

Volume & Turning Movements Project

Steering Committee Meeting #9

August 27, 2019



Housekeeping Items

- Please call xxx-xxx-xxxx for difficulties with the web or audio application
- This is a virtual meeting experience
 - Please keep your phone muted until asking a question or speaking (press *6 to mute/unmute individual phone lines)
 - Please do not place call “on hold” as your hold music will be heard by the group
- Speakers will answer questions at the end of their presentation
- The audio from this meeting is being recorded
- All materials & contact information will be available to participants after the webcast



Attendees

Agencies and Organizations

| | |
|-------------------------|------------------------|
| Capitol Area MPO | North Carolina DOT |
| Colorado DOT | NREL |
| FHWA | Pennsylvania DOT |
| Georgia DOT | South Carolina DOT |
| I-95 Corridor Coalition | Streetlight Data |
| INRIX | TomTom |
| Maryland DOT | UMD CATT |
| Maryland SHA | University of Kentucky |
| MWCOG | Virginia AOT |
| NJTPA | Virginia DOT |



Speakers



Denise Markow, PE
TSMO Director
I-95 Corridor Coalition
dmarkow@i95coalition.org



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



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



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



Please confirm that
your line is **muted**

***6**

Thank you!



Agenda

| | | Topic | Speaker |
|---|--------|---|--|
| 1 | 5 min | Introductions & Welcome | Denise Markow, I-95 Corridor Coalition |
| 2 | 15 min | Project Recap from Phase I (VTM Phase I) Objectives of Phase II | Stan Young, NREL |
| 3 | 10 min | Scope of work of UMD with Maryland & preliminary results | Kaveh Sadabadi, UMD CATT |
| 4 | 10 min | Scope for Work for DOE Funded Work <ul style="list-style-type: none">• Colorado – TomTom• Chattanooga, TN – TomTom• North Carolina – TomTom Phase II Joint Work - Pennsylvania – TomTom/INRIX | Venu Garikapati, NREL |
| 5 | 10 min | Opportunity for Additional State Demos FHWA pooled fund | Stan Young, NREL |
| 6 | 5 min | Discussion & set expectation for next call (November) <ul style="list-style-type: none">- North Carolina Results (NREL)- Optimal Traffic Monitoring in a New Data Age | Denise Markow & Stan Young |





Department of Energy

Volume Turning Movements Phase II

“Taking it from the Lab to the Streets”

Stan Young

Aug, 2019

Background / Motivation / History

1

- Why are we doing this -

2

- Phases of R&D -

3

- The Team Members and their Roles

4

- Getting Involved

Why do we need better
volume?

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



Planning



Operations



Performance Measures



Forecasts (special events)

Ubiquitous Traffic Volumes



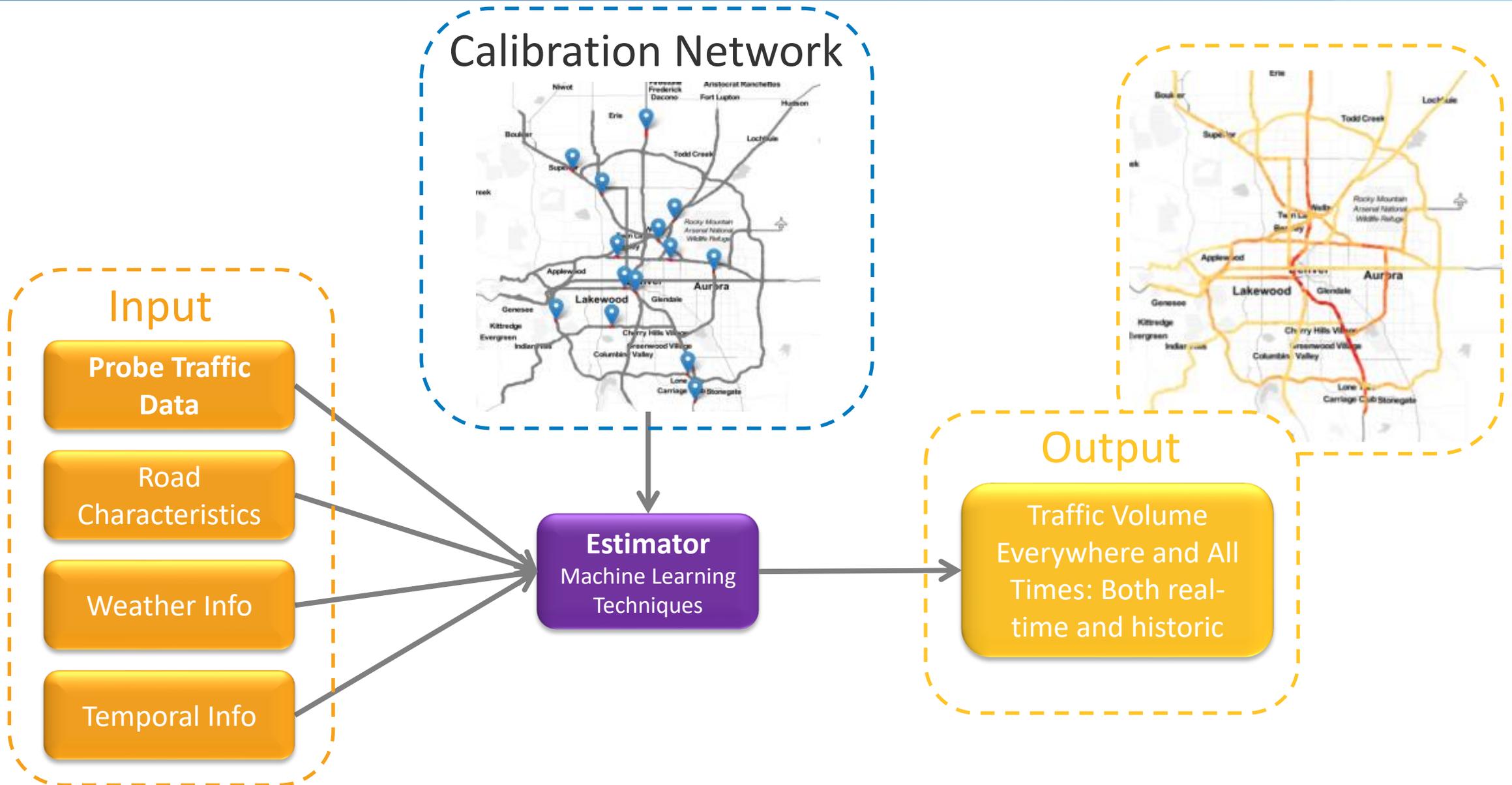
Ubiquitous network observability

- 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 each and every link of the road network

Proposed Solution



How Good is Good Enough?

MAPE is Volume Dependent!

| AADT Range | Acceptable % Change | |
|----------------|---------------------|------------------|
| | Decreasing (-) | Increasing (+) |
| 0 - 19 | -100% | 400% |
| 20 - 49 | -40% | 50% |
| 50 - 99 | -30% | 40% |
| 100 - 299 | -25% | 30% |
| 300 - 999 | -20% | 25% |
| 1,000 - 4,999 | -15% | 20% |
| 5,000 - 49,999 | -10% | 15% |
| 50,000+ | -10% | 10% |

MNDOT Example

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

- R² 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}

The 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 – FHWA Pooled Fund Study – Path forward for AADT

The Phases of Research

Phase I – I-95 Volume-Turning Movement Research Initiative

Phase II – Technology Commercialization

Phase III – FHWA Pooled Fund Study

Phase 1 Research - VTM

- Objective – Proof-of-concept of accurately estimating traffic volume using commercial probe data AND other data sources
- Funding - ~\$500K, from USDOT through I-95 Corridor Coalition
- Timeframe - 3 years (2016-2018)
- Partners
 - I-95 CC – convene and manage
 - UMD & NREL, primary R&D (shout out to TTI)
 - **I-95 CC Steering Committee, feedback and direction**
 - Colorado DOT, Maryland SHA, Florida DOT, NH DOT (Demo Sites)
 - Industry – TomTom & INRIX, primary data providers
 - Collaboration with HERE, Streetlight Data and others
- Success! – Concept produced volume estimates

Phase I - Summary / Conclusions

- Volume estimation can be supported with a combination of:
 - Commercial Probe Data (Probe count & Speed/Travel Times)
 - Other road attribute data and weather
 - High confidence ground truth sensor for calibration & validation
- Works on **Freeways** as well as **Lower Class Roads**
- Machine learning provides rapid & sustainable calculation methods
- Probe data has significant impact on volume estimation accuracy
- Can be applied for both historical and real-time

Phase I – Caveats / Follow-ons

- Answered *'How good is good enough?'*, for several applications
- Phase I emphasized operations (nature of I-95)
 - Planning AADT emerged as early win
 - Truck volumes emerged as valuable addition
- Low-volume roads / off-freeway critical
- Path toward sustainable commercialization
 - Large data sets / Black box / Verifiable accuracy
- Turning Movements yet to be addressed

The Phases of Research

Phase I – I-95 Volume-Turning Movement Research Initiative

Phase II – Technology Commercialization

Phase III – FHWA Pooled Fund Study

Phase II VTM Research – From the Lab to the Streets

- Objective – Out of the lab and into the streets! – Kick the tires, see if it works
- Funding -
 - NREL funded through Department of Energy Tech Commercialization Grant
 - UMD funded through Maryland SHA & Other Opportunities
- Timeframe - 2019-2021
- Partners - Similar
 - I-95 CC – convene stakeholders
 - NREL and UMD primary R&D, supported by TTI
 - **Multi-State Steering Committee, feedback and direction**
 - Colorado DOT, Maryland SHA, PennDOT financial support
 - Collaboration with many states
 - Industry – TomTom, INRIX primary data providers

Phase II Research –

- DOE Tech Comm Grant – Pathway for commercialization of volume estimates from TomTom Data.
 - Funding - ~\$1M, ~\$500K DOE Tech Comm Grant
 - In kind contributions from Industry (TomTom - ~\$450K),
 - State DOTs \$50K(Colorado DOT, PennDOT, and growing)
- Maryland SHA – Enable Statewide Volume Estimates for integration into Maryland SHA business processes
 - Multi-year INRIX data purchases for entire state

Overview of UMD Phase II Effort & Initial Results

Kaveh Sadabadi



Traffic Volume Estimation using GPS Traces: Maryland 2018 dataset

Analysis Performed by:
Przemyslaw Sekula, Zachary Vander Laan

Presented by:
Kaveh Sadabadi

VTM Steering Committee Meeting
August 27, 2019



Today's Presentation

- Research objectives & background
- Data-Driven approach
- Phase 1: Brief summary of accomplishments
 - Maryland
 - Florida
 - New Hampshire
- Phase 2: Maryland 2018 dataset
 - Data description
 - Initial results
- Summary & next steps

August 27, 2019



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

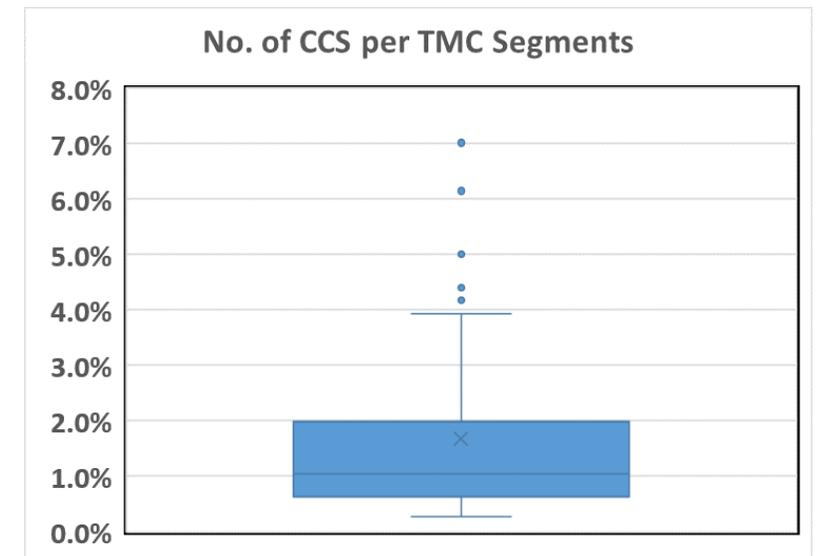
Definition of ubiquitous

: 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



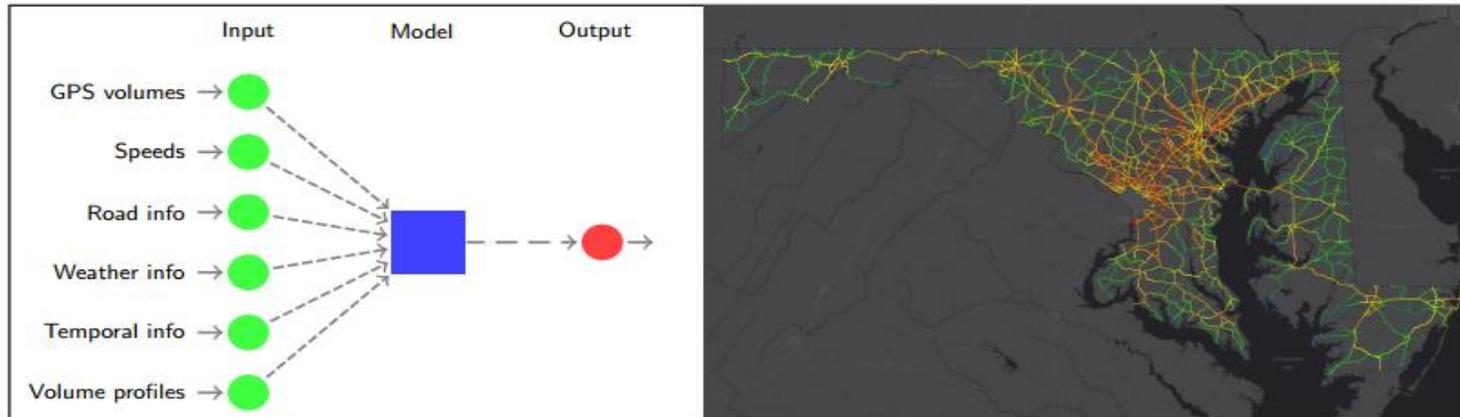
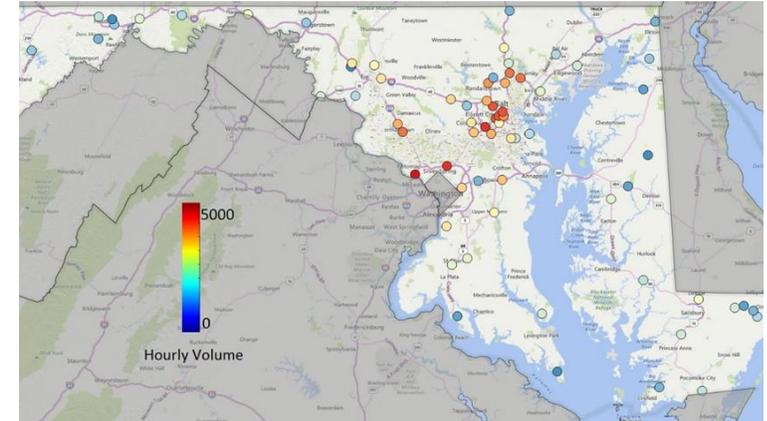
FHWA, Traffic Monitoring Guide, October 2016. (https://www.fhwa.dot.gov/policyinformation/tmguide/tmg_fhwa_pl_17_003.pdf)

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

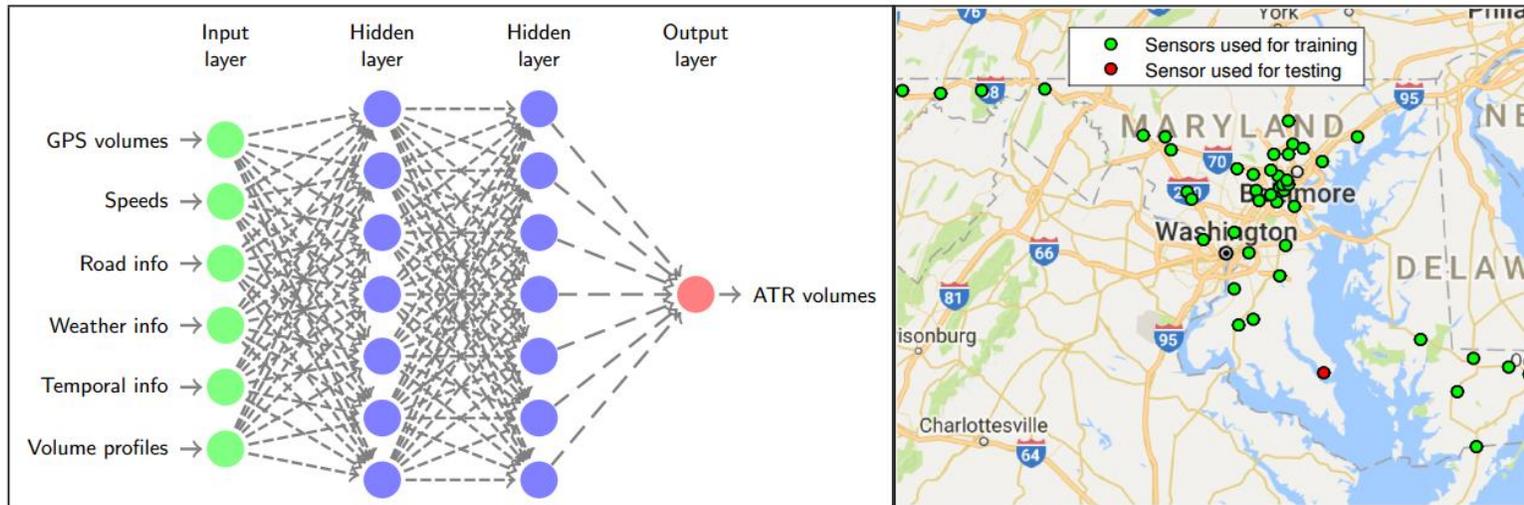
Phase 1 Recap: Research Problem Statement

- 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)*

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

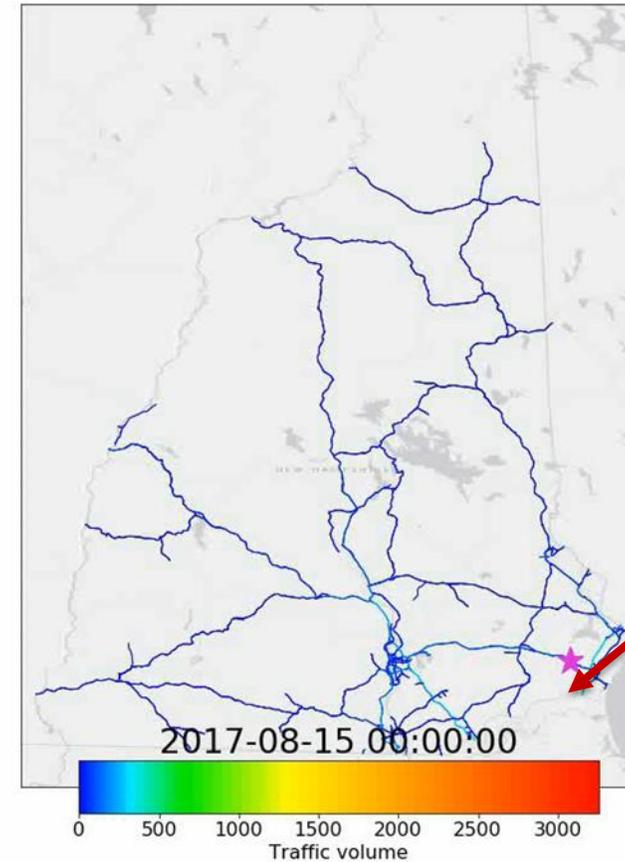
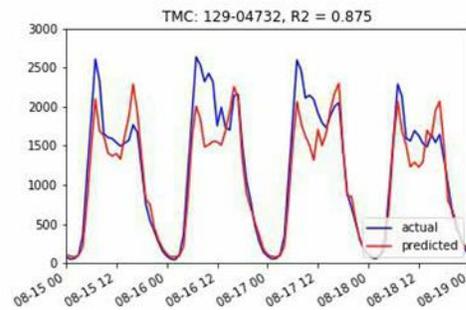
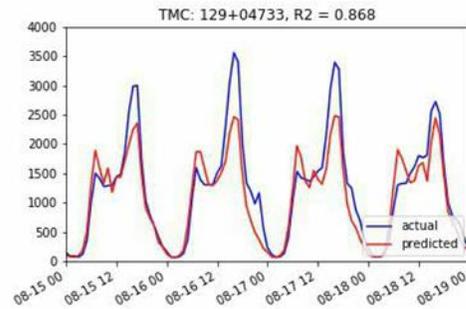
- Coefficient of Determination (R²)

- 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

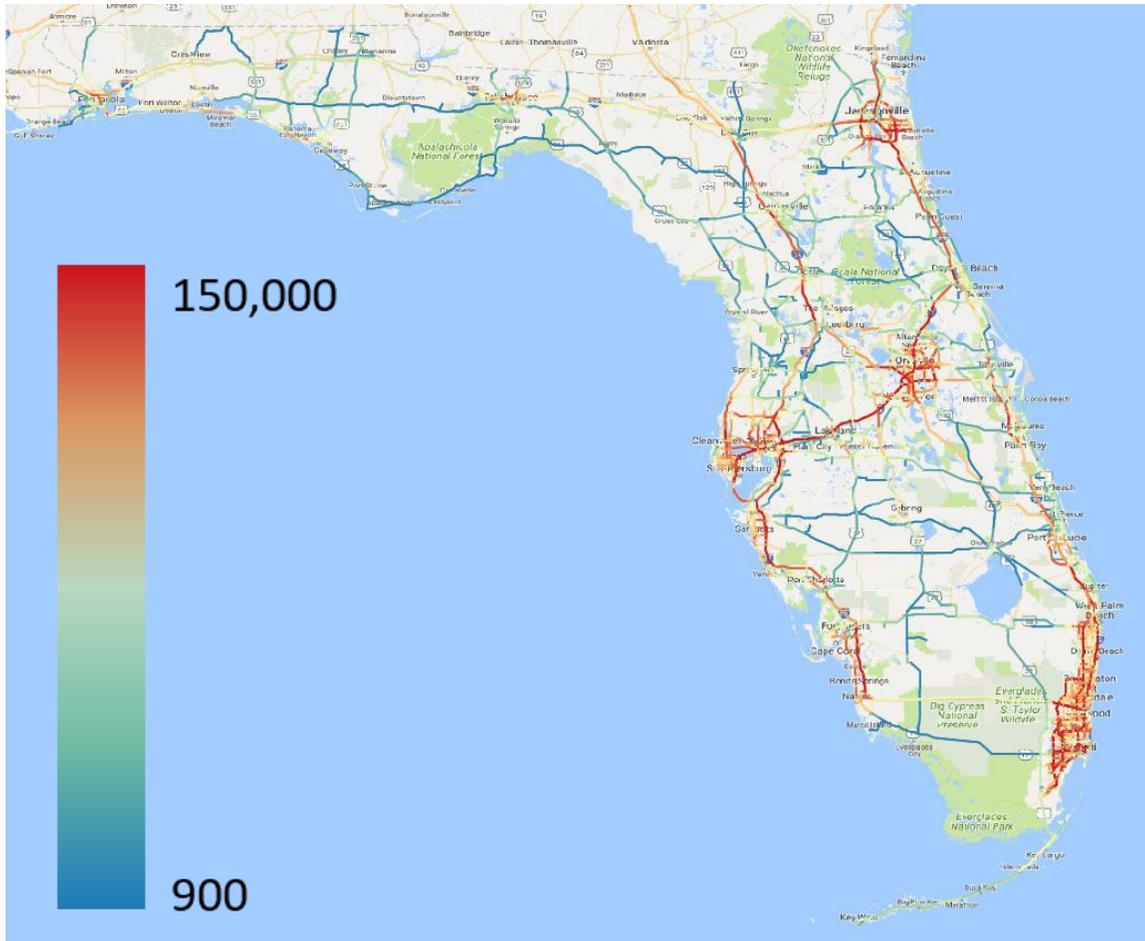
- Statewide models have been prepared for Florida, New Hampshire and Maryland
- Example: New Hampshire statewide model



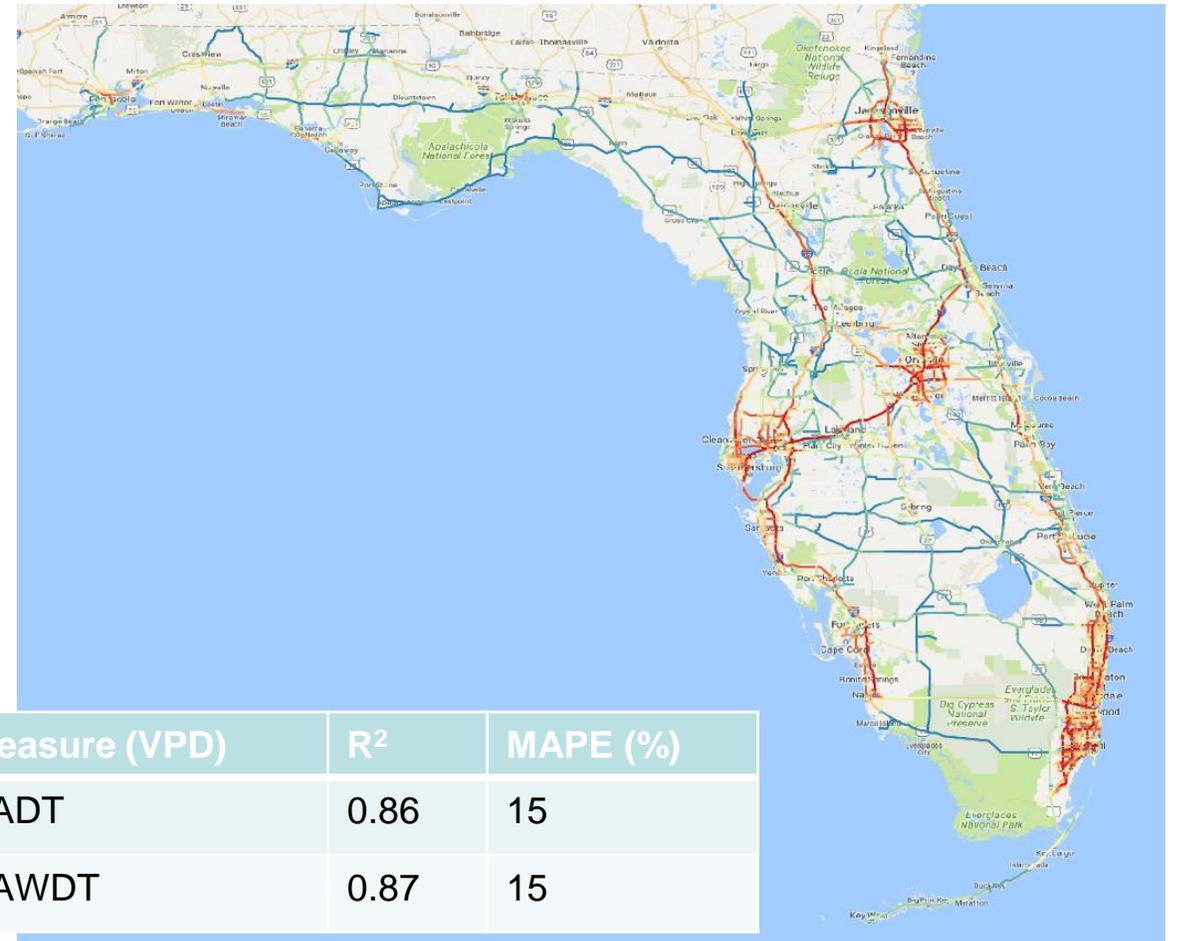
Continuous count station
selected that exhibits
typical (median) model
performance

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 |

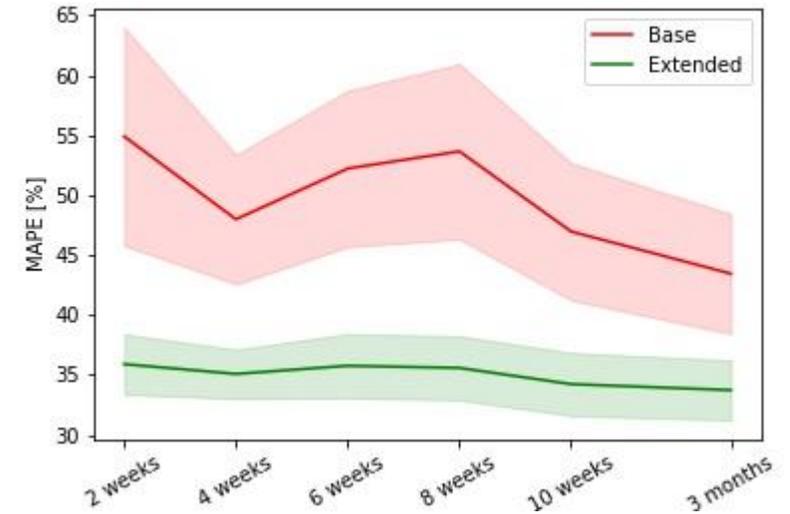
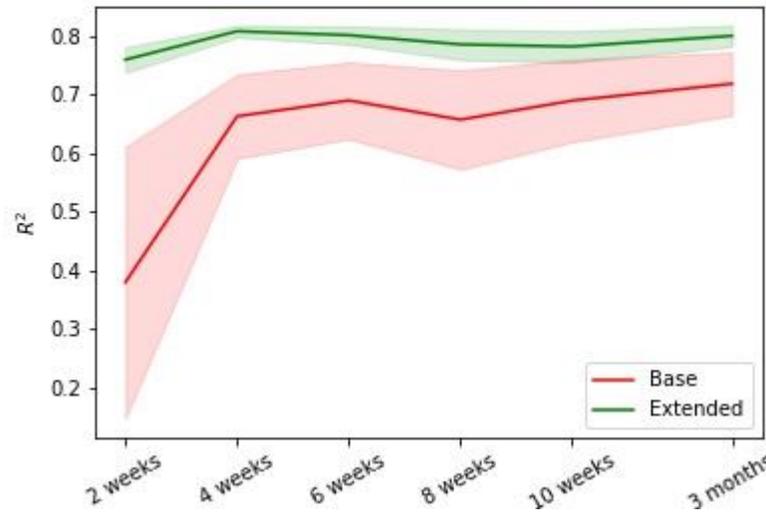
Heavy Trucks

| 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

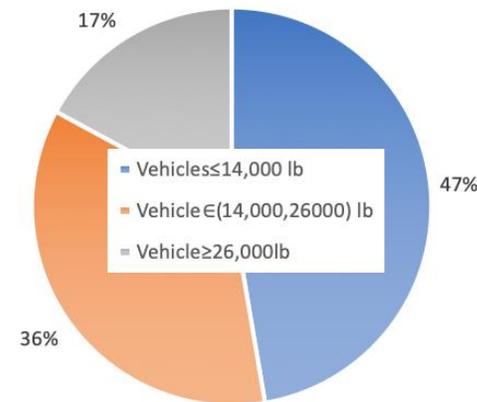
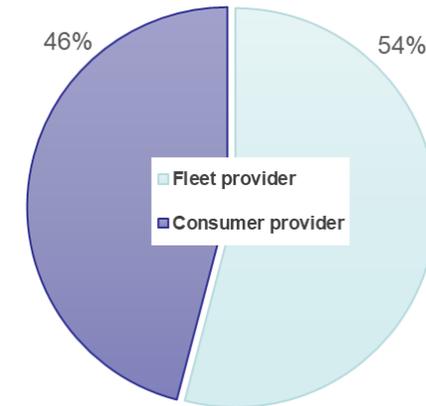
- When Florida dataset is combined with only 2 weeks of New Hampshire data for training, model predictions are reasonably good!
- Over time, only small amount of data for each new geography is needed to create powerful models!
- Potentially, the statewide traffic count data collection (and its associated 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 = 0.77
- 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) | 0.86 | 19 | 6.6 | 708,649 |
| FRC 2 (Other Freeways & Expressways) | 0.70 | 37 | 8.6 | 802,877 |
| FRC 3 & 4 (Other principal & minor arterials) | 0.49 | 43 | 10.2 | 173,448 |

| Hourly Volume (vph) | R2 | MAPE (%) | EMFR (%) | Obs |
|---------------------|------|----------|----------|---------|
| 0-1k | 0.68 | 41 | 8.3 | 891,706 |
| 1k-2k | 0.71 | 29 | 9.1 | 196,097 |
| 2k-3k | 0.85 | 25 | 7.1 | 231,518 |
| 3k+ | 0.92 | 15 | 5.9 | 365,653 |

| Avg probe counts / hr | R2 | MAPE (%) | EMFR (%) | Obs |
|-----------------------|------|----------|----------|---------|
| “Low” [0-19] | 0.64 | 40 | 9.2 | 549,734 |
| “Medium” [19-48] | 0.77 | 34 | 7.6 | 555,474 |
| “High” [48-178] | 0.89 | 17 | 6.2 | 579,766 |

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!
 - Existing data pipelines have reduced the data ingestion burden
 - Working on data pipelines to ingest short-term counts
 - 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, more importantly to realize promised cost-savings in Maryland DOT data collection program

Questions

Contact Information

Kaveh Farokhi Sadabadi

kfarokhi@umd.edu

Przemyslaw Sekula

psekula@umd.edu

Zachary Vander Laan

zvanderl@umd.edu

August 27, 2019

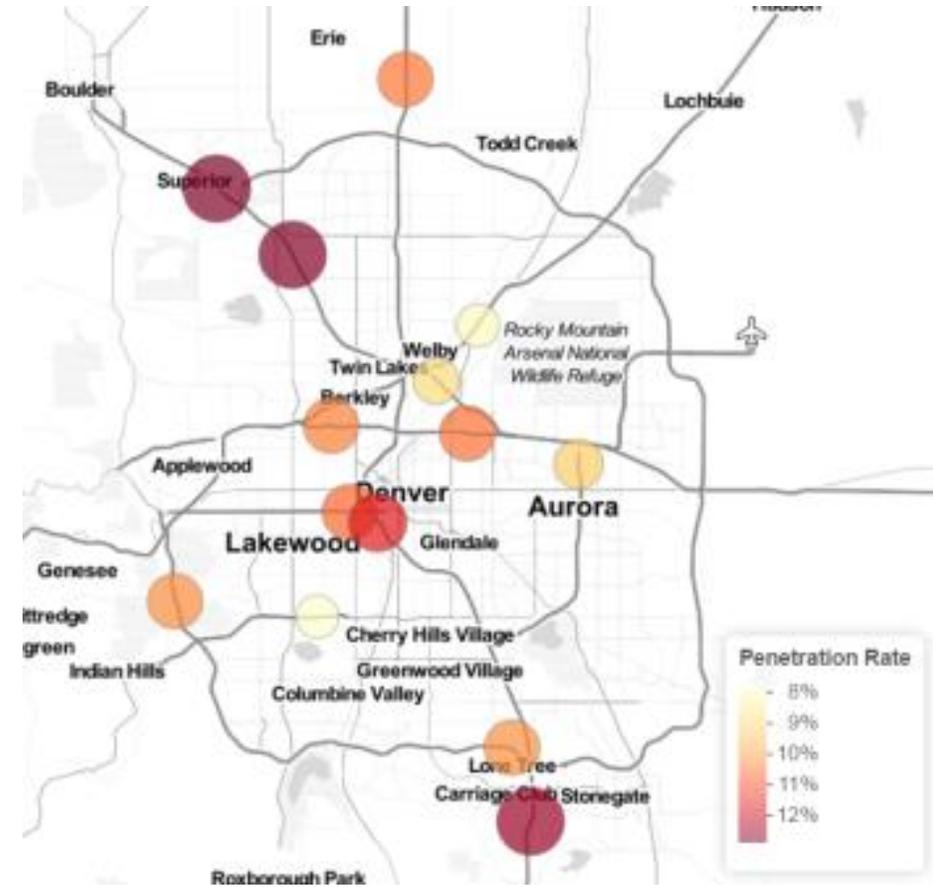


Overview of NREL Phase II DOE Technology Commercialization Grant

Venu Garikapati

Phase I – Freeway Volume Estimation

Probe vehicle penetration ranges from 8%-12%



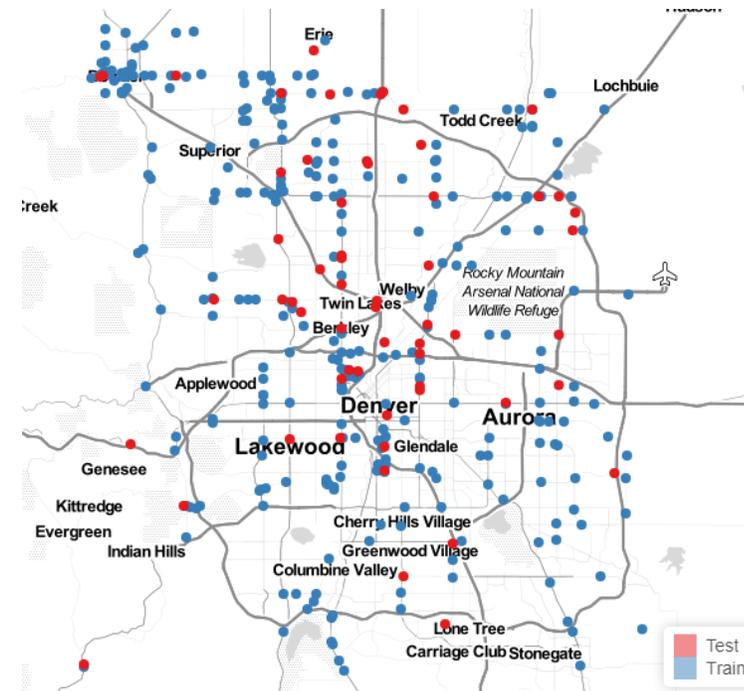
Estimation Model

- 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

Volume Estimation on Lower Functional Class Roads

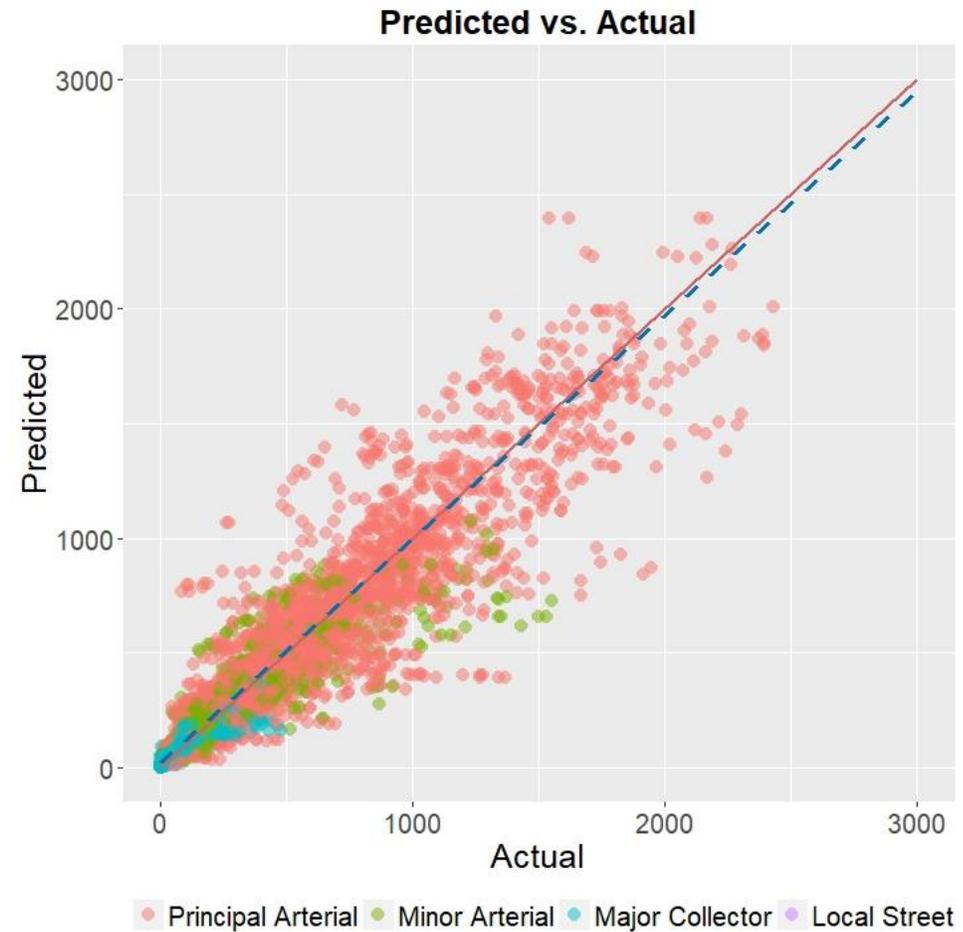
| | Lower Class Roads | Freeways |
|-----------------------------------|---------------------------------|---------------------------------|
| Volume data source | 48-hour short-term count | Continuous count stations |
| Number of locations / Data points | 359 / ~35,000 | 14 / ~52,000 |
| Data collection 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



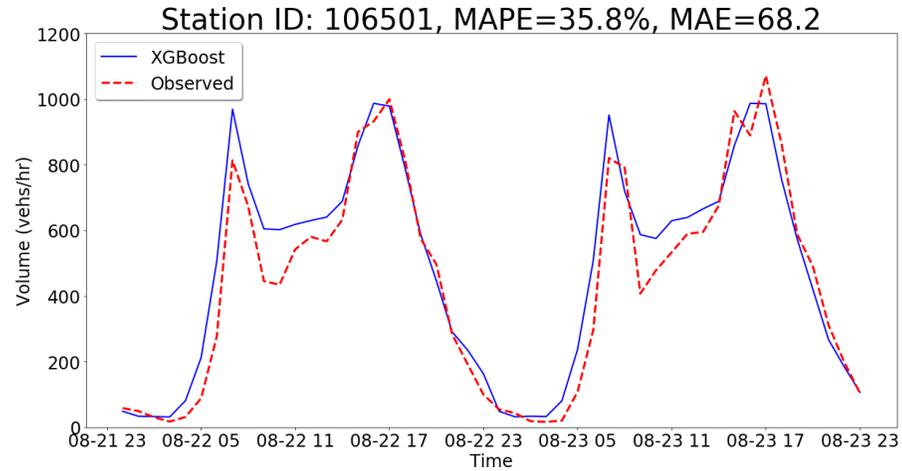
Model Results

| Roadway Type | Model | MAE | EMFR | R2 |
|--------------|---------|-----|-------|------|
| Off-Freeway | Linear | 153 | 29.5% | 0.80 |
| | XGBoost | 89 | 13.2% | 0.88 |
| Freeway | XGBoost | 357 | 5.3% | 0.91 |

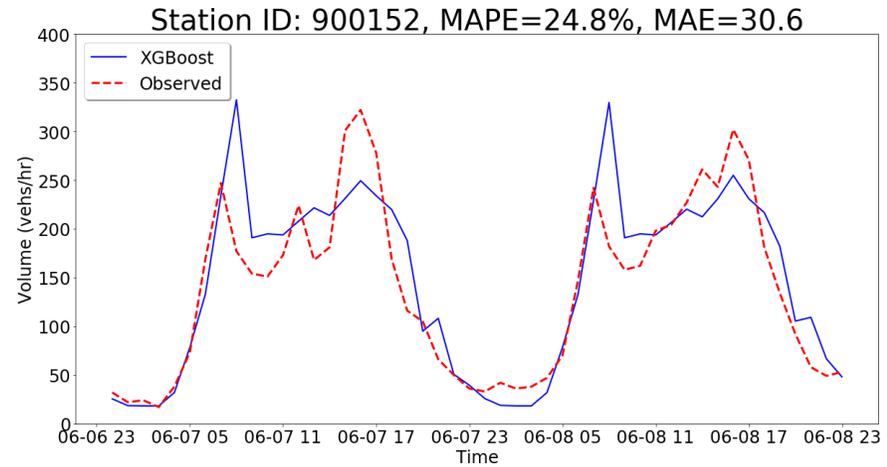


Model Performance

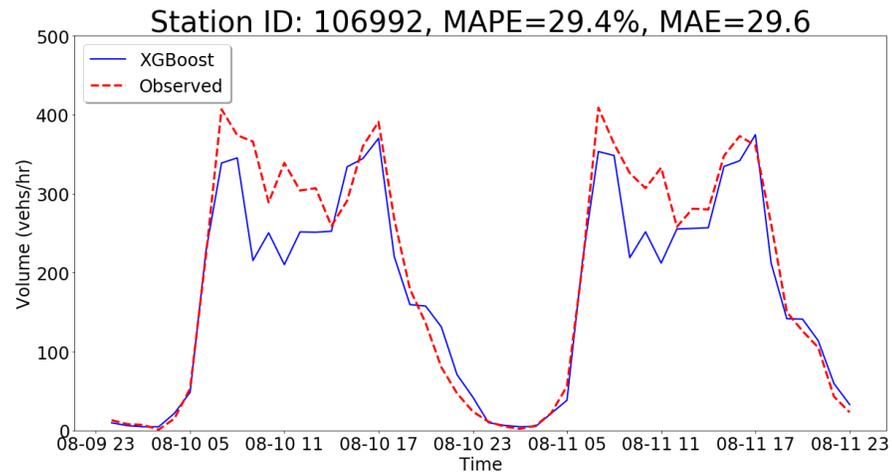
Principal Arterial



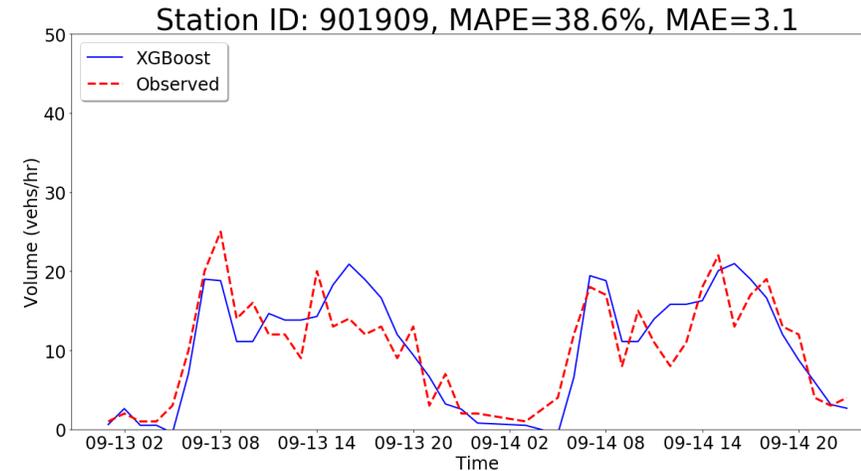
Minor Arterial



Major Collector



Local Street



Phase II Research – NREL

DOE Technology Commercialization Fund (TCF)

*“The TCF is a nearly **\$20 million funding opportunity** that leverages the R&D funding in the applied energy programs to mature promising energy technologies with the potential for high impact.*

*It uses 0.9 percent of the funding for the Department’s applied energy research, development, demonstration, and commercial application budget for each fiscal year from the **Office of Electricity, Office of Energy Efficiency and Renewable Energy, Office of Fossil Energy, and Office of Nuclear Energy.***

*These funds **are matched with funds from private partners** to promote promising energy technologies for commercial purposes.”*

TCF Objectives

- This project aims to **bring to market** a first of its kind traffic volume data product that provides accurate estimates of traffic volumes across the entire road network for all times (**24x7**) and all locations (**100% coverage**).
- The team will **demonstrate the product** to various regional, state, and federal level transportation operations and planning entities, and solicit their feedback.
- The feedback will be incorporated into a final version of a laboratory **prototype** to be integrated into our industry partner's (TomTom) product deployment environment.

Proposed Commercialization End State

- A **web-based traffic volume estimation product** that provides hourly traffic volumes across all the roads in a state or region
- Volumes are accessed either visually (as in a thematic maps) or through an API call for a specific roadway and time for **machine to machine communication**.
- Mechanisms built in for **periodic self-calibration & validation** (every week, every month etc.,) based on reference data available at high-confidence, publicly maintained traffic volume sensor locations.
- Volume estimates **integrated into existing transportation and energy applications** at DOTs and energy planning offices.

Project Plan

| Project Phase | Focus |
|---------------------------------------|--|
| Estimation Methodology | <ul style="list-style-type: none">➤ Algorithmic refinements to account for varying probe data density in space and time➤ Machine Learning Extrapolation issues➤ Confidence measures and automated confidence scoring➤ Transferability➤ Web prototype |
| Productization and Market Research | <ul style="list-style-type: none">➤ Feedback from state and local DOTs➤ Revise algorithms and prototype➤ Commercialization plan |
| Product Integration and Demonstration | <ul style="list-style-type: none">➤ Transfer prototype to TomTom platform➤ Product testing |

The Phases of Research

Phase I – I-95 Volume-Turning Movement Research Initiative

Phase II – Technology Commercialization

Phase III – FHWA Pooled Fund Study

Phase II Research – States in the Pipeline So Far

| State | Application | Research Lead |
|---------------------------------------|---|---------------|
| Colorado DOT | Statewide Volume Estimation (both for Freeways and Off-Freeways) | NREL |
| Maryland DOT | Statewide Volume Estimation (both for Freeways and Off-Freeways) | UMD |
| Pennsylvania DOT (Harrisburg) | City level volume estimation Mid-sized city representative of larger potential for state | NREL and UMD |
| North Carolina DOT | Off-freeway volume estimation | NREL |
| Tennessee DOT (Chattanooga) | City level volume estimation Small and rural city | NREL |
| Other State Opportunities | | |

Opportunity for additional State DOT Involvement

- Participating in Phase II –
 - Opportunity to demo volume estimates for a mid-size metro area (i.e. Harrisburg)
 - Expectation of States
 - Provide data for calibration of volume estimation
 - Evaluate volumes for either operations, planning, or performance measures
 - Produce a report on use and adequacy
 - Participate in steering committee

The Phases of Research

Phase I – I-95 Volume-Turning Movement Research Initiative

Phase II – Technology Commercialization

Phase III – FHWA Pooled Fund Study

Non-Traditional AADT Volume Estimates

- Objective – Provide guidelines for non-traditional volume data collection for AADT and planning purposes
- Funding -
 - FHWA Pooled Fund Study - ~ \$1.3 M
- Timeframe - 2019-2021
- Partners -
 - Pooled fund study (insert website)
- Status
 - Primary research firm under procurement
 - Will contract for validation (referee)
 - Should know more into fall of 2019

Final Questions



November Steering Committee Meeting

- North Carolina volume estimate results – NREL
- Machine Learning Extrapolation addressed
- *Optimal Traffic Monitoring in a New Data Age* –
– An I95 Sponsored White Paper
- Updates on other Phase II activities

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

