the data ingestion burden
- Model calibration and evaluation is streamlined
- Initial results indicate 15 minute flow rates are achievable!
- Improvements in hourly, AADT, and freight count results are expected
- Optimization of CCS and short-term count locations are underway to improve model performance; and, to realize promised cost-savings in data collection programs

Questions

Contact Information

Kaveh Farokhi Sadabadi

kfarokhi@umd.edu

Przemyslaw Sekula

psekula@umd.edu

Zachary Vander Laan

zvanderl@umd.edu



Questions & Wrap Up



Denise Markow, PE

I-95 Corridor Coalition

TSMO Director



Final Questions



Thank You!

For Questions, please contact:

- UMD PI – Kaveh Sadabadi (UMD-CATT) 301-405-1352 or kfarokhi@umd.edu
- NREL PI – Venu Garikapati (NREL) 303-275-4784 or Venu.Garikapati@nrel.gov
- Coalition Coordinator – Denise Markow (I-95 Corridor Coalition) 301-789-9088 or dmarkow@i95coalition.org
- Phase II Program Manager – Stan Young (NREL) 301-792-8180
stanley.young@nrel.gov
- Logistics –Justin Ferri (KMJ Consulting, Inc.) 610.228.0759 or jferri@kmjinc.com

