Final Project Presentation

Current Status of Transportation Data Analytics and Pilot Case Studies Using Artificial Intelligence (AI)

Principal Investigators

Yuanchang Xie, University of Massachusetts (UMass) Lowell

Danjue Chen, UMass Lowell

Tingjian Ge, UMass Lowell

Ali Shirazi, University of Maine (UMaine)

February 20, 2024

OUTLINE

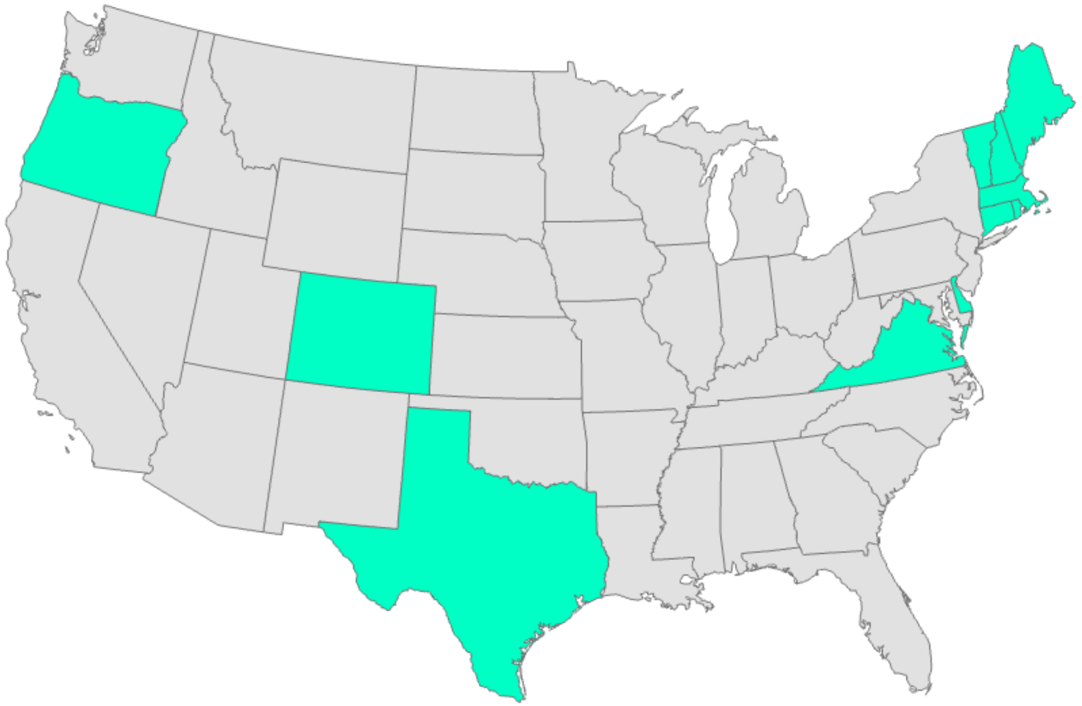
- **Phase I:**
  - Review and summarize the current data practices related to TSMO at New England state DOTs
  - Analyze the current and future data needs and sources
  - Provide recommendations to improve such practices

- **Phase II:**
  - Demonstrate the power of utilizing advanced data analytics tools to make the best use of existing data owned by New England state DOTs
  - Summarize and document the entire study

PHASE I

METHODOLOGIES

- Review of literature
- Interviews via email/zoom
  - FHWA
  - The Eastern Transportation Coalition
  - Delaware
  - Virginia
  - Texas
  - Oregon
  - Colorado
  - Six New England states



SUMMARY OF DATA AND SOURCES

<table>
<tr><th></th><th>Traditional Data & Data Source</th><th>New and Emerging Data and Data Source</th></tr>
<tr><td>Highway</td><td>Loop detectors, microwave detectors, traffic cameras, Bluetooth/Wi-Fi MAC address readers, weather stations, weigh-in-motion stations
Occupancy, delay and travel time, spot and segment speeds, volume, vehicular OD</td><td rowspan="7">Drone, Mobile LiDAR

Crowdsourced Data (e.g., Waze)

Fleet data (DOT vehicles, commercial vehicles)

Transportation Network and Logistics Companies (e.g., Uber Movement)

Connected Vehicle Pilot Deployment Program

TomTom, HERE, WeJo, StreetLight, INRIX, AirSage, SkyHook, Cuebiq, SafeGraph, Google, Apple

Mobile Carrier (e.g., AT&T, Verizon)

Cell tower triangulation, Cell Phone (or vehicle) GPS

Social Media (e.g., X, formerly Twitter, Facebook)

Data from Autonomous Vehicles (e.g., Lyft, Waymo)</td></tr>
<tr><td>Incidents and Crashes</td><td>Incident/crash records (e.g., location, time, duration), highway safety patrol records, 511 phone records</td></tr>
<tr><td>Arterial</td><td>Traffic signals, vehicle detectors, cameras, data from Automated Traffic Signal Performance Measures (ATSPM) system, queue length from drone.</td></tr>
<tr><td>Transit</td><td>GTFS, transit fare collection data (e.g., smart card, Mobile ticket), traffic cameras, APC data, ridership, etc.</td></tr>
<tr><td>Parking</td><td>Static (e.g., location and # of lots) and dynamic data (e.g., parking duration), parking fee data, Mobile parking app data</td></tr>
<tr><td>Assets</td><td>Highway: conditions of traffic sign, pavement, marking, guardrail, bridges, tunnels, etc.
ITS: conditions of variable message signs, sensors, communication devices, traffic controllers, etc.
GIS maps (e.g., highway geometry), speed limits</td></tr>
<tr><td>Maintenance & Work Zone</td><td>Maintenance: real-time locations and speeds of plow trucks, National Weather Service Data
Work Zone: smart work zone data, location, duration, configuration, etc.</td></tr>
</table>

RECOMMENDATIONS ON DATA NEEDS

ID	Data Needs	Recommendations
1	• Incident detection; Traveler Information Systems (TIS); Travel time estimation	The existing probe data (e.g., TomTom, INRIX) in general provides a good coverage of highways.
2	• Vehicle trajectories	Vehicle trajectories from connected vehicles (e.g., Wejo, Otonomo) cover a large area but only a small sample of all vehicles. Roadside sensors (e.g., high-resolution Radar, camera, LiDAR) cover a short road segment but can capture all passing vehicles.
3	• Passenger and freight OD	Data from mobile device GPS (e.g., location-based service data) and various vehicle ReID technologies make it possible to derive traffic OD for a large geographic area.
4	• Traffic volume and capacity	DOTs should expand the station network using roadside sensors. Such sensors may also be used to provide vehicle trajectory data for safety analysis, vehicle OD, and detailed vehicle classification data (see below).
5	• Detailed vehicle classification and ReID data	AI technologies make it possible to detect, track, and classify vehicles reliably from RGB camera, thermal camera, Radar, LiDAR, and traditional loop detectors. This can generate commodity type data such as flatbed, dry goods semitrailer, tankers, refrigerated trucks, and recreational vehicles.
6	• Travel time	Consider installing Bluetooth sniffers/readers to collect travel time data (e.g., read E-ZPass transponders).
7	• Corridor freight data	Parking information along major corridors such as I-95 is important for truck drivers. DOTs may use camera + AI + edge computing + 4G technologies to collect and share such information.
8	• ITS asset condition data	Detailed and real-time condition information about ITS assets is critical. This is especially true for traffic controllers (e.g., ATSPM) and ITS assets that provide real-time traffic data.

RECOMMENDATIONS ON EMERGING DATA SOURCES

ID	Emerging Data Sources	Recommendations
9	• Connected vehicles and travelers	Connected vehicle datasets, such as Wejo and Otonomo, include vehicle trajectories as well as event logs (e.g., wiper speed and activation/deactivation). Useful information can be derived from such data sources, including OD, route and mode choice, driver behavior, and safety issues associated with highway geometric designs.
10	• Sensors powered by AI and edge computing: thermal and RGB cameras, loop detectors, LiDAR, Radar, E-ZPass transponder	Thermal and RGB cameras can detect, track, and classify vehicles, pedestrians, and bicycles. They can detect lane changing activities, vehicles stopped in the emergency lane, bus lane violations, reidentify vehicles at different locations, etc. High-resolution LiDAR and radar can generate more accurate vehicle speed and location information than cameras and cover larger areas. Vehicle signatures from retrofitted loop detectors can be used to classify and reidentify vehicles. DOTs are encouraged to explore the potential of traditional and new sensors mounted on portable platforms. These portable platforms can (1) collect trajectory data for safety studies, and (2) collect speed and travel time data.
11	• Automated vehicle data	Car manufacturers such as Tesla are collecting a vast amount of data (e.g., videos, vehicle control parameters) from vehicle owners. The data covers driver behavior and the surrounding environment. For example, Tesla uses such data to calculate safety scores for drivers. Such data can also be used to detect road debris, pavement cracks, pavement marking conditions, damaged traffic signs, problematic highway geometric designs, etc.

RECOMMENDATIONS ON DATA PROCESSING AND ANALYTICS

ID	Data Processing and Analytics	Recommendations
12	• Data quality validation	Monitor the quality of probe and connected vehicle data, particularly for rural areas where the penetration rates might be low.
13	• Data integration and conflation	Data conflation is a major issue faced by many DOTs and should be given enough attention.
14	• Detailed incident data analysis	Duration, queue length, clearance time, and effects on secondary incidents
15	• Connected vehicle data analysis	Many auto makers have already been collecting data using their new cars. Such datasets can be utilized to estimate crash risk and identify safety issues due to inappropriate highway geometric designs.
16	• Effective utilization of existing data	Existing datasets are not effectively utilized or explored. Data from loop detectors are often not streamed to highway operations center in real time. Traffic cameras are only used for incident verification and traffic videos are reviewed manually.
17	• ATSPM data analysis	ATSPM systems generate high-resolution (e.g., every 1 second) detector and signal controller data (e.g., detector on/off, green light on). How to effectively utilize such data beyond calculating signal performance measures is a very interesting question.
18	• Innovative data analysis methods	Emerging data sources such as probe vehicles, connected vehicles, and ATSPM require innovative data analysis methods.
19	• Data sharing and brainstorming	DOTs are encouraged to share data with the public when applicable. This may help to generate new application ideas. For example, MBTA makes real-time GTFS data public, based on which many mobile Apps have been developed without costing MBTA anything.
20	• AI + Edge computing for data analysis and reduction	DOTs are encouraged to explore AI and edge computing technologies to speed up the processing of images and videos. This will reduce the amount of data that needs to be transferred and stored.
21	• Road Weather Information	More still needs to be done to analyze the collected road weather data. For example, such data can be used to estimate the optimal amount of deicing materials to be applied.

OTHER RECOMMENDATIONS

ID	Others	Recommendations
22	• Collaboration among DOTs	Leaders from TSMO divisions get together regularly to share best practices, experience, and issues encountered.
23	• Organizational changes	Have a central office to handle data related issues. This will allow data analytics to be done more efficiently and professionally (in terms of data safety, retention, sharing, etc.).
24	• Data storage and sharing among different DOT divisions	Move data to the cloud when applicable, which will make it easy to share data and help to ensure data safety, security, privacy, and integrity. Promote and facilitate data sharing among different divisions of DOTs and different agencies (e.g., Transit vs. Highway; Turnpike vs. TSMO).
25	• Workforce	Many DOTs are creating data scientist/analyst positions and they are encouraged to continue doing this as needed. It is important for DOTs to understand what is being done by private companies.
26	• Personalized TIS	In the future, personalized data sharing with travelers would be important. Variable message signs most likely will be phased out. Instead, DOTs need to provide dynamic traffic information in digital formats that can be easily and precisely interpreted by CAV.
27	• Drone as a data collection platform	Investigate the potential of AI + drones (e.g., drone-in-a-box solution) for post-disaster roadway condition assessment.
28	• Relying on data vendor vs. investing in data collection infrastructure	Compare the life-cycle costs of relying on data vendors and their own data collection infrastructure. Invest in portable data collection units (similar to portable variable message signs) for areas that are not well covered by probe data. These portable units can also be used to collect trajectory data for safety studies. Invest in retrofitting existing traffic cameras and loop detectors using AI and edge computing technologies to expand the capacities of these traditional sensors. Develop data and communication interface standards for vendors.

PHASE II

CASE STUDIES

C1 - Speed Behavior on Highway Horizontal Curves

C2 - Speed and Lane Changing Behavior
Prior to Highway Work Zone

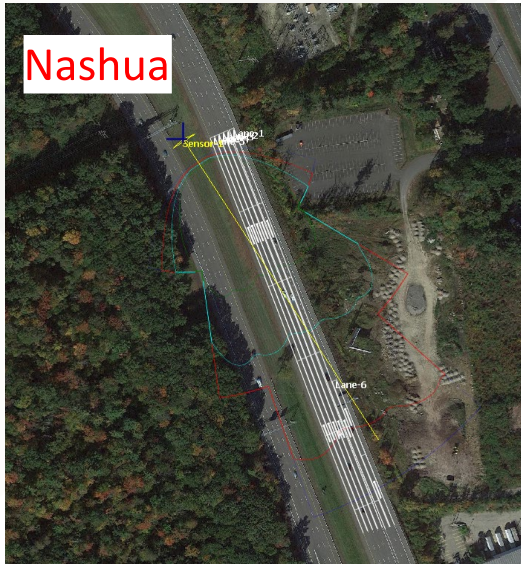
C3 - Network-Wide Speeding Activity Analysis
Using Probe Vehicle Data

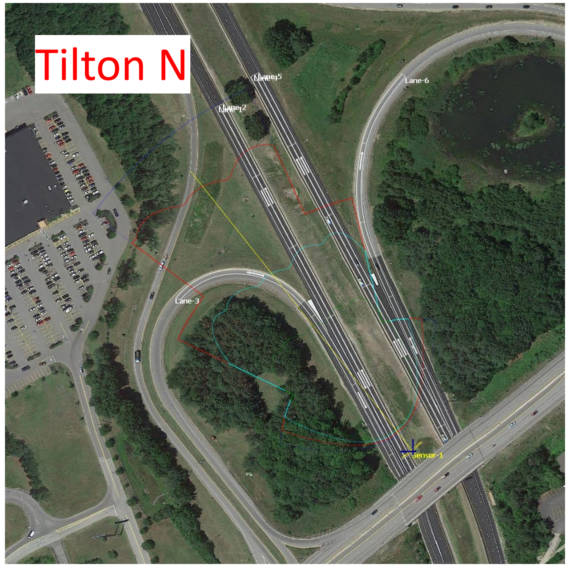
Portable sensors, trajectories, and AI (Computer Vision) at individual sites for safety and operations: C1 and C2

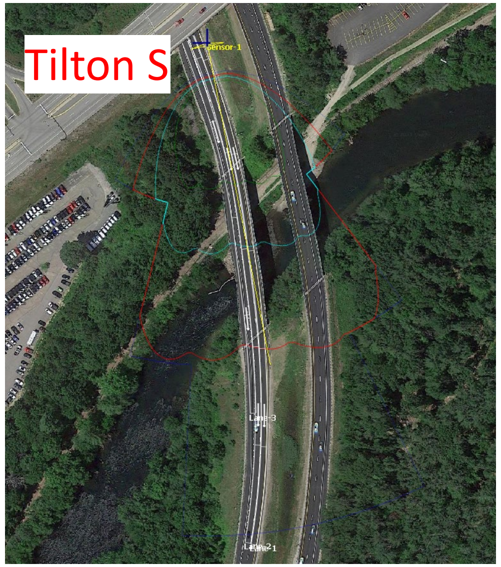
Network-wide crowdsourced (connected vehicle) data for safety and site screening: C3

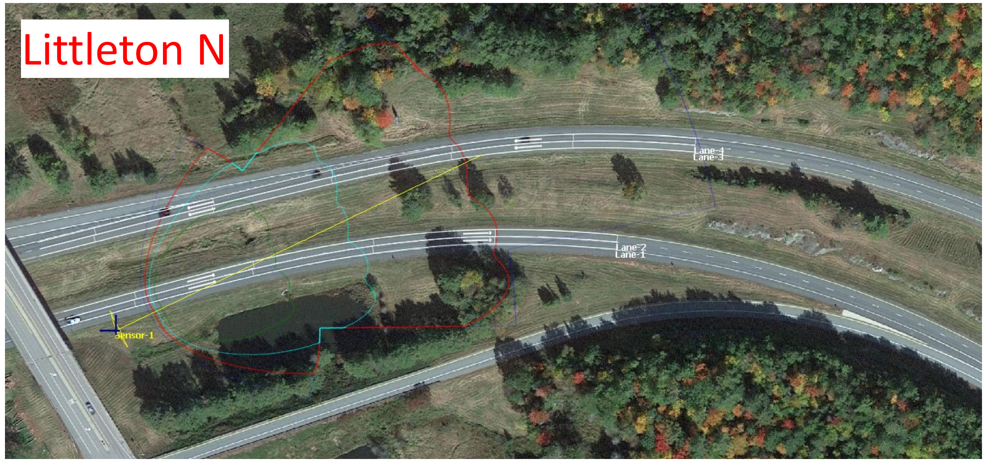
Image generated by ChatGPT4

C1 - SPEED BEHAVIOR ON HIGHWAY HORIZONTAL CURVES





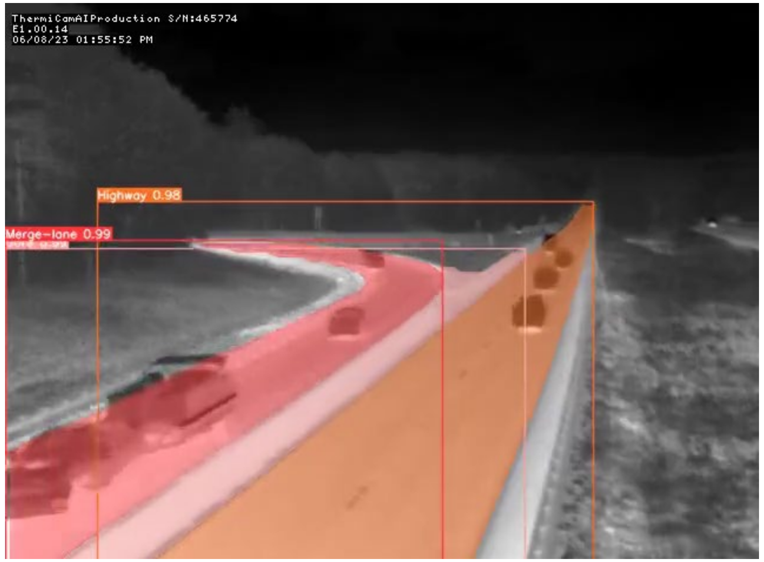




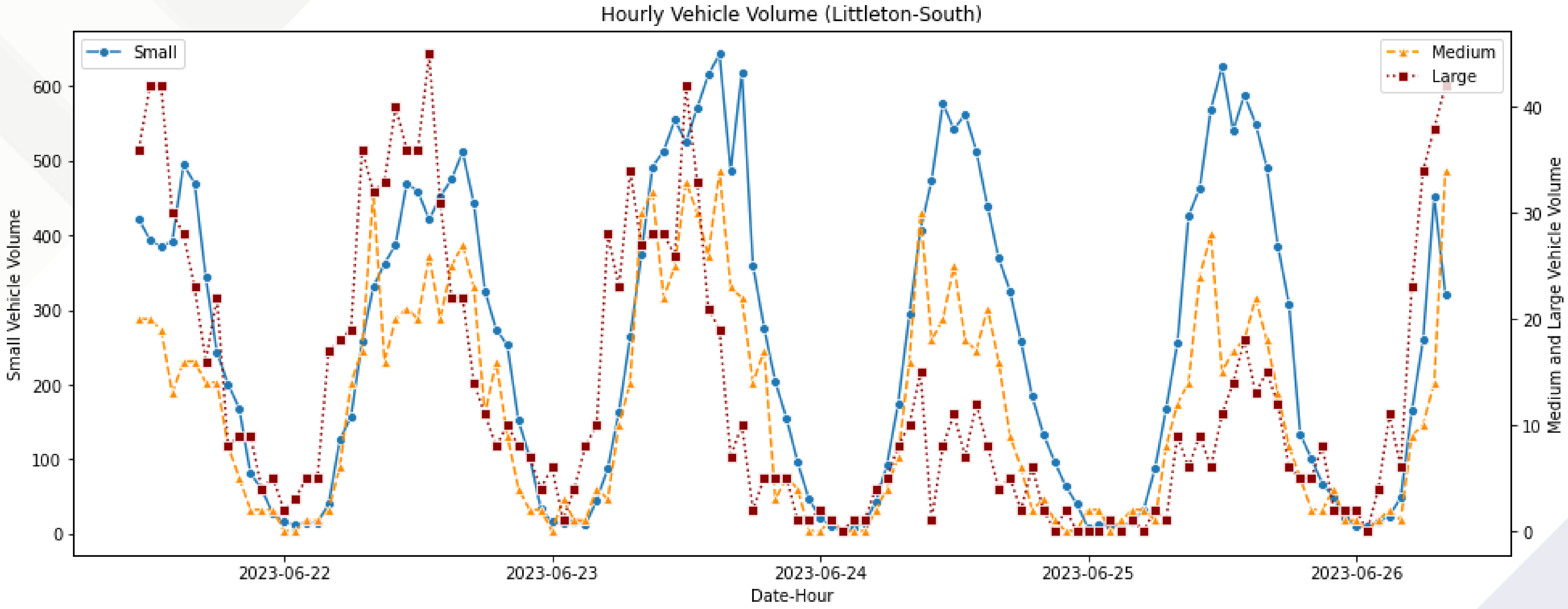
Site #	Coordinate	Site Name	Start Date	End Date
1	42.7316584, -71.4535208	Nashua	4/11/23	4/16/23
2	43.4534073, -71.5710513	Tilton North	6/8/23	6/17/23
3	43.4529169, -71.5707403	Tilton South	6/8/23	6/17/23
4	44.3247220, -71.8052780	Littleton North	6/21/23	6/26/23
5	44.3064868, -71.7982047	Littleton South	6/21/23	6/26/23

DATA COLLECTION AND ANALYSIS

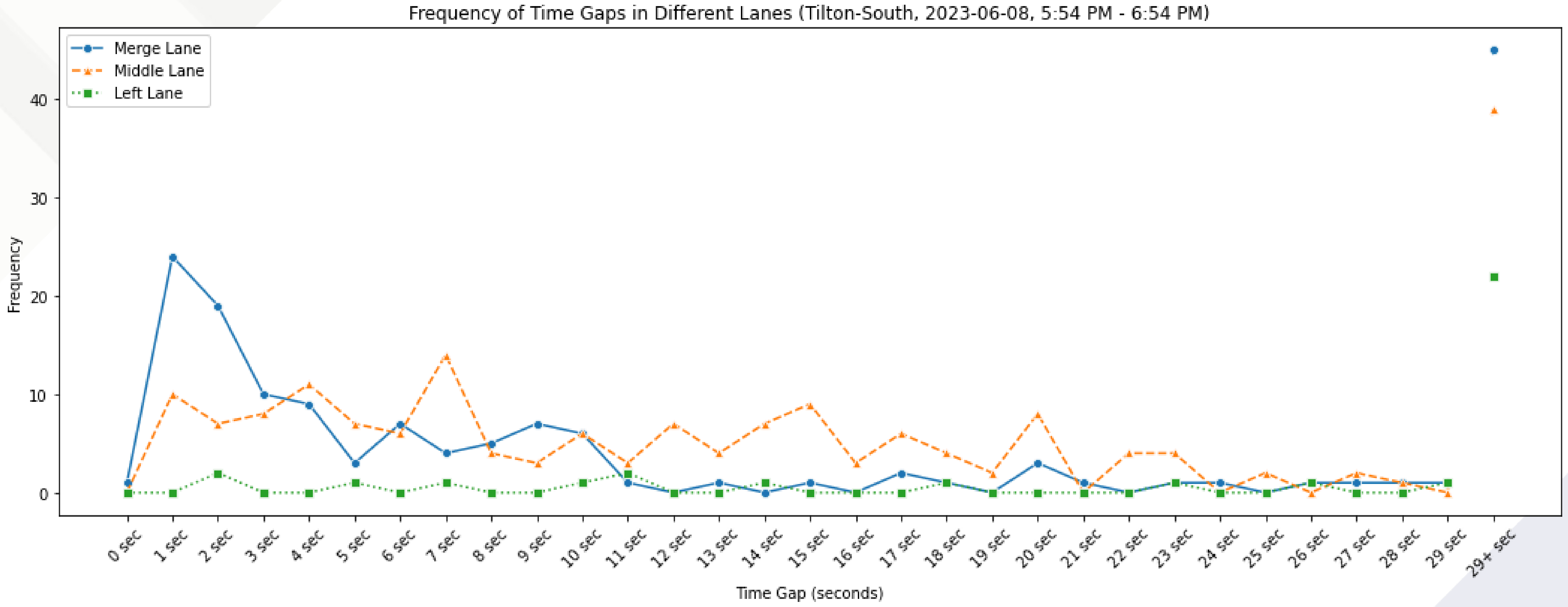
- Thermal camera to capture traffic videos
  - YOLOv8 and ByteTrack models to generate vehicle trajectories
  - Trained a dedicated model to recognize highway features
- Advanced radar to track individual vehicles and generate speeds and trajectories



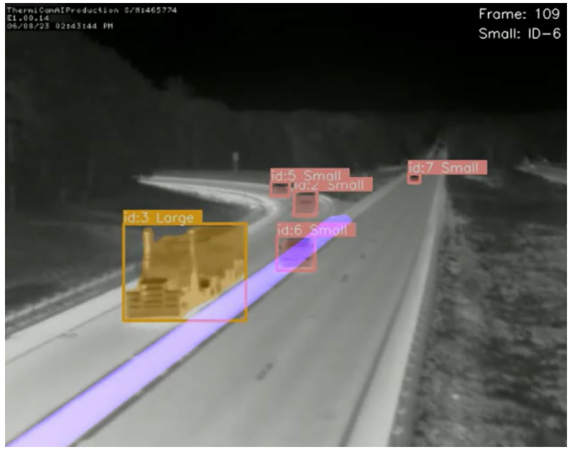
VEHICLE COUNTS

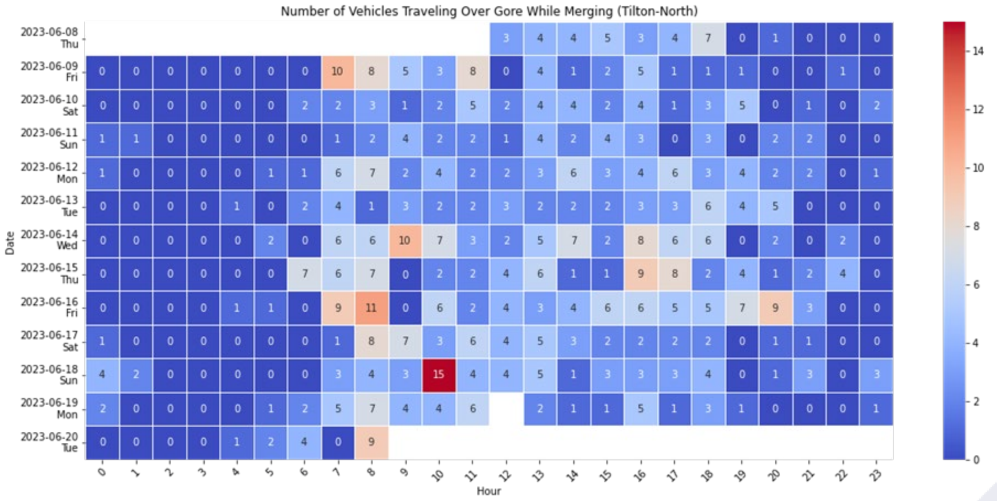


VEHICLE TIME HEADWAY DISTRIBUTION

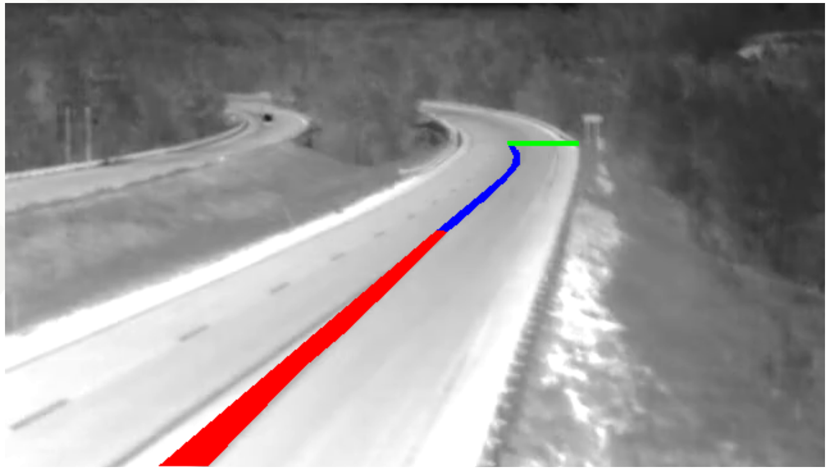


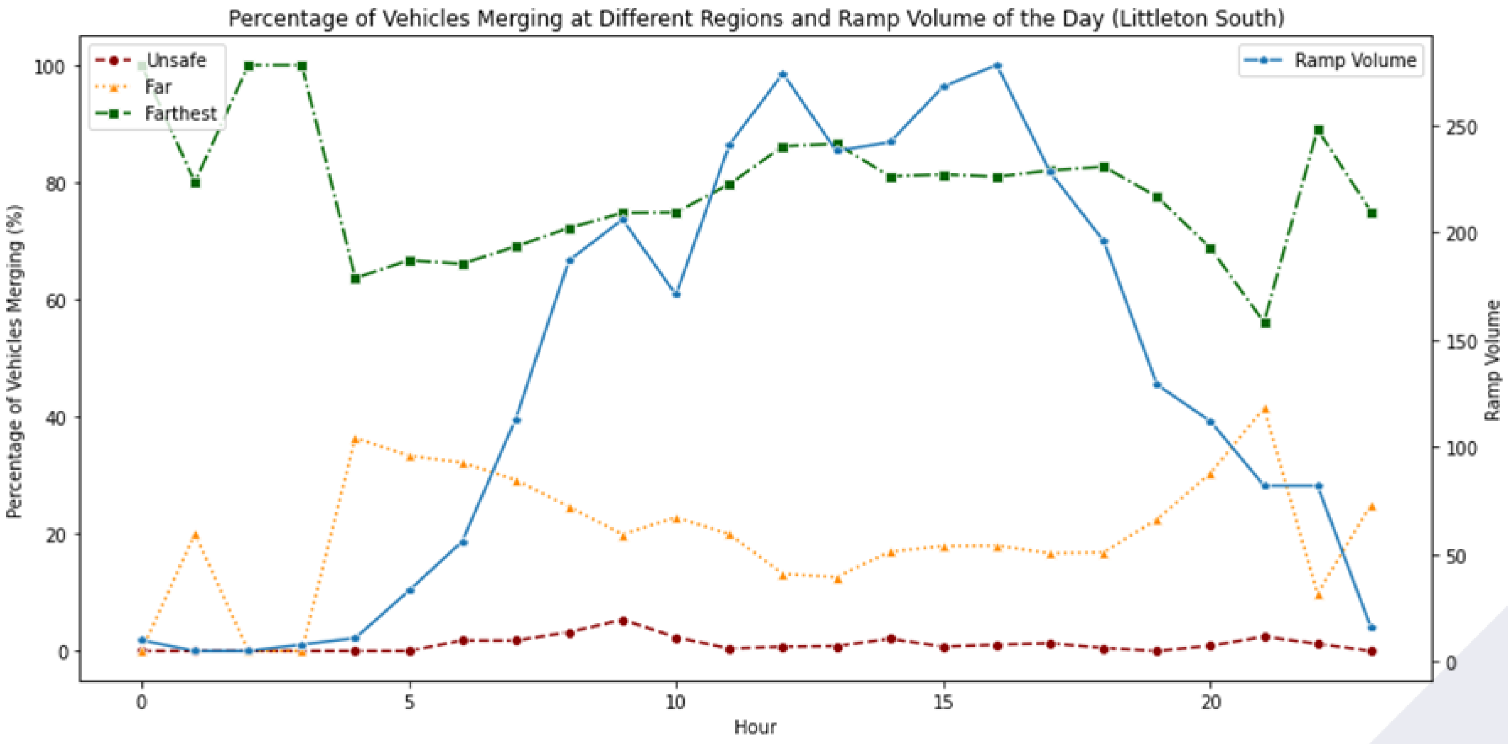
RISKY LANE CHANGES



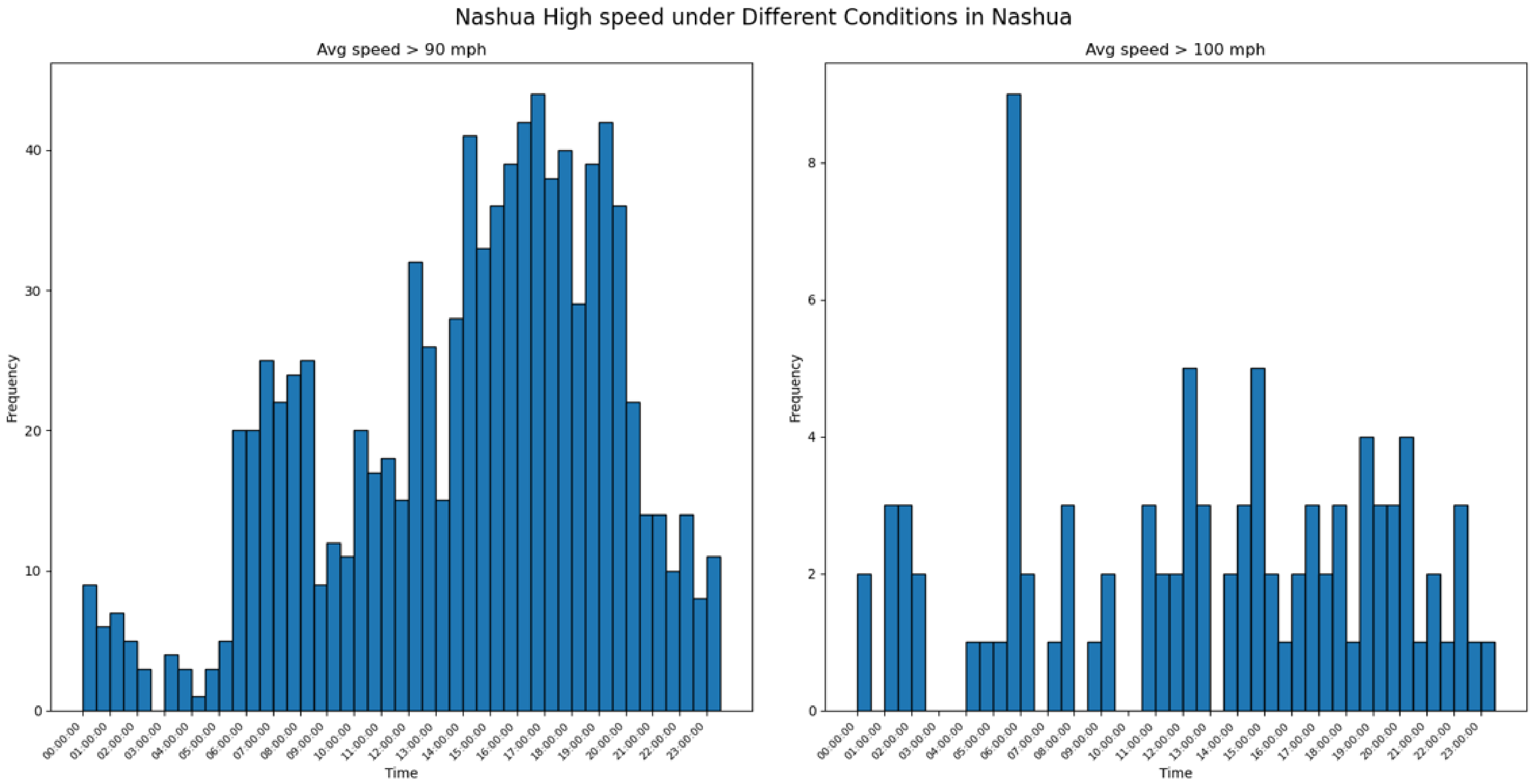


MERGING POINT ANALYSIS

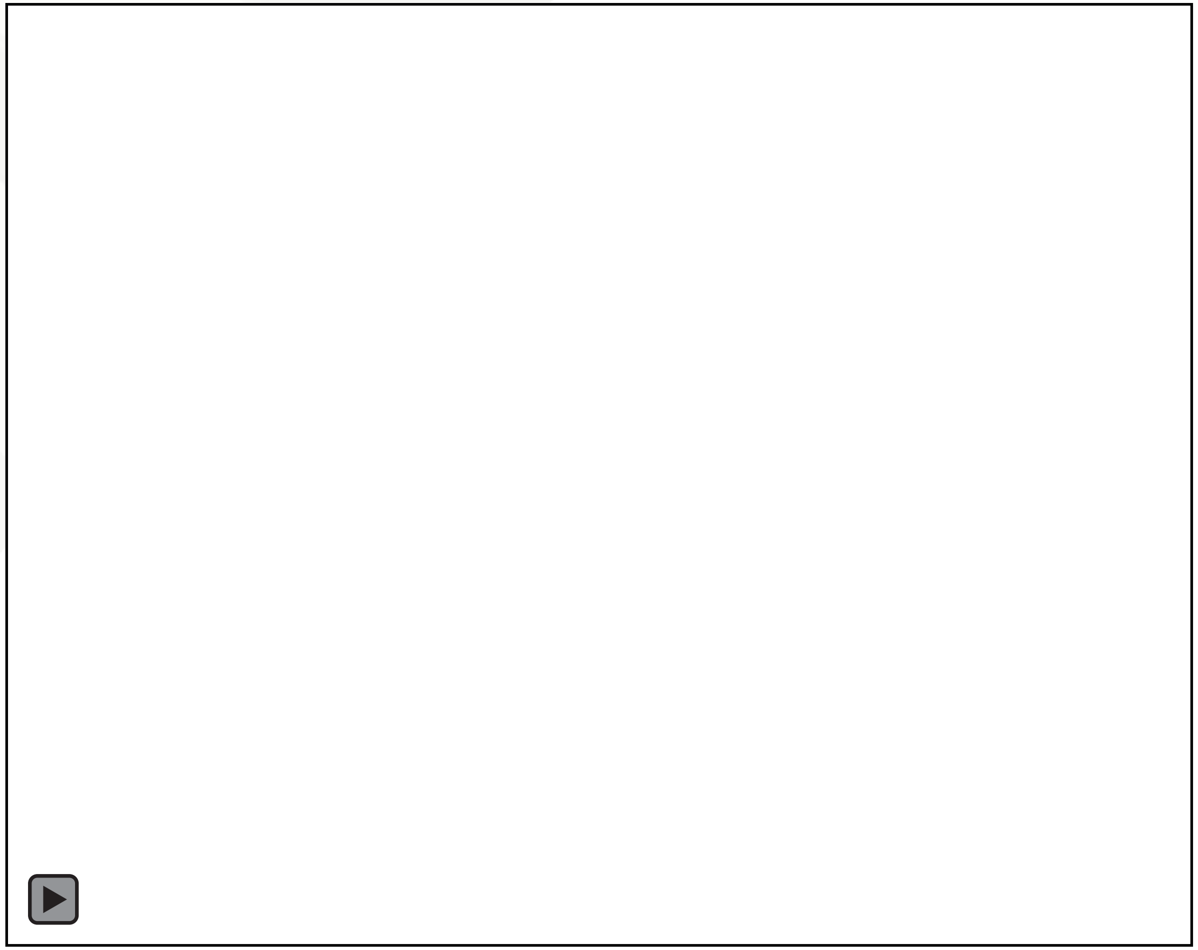


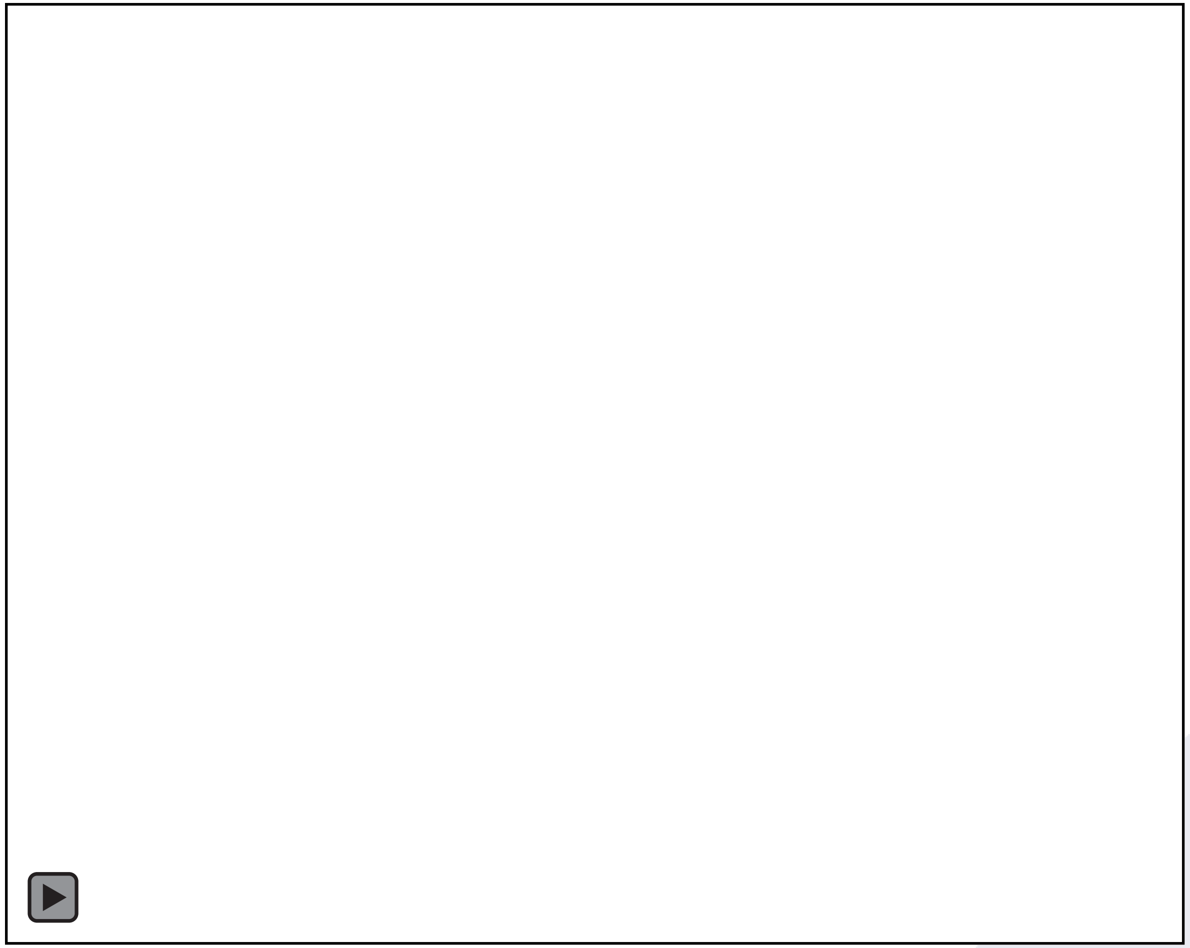


SPEED ANALYSIS

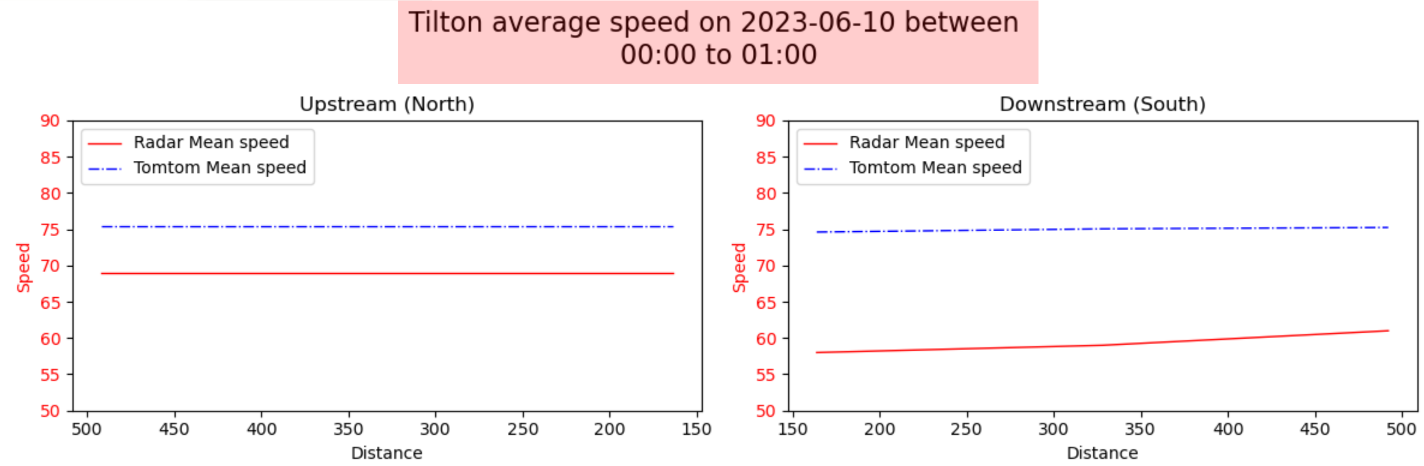


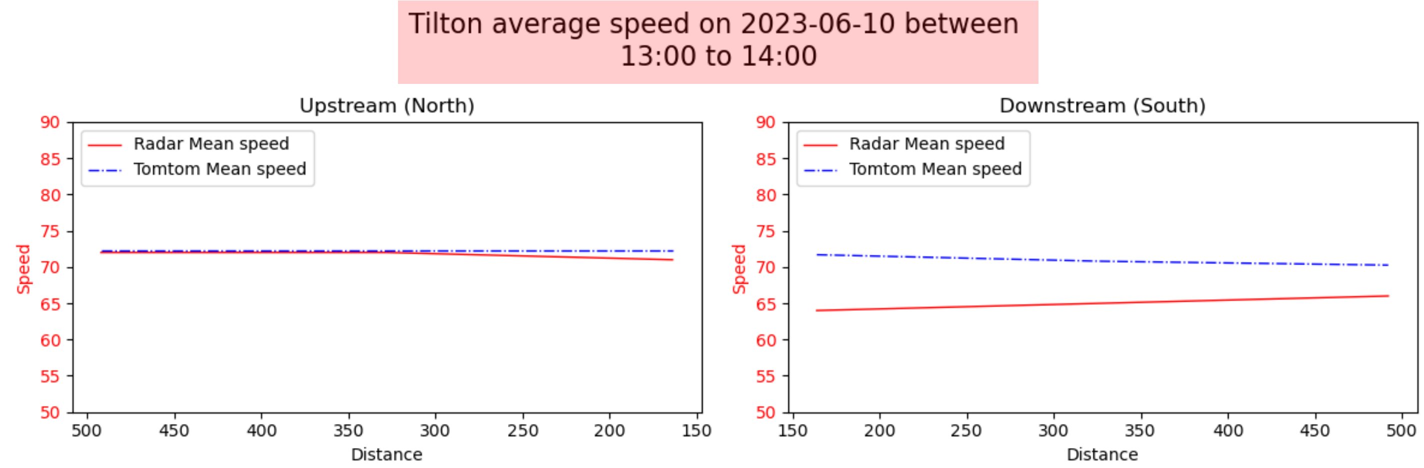
SPEED ANALYSIS (CONT.)

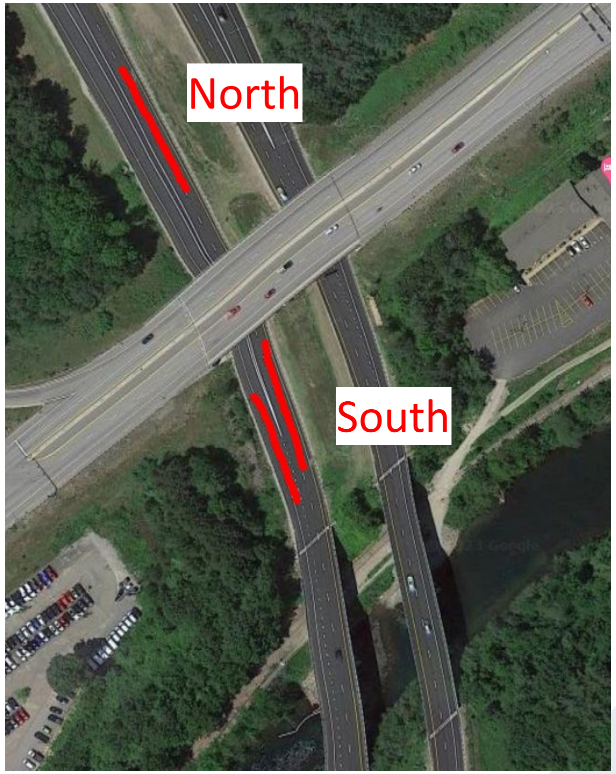




COMPARISON OF RADAR AND TOMTOM SPEED DATA



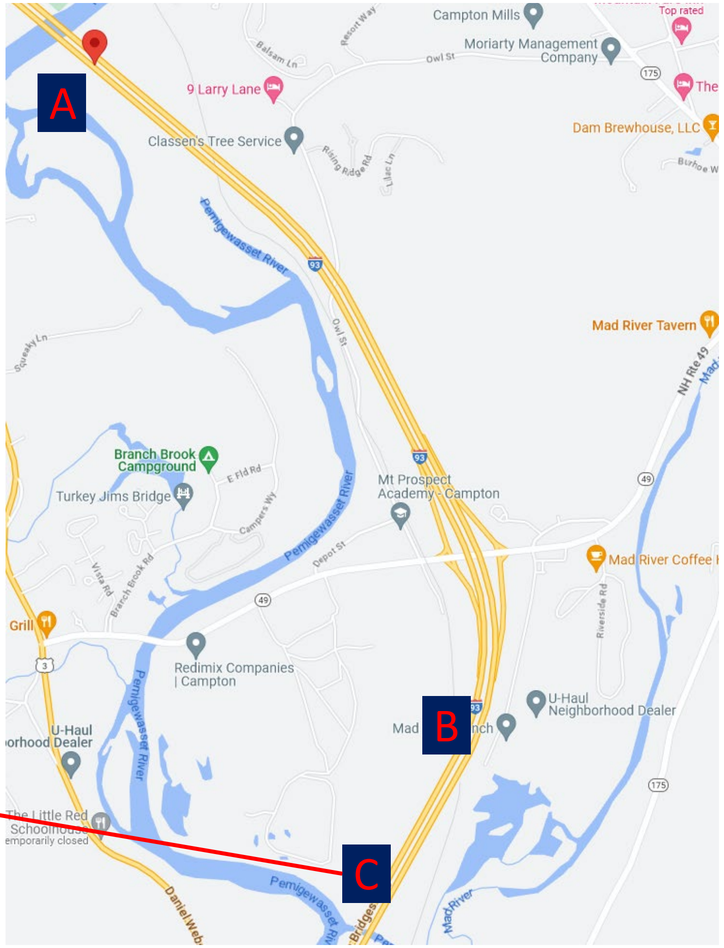




C2 - SPEED AND LANE CHANGING BEHAVIOR PRIOR TO HIGHWAY WORK ZONE

WORK ZONE IN CAMPTON NH

Location: 43.841677, -71.648858



- A: One portable changeable message sign (PCMB) in the median showing

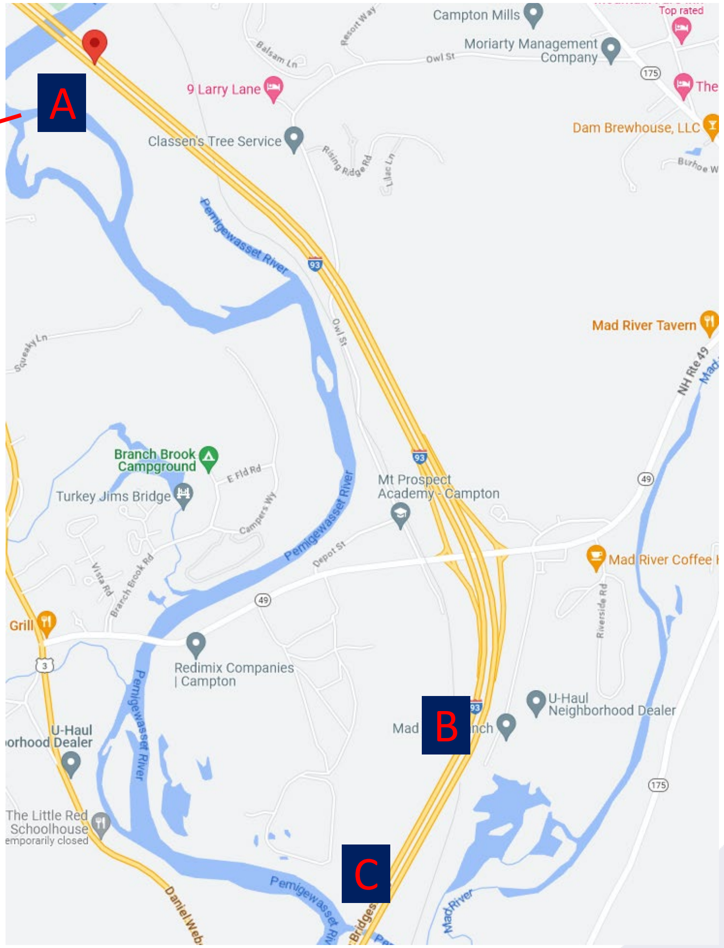
1	2
LEFT LANE CLOSED	POSSIBLE SLOW OR STOPPED
MM 86.4 MERGE EARLY	TRAFFIC AHEAD BE AWARE

- B: A flashing speed limit sign on each side of the road
- 2 miles between A and B
- 0.2 miles between B and C

WORK ZONE CONTROL PLAN

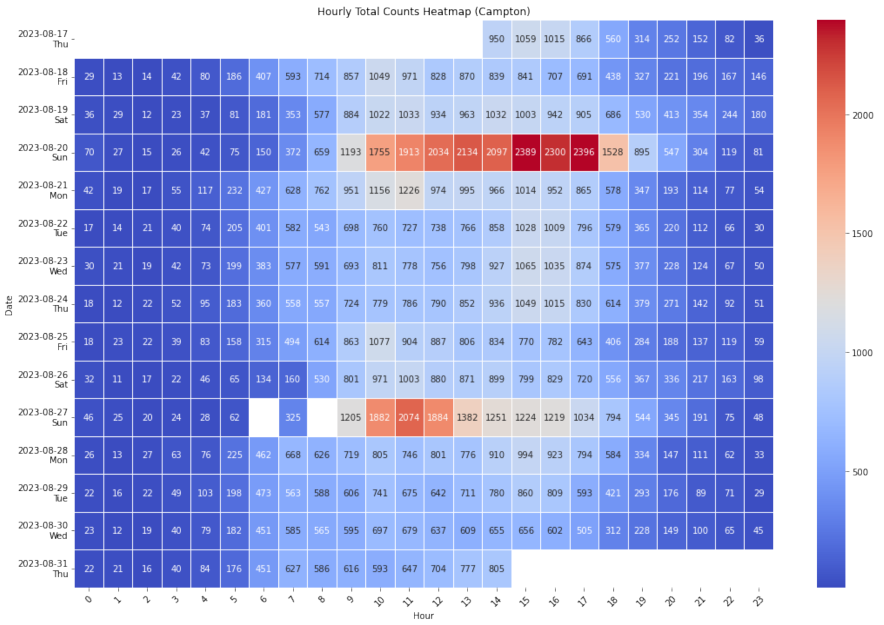
Table 5-1. Work Zone Control Strategies

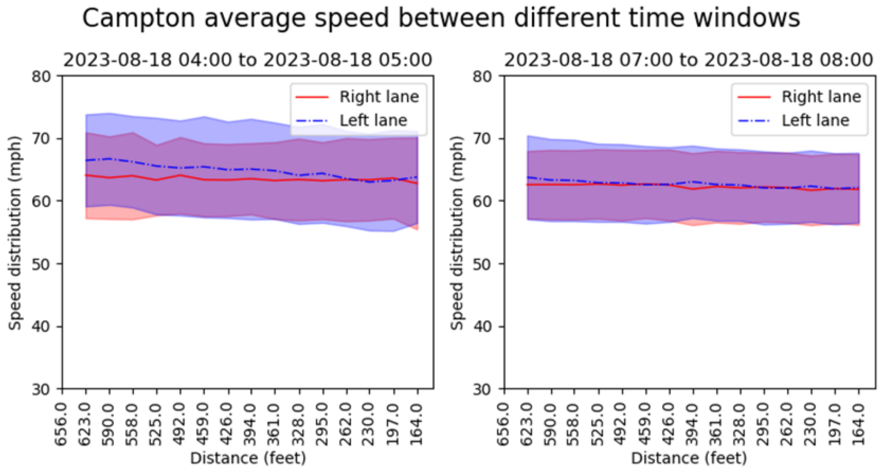
Date	Flashing Speed Limit Signs			PCMB Messages	
	UP/ON	UP/OFF	DOWN/OFF	LEFT LANE CLOSED MM 86.4 MERGE EARLY	POSSIBLE SLOW OR STOPPED TRAFFIC AHEAD BE AWARE
8/17/2023	0600	1330	NO	ALL DAY	NO
8/18/2023	0600	NO	0900	UNTIL 1230	1230
8/19/2023	NO	NO	ALL DAY	NO	ALL DAY
8/20/2023	NO	NO	ALL DAY	NO	ALL DAY
8/21/2023	0600	1330	UNTIL 0600	1300	UNTIL 1300
8/22/2023	0600	1330	NO	ALL DAY	NO
8/23/2023	0600	1330	NO	ALL DAY	NO
8/24/2023	0600	1330	NO	ALL DAY	NO
8/25/2023	0600	NO	1300	UNTIL 1500	1500
8/26/2023	NO	NO	ALL DAY	NO	ALL DAY
8/27/2023	NO	NO	ALL DAY	NO	ALL DAY
8/28/2023	0600	1400	UNTIL 0600	NO	ALL DAY
8/29/2023	0600	1300	NO	NO	ALL DAY
8/30/2023	0600	1300	NO	NO	ALL DAY
8/31/2023	0600	1730	NO	NO	ALL DAY

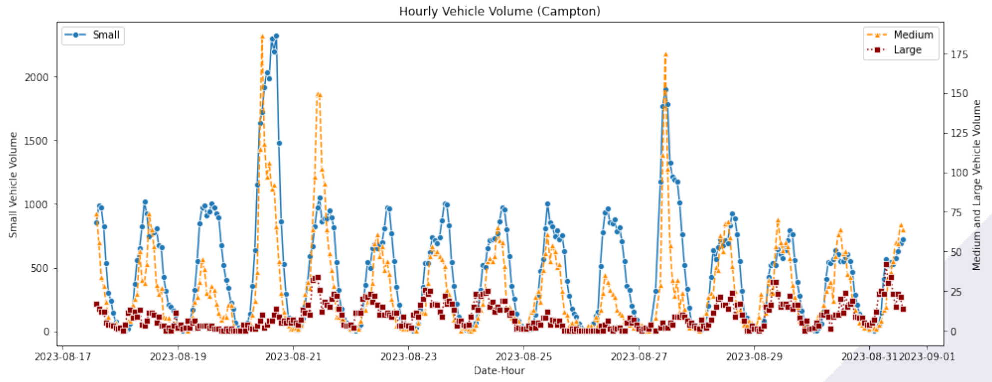


DATA COLLECTION AND PROCESSING

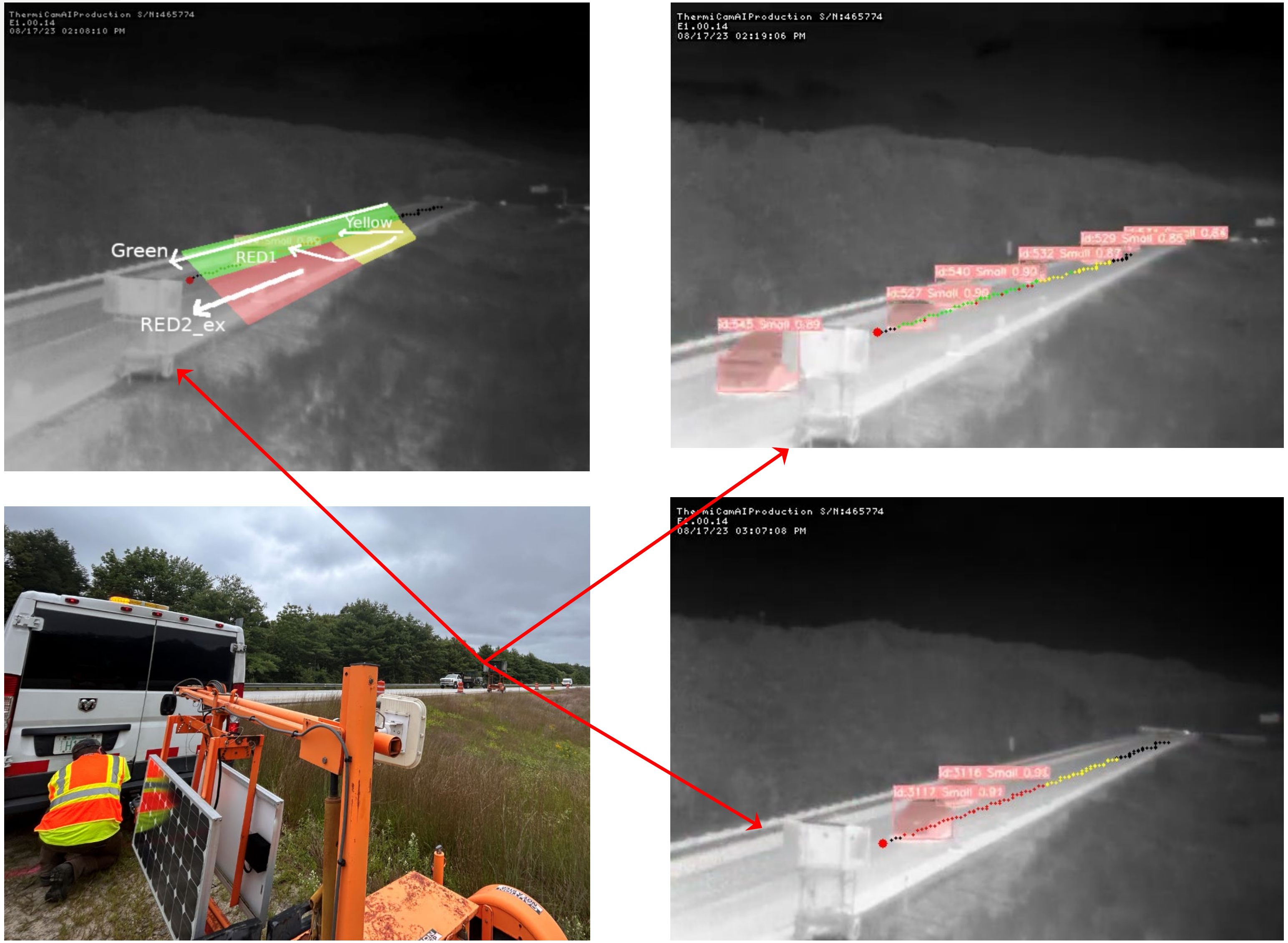
- **Thermal videos**
  - Traffic counts
  - Vehicle classifications
- **Radar data**
  - Speeds of individual vehicles
  - Vehicle trajectories





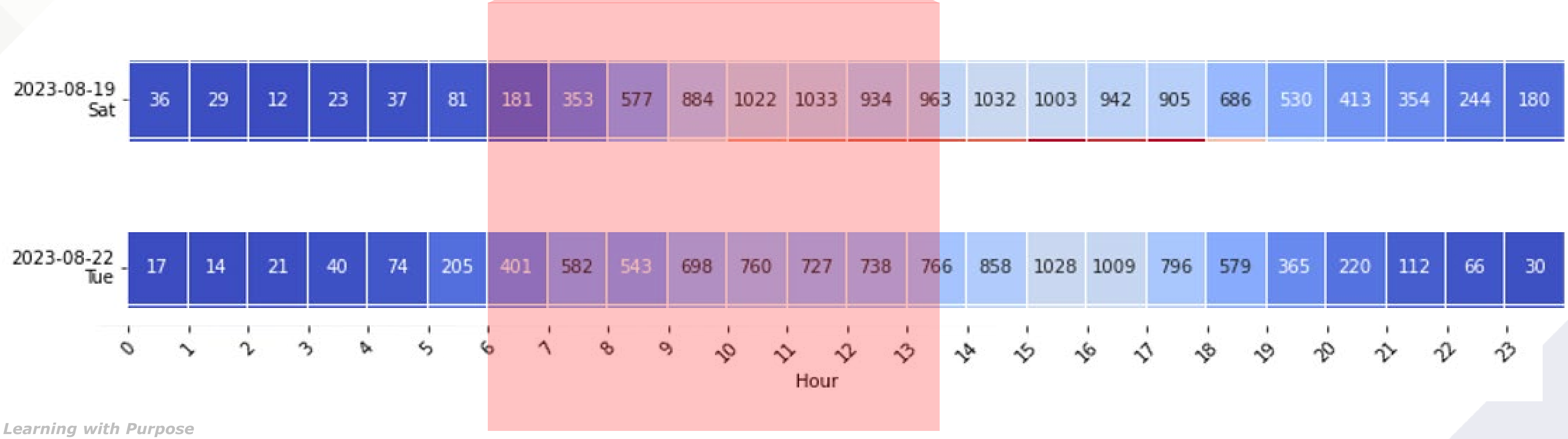


MERGING POINT ANALYSIS



MERGING POINT ANALYSIS (CONT.)

	Date	Flashing Speed Limit Signs			PCMB Messages	
		UP/ON	UP/OFF	DOWN/OFF	LEFT LANE CLOSED MM 86.4 MERGE EARLY	POSSIBLE SLOW OR STOPPED TRAFFIC AHEAD BE AWARE
Baseline	8/19/2023	NO	NO	ALL DAY	NO	ALL DAY
New control	8/22/2023	0600	1330	NO	ALL DAY	NO

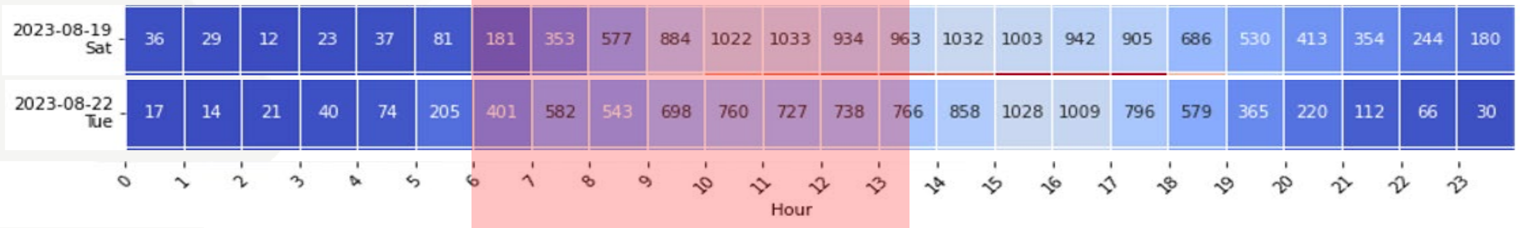


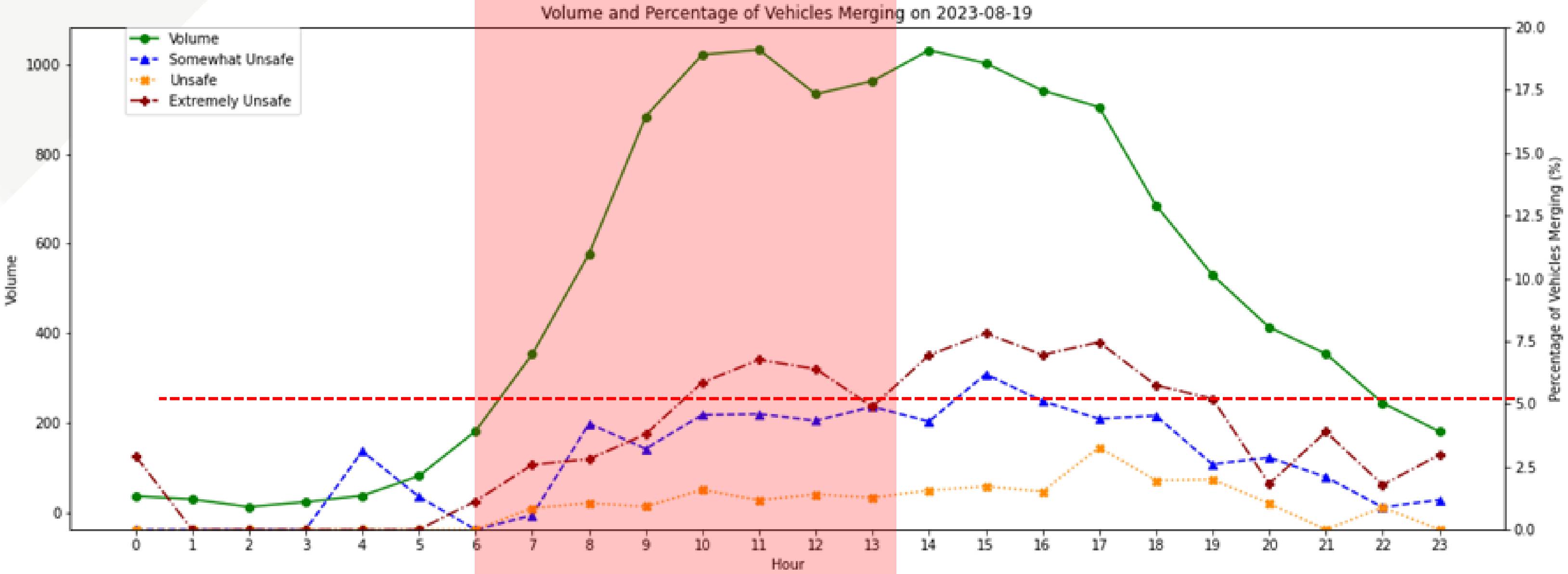
Baseline

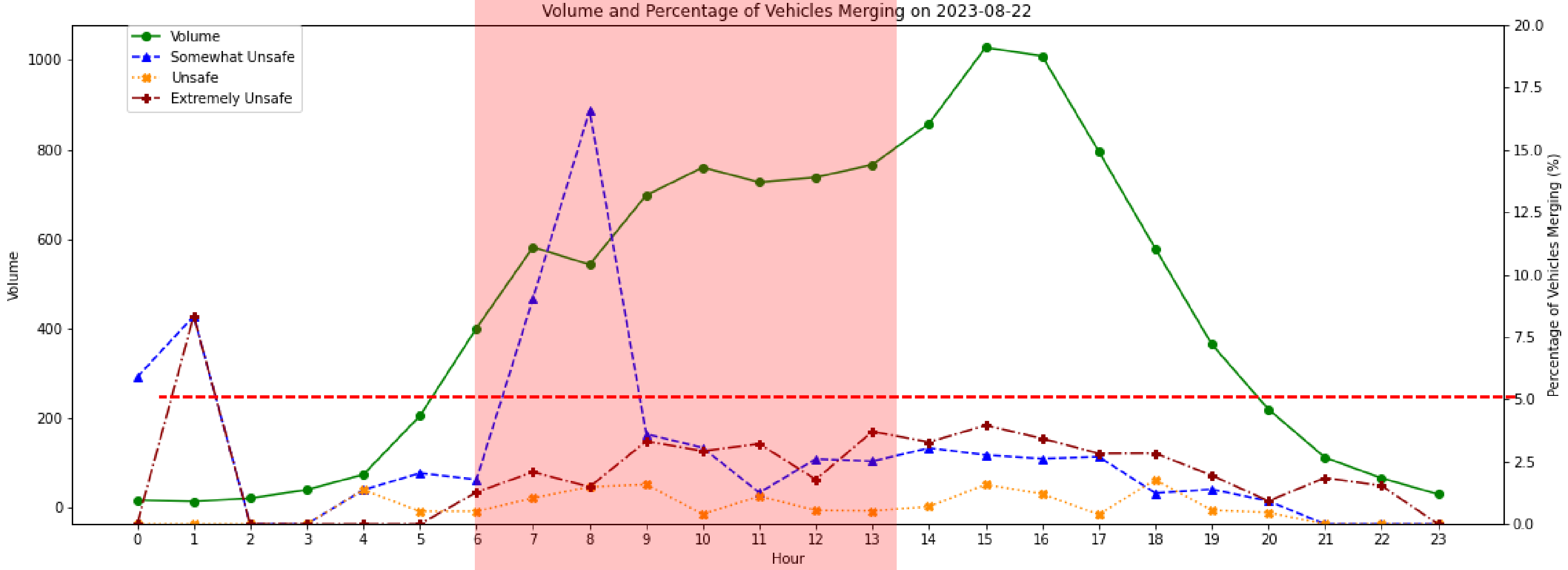
All Day – “Possible Slow or Stopped” & “Traffic Ahead Be Aware”

6am to 1:30 pm Flashing Speed Limit Sign

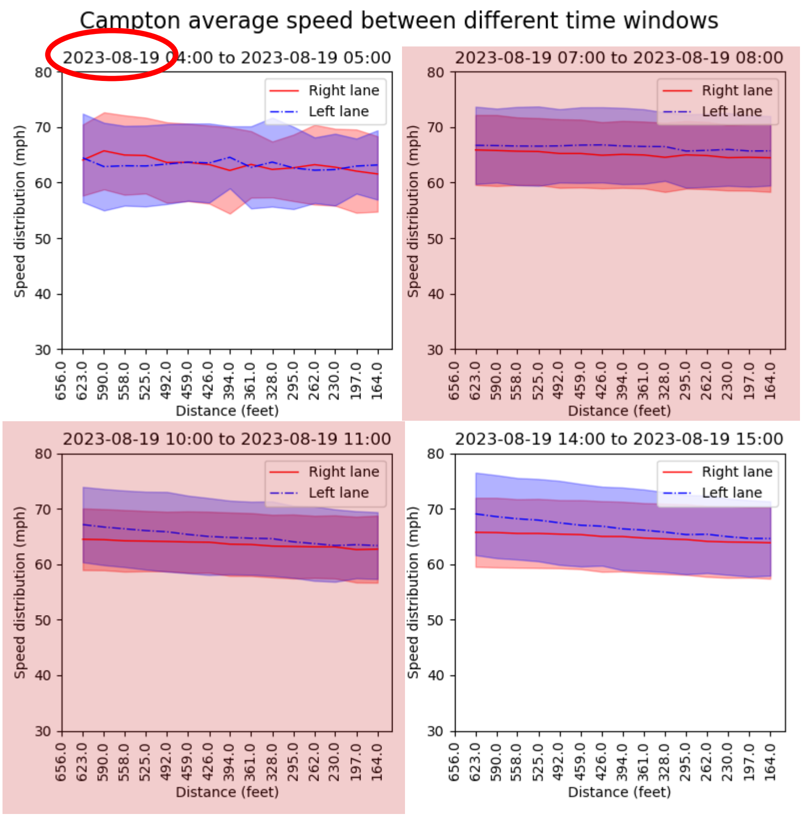
All Day – “Left Lane Closed” & “Merge Early”

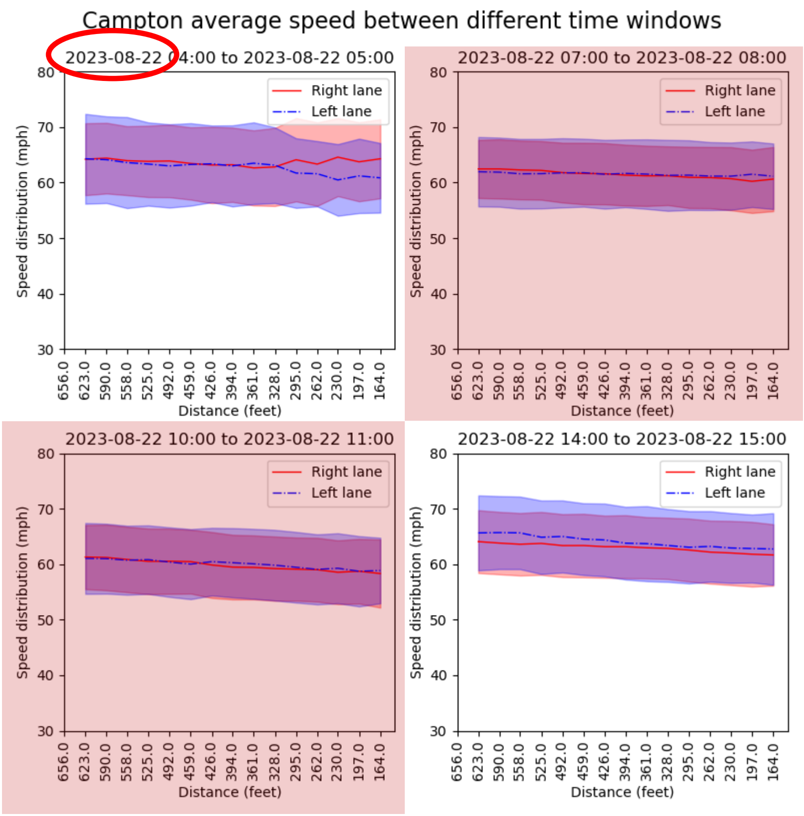






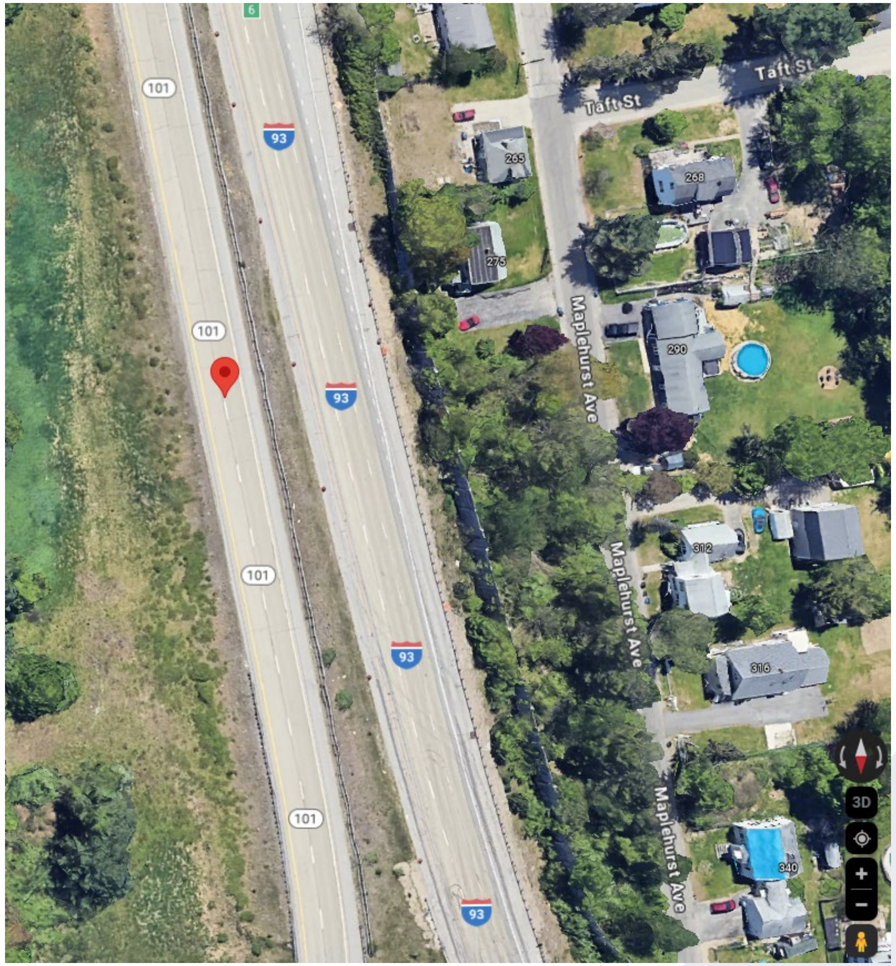
SPEED ANALYSIS





Date	Flashing Speed Limit Signs			PCMB Messages	
				LEFT LANE CLOSED	POSSIBLE SLOW OR STOPPED
	UP/ON	UP/OFF	DOWN/OFF	MM 86.4 MERGE EARLY	TRAFFIC AHEAD BE AWARE
8/19/2023	NO	NO	ALL DAY	NO	ALL DAY
8/22/2023	0600	1330	NO	ALL DAY	NO

ONGOING WORK



Location: 42.992511, -71.417054

C3 - NETWORK-WIDE SPEEDING ACTIVITY ANALYSIS

- From StreetLight, **traffic volume**, **segment speed**, and **speed distribution** were collected **every hour** for the duration of three years spanning from 2017 to 2019.

- Other variables include
  - Area type (urban/rural)
  - Traffic density (derived from volume and speed)
  - Time of the day, week, and year
  - Speed limit
  - Radius
  - Superelevation
  - Arc angle
  - Lane width

- Question
  - +10 mph, +15 mph, and +20 mph speeding

C3 - NETWORK-WIDE SPEEDING ACTIVITY ANALYSIS

- Increased chance of speeding
  - on curves and ramps located in rural compared to urban areas
  - during the morning and evening peak hours, and weekends
  - on less congested interstate horizontal curves and ramps (e.g., low-volume roads)
  - on roads with superelevation greater than 3%
  - sharp curves
- Speeding decreases significantly during
  - the peak of winter seasons with frequent snowfall, cold weather, and icy and frozen roads

TECHNOLOGY TRANSFER ACTIVITIES

- Liu, Q., & Ge, T. (2022, August). RL2: A Call for Simultaneous Representation Learning and Rule Learning for Graph Streams. In Proceedings of the 28th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 1109-1119).
- Liu, R., Liu, Q., & Ge, T. (2023, August). Fairness-Aware Continuous Predictions of Multiple Analytics Targets in Dynamic Networks. In Proceedings of the 29th ACM SIGKDD Conference on Knowledge Discovery and Data Mining (pp. 1512-1523).
- Vergara, E., Aviles-Ordonez, J., Xie, Y., & Shirazi, M. Understanding Speeding Behavior on Interstate Horizontal Curves and Ramps Using Networkwide Probe Data, Journal of Safety Research, accepted for publication.

THANK YOU

Image generated by ChatGPT4

Contact Information for Additional Questions
Email: Yuanchang_Xie@uml.edu