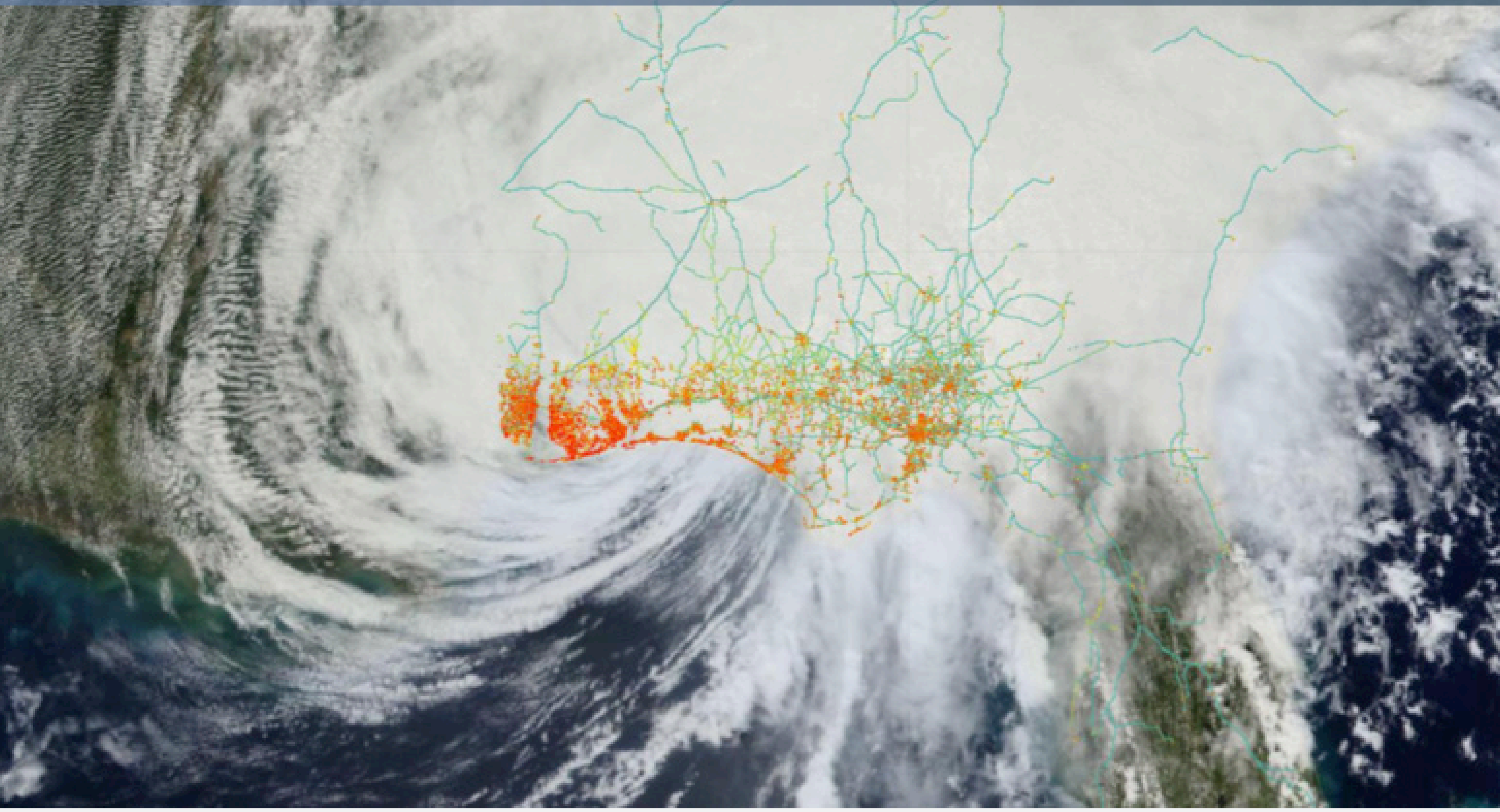


HURRICANE PROOF OF CONCEPT RESULTS

STATES' EXPERIENCE WITH REAL-TIME CONNECTED VEHICLE DATA

THE EASTERN TRANSPORTATION COALITION

CONNECTING FOR SOLUTIONS

ACKNOWLEDGEMENTS

The TETC would like to thank its technical partners and the state member agencies who participated in the preparation and implementation as well as providing invaluable feedback during the proof of concept. The TETC looks forward to the mutual collaboration and development effort that will result in valuable tools, data and information for TETC members to apply in their duties to operate, manage and plan for an effective transportation system.

THE TECHNICAL TEAM

Stan Young
NREL
Advanced Transportation & Urban Scientist

Kaveh Sadabadi
University of Maryland CATT
Senior Faculty Specialist

Zachary Vander Laan
University of Maryland CATT
Faculty Research Assistant

Peter Carnes
Traffax
Chief Executive Officer

Jaap Van Den Hoek
Wejo
Director of Solutions Engineering

Eimar Boesjes
MoonShadow
Chief Executive Officer

Wander Boesjes
MoonShadow Mobile, Inc
Training Coordinator

AGENCIES:

Maine DOT

Virginia DOT

North Carolina DOT

Georgia DOT

Florida DOT

Alabama DOT

Tennessee DOT

EXECUTIVE SUMMARY

This pilot proof concept started as a basic question – "Could Connected Vehicle Data (CVD) be used as a real time data source to estimate volumes and if so, what would be needed to make this possible?

So, The Eastern Transportation Coalition assembled a technical team and procured three months of Connected Vehicle Data along with access to an analytics platform to assess the viability of delivering and visualizing a real time CV data feed as well as devising a process that could validate the accuracy of the traffic volume estimates.

The visualization of regional CV data as individual vehicles in a live data stream brings amazing intuitive value. CV data is growing in size and velocity and we are blazing new ground to visualize real time volume and O&D. Six southeastern states participated in the pilot, hence the volume of data was huge. There were over 75B data points per month which came to over 230B data points in the 3-month pilot study period for the six states. It showed that managing CV data at scale is challenging for the industry and that sheer size and velocity of the data will require efficient processing architecture, as well as streamlined calibration, calculation and conflation techniques.

For the validation process that is detailed in the report, simple scaling factors based on roadway class, time-of-day, day-of-week, and state were demonstrated to provide a workable and sufficiently accurate traffic volume estimate. Including additional factors, and leveraging more advanced machine learning may improve accuracy.

Overall, the pilot clearly points to a viable path toward real-time traffic volume estimates based on CV data.

INTRODUCTION

There has been a long history of collaborative research between The Eastern Transportation Coalition (TETC), the University of Maryland (UMD) and the National Renewable Energy Laboratory (NREL) in looking at how to derive accurate volume data from vehicle probe data. There have been several past studies working with the states of Florida, Maryland, Colorado, New Hampshire, Pennsylvania, Massachusetts, Tennessee – all with the objective to validate historical volume estimates from various types of probe data.

This past research has shown that we can estimate volumes historically with adequate accuracy to be helpful to planners, however, for operational benefits, there is still a real need to estimate traffic volumes in real time. Traffic volumes are the second dimension (alongside travel time and speed) for full operational awareness. Hence, the overall goal of this project was to evaluate the viability of using real time connected vehicle data to estimate real time volumes.

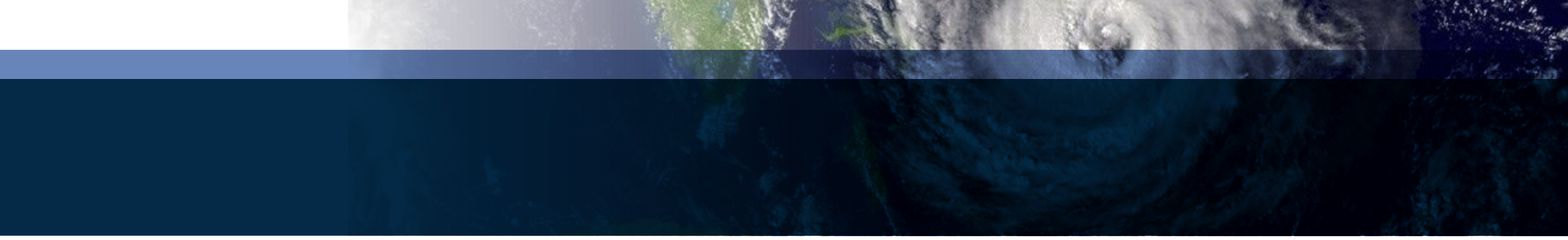
In 2019, Connected Vehicle (CV) data appeared in the marketplace for the first time. Connected Vehicle Data (CVD) originates from automobile manufacturers. Since approximately 2014, manufacturers began to embed broad band communications directly into the vehicle electronics, allowing the vehicle to communicate with the manufacturer. This communications ability along with built in GPS for location referencing has steadily grown in the fleet so that by 2020 the density and richness of data provided an opportunity to test real time CVD for volume estimation purposes. In 2020, the Coalition partnered with Wejo as the data source that was chosen for this hurricane pilot study. Many questions arose with the Proof of Concept:

- Could CVD that comes directly from OEM manufacturers from all across the world be used as a real time source? The intent was to evaluate this data source as a viable real time traffic feed in the context of the hurricane season which is growing in importance especially with global warming concerns.
- Given that the data feed was viable, could the traffic data be processed and visualized in real time, and if so, what would the issues and challenges be working with this type of data source?
- Could we use CV data to generate traffic volume estimates in real time so that this data source could be used to monitor hurricane evacuations broadly across a multi-state region?
- Lastly, could we provide meaningful feedback to the operators on travel patterns related to hurricane preparedness and evacuation. This includes people traveling to get outside of the path of the storm or taking appropriate precautions (sheltering in place for example).

STAKEHOLDERS

The stakeholders that participated in the pilot POC are identified below.

PLAYERS	FRAMEWORK
The Eastern Transportation Coalition	Provided overall coordination and funding
Traffax	Traffax, Inc. and staff, known for their validation work with the Coalition's marketplace, brought the data together, processed it and provided insights on its accuracy.
State DOT ETC Members	FL, AL, GA, TN, NC, VA - some of the agency's best data scientists were involved in providing feedback.
Wejo/MoonShadow	CV data source and analytics platform partners
National Renewable Energy Lab	Collaborator & Technical Partner for ongoing research providing expert feedback and guidance



TESTING A SOLUTION

WHAT?

- Wejo CV Data representing about 3% or more of all vehicle trips on the roadway
- The pilot received data from each probe vehicle every 3 seconds
- 21 seat licenses were distributed to the six participating states to view live stream vehicle movement on the MoonShadow analytics tool
- The Traffax team estimated real time volumes based on basic linear scaling factors due to the short testing period. If the process proved viable, more sophisticated modeling tools could be applied later.

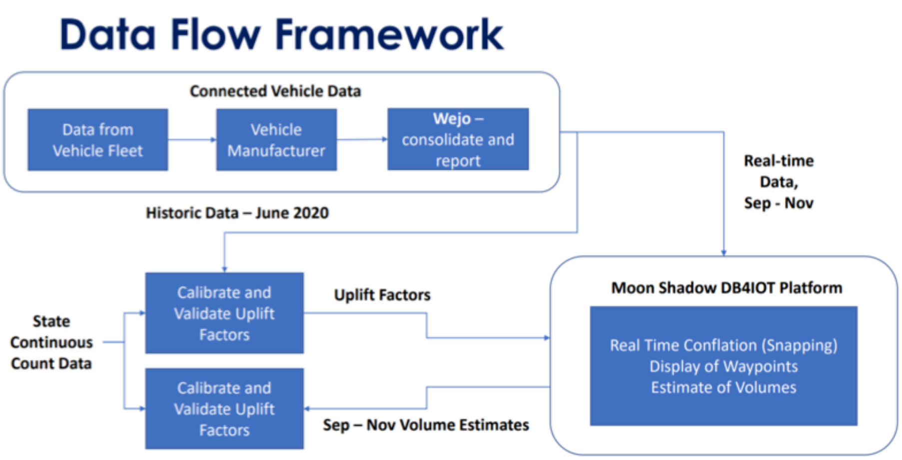
WHERE?

- The southeastern states of Alabama, Florida, Georgia, North Carolina, Tennessee, Virginia, states impacted by potential hurricane events were invited to participate
- The top road classes (Interstates, Turnpike and the NHS system) were the target of the proof of concept. The project used Open Street Maps as a common mapping system across the six participating states.
- These roadways correlated well with the FHWA Functional Roadway Classes (FRC) 1,2 & 3.

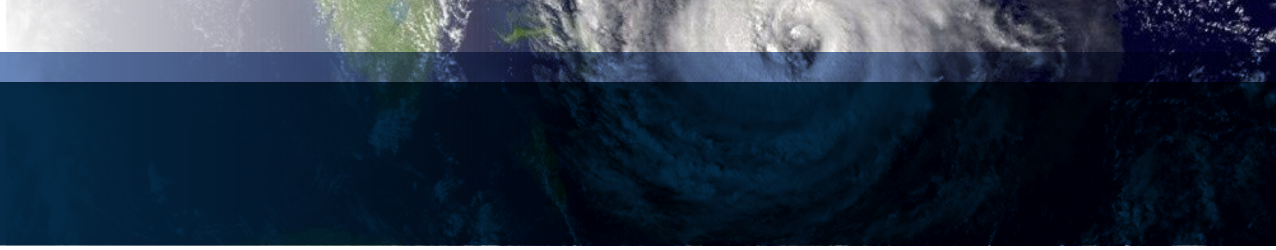
WHY?

- Are we ready to take what we know works about real volume estimate for historical purposes and apply it to a real time data feed to support operations?

A conceptual data and processing flow diagram, shown to the right, depicts the information flow for the proof of concept. Data from vehicles are reported back to Original Equipment Manufacturers (OEM) who package, sanitize and send that data to Wejo who consolidates and communicates it in real time. The maximum latency in the process is 60 seconds, which means 30-60 seconds from the time it leaves the vehicle to the time it is ready for applications.



Connected Vehicle Data Pilot architecture



Both historic data and a real time data were utilized in the proof of concept. Historic CV data from June of 2020 from Wejo along with data from participating states' continuous count stations were used to develop uplift factors and estimate accuracy. From September through November of 2020, a real time data feed, literally a firehose of data, was processed it in real time, snapped to Open Street Maps, and volume estimates derived. The MoonShadow analytics platform, which displayed the raw data in real time, provided querying tools to view volume estimates, and overall was the experimental platform to explore how the information could be displayed and managed in an implemented system.

GOALS AND OBJECTIVES

There were four goals in the pilot study:

- **Goal #1:** Demonstrate that real-time connected vehicle data across all six states can be delivered and visualized in real time.
- **Goal #2:** Process observed CV trips in real time, assigning them to appropriate roadway segments, and obtain counts of probe vehicles from which to estimate traffic volume.
- **Goal #3:** Validate the accuracy of Traffic Volume Estimates derived from CV data applied to a live data stream by calibrating to known journey penetrations rates, and observing consistency of probe penetration rates during the proof of concept.
- **Goal #4:** Show that the mobility patterns of people, where they go and when, change significantly in the event of a major storm.

CONNECTED VEHICLE DATA

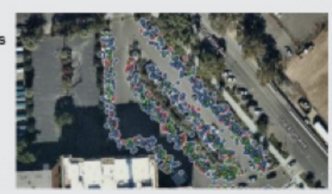
Why are organizations considering connected vehicle data for traffic operations?

- Greater spatial accuracy – OEM telematics typically have antennas (both GPS and cellular) integrated into the vehicle for more reliable telemetry, resulting in higher quality data.
- More data attributes – in addition to location, CV data typically has information about the vehicle operations such as vehicle start/stop, windshield wiper status, seat belt usage, and hard stops.
- Consistent higher frequency – data is reported every few seconds
- Leads to better city and infrastructure design
- Supports new or modified regulations
- Underpins new mobility propositions

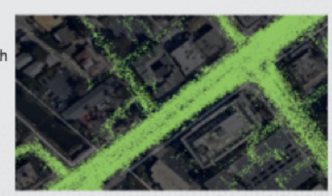
The Power of CVD: Unlocking Real Value

3 second capture rate with 30 second latency

Accurate to 3 meters, which helps identify highway lanes and parking spots



High volume of journeys tracked: 1.3 billion per month



Data from historical events gives insights about incident hotspots, harsh breaking or acceleration, speeding and more

The first objective was to demonstrate that real time connected vehicle data across all six states could be delivered and visualized in real time for use in estimating traffic volumes. The table below gives an idea of the daily data ingestion rate of the CV data from the six state region used in the pilot study. The MoonShadow DB4IoT was the visualization platform utilized to report on these metrics.

WEJO METRICS CAPTURED	PER DAY
Waypoints/Day	2,500,000,000
Trips/Day	7,500,000
Vehicles/Day	2,500,000
Peak Hour Vehicle Updates per Second	100,000
Files per Hour	40,000
Gigabytes/Day	50
Average Vehicle to User Map Latency (seconds)	40
Average User Map Update Frequency (seconds)	
Vehicle Update Frequency (seconds)	

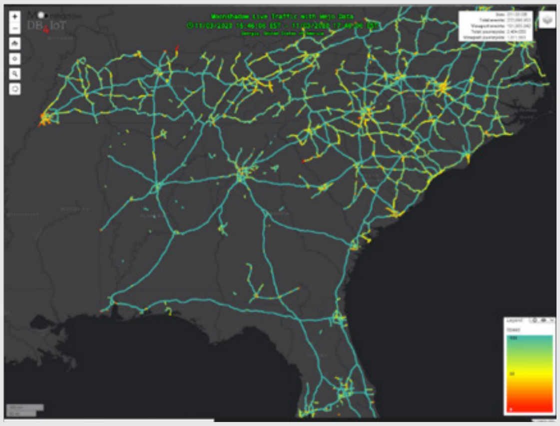
Some of the lessons learned as well as the challenges that were seen in working with CV data are noted below.

- The volume of data was huge. Wejo data is a true Big Data (>TB) Real Time Data Source
- There were over 75B data points per month which came to over 230B data points in the 3-month pilot study period for the six states
- Real time data from six states that flowed from Wejo to the MoonShadow analytics platform displayed individual vehicles moving in real time
- Establishing streaming CV data connections from Wejo to MoonShadow in a fairly short amount of time was challenging
- Real time visualization across six state was not a trivial challenge, and required significant software engineering to meet the challenge during this POC
- Base Data provided in geodetic (lat-lon) from Wejo was map agnostic
- Before the end of the POC. Wejo and MoonShadow succeeded in generating traffic volume estimates in near real time, in 15-minute time intervals based on a 3% probe penetration rate
- Overall, the size and velocity of data requires planning and the appropriate network, storage and processing architecture

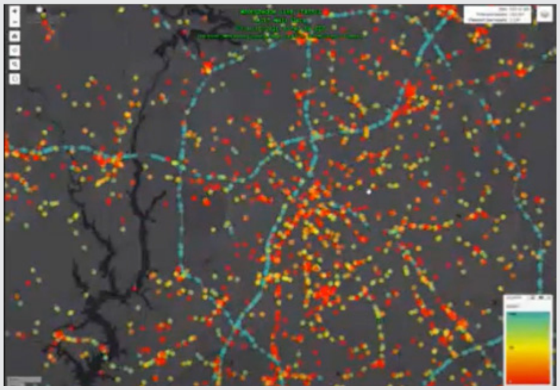
The second objective was to a) process observed CV trips in real time, b) assign them to appropriate roadway segments, and c) obtain counts of probe vehicles from which to estimate traffic volume. As previously noted, the team worked with the MoonShadow DB4IoT platform which ingested and processed the CV data. This figure shows a representation of the real time CV probe data trips in this platform. Each dot is the location of a moving vehicle updated every 3 seconds. A major comment from participants about the pilot is that this visualization gives one an intuitive sense of flow on the network even before the data is processed to volume estimates. The visualization of raw Wejo CV data could be used to confirm contraflow traffic operations, or whether a road is open/closed/passable during a hurricane event.

However, CV data typically lacks traffic performance informational content such as speed or volume of the traffic stream in its native form. CV data needs to be aggregated/assigned to appropriate roadway segments to obtain such information.In order to capture probe counts (the basis data for volume estimation), CV data needs to be snapped or conflated to a base map. This pilot used Open Street Maps (OSM) as a common source of mapping data across the six state region. Some of the processes and lessons learned in this process are noted below.

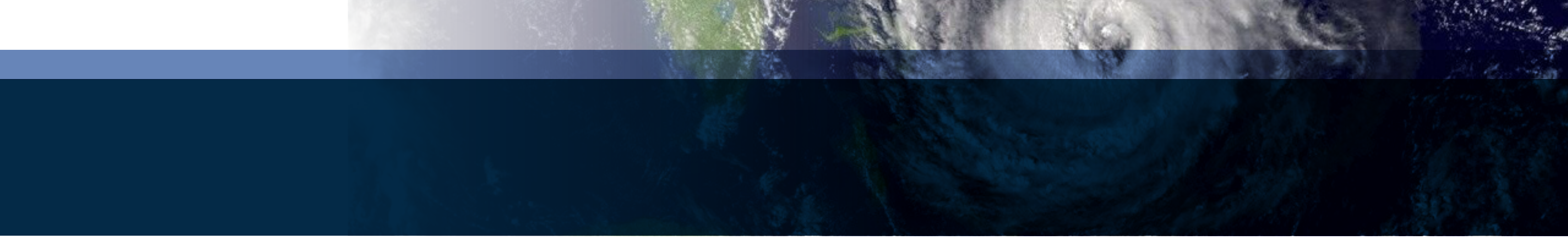
- Real time CV data is intuitive but volume estimates require aggregation to segment levels
- Decide on a stable base map version and stick with it.
- Append OSM 'Way IDs' (segment identifier) to each individual data point coming in
- Conflating data to map segments (OSM) in real time is challenging but doable
- The data was successfully snapped to road segments in real time
- Segment aggregation was then made available through querying tools within the analytics platform.
- Road snapping required approximately another 60 seconds in processing lag, resulting in greater latency for the raw CAV data.
- Display of Wejo raw CV data averaged approximately 30 second measured latency between data capture (in the vehicle) and display in the MoonShadow platform



MoonShadow DB4IoT Analytics Platform Visualization



MoonShadow DB4IoT Analytics Platform Visualization



The third objective was to see if the team could assess meaningful volume measures in real time from the live data streams by calibrating to known journey penetrations rates. This was the heart, and central objective of the pilot. Due to the compressed timeframe of the project, the technical team could not put an end to end calibration/validation process in place, but rather produced as much evidence as possible within the given resources that real time accurate volume estimates were achievable.

THE UPLIFT PROCESS

Uplift is a process where the historical relationship between the counts of CV data detections to the proximate count-station counts is applied to all similar road types to let the end user think about roadway volume as number of vehicles, an intuitively meaningful way, vs. number of CV detections. This uplift process allows the system to manage any variability in this relationship whether by road type, day of week, hour of day, etc. or combinations of factors.

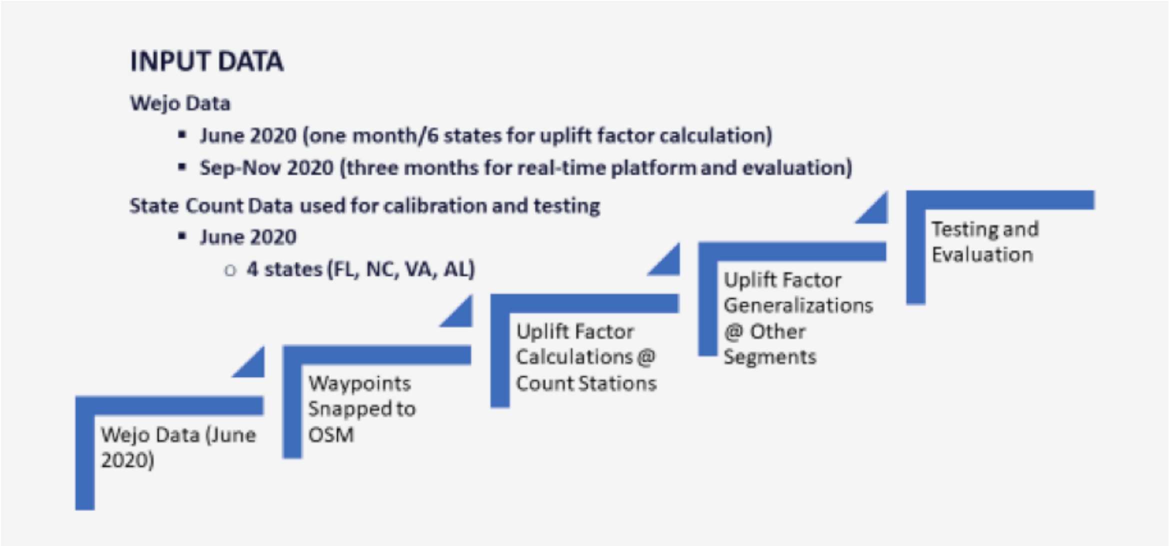
The calibration team received June 2020 CV data from Wejo to use for the calibration process. Waypoints were snapped to common OSM segments, uplift factors were calculated at the continuous counting stations and the nexus was to generalize these factors to the rest of the segments in the network. Calculations were performed for Florida, North Carolina and Virginia focusing on Functional Road Classes (FRC) 1,2 3 as shown below.

FRC 1

FRC 2

FRC 3

As the process started, it became apparent that important factors like Time of Day (TO&D) and Day of Week (DOW) needed to be included due to the observed fluctuations in the density of probe vehicle observations. Fifteen-minute time intervals for 7 days of the week for a total of 96 time intervals were captured for each day.

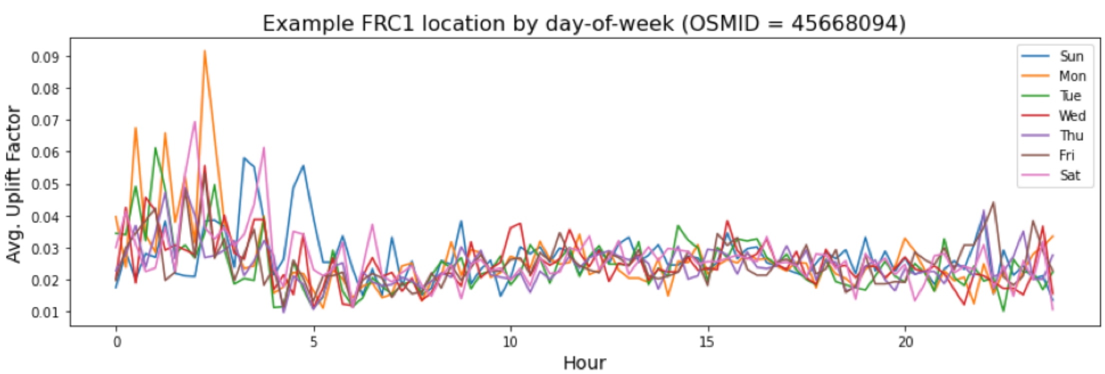


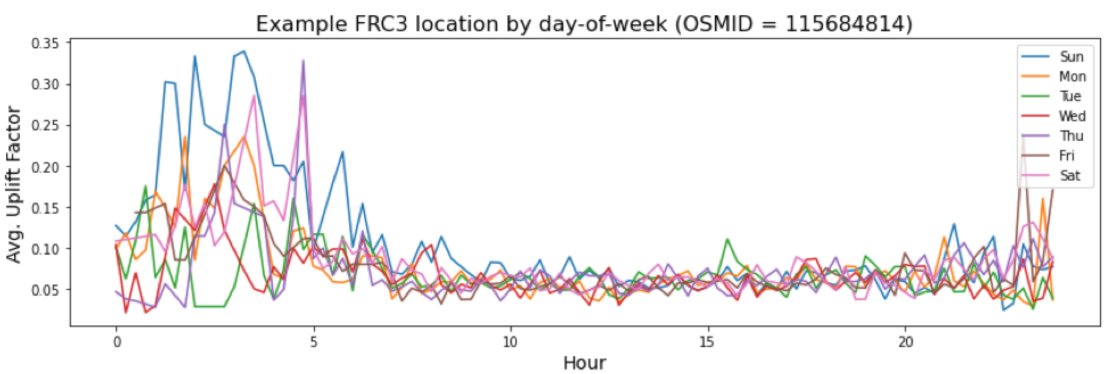
The technical team noted some adjustments that had to be addressed with the input data.

- The real time counts were aggregated into 15-minute windows. The historical data was provided in one hour formats so an additional task was needed to realign the data into those 15 minute intervals
- Some of the continuous count stations reported bidirectional counts (combining counts from both directions of traffic flow), and some reported unidirectional. To match with the directional CV data, an allocation of combined counts was made when necessary.
- When the reference count data was conflated to OSM, there was some error noted in snapping of data during the pilot. The conflation process needs streamlining for future implementation
- Time harmonization is needed when count data and real time waypoint time reference spanned multiple time zones as it did in this POC. Best practice is to use GMT for all calculations, and convert to a local time reference before display.

State	Road Class	All Locations		Count Locations		
		# OSM Segments	Total Miles	# OSM Segments	Total Miles	# Hours of Data
Florida	FRC 1	9,426	4,754	134	547	52,723
	FRC 2	1,432	478	8	7	2,429
	FRC 3	37,606	12,818	210	525	84,356
North Carolina	FRC 1	7,095	4,648	81	267	49,361
	FRC 2	7,517	3,825	24	69	12,472
	FRC 3	12,041	5,582	7	22	3,796
Virginia	FRC 1	5,874	3,339	442	1,090	388,665
	FRC 2	5,588	3,726	109	326	81,472
	FRC 3	10,039	4,704	86	393	51,330

The chart at the right showcases the number of OSM segments and count locations that were used in the study as compared to the overall number of segments and locations for the functional road classes (FR1,2,3) for the three states of Florida, North Carolina and Virginia.





The graphs above present example data from a FRC 1 and FRC 3 segment in Florida. These compare the uplift factors over a 24-hour period for each day of the week. The results point out that there are larger variations during the overnight hours (where there are typically low volumes) but these variations become fairly consistent during the day for all days of the week. This key finding indicates that CV data can serve as a reliable data source to estimate volumes.

The original goal was to deliver a table of uplift factors to MoonShadow that would be applied by segment identification number, time of day in 15 minute intervals, and day of week, and to apply such uplift factors in real time. This means that for each segment of roadway (as designated by the OSM Wayid) there would be 96 uplift factors per day (24 hours * 4 fifteen minute periods/ hour) or 672 per week. This design allowed for consistency in implementation. A table was created to support this process but due to time and overall project complexity it was not implemented in the MoonShadow analytics software. The traffic volumes calculated in the analysis platform, used an uplift table of similar structure, but all uplift factors were based on an assumption of 3% probe penetration rates.

Analysis of the relationships between the Wejo data probe counts at state maintained continuous count stations identified key attributes, such as time of day and road class to determine, to best calculated uplift factors. For example, the analysis indicated that road class was a key differentiating factor, which would warrant different uplift factors, a finding consistent with historical traffic volume estimation research. Through this methodology and application of uplift factors to Wejo CV data probe counts on like roadways, states would be able to observe estimated traffic volumes on roadways that did not have continuous counters. Whereas this effort used primarily three differentiators (road class, time of day, and day week), there is opportunity to bring in other data types like weather conditions or urban versus rural for even greater fidelity.

The fourth and last objective was to show that the mobility patterns of people, (i.e. where they go and when), change significantly in the event of a major storm. Having this ability addresses such questions as 'Are people getting out of the way?' or 'Are they preparing? '(as evidenced by more frequent trips to grocery stores, home improvement store or fueling stations), or 'Are they sheltering place?' Thankfully, there were no major hurricanes during the pilot period. The technical team did however look back at a minor storm that made land fall in Alabama at the onset of the POC. The analysis of the storm as well as overall analysis of the Wejo data showed …

- O&D patterns were clearly visible in the data and typically repeatable every day during normal conditions
- Storm O&D patterns/case studies revealed people getting out of the way and sheltering in place
- In the Alabama case study, measuring roadway traffic across coastal causeways were noticeably different than during a normal day
- The team hypothesized that Trips to home improvement centers (such as Home Depot and Lowes) would escalate in the days prior to a storm event, showing evidence of storm preparedness through changes in mobility patterns. Such behavior was observed for the Alabama case study.
- There is no precedent for Real Time O&D visualizations, so this concept is blazing new ground.
- Future work would continue to look deeper into these patterns

This forth objective received less attention and resources due to expenditure of project resource resources to tackle the challenges of the first three objectives. Even so, the ad-hoc analysis performed in cooperation with MoonShadow showed positive results, but will require additional research and resources to bring to prepare for a full system implementation, thus … 'to be continued.'

STATE OUTREACH AND PARTICIPATION/FEEDBACK

Due to the short term nature of the pilot study, state representatives were enlisted and involved from the outset with a hands-on initiative to gain immediate development feedback and to generally 'kick the tires' on the platform application as it was being integrated with CV live streaming data. Several power users/subject matter experts from participating states were identified and granted licenses to access the MoonShadow DB4IoT analytics platform. Twenty-one (21) licenses were provided to a set of state representative users. One on one training sessions were held with DB4IoT staff. A key finding is that after the individual training, some of the initial license holders transferred their licenses to others they identified with data science expertise, as being a better fit to participate in the platform feedback and development. Typically, one user from each participating state became more of a power user, and provided in depth feedback to the team. Thank you to the users listed below for their honest and detailed feedback. See Appendix for additional state feedback.

Virginia DOT was asked to present its feedback and findings relative to the CV data and the MoonShadow DB4IoT analytics platform during the TETC webinar that presented the findings of the project. The useful features and agency challenges of the POC analytics are noted below

- The volume, freshness and spatial coverage of the data was impressive
- The data set was rich in timestamp/location/route detail
- It is an emerging type of data, and there is evidence that it can provide new insights and meet current needs for the agency (ubiquitous volumes, O&D matrices, vehicle trajectories)
- It is a totally different scale (ubiquitous and expansive) than currently available volume information
- It complements currently available VDOT data
- The data brings a real-time perspective that is valuable for Traffic Operations Centers especially with weather event operations
- CV data provides a capability of observing and understanding more/less frequently used roads
- The data can provide compliance with designated evacuation routes
- There is the possibility of obtaining data from border regions with other states. VDOT is part of the Capital Region so cross border information is critical.
- Real time volumes are a big DOT need. The challenge is to be able to scale up journeys to volumes in an analytics platform.
- There is a need to build trust in this type of data at the DOT level which can be accomplished by validation work and state specific customization.
- With CV probes, there may be bias towards newer vehicles equipped with CV tech.
- Privacy Issues arise when there is accessibility to geodetic data which is often sufficient to identify one's house.
- For Map matching, lat-lon data is not always sufficient to determine direction of travel, so the heading data element is very useful

In addition to providing a thorough review of the data and platform, VDOT also allowed to download one week of historical O&D data to create its own matrix which was set up by county. VA was able to affirm the validity of these numbers by comparing to known agency data.

Unique Journey ID O&D Matrix by District: Virginia, Sep 13 – 19, 2020

Destination District \ Origin District	Bristol	Culpeper	Fredericksburg	Lynchburg	Hampton Roads	Northern Virginia	Richmond	Salem	Staunton	Out of State	UKN
Bristol	49,844	128	15	56	157	26	104	2,735	391	1,848	8,666
Culpeper	133	187,676	4,908	1,370	331	8,912	3,187	343	3,781	137	6,948
Fredericksburg	10	4,894	265,795	165	4,365	9,579	9,324	47	382	596	11,694
Lynchburg	63	1,387	179	144,236	988	110	2,894	6,791	1,240	2,163	4,222
Hampton Roads	298	362	4,644	897	921,233	593	7,646	3,557	521	8,517	22,972
Northern Virginia	11	9,006	9,460	102	449	705,610	1,069	82	3,314	255	70,570
Richmond	133	3,010	9,602	2,910	7,587	1,143	582,944	480	1,070	2,591	13,324
Salem	2,495	314	43	6,784	3,381	109	482	251,139	2,594	1,794	9,494
Staunton	313	3,885	400	1,186	450	3,315	1,151	2,258	246,676	779	19,352
Out of State	2,517	171	803	2,326	9,341	370	2,777	2,167	987	9,489	3,397
UKN	14,751	13,310	20,680	8,670	39,711	121,291	24,340	17,608	33,524	8,970	446,266

In follow up communication with state representatives, there were several consistent and telling comments as to the future benefits of this data. Refer to Appendix A for more detailed feedback from the participating states. A summary of these insights include:

- VDOT's only data gap is volume estimates across the entire roadway network, so VA is very interested in this data and traffic volume estimate methodology
- VDOT recommended that data be segmented and processed in smaller spatial parcels for potential performance enhancement. When creating smaller regions, such as state boundaries, small buffers across state lines are desirable to see approaching traffic issues
- For North Carolina, the ability to capture O&D patterns caught our attention; it is a wonderful source in tracking O&D patterns or potential changes in traffic volumes
- The visualization tool is not intuitive for transportation engineering, requiring significant data skills. More developmental is needed to make it more user friendly and appeal to a broader user group.
- A benefits of this data is that it gives us a different lens to observe network traffic condition, including O&D patterns, volumes, and individual vehicle movements.

One of the overarching goals of this initiative is to prepare for a real time volume estimate reporting tool for major events such as hurricane evacuation. The sample table provided by North Carolina provides an example of useful traffic information needed when reporting real time volumes for major weather/traffic conditions. This is the "holy grail" of real time volumes.

			Counted	Historical	Get from NCDOT	Calculated
Day of Week	Hours	Route	Today's Wejo Volume	Normal Wejo DOW HOD Vol	Normal NCDOT DOW HOW Vol	Estimated Total Volume
Monday	12 noon to 4 pm	I-40	3000	1000	30,000	90,000

State feedback was positive for the most part, especially in regards to the richness of the data. The analytics platform initially had responsiveness issues (due to the high volume of data) but then stabilized as MoonShadow was able to provide software engineering adjustments, and performed well thereafter for the remaining month of the proof of concept period (September through November). The ability to observe the real time trace of individual vehicles, and the relative density of the probe vehicle provided valuable intuitive feedback previously not available. Such visualization reinforced confidence that CV data can be used for routing, confirming contraflow operation, and to detect, confirm, and notify concerning incidents and slowdowns. The pilot provided ample evidence that real time probe data and associated applications such as real time volume estimation are within reach.

In summary,

- CV data is viable now and will only grow in size and velocity; we are blazing new ground to visualize real time volume and O&D
- Managing CV data at scale is challenging for the industry; sheer size and velocity of the data will require efficient processing architecture, as well as streamlined calibration, calculation and conflation techniques
- Visualization of CV data (individual vehicles) brings intuitive value
- A recommended change for future projects would be to segregate the system spatially by state, but still have state overlap with buffer zones for continuity. This would both reduce the latency and increase the system responsiveness.
- Simple scaling factors based on roadway class, time-of-day, day-of-week, and state were demonstrated to provide a workable and sufficiently accurate traffic volume estimate. Including additional factors, and leveraging more advanced machine learning may improve accuracy.
- The pilot clearly points to a viable path toward real-time traffic volume estimates based on CV data
- The pilot provided valuable insight for the next steps moving forward such as -
- A need to develop specifications for operational systems in terms of use and functionality
- A need to bridge the language divide between Traffic Engineering, Data Science and the Information Technology industry to bring about a responsive and easy to use interface

OVERALL – the POC indicated that real-time CV data is sufficient to provide workable real time Volume Estimates!

Final Quote from GDOT: "If only all research could be done this way"

FUTURE PLANS

It was the intention of the TETC to use the results from this POC as the basis for future plans concerning a real-time traffic volume service for its members. With the encouraging results and the lesson learned, the TETC is planning for an initiative to begin later in 2021. The initiative will outline the framework for a system that can:

- Ingest connected vehicle data
- Organize it appropriately via jurisdictional boundaries
- Marry the data with roadway network information
- Aggregate the data with respect to probe vehicle counts, O&D and associated land use, as well as traditional speed and volume estimate
- Display real time raw CV data in a responsive, intuitive and easy to use interface
- Calibrate and validate uplift factors for accurate traffic volume estimates
- Display traffic volumes appropriate to operations purposes

TETC intends to convene a steering committee as well as seek member support in 2021. CV data is included in the TETC Mobility Data Marketplace as a foundation from which to developed a traffic volume estimate resource.

The Coalition has a long history of collaborative research between TETC and the University of Maryland (UMD) and the National Renewable Energy Laboratory (NREL) in looking at how to derive accurate volume data from vehicle probe data. Several past studies working with the states of Florida, Maryland, Colorado, New Hampshire, Pennsylvania, Massachusetts, Tennessee – all with the objective to validate historical volume estimates from various types of probe data have enhanced the foundational knowledge in this area. The research and collaboration is coming to fruition both in historical volume estimates procured through the Mobility Data Marketplace, and valuable lessons learned in application to the hurricane real time POC, and subsequent development efforts moving forward.

The TETC would like to thank its partners (Wejo/MoonShadow, Traffax, NREL) and in particular the state member representatives that provided invaluable feedback during the proof of concept. The TETC looks forward to the mutual collaboration and development effort that will result in valuable tools, data and information for TETC members to apply in their duties to operate, manage and plan for an effective transportation system.

Appendix A

State Interview Notes – License holders

ALABAMA

Primary Beta User – Alex Hainen, University of Alabama

Alex Hainen:

- Performance a challenge – the platform will work on Wander's computer but not on others at times.
- Ultimately, we are looking for volumes
- There is more work to be done to create a scaling/uplift factor that is needed for a volume estimation - Stan Young knows what he is doing here
- Alex is being the beta tester for feedback for Alabama
- We must realize that the DB platform may be the only one out there but it still needs development
- The Wejo data is awesome – but we want more reliable performance of the live data platform
- It may be that the quantity of data is overwhelming the system and that maybe reducing the dataset will help performance
- Alex has seen the moving data but now needs to look for a range and must do so in MoonShadow but is not sure he can do that yet – needs the scaling factor.
- Agrees that it is best to term as a beta test
- Requesting some sample of the raw data not inside the MoonShadow platform. (Wejo can obfuscate the vehicle ID or trip ID) would greatly enhance the evaluation for Alabama.

Chris Hilyer:

- The volumes over time did not seem much different, but that could be a sample size issue;
- The speed is unclear as to mph or kph, some of these vehicles are moving extremely fast based on the speed key;
- Often the dots do not align with the roadway, so it is difficult to pin down any affects or contributing factors

FLORIDA

Raj Ponnaluri (Connected Vehicle and Arterial Management Manager)

- Florida has had previous conversations with Wejo (about 8 months ago) but was unable to acquire much data due to cost and the small amount that was made available to them was not enough to truly evaluate.
- With this data, they could see a snapshot of the data and a range of dates

FLORIDA CONTINUED

- Data analysis for O&D could be used for planning
- They could not see any real-time use for it did not get into roadway volume (segment based) –it was only journey id's.
- Also it was a holistic system (aggregated data) not raw data
- Directional delay was not available
- The visualization tool is not intuitive for a transportation use
- Using kph and not mph for the speed legend
- Questioned how the tool will get us to a goal of improving Safety and Mobility
- Volume were not easy to find
- Roadways are not searchable
- Travel time delay was not there
- FDOT did not differentiate between Wejo and MoonShadow – they always referenced as one unit (Wejo) – This is an important point

GEORGIA

Tom McQueen (Asst Director of Planning):

- There is a lot of interesting data – we are seeing how we can utilize the data
- We are doing a lot of brainstorming- very data focused
- Can we use this dataset to correlate to other datasets?
- Benefits of this data is that it gives us a different lens to look at data other than O&D data
- Data distribution pattern is different than average distribution (i.e. seeing newer cars and younger drivers)
- It is almost like there is a little bit of bias introduced toward newer cars
- Data can help tell the story; we can associate with different users – call it driver distribution
- Maybe helpful when looking at air quality & emissions
- Maybe data could be used for modeling; Note similarities w/Synchro & SimTraffic (macro vs micro) modeling
- Maybe can use the CV data for identifying bottlenecks
- Could do historical analyses
- Can we leverage the data trends to understand where we should invest?
- This could help with any ROI's we do
- If we can see historical trends and they are growing tremendously, this will give us a lot of value
- We should spend more time on how to use as it is a technical platform - application
- Lastly, suggested what they called a virtual improv – ask people suggestions online and chat with them online – asking what worked and what did not work

GEORGIA CONTINUED

Brad Mann (Special Studies Engineer)

- Initially had trouble seeing the data
- Legend is kph and not mph
- Performance was originally slow – originally thought I was doing something wrong
- Data downloads go into a zip file and had to install an app to export the data into notepad
- Used a spreadsheet to look at data – better to be able to export data directly into excel!!!
- When pulling data from two roads, hard to tell which road is which
- What does "event" mean?

NORTH CAROLINA

Dominic Ciaramitaro
Kelly Wells
Nathan Webster

- Performance of the MoonShadow platform is clunky– they are working to improve the performance — considering reducing the size of the data they are integrating to help performance
- Having real time volume data to report on during hurricane events is the ultimate goal. See sample spreadsheet provided by Kelly.
- Requesting to change from kph to mph – English units (addressed)
- Having historical data during a period of before and after Hurricane Isaias (August 4, 2020) would be more beneficial to the states of Georgia, NC and Virginia.
- Per Kelly, the key for us is having that info and having it quickly

Sample Table from NC of info request

			Counted	Historical	Get from NCDOT	Calculated
Day of Week	Hours	Route	Today's Wejo Volume	Normal Wejo DOW HO&D Vol	Normal NCDOT DOW HOW Vol	Estimated Total Volume
Monday	12 noon to 4 pm	I-40	3000	1000	30,000	90,000

NORTH CAROLINA CONTINUED

Researcher:
Dr. Srivinas Pulugurtha

- It is a wonderful set of real time data (connected vehicle, GPS-based)
- No medium or heavy duty vehicles
- There was not much historical data – maybe a 24 hour look back feature
- The data seems to update every 3 seconds
- The data resolution was very good
- One could get data about each connected vehicle (for example, spacing indicates how fast the vehicle is moving) – it was not just a summary
- Ability to capture O&D patterns caught my attention; wonderful source in tracking O&D patterns or potential changes in traffic volumes
- Using Wejo data to integrate with loop data or other software such as Vissim seems doable
- Wish there could be access to non-cv vehicles as well as cv – it would be good to know how those vehicles compare to Wejo data
- Could we look at specific days and evaluate how many vehicles have been detected over a given period of time?
- Compare Wejo data with other probe data (connected vehicle compared to traffic stream performance)?

TENNESSEE

David Lee
Michelle Nickerson

- Minimal use of licenses – no feedback to generate
- Difficulty in getting access – system would not work – can't get into the site
- Discussing transferring license to a researcher for better analysis (like NC and Alabama)

VIRGINIA

VDOT evaluation following page – very complete and thorough analysis of MoonShadow functionality

- Recommended for potential performance enhancement to create smaller regions with small buffers across state lines such as what Inrix does for VA, DC and MD
- County level is a big deal for VDOT
- Not liking the OSIM map
- Speed legend is backward for coloring; we don't need "blue coloring"; need the English system (addressed)
- VDOT's only data gap is volume estimates so very interested in this data (Bingo!)
- Connection makes a difference in platform responsiveness