I-95 CC – Volume & Turning Movement Project
Steering Committee Meeting #5
July 27, 2017

Agenda:

	Topic	Speaker
1	Welcome & Project Status Update	Denise Markow, I-95 Corridor Coalition Stan Young, NREL
2	Spotlight Presentation: Real-time Volume Estimation with TomTom Probe Data – Denver Area	Yi Hou, NREL
3	Polling Questions for Steering Committee Input	Steering Committee members
4	Validation Framework	Shawn Turner, TTI
5	Next Steps & Wrap Up	Stan Young, NREL Denise Markow, I-95 Corridor Coalition

Next Steering Committee Meeting: Thursday, Nov 9, 2017 (10:30 am – 12:00 pm EST)

Meeting Notes:

- ***Welcome – Denise Markow, I-95 CC***
  - Opening remarks and highlighted the year in review
    - User Survey – Complete
    - Preliminary Data Analyses
      - Maryland - Complete
      - Rhode Island – target for next analysis
      - Florida – target for next analysis
      - Colorado (Denver Area) – complete
    - Calibration (FHWA TMAS)
    - Validation (TTI)
- ***Project Tasks Status Update – Stan Young, NREL***
  - Reviewed the project goals, objectives, and status
  - Stan briefly explained the status of currently available volume data and its limitations. The goal and objectives were reviewed as well as where the effort currently stands related to these objectives.
  - Stan presented the project flow chart and explained how it is evolving and how roles have changed. UMD is working with the INRIX data and NREL is working with the TomTom data and HERE data is being evaluated for suitable to the volume estimation task. TMAS is anticipated to be the heart of the calibration testbed and TTI will take the lead on developing a validation framework. The products of the research initiative remain the same even though roles have evolved. He also noted that all base data sets from the vendors are different in nature.
  - The status of the project elements was reviewed
    - Work by UMD and NREL are being formalized, and targeting publication opportunities toward the end of the year (TRB, IEEE, ITSWorld)
    - UMD is evaluating samples of HERE and Streetlight data – ongoing (note the nature of data brought by each is different).

- The FHWA TMAS system serves as the calibration test bed. It works well for archive data; however, the research team needs to work with states for real-time feeds, or low latency volume data for calibration. NREL is currently in the process of working with CO for such data.
- Shawn Turner, who will speak later from TTI, has initialized a validation framework.

- ***Real-time Volume Estimation with TomTom Probe Data–Denver Area - Yi Hou, NREL***
  - Problem: previously all volume data was collected by stationary sensors – how can we obtain an accurate estimate of ubiquitous real-time traffic volume data in road networks?
  - Proposed Solution: Build a model based on probe data, road characteristics, weather information, and temporal information to predict traffic volumes
  - Data sources for the model:
    - FHWA TMAS (ATR station/road characteristics)
    - CDOT ATR Counts (temporal and volume information)
    - Weather (weather underground, weather info)
    - TomTom Probe Data (probe vehicle data)
  - Procedure:
    - Match ATR locations with TomTom segments
    - Merge all data sources into a table
    - Determine which variables correlate with volume (either positive or negative)
      - Positively Correlated: average hourly probe counts of past 6 hours; number of lanes; capacity
      - Negatively Correlated: average hourly average speeds of the past 6 hours; speed limits
    - Build model to predict volume data on locations that don't have any data based on input variables
  - Selected several input variables from each data set for model
    - Average speed, probe count, temperature, visibility, road type, speed limit, number of lanes, capacity, month, day of week, hour of day
  - Used three machine learning models (NREL selected models which don't require detailed mathematical forms/assumptions, have few parameters, and are suitable for the study)
    - Random Forest (RF)
    - Gradient Boost Machine (GBM)
    - Extreme Gradient Boost (XGBoost)
  - Trained the models on 14 ATR locations (13 for training, 1 for validation) and repeated 14 times for each 14 stations
  - Evaluated the model with three criteria (all performed significantly better than linear regression)
    - Coefficient of Determination (R^2) - how well the model can explain data variance
    - Mean Absolute Percentage Error (MAPE) - error measure when comparing estimated volume with ground truth
    - Error to Theoretical Capacity Ratio (ETCR) - error measure when comparing estimated volume with roadway capacity

- Conclusions:
  - Models are much better than linear regression
  - Probe vehicle data has significant impact on volume estimation accuracy
  - Model performs poorly on unique cases or outliers (e.g. testing on roadway with double the volumes of those in training data set)
  - Accuracy can be improved with more training data
  - Can repeat the procedure with more data or possibly deep learning to see more results
- Gave a demo of webapp framework for the model
- Questions regarding this presentation included:
  - Daivamani Sivasailam (MWCOG) asked if the GBM model was the best based on error calculations. Yi agreed that it seemed to perform best, however the XGBoost has a much lower computation time, so while it was a little less accurate, it would be more efficient.
  - Nick Cohn (TomTom) asked Yi to elaborate on outliers. The models are trained on several data sets and then validated on a different data set. Data sets that have significantly higher volumes are difficult for the model because they haven't experienced them yet – increasing the amount of training sets will improve model output. The model may also differ with road type (mostly because it has trouble with very different road volumes).
  - Stan Young (NREL) asked if model constraints were added to current methods (machine learning), would it break the model or would it improve? Yi remarked that it would be difficult to incorporate in current machine learning models and he will have to explore more methodology.
- Polls
  - #1 - Having heard the descriptions of the error measures below, which is most interesting to you? (Select all that apply)
    - R^2 – 6/16 (37.5%)
    - MAPE – 14/16 (87.5%)
    - ETCR – 5/16 (31.25%)
  - #2 - Having heard the descriptions of the error measures (R^2, MAPE, ETCR) should the research team consider other error measures?
    - Yes – 6/15 (40%)
    - No – 9/15 (60%)

- ***Validation Framework – Shawn Turner, TTI***
  - Shawn Turner explained that TTI is involved in this project as a 3rd party independent validator since UMD and NREL have now become data providers.
  - Two-pronged validation – "trust but verify"
    - Trust: self-reported info about model calibration process; high level descriptions of sample scaling approach, goodness-of-fit or other calibration statistics
    - Verify: 3rd party independent validation similar to I-95 CC VPP travel time validation; use existing DOT continuously count stations for benchmark data; standardized accuracy measures and measure categories
  - Self-reported information: TTI wants transparency from vendors in how data and models are set up, how they scale, and how models are calibrated (and with what data and training sets)

- TTI will verify via benchmarks for volume data from DOTs (that is not yet publicly accessible). TTI needs DOTs who can provide bench mark data soon after it is collected (and before it is released to the public).
- Accuracy measures: MAPE is good, but at high volumes it's not a great measure (10% error at 200 vehicles is different than 10% at 20 vehicles). The proposed solution is to make a sliding scale for volume estimates (i.e. higher % error is more acceptable at lower AADT).
- TTI is looking to standardize accuracy measures and measure categories. Categories might be volume categories or percent of capacity categories binned into "high, medium or low"
- Volume category accuracy measures being considered include: MAPE, ETCR, Mean Signed Error – Bias (does the data consistently skew toward higher/lower?)
- Percent of capacity accuracy measures being considered include: Error-to-Maximum-Flow Ratio (EMFR)
- TTI initially considered other measures (including GEH statistic, some other error formulations) but didn't recommend them because the statistic results aren't intuitive
- Next Steps:
  - Gather initial feedback (today)
  - Identify available benchmark DOT continuous data count (must be provided to TTI before publicly available)
  - Develop more detail in concept paper
  - Gather feedback on concept paper on I-95CC and volume producers in September/October
  - Develop consensus on validation framework by November
- Questions:
  - Stan Young (NREL) asked if the validation results should be presented in bins, similar to the I-95CC VPP validation – either by absolute volume bins or % capacity bins. Shawn agreed that yes, he would suggest providing the results in that manner. Right now, they're looking at a MNDOT project and it shows that volume levels matter and matter quite a bit. Level of congestion also matters in terms of accuracy results too. They can certainly aggregate volume levels and report a single error or accuracy statistic, but more insight from discrete volume categories. The question is how they should be binned. Also consider binning by signalized or not signalized, rural or urban, etc.

- ***Wrap up – Stan Young, NREL***
  - We can only develop spatially ubiquitous, omni-present volume data for both real-time and historic applications in conjunction with traditional continuous counters for calibration and validation. This project is not intended to replace the existing network of high quality count stations, only augment in space and time.
  - The evidence from the R&D suggests that probe data is approaching a point to support volume estimation everywhere and at all times.
  - Challenges for the future:
    - Variation in penetration rates (8-12%) presents issues
    - The Colorado results indicate that the calibration data set must span the full range of volumes expected to be experienced on the network. Or in

other words, calibration data sets have to have a large breadth of conditions.
    - Large fluctuations in estimated volume along a roadway are a concern, and are being investigated
    - Making a constrained model will require more research/considerations

- ***Closing Remarks – Denise Markow*** (I-95 Corridor Coalition)
    - Denise thanked all members for their participation and reminded them about the **next steering committee meeting - Thursday, Nov 9, 2017 (10:30 am – 12:00 pm EST)**

Presenter Contact Info:

Real-time Volume Estimation - Denver Area:
Yi Hou, PhD, National Renewable Energy Laboratory (NREL) yi.hou@nrel.gov
Validation Framework:
Shawn Turner – Texas A&M Transportation Institute (TTI) s-turner@tti.tamu.edu
Project Contact Info:
PI – Kaveh Sadabadi (UMD-CATT) 301-405-1352 or kfarokhi@umd.edu **Co-PI** – Denise Markow (I-95 Corridor Coalition) 301-789-9088 or dmarkow@i95coalition.org **Co-PI** – Stanley Young (NREL) 301-792-8180 or Stanley.Young@nrel.gov **Logistics** – Joanna Reagle (KMJ Consulting, Inc.) 610.228.0760 or jreagle@kmjinc.com

Action Items:

#	Action Item	Whom	Status
	VTM Steering Committee Meeting – July 27, 2017		
1	DOTs willing to submit traffic count data for validation (specifically continuous traffic counter data shortly after it is collected and prior to publishing to public) should contact Shawn Turner, TTI	Member agencies	
	Previous VTM Steering Committee Meetings		
2	Determine if a report documenting the methodology used by UMD to compare traditional AADT profiles to estimate volumes is available. If so, to provide it to Wenjing Pu and the rest of the Steering Committee.	Kaveh Sadabadi	Currently being prepared. Once this report is complete it will be shared with the Steering Committee.
3	Verify the values shown on the performance evaluations with the analysts and provide clarification to the Steering Committee.	Kaveh Sadabadi	Ongoing. A brief report is being prepared to address the questions and to provide more insight into the model evaluation results.

Participants:

Project Team:
Denise Markow, Marygrace Parker, I-95 Corridor Coalition Stan Young, NREL Kaveh Sadabadi, UMD CATT

Steering Committee:	
Erik Sabina	Colorado DOT
Jason Tao	District DOT
Jimmy Chu	FHWA
Joe Guthridge	HERE
Amy Lopez	INRIX
Glenn McLaughlin, Abhay Nigam, Jeff Scruggs (JMT)	MDOT SHA
Daivamani Sivasailam, Patrick Zilliacus	MWCOG
Keith Miller	NJTPA
Scott Benedict, Steve Gault (Michael Baker)	Pennsylvania DOT
Steve Brown	Port Authority New York & New Jersey
Michael Dennis	South Carolina DOT
Nick Cohn	TomTom
Shawn Turner	TTI
Mena Lockwood	VDOT
Nikola Marković, Przemek Sekula	UMD CATT

Consultant Support Staff:
Joanna Reagle, Justin Ferri, KMJ Consulting, Inc.