TECHNICAL REPORT DOCUMENT PAGE

1. Report No. NETCR124	2. Government Accession No.	3. Recipient's Catalog No. N/A
4. Title and Subtitle Current Status of Transportation Data Analytics and Pilot Case Studies Using Artificial Intelligence (AI)		5. Report Date December 2023 6. Performing Organization Code N/A
7. Author(s) Yuanchang Xie, Ruifeng Liu, Zubin Bhuyan, Danjue Chen, Tingjian Ge, Mohammadali Shirazi, Eduardo Vergara, and Juan Aviles-Ordonez		8. Performing Organization Report No. N/A
9. Performing Organization Name and Address University of Massachusetts Lowell 1 University Avenue Lowell, MA 01854		10. Work Unit No. N/A 11. Contract or Grant No. NETC 20-2
12. Sponsoring Agency Name and Address New England Transportation Consortium C/O Jeffrey Pulver Maine Department of Transportation 24 Child Street Augusta, ME 04330		13. Type of Report and Period Covered Final Report (05/2021-12/2023) 14. Sponsoring Agency Code NETC 20-2
15. Supplementary Notes		

16. Abstract

Data is becoming increasingly important for state Departments of Transportation (DOTs) in making strategic and day-to-day decisions. This research offers a comprehensive review of both existing and emerging data sources for Transportation Systems Management and Operations (TSMO). It discusses and summarizes the pros and cons of each data source. Additionally, interviews were conducted with federal and state DOT employees to gather their insights on future data sources and data needs, data integration and analysis, data archiving, sharing, security, and privacy, as well as stakeholders and workforce development. Based on the review and interview results, recommendations are provided regarding future data needs, emerging data sources, data processing and analytics, etc. The research also conducts three case studies to showcase the potential of using emerging data and AI technologies to address TSMO needs. These studies utilize advanced radar and thermal camera sensors, along with probe data, to model vehicle speed when approaching highway horizontal curves, examine how traffic signs may impact vehicle speed and lane-changing behaviors at highway work zones, and explore factors influencing speeding at highway horizontal curves and ramps. These studies demonstrate the benefits and necessity of combining data from different sources to meet TSMO needs and highlight the potential of AI.

17. Key Words Transportation data analytics; thermal; microwave; radar; deep learning; computer vision; work zone; safety; ramp; probe data; artificial intelligence, AI		18. Distribution Statement No restrictions.	
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 188	22. Price

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

SI* (MODERN METRIC) CONVERSION FACTORS				
APPROXIMATE CONVERSIONS TO SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1,000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2,000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa
APPROXIMATE CONVERSIONS FROM SI UNITS				
Symbol	**When You Know**	**Multiply By**	**To Find**	**Symbol**
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2,000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	2.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

*SI is the symbol for International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380. (Revised March 2003)

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

Final Report

Prepared By:

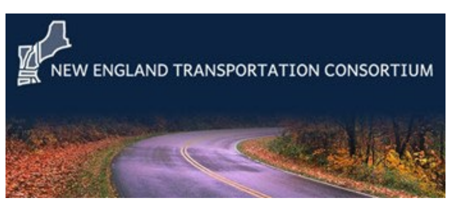
Yuanchang Xie[1], PhD, PE, Professor
Ruifeng Liu[2], Graduate Research Assistant
Zubin Bhuyan[2], Graduate Research Assistant
Danjue Chen[1], PhD, Associate Professor
Tingjian Ge[2], PhD, Professor

[1] Department of Civil and Environmental Engineering
[2]Department of Computer Science
University of Massachusetts Lowell

Mohammadali Shirazi, PhD, Assistant Professor
Eduardo Vergara, Graduate Research Assistant
Juan Aviles-Ordonez, Graduate Research Assistant
Department of Civil and Environmental Engineering
University of Maine

Prepared For:

New England Transportation Consortium (NETC)



December 2023

This page left blank intentionally.

Acknowledgements

The project team extends sincere thanks to all members of the project technical committee for their support and guidance. We are especially grateful to the New Hampshire Department of Transportation (NHDOT) for their generous assistance in facilitating radar and thermal video data collection. Our heartfelt appreciation goes to Susan Klasen, David Chase, Kevin Carlson, Patrick Sweet, Michael O'Donnell, Patrick Doughty, Christopher Moore, and Boyang Xiang for their support. Lastly, we express our gratitude to all the respondents who kindly participated in our interviews.

Disclaimer

The contents of this report solely represent the views of the authors, who are responsible for the accuracy of the presented facts and data. The mention of specific commercial products is included in this report to enhance comprehension and does not imply any recommendation or endorsement by the research team or the New England Transportation Consortium (NETC). The contents do not necessarily reflect the official views or policies of the NETC. This report does not constitute a standard, specification, or regulation.

This page left blank intentionally.

Executive Summary

Data is becoming increasingly important to state Departments of Transportation (DOTs) in making strategic and day-to-day decisions. Different DOT divisions have specific data needs. These needs sometimes overlap with each other but are not always addressed in a holistic manner or by a central office. The growing volume and complexity of data is another major challenge to state DOTs. Each year, various DOT divisions generate substantial amounts of data using LiDAR, drones, radar and loop detectors, weather stations, Bluetooth readers, traffic cameras, Automated Traffic Signal Performance Measure (ATSPM) systems, General Transit Feed Specification (GTFS) systems, pavement, and rail track inspection vehicles, among others. Data storage, security and privacy are also critical aspects of transportation data analytics. Additionally, with recent technological advancements, third-party data vendors using non-traditional data sources are playing an increasingly important role in supplying transportation data. DOTs need to carefully assess the pros and cons of relying on third-party data vendors versus their own data collection infrastructure to achieve a balance.

The objectives of this research are three-folds:

(1) Providing a clear and comprehensive picture to the six New England state DOTs regarding their data assets and needs, data modeling and workforce requirements, emerging data sources, as well as data collection, analysis, utilization, storage (e.g., security), and sharing (e.g., security) practices related to traffic operations.
(2) Offering strategic and practical recommendations to prepare New England DOTs for future data-driven transportation system analytics, considering emerging sensing and analytical technologies such as connected vehicles, the Internet of Things (IoT), and Artificial Intelligence (AI).
(3) Conducting case studies using AI techniques and/or emerging data sources for improving traffic operations and safety.

A comprehensive review of existing and emerging data sources for Transportation Systems Management and Operations (TSMO) has been conducted. The pros and cons of each data source are discussed and summarized. Additionally, interviews with domain experts have been undertaken to gather their insights on future data sources and needs, data integration and analysis, data archiving, sharing, security, and privacy, as well as stakeholders and workforce development. The results of the review and interviews are provided in Chapter 2. Based on the findings in Chapter 2, recommendations for future data needs, emerging data sources, data processing and analytics, etc. are presented.

This study also identifies several research topics to demonstrate the potential of using emerging data and AI technologies to address TSMO needs. The first case study utilizes an ultra-high-definition radar and a thermal camera to study driver behavior on horizontal curves. For the five selected horizontal curves, the collected vehicle trajectories suggest that drivers do not significantly alter their speeds when approaching horizontal curves. The collected speed data is compared with TomTom data at one site. Overall, the radar speed data and TomTom speed data match well. However, the radar sensor clearly provides more samples early in the morning, and it seems that TomTom data overestimates vehicle speeds during this period. Algorithms have been developed to identify speeding activities and risky drivers from the radar data. The corresponding thermal video clips have also been filtered out to confirm the radar results. The combined data provides a useful tool to monitor traffic and identify periods with more risky events for targeted law enforcement. AI algorithms have been developed to process the collected thermal videos, generating vehicle counts, vehicle time headways, and detecting and counting risky driving behaviors. This case study suggests that the radar and camera sensors ideally should be mounted on a fixed structure to mitigate the negative impacts of vibration. Also, both sensors should be mounted directly above the traffic to accurately capture vehicle lane-changing activities.

The second case study focuses on highway work zone safety. A work zone on I-93 in Campton, NH, has been selected. This work zone is equipped with both flashing speed limit signs and changeable message signs. The purpose of this case study is to find out how such signs can affect vehicle speed and merging behavior. One week of radar data and two weeks of thermal videos are collected from this work zone. AI algorithms have been developed to process thermal videos and count vehicles that merge at different distances from the beginning of the work zone lane closure taper. The results suggest that flashing speed limit signs are helpful in reducing vehicle speed when approaching a work zone as well as prompting drivers to merge earlier. The first two case studies successfully demonstrate the benefits of using portable sensors to collect detailed vehicle trajectories for both TSMO and safety purposes.

The third case study utilizes location-based service (LBS) data, specifically StreetLight data, to investigate vehicle speeding activities at highway horizontal curves. Lane departure collisions contribute significantly to roadway fatalities in the United States, with many occurring on horizontal curves or ramps due to speeding. This case study explores factors influencing speeding on Interstate horizontal curves and ramps, utilizing two unique data sources. The first database incorporates comprehensive curve and ramp characteristics from MaineDOT, while the second includes volume, average speed, and speed distribution data from StreetLight Insight®. The evaluation considers factors such as level of service (LOS), time of day (morning, evening, and off-peak hours), day of the week (weekdays and weekends), and month of the year (January-December), along with geometric characteristics like curve radius, arc angle, and superelevation. The findings indicate increased odds of speeding at horizontal curves with improved LOS, larger radii, and greater superelevation. Conversely, speeding decreases on curves with larger arc angles and during winter months. Similar trends are observed in ramp models, except for ramp radius, which is found to be an insignificant factor. These results underscore the significance of speed enforcement and other countermeasures to reduce speeding on curves and ramps with low traffic volumes, high speed limits, and large radius and superelevation, particularly in rural areas. The results can inform the prioritization of locations for implementing speed countermeasures or signage, such as advisory speed signs, as well as deploying enforcement resources to high-priority locations and times.

Table of Contents

.

List of Tables

List of Figures

This page left blank intentionally.

List of Acronyms

Acronym	Expansion
AADT	Annual Average Daily Traffic
AE	Autoencoder
AI	Artificial Intelligence
AIC	Akaike Information Criterion
AID	Automated Incident Detection
AI-ITMS	Artificial Intelligence - Intelligent Traffic Management System
APC	Automated Passenger Counter
ARAN	Automatic Road Analyzer
ATMS	Advanced Traffic Management Systems
ATCMTD	Advanced Transportation and Congestion Management Technologies Deployment
ATSPM	Automated Traffic Signal Performance Measure
AV	Autonomous vehicle
AVL	Automatic Vehicle Location
AWS	Amazon Web Services
BIC	Bayesian Information Criterion
CAD	Computer-Aided Dispatch
CAVs	Connected and Automated Vehicles
CCTV	Closed-Circuit Television
CSP	Connecticut State Police
CTDOT	Connecticut Department of Transportation
C-Value	Confidence Value
DelDOT	Delaware Department of Transportation
DOT	Departments of Transportation
DSRC	Dedicated Short-Range Communications
DSS	Decision Support Systems
DTC	Delaware Transit Corporation
ESRI	Environmental Systems Research Institute
FDOT	Florida Department of Transportation
FHWA	Federal Highway Administration
GIS	Geographic Information System
GPS	Global Positioning System
GTFS	General Transit Feed Specification
HCM	Highway Capacity Manual
HPMS	Highway Performance Monitoring System
HOC	Highway Operations Center
HPMS	Highway Performance Monitoring System

IMO	Integrated Mobile Observations
IoT	Internet of Things
IT	Information Technology
ITS	Intelligent Transportation System
ISO	International Organization for Standardization
KBES	Knowledge Based Expert Systems
KDE	Kernel Density Estimation
LBS	Location Based Service
LiDAR	Light Detection and Ranging
LRS	Linear Referencing System
LOS	Level of Service
MAC	Media Access Control
MaineDOT	Maine Department of Transportation
MassDOT	Massachusetts Department of Transportation
MBTA	Massachusetts Bay Transportation Authority
MDSS	Maintenance Decision Support System
MPO	Metropolitan Planning Organization
MRWIS	Mobile Road Weather Information System
NDS	Naturalistic Driving Study
NETC	New England Transportation Consortium
NHDOT	New Hampshire Department of Transportation
NHS	National Highway System
NLP	Natural Language Processing
NOAA	National Oceanic and Atmospheric Administration
NPMRDS	National Performance Measurement Research Data Set
OD	Origin-Destination
ODOT	Oregon Department of Transportation
PCMS	Portable Changeable Message Sign
PDO	Property-Damage-Only
PeMS	Performance Measurement System
PII	Personally Identifiable Information
PTZ	Pan-Tilt-Zoom
ReID	Re-IDentification
RIDOT	Rhode Island Department of Transportation
RIPTA	Rhode Island Public Transit Authority
RGB	Red, Green, and Blue
RWIS	Road Weather Information System
SPaT	Signal Phase and Timing
TB	Terabytes
TIS	Traveler Information Systems

TMC	Traffic Management Center
TSMO	Transportation Systems Management and Operations
TxDOT	Texas Department of Transportation
UAV	Unmanned Aerial Vehicles
USDOT	United States Department of Transportation
VAE	Variational Autoencoder
VDOT	Virginia Department of Transportation
VTrans	Vermont Agency of Transportation
WIM	Weigh-in-Motion
Wi-Fi	Wireless Fidelity
WZDx	Work Zone Data Exchange
XD	eXtreme Definition

This page left blank intentionally.

1.0 Introduction

Data is becoming increasingly important to state Departments of Transportation (DOT) for strategic and day-to-day decision-making. Different DOT divisions have specific data needs, which sometimes overlap but are not consistently addressed in a holistic manner or by a central office. Intuitively, close collaborations among different divisions (e.g., planning, operations, safety, transit, maintenance, asset management) would lead to more integrated, coordinated, and cost-effective efforts to make the best use of data and DOT resources. By integrating data from different sources and better understanding the needs of various DOT divisions, duplication of data collection and analysis efforts can be avoided, and the value of data can be fully understood and exploited.

The increasing volume and complexity of data pose another major challenge to state DOTs. Each year, various DOT divisions generate a substantial amount of data using LiDAR, drones, radar and loop detectors, weather stations, Bluetooth readers, traffic cameras, Automated Traffic Signal Performance Measure (ATSPM) systems, General Transit Feed Specification (GTFS) systems, pavement, and rail track inspection vehicles, among other sources. Some critical problems include:

- How to transform data of different types, quality, varieties, etc. into useful information for decision-making?
- How to manage such a vast amount of data, sometimes unstructured, and archive only the essential information with long-term value?
- How to train/recruit engineers with expertise in both transportation domain knowledge and advanced models to analyze the data?

Data storage, security, and privacy are also critical aspects of transportation data analytics. For instance, many DOTs simply use cameras to monitor traffic and do not further analyze or archive traffic videos due to concerns about data storage and privacy, even though Artificial Intelligence (AI) methods now are sophisticated enough to extract accurate and useful information from videos. Additionally, data generated by drone cameras, transit fare collection systems (some based on mobile apps), license plate recognition systems, cell phone towers, cell phone GPS, and Wi-Fi/Bluetooth readers contain valuable information for improving traffic operations but may also include personally identifiable information. Therefore, it is important to examine and update existing data storage, security, and privacy policies and practices, so that such datasets can be effectively and safely utilized.

Most DOTs rely on both agency-owned data collection infrastructure and third-party data vendors, such as INRIX. The conditions of the data collection infrastructure are often overlooked during data inventory. It is important to keep track of them to ensure that DOTs' data collection infrastructure is in a state of good repair. Such a database is also important for DOTs to develop short- and long-term data infrastructure investment plans.

With recent technological advancements, third-party data vendors using non-traditional data sources are playing an increasingly important role in supplying transportation data. DOTs need to carefully assess the pros and cons of relying on third-party data vendors versus their own data collection infrastructure to achieve a balance. While doing this, it is important to consider emerging data sources (e.g., connected and automated vehicles), the quality and reliability of third-party data sources, future data needs, life-cycle costs of each option, etc. For example, can third-party vendors still provide reliable information under severe weather conditions? Can their data cover rural and urban areas equally well to ensure fair treatment for rural travelers?

The objectives of this research are three-folds:

- Providing a clear and comprehensive picture to the six New England state DOTs regarding their data assets and needs, data modeling and workforce requirements, emerging data sources, as well as data collection, analysis, utilization, storage (e.g., security), and sharing (e.g., security) practices related to traffic operations.
- Offering strategic and practical recommendations to prepare New England DOTs for future data-driven transportation system analytics, considering emerging sensing and analytical technologies such as connected vehicles, the Internet of Things (IoT), and Artificial Intelligence (AI).
- Conducting case studies using AI techniques and/or emerging data sources for improving traffic operations and safety.

While the primary focus of this research is data analytics for traffic operations, we will also address related topics, including safety. The results will serve as an example for state DOTs to expand this research and cover other application areas (e.g., engineering, planning, asset management) of transportation data analytics.

This project is divided into two phases. Phase I focuses on reviewing current DOT needs and practices related to data and TSMO, while Phase II involves conducting case studies using AI techniques and/or emerging data sources to improve traffic operations and safety. The results of Phase I are presented in Chapters 2 and 3, based on which several candidate topics for Phase II are identified. After consulting with the project panel, the research team eventually conducts three case studies. The first and second case studies aim to demonstrate the advantages of using portable sensors to collect detailed vehicle trajectory data for studying driver behavior under different circumstances. They also showcase the capability of using AI to analyze and reduce the collected trajectories, generating meaningful conclusions. Specifically, the first case study focuses on driver speed choices on horizontal curves. The second case study centers on driver speed choices and lane-changing behavior when approaching a highway work zone. While the first two case studies focus on collecting and modeling detailed trajectory data, the third case study aims to show the power of combining location-based service (LBS) data and traditional road inventory data to study driver speeding activities at a network scale.

The rest of this report is organized as follows:

- Phase I results are presented in Chapters 2 and 3, where Chapter 2 provides a comprehensive review of data and data sources for TSMO purposes, along with presenting the interview results with federal and state DOT employees.
- Building on the findings of Chapter 2, Chapter 3 offers recommendations regarding data needs, emerging data sources, data processing and analytics, and others, to state DOTs in the New England region.
- Phase II results are provided in Chapters 4 through 6, which describe the details for the three case studies, respectively.
- Chapter 7 summarizes this entire research and highlights important findings.

This page left blank intentionally.

2.0 Review of Data and Data Sources

This chapter aims to provide a clear and comprehensive picture to the six New England state DOTs regarding their data assets and needs, data modeling, and workforce requirements, along with emerging data sources. It also covers data collection, analysis, utilization, storage (e.g., security), and sharing (e.g., security) practices related to traffic operations. The chapter begins with an analysis of data and data sources related to Transportation Systems Management and Operations (TSMO) (Section 2.1), followed by potential applications of AI techniques in modeling TSMO data (Section 2.2). Interviews with domain experts regarding TSMO data and data analytics are conducted, and the results are summarized in Section 2.3.

2.1 Analysis of Data and Data Sources

The analysis of data and data sources consists of two sections: the first focuses on traditional data and sources, and the second on new and emerging data and sources.

2.1.1 Traditional Data and Data Sources

2.1.1.1 Highway Data

Traditional highway data sources mainly include loop detectors, microwave detectors, CCTV traffic cameras, and Bluetooth/Wi-Fi MAC address readers. From these sources, agencies can obtain occupancy, delay and travel time, spot and segment speeds, and volume data. Many agencies also acquire data from weather stations and weigh-in-motion stations. Since such data are quite different from traffic flow parameters, they are discussed separately in Sections 2.1.1.2 and 2.1.1.4, respectively.

2.1.1.1.1 Inductive loop detector

Inductive loops are one of the most common data sources for highways. They have been widely adopted by state DOTs to collect traffic count, speed, length (if dual loop detector), occupancy, etc. on highways. They have also been extensively used at intersections to provide input data to traffic signal controllers.

These loop detectors are less sensitive to the environment (e.g., temperature, lighting, snow, strong wind, vibration) and provide robust traffic measurements. However, since they are installed underneath the pavement, it is difficult to repair them if broken. Another major issue is that such detectors often are used to generate Annual Average Daily Traffic (AADT) data to meet the Highway Performance Monitoring System (HPMS) reporting requirements. The generated traffic measurements typically are not streamed in real time to Highway Operations Centers (HOC) or Traffic Management Centers (TMC), making them unsuitable for real-time applications such as incident detection and response.

Additionally, these detectors are installed at limited locations on major highways and intersections. Therefore, they can only provide situational awareness for highway segments near those locations. For traffic incidents that happen far away from those locations, they will not be detected in a timely manner, which is critical to emergency response. Even if an incident is detected, it is difficult to accurately estimate its location with a sparse inductive loop detector network. Again, knowing the location of an incident is very important for efficient emergency response operations.

The above issues can be addressed by adding more loop detectors and investing in communication and IT infrastructure. For example, Caltrans maintains a PeMS (Performance Measurement System), which consists of about 40,000 detectors covering freeways across all major metropolitan areas of California, providing both real-time and historical traffic data. However, the cost for doing so can be prohibitive, especially for many New England state DOTs with a significant portion of their highways in rural areas.

2.1.1.1.2 Microwave sensor

Similar to inductive loop detectors, microwave sensors are installed at limited locations. Also, the collected data sometimes are not streamed in real time to HOC or TMC. In this sense, microwave sensors share the abovementioned limitations of inductive loop detectors. However, compared to inductive loop detectors, microwave sensors are easier and less expensive to install and maintain. They are installed on roadside poles and the installation and maintenance cause no or little impacts on traffic. Some new microwave sensors can each cover more than 20 lanes and over 200 ft of road segments. They can also track individual vehicles, detect lane changes, and measure vehicle length in these segments, while inductive loop detectors can only measure traffic at a single point or over a very short segment (e.g., 20 ft).

2.1.1.1.3 CCTV camera

All New England state DOTs operate and maintain a CCTV camera network. For example, RIDOT has about 200 cameras (many of them are around rest stops) and plans to add more. These cameras provide important video feeds for identifying and confirming traffic incidents. However, in most state DOTs such CCTV traffic videos are reviewed manually to confirm traffic incidents (detected/reported using other methods) and provide traffic situational awareness. Such traffic videos typically are not recorded. They are not utilized to automatically detect traffic incidents, although technically it is possible to utilize video image processing algorithms to process live CCTV camera feeds and generate data such as vehicle count, speed, and density for data collection and detecting incidents.

Like inductive loop and microwave detectors, CCTV cameras are deployed at limited locations, although they are getting increasingly popular. One concern with CCTV is privacy, especially for high-definition cameras. Such a problem can be addressed in many ways. One solution is to utilize edge computing devices to process videos in the field without saving them (i.e., only keep and stream the extracted traffic measurements). With the wide deployment of CCTV cameras and adoption of Artificial Intelligence (AI) based video processing algorithms, CCTV cameras may potentially become a major source in the future for traffic data collection and incident detection.

Some toll road authorities are using high-definition CCTV cameras for toll by plate purposes. This application can generate Origin-Destination (OD) and segment travel time data beyond count, speed, and density. Such travel time data allows HOC operators to identify congested segments. However, it cannot provide much useful information related to location (e.g., where a congested segment starts and ends), unless the distance between upstream and downstream cameras is sufficiently short.

Some New England state DOTs have investigated the possibility of turning existing cameras into traffic sensors for traffic data collection and incident detection. A few issues they encountered include the low resolution of existing cameras, Pan-Tilt-Zoom (PTZ) cameras making it difficult to calibrate them, preferring a central video processing solution than adding video processing hardware to individual cameras, etc.

2.1.1.1.4 Bluetooth data

Bluetooth technology has been widely used in collecting travel time data. It detects the Media Access Control (MAC) addresses of Bluetooth devices on passing by vehicles and matches upstream and downstream MAC addresses to derive travel time. This is like matching upstream and downstream license plate numbers as some toll road authorities are doing (see discussion in Section 2.1.1.1.3 above) to determine the toll rate for charging users. A main difference is that Bluetooth readers are less expensive and do not require sophisticated data processing algorithms (e.g., AI algorithms for detecting and recognizing license plates).

Portable Bluetooth readers have been developed and can be easily deployed as needed. The collected data can be either stored locally or transmitted to TMC in real time via a 4G cellular network. Given that most new cars are equipped with Bluetooth, this data source is becoming increasingly important and reliable. However, there are several major limitations for Bluetooth data. First, the data sample is often biased. It is not uncommon to have several people (i.e., multiple Bluetooth devices) in one vehicle. This can lead to biased travel time measurements. Second, like all previously discussed sensors, the coverage of Bluetooth readers is still limited. A dense network of Bluetooth readers is needed, especially for quickly detecting incidents and accurately estimating their locations. Third, Bluetooth cannot provide information for individual lanes like what inductive loops, microwave, and cameras can do. Finally, some mobile devices can randomize their MAC addresses to avoid being tracked.

Another potential application of Bluetooth sensors is to derive OD and driver route choice data, which are very useful for TSMO. With such information, DOTs can better understand how drivers respond to congestion (e.g., messages displayed on changeable message signs) and make route choice decisions. This may help TMC operators develop effective traffic management and control strategies. However, deriving accurate OD and driver route choice data is not a trivial task, especially given some of the limitations of Bluetooth sensors.

MassDOT has a Bluetooth travel time system called GoTime and seems to be satisfied with its performance. NHDOT and VTrans have tested Bluetooth travel time systems but are not very satisfied with their performance.

2.1.1.1.5 Summary

Data from loop detectors often are not streamed to HOC in real time and are mainly used for HPMS reporting purposes. Such data are saved in roadside devices and are manually downloaded. Some New England DOTs (e.g., NHDOT) are thinking about connecting these loop detectors wirelessly to Traffic Management Centers (TMC) so that data can be downloaded in real time and remotely. Most DOTs are moving away from loop detectors due to the high installation and maintenance costs. Installing or repairing loop detectors requires setting up temporary work zones, which is expensive and creates safety risks.

Microwave sensors are widely used by New England DOTs. Data from microwave sensors (e.g., volume, speed, and occupancy by lane, vehicle length, vehicle type) are typically streamed to TMC in real time. Compared to cameras, microwave sensors are not affected by lighting conditions. Although thermal cameras can address the lighting issue, they are more expensive, tend to have lower resolutions than regular cameras in the same price range, and are not deployed at large scales.

Some retrofit loop detectors can generate inductive vehicle signatures, which can be used to re-identify vehicles at different locations. Similarly, Bluetooth/Wi-Fi sensors and CCTV camera-based vehicle re-identification techniques can match vehicles at different locations, thus generating vehicle Origin-Destination (OD) information. The vehicle signatures generated by loop detectors are usually only effective for matching vehicles at nearby locations, while MAC addresses from Bluetooth/Wi-Fi readers and license plate numbers (or vehicle video signatures) generated by CCTV cameras are generally more accurate. Research is still needed to derive accurate OD information from MAC addresses due to issues such as randomization, sampling rate, and bias. For the license plate method, a wide deployment of CCTV cameras is required, and its performance can be affected by lighting and camera angle factors.

2.1.1.2 Road Weather Information System (RWIS) and Winter Maintenance

State DOTs are using remote weather stations to monitor road surface conditions under different weather. Also, DOTs are interested in integrating data from weather forecasts, weather stations, and sensors installed on vehicles (e.g., plow trucks). Some states have plow trucks equipped with AVL and sensors, which provide real-time locations and speeds of plow trucks, material types and application rates, pavement and air temperatures, engine diagnostics, dashcam images, surface friction, and humidity. Information from weather stations, probe vehicles, and weather stations are critical to TSMO under severe weather conditions.

CTDOT installed an Integrated Mobile Observations (IMO) system in maintenance vehicles. Approximately 210 plow trucks (with the eventual goal of all plow trucks) are equipped with forward looking video camera, GPS, and temperature and relative humidity sensors. The collected data is sent to a vendor for CTDOT. The data is fed into a local weather forecasting system, and the processed data is then used to inform a Maintenance Decision Support System (MDSS), which recommends which and when roadway segments should be plowed and how much salt should be used in the winter to improve traffic safety.

The MDSS also takes data from a Roadway Weather Information System (RWIS) deployed at 40 fixed locations, which collects pavement friction, wet, dry, icy, salinity surface and sub-surface pavement temperature as well as atmospheric temperature information. A mobile version of this sensor suite is under development (MRWIS).

MassDOT has remote sensors to monitor roadway conditions, real-time locations and speeds of plow trucks, National Weather Service Data, smart work zone data (e.g., locations, durations, configurations).

Besides NOAA data, NHDOT also collects many data elements (e.g., pavement temperature, visibility, precipitation, water, ice, friction factor) from about 30 weather stations that are mainly on interstate highways and the turnpike. The weather station data are used to guide snow plowing and salting activities. If the friction factor is below standard, a text message will be generated by the weather station and sent to maintenance crews.

The plow trucks owned by NHDOT are equipped with the AVL system, which provides real-time information such as speed, location, plow up, plow down, spreading rate, etc. They are planning to add mobile RWIS to these plow trucks. NHDOT also has some maintenance trucks (not plow trucks) equipped with air and pavement temperature sensors. These trucks are driven by highway patrol foremen after snowstorms to determine how road segments should be treated. However, these maintenance trucks do not have the AVL system. They are thinking about integrating the data from weather stations, maintenance trucks, plow trucks, and weather forecast and presenting them in a simple but meaningful format for decision making.

NHDOT currently does not have a Maintenance Decision Support System (MDSS®). A consulting company took all the weather data from NHDOT and tried to develop a system to predict temperature at any point in the road network. The predictions sometimes were accurate, but the accuracy was not stable.

NHDOT, MaineDOT, and VTrans all divide their states into zones (6 for NHDOT, over 100 for Maine, and about 10 for VTrans). They use weather stations, trucks equipped with weather sensors, sensors at highway maintenance sheds, cameras, speed data, and NOAA data to estimate roadway conditions and publish the results on the regional 511 website.

2.1.1.3 Work Zone
Smart work zone technologies have been widely used by New England state DOTs. These smart work zones utilize sensors such as microwave, camera and Bluetooth to monitor traffic and collect data such as travel time, speed, delay, and queue. MaineDOT uses Linear Referencing System (LRS) to manage their work zone related information in ATMS, although the information is not updated in real time and requires data standardizations. FHWA established the Work Zone Data Exchange (WZDx) program a few years ago, and RIDOT is working on sharing smart work zone data using WZDx.

2.1.1.4 Weigh-in-Motion (WIM)
State DOTs also collect data from WIM stations, including traffic volumes by vehicle classification and weight, date, time, vehicle length by axle spacing, speed, and axle weight.

2.1.1.5 Tolling Data
Since toll roads are typically operated by private companies, DOT TSMO divisions often do not have full/direct access to tolling data. They must request such data through their turnpike authorities. Two types of tolling data can be useful for TSMO purposes: E-ZPass and license

plate records. E-ZPass data is similar to Bluetooth data. The main difference is that E-ZPass uses the Dedicated Short-Range Communications (DSRC) technology to read transponders in individual vehicles instead of MAC addresses. Since each vehicle has a unique transponder ID, the travel time derived from E-ZPass data is less biased than Bluetooth data. From E-ZPass records, time-dependent OD can be easily derived.

Some turnpikes also allow vehicles not equipped with E-ZPass to use by tracking their license plate numbers. By matching license plate numbers observed at entrances and exits, accurate OD and travel time information can be obtained. A clear limitation with the E-ZPass and license plate number data is that they are only available for toll roads.

2.1.1.6 Incident and Crash

All state DOTs have a database for historical incidents and crashes. Such data include incident/crash location, time, duration, etc. Some state DOTs also keep track of highway safety patrol records (e.g., MaineDOT) and 511 phone call records. NHDOT does not have a 511 system anymore but has access to highway safety patrol records and 911 calls related to traffic accidents. Their safety patrol records are in paper format and are entered into a database by TSMO staff. NHDOT uses the incident data information to optimize (based on human intelligence not automated algorithms) safety patrol schedules.

NHDOT's ATMS can take state police inputs to show crash alarms. It also allows traffic operators to define speed thresholds to display roadway segments in different colors based on their speeds. NHDOT finds this to be very useful for detecting incidents. The speed data used in the application come from both TomTom and microwave sensors on the roads. One issue with TomTom (and other similar products) is that DOTs do not have control over the segment length and only average speed data for the entire segment are provided. If an incident happens on a long segment, its impact may take a long time to be reflected in the average segment speed. Therefore, NHDOT uses DOT owned microwave sensors to complement TomTom data for incident detection on long segments.

In Connecticut, some highway patrol vehicles are equipped with data collection devices developed by a private company. CTDOT takes data from both highway patrol and Waze for incident detection. They have access to the incident information entered into the Connecticut State Police (CSP) Computer-Aided Dispatch (CAD) system. CSP dispatchers also notify the CTDOT of incidents on state roadways via dedicated telephone lines. CSP is in the process of revising its CAD time logs to provide additional scene clearance information for improved analysis of incident clearance times. CTDOT found the CSP CAD records to be a timely and reliable data source for incident detection, since almost all drivers have a cell phone. Overall, CTDOT is satisfied with the performance of the CSP CAD and Waze reports for incident detection.

For incidents reported by Waze, CTDOT uses CCTV cameras to further verify them. The incoming Waze reports of the same incident are automatically aggregated so that one incident will not generate many alerts that overwhelm HOC operators. CTDOT can also update Waze incident data to remove false alarms and incidents that have already been cleared. This two-way communications between CTDOT and Waze help provide reliable information to travelers and improve incident response.

2.1.1.7 Arterial

Loop detectors, traffic cameras and microwave sensors are commonly used at arterial intersections for sensing and data collection. Some state DOTs (e.g., MassDOT) also experimented with drones to collect traffic condition data (e.g., queue length) at intersections. Several state DOTs (e.g., VTrans, MassDOT) have deployed Automated Traffic Signal Performance Measures (ATSPM) systems, which provide real-time data on traffic detector state (e.g., occupied vs. unoccupied) and health condition, traffic control (e.g., which signal head is currently in green), turning movement counts, queue length, and speed. For intersections not equipped with ATSPM, the above data sometimes are also captured but are often not streamed to TMC. The ATSPM data can have many applications other than monitoring the health conditions of traffic detectors and controllers but have not been fully utilized yet. Since not every state has ATSPM, it is considered as an emerging data source and is further discussed in Section 2.1.2.3.

2.1.1.8 Transit

Transit agencies also collect many data that can be used for TSMO purposes, including General Transit Feed Specification (GTFS) data, transit fare collection data (e.g., smart card, Mobile ticket), CCTV camera videos, Automated Passenger Counter (APC) data, and ridership. For example, GTFS data can be used to estimate link travel time on urban arterials. However, the GTFS data is only widely available in major cities with many bus routes like Boston, not on state-maintained highways. Also, such data are owned by transit agencies, and are not directly accessible by TSMO division. For example, in Rhode Island, Rhode Island Public Transit Authority (RIPTA) is an agency separated from RIDOT. Sharing data across agencies is important, but can be difficult, especially for real-time data sharing. This partially explains why none of the six New England state DOTs explicitly utilize transit data for TSMO. For example, NHDOT TSMO does not utilize any transit data.

Delaware Transit Corporation (DTC) supplies fixed route and paratransit services statewide. DTC is an agency under DelDOT. DTC's automated fleet management system is integrated in DelDOT's AI-ITMS. Future AI-ITMS development could include transit system status information.

2.1.1.9 Parking

Static (e.g., location and # of lots) and dynamic parking data (e.g., parking duration), parking fee data, and mobile parking app data can also be useful for traffic management and control. For example, The Boston Central Transportation Planning Staff (CTPS) has done license plate surveys at commuter rails stations to derive passenger OD information, which is important for multimodal corridor transportation management.

TSMO division often does not have direct access to parking data. One reason is that many of the parking facilities are owned and operated by private companies. NHDOT provides support to allow third-party Apps to show the availability of parking spaces. However, the data is not utilized for TSMO purpose at this moment. NHDOT is only concerned about parking in a very limited number of areas (e.g., White Mountains).

2.1.1.10 Assets

Asset condition data are handled by different DOT divisions mostly in GIS format, including pavement, bridges, speed limits, traffic signs and markings, tunnels, Intelligent Transportation System (ITS) equipment, etc. Other than the conditions of bridges and tunnels, state DOTs are also collecting the condition data for ITS assets. NHDOT maintains a detailed database for ITS equipment such as variable message signs, sensors, communication devices, and traffic controllers.

NHDOT is in the process of loading all their ITS assets into a comprehensive GIS database. Previously NHDOT only tracked the locations of ITS devices (e.g., cameras, variable message signs), not the detailed condition information for each asset, for example, the cabinet for a camera has a modem and a server rack. They are also working on a work order system to ensure the state of good repair for ITS assets. The goal is to collect detailed asset condition and configuration data (e.g., modem type, maker) and connect them to asset locations managed by GIS. Also, such a system will be integrated with the work order system (e.g., Assetworks), so that NHDOT can track when and where a device is replaced or repaired. Maintenance is important for ITS assets. To maintain a state of good repair, DOTs need to know how much funding is needed for the next five or ten years for ITS asset maintenance, which will benefit from a detailed and accurate asset inventory system.

Laser, LiDAR (Light Detection and Ranging), and camera (mounted on vehicles and drones) have been extensively used by state DOTs to collect asset inventory and condition data. Such sensors have generated an enormous amount of data. More research is needed to reduce such datasets and explore how they may be used for TSMO purposes.

2.1.2 New and Emerging Data and Data Sources

2.1.2.1 Drone

Drones or Unmanned Aerial Vehicles (UAV) have been used as a popular platform for collecting highway data. RGB and infrared cameras and LiDAR have been mounted on drones for various applications. Those related to TSMO include providing situational awareness at incident/crash scenes and traffic monitoring. Traffic monitoring via drones can overcome the limitations of traditional methods of monitoring due to their simplicity, mobility, and ability to cover large areas. A recent paper presents a good review of research efforts that use drones in relation to online and offline extraction of traffic parameters from video data [*7*]. MassDOT is working on establishing a drone-based emergency response network and has used drones to monitor queue length at signalized intersections. Drones have been extensively used by RIDOT for construction sites monitoring.

2.1.2.2 LiDAR

LiDAR has also attracted significant attention in the past decade. MassDOT used LiDAR to scan all state-maintained highways, resulting in several hundred terabytes of data. They extracted useful information, such as traffic signs, from the LiDAR data.

Drones and LiDAR have generated a vast amount of data. How to extract useful information from such datasets and share and store them has now become a major issue. DOTs certainly do not want to discard such datasets and later find that valuable information could have been extracted from them. A good idea is to share such datasets (when possible) with universities, private companies, and the public, allowing them to come up with innovative ideas to analyze and utilize the data.

Report/Feature	Open Source		Econolite		Trafficware		Miovision	
Phase Termination Metric(s)	●		●		●		◐	(1)
Progression Quality Metric(s)	●		●		●		●	
Split Failure Metric(s)	●		◐	(2)	◐	(2)	●	
Delay Metric(s)	●		●		●		●	
Volume Metric(s)	●		●		●		●	
Yellow and Red Actuations Metric(s)	●		○		●		○	
Pedestrian Metric(s)	●		●		●		●	
Preemption Metric(s)	●		●		●		◐	(3)
Speed / Travel Time Metric(s)	◐	(4)	○		○		●	
Chart Customizations (e.g., Axis Min/Max, Data Filters)	●		◐	(5)	●		●	
Query Multiple Days on a Single Chart	●		◐	(6)	○		●	
Filter Data by Day of the Week	◐	(7)	◐	(8)	○		●	
Historical Data Comparison	○		◐	(9)	○		◐	(10)
Query Multiple Intersections on a Single Chart	○		○		○		◐	(11)
Dashboard Metric(s) for Multiple Intersections (Corridor / Network)	○		●		○		●	
Summary Tables	◐	(12)	●		○		●	
Highlight "Hot Spots"	○		●		○		●	
Programmable Alerts	●		●		○		●	
Optimization Features (e.g., Cycle Length, Split, Offset)	◐	(13)	●		○		●	
Process Data from Different Vendors	●		○		○		●	
No External Hardware Required	●		●		●		●	(14)
Integrate with Non-Linux-Based Controllers (ATC or 2070 with 1C CPU)	○		○		○		●	
Access to Raw High-Resolution Data	●		◐	(15)	◐	(15)	●	
Ability to Customize Reports	●		○		○		○	
Guidance Documentation	●		●		○		●	
Legend	● Available		◐ Partially Available		○ Not Available			

* Note: Reports and features are under development for all ATSPM systems. Evaluation reflects available reports and features as of 5/16/18.
(1) Phase duration information available. Phase termination type not available (i.e. max out, force off, gap out, skip).
(2) Green and red occupancy information available. Split failures not identified based on occupancy threshold.
(3) Preempt alert monitoring available. No preemption details available.
(4) Speed metric available only for Wavetronix radar detection.
(5) Ability to zoom in and out of charts. No available axis settings or data filters.
(6) Ability to overlay data from multiple days for some metrics (i.e. phase termination, progression quality, delay, volume, pedestrian) in comparison charts. Not available for all metrics.
(7) Link pivot arrivals on green can be filtered by day of the week. Not available for all metrics.
(8) Ability to filter data by day of the week for some metrics (i.e. phase termination, progression quality, delay, volume, pedestrian) in comparison charts. Not available for all metrics.
(9) Ability to overlay data from two date ranges for some metrics (i.e. phase termination, progression quality, delay, volume, pedestrian) in comparison charts. Not available for all metrics.
(10) Ability to add historical data to some charts (i.e. delay, volume, pedestrian, speed / travel time). Not available for all metrics.
(11) Multiple charts can be displayed together for progression quality and travel time / speed metrics. Additionally, all metrics can be viewed on the same reporting canvas.
(12) Summary tables available for turning movement counts and link pivot arrivals on green.
(13) Link pivot available for offset optimization.
(14) Software-only solution available. External hardware provides travel time data, cellular communication, and ability to collect high-resolution data if there is a non-Linux-based controller.
(15) Raw high-resolution data available, but must be requested from vendor.

Figure 2-1. Data elements collected by well-known ATSPM systems [*2*]

2.1.2.3 ATSPM

Utah was among the first few states to invest in the Automated Traffic Signal Performance Measures (ATSPM) system starting in 2013 [*1*]. Several New England state DOTs now have also deployed the ATSPM, which collects very detailed traffic signal performance measures every 1/10 seconds and stream the data in real time [*2*] to TMC. The data elements collected by well-known ATSPM systems are listed in Figure 2-1.

NHDOT currently has very few intersections that are connected via fiber to TMC. Over the next few years, many intersections will be connected to TMC either by fiber or wireless network. The ATSPM data can be used for improving traffic safety and operations at intersections. Since this system is relatively new, researchers and practitioners are still trying to find out how to effectively utilize the generated data. A potential challenge for example is that traffic signals are not directly under the TSMO bureau at NHDOT. Traffic signals, pavement marking, and signs are under the bureau of traffic at NHDOT. The NHDOT TSMO bureau does not have traffic signal engineers or technicians on staff. This organizational structure is typical for other DOTs. Overall, how to make full use of ATSPM data seems to be an interesting and timely topic for DOTs.

2.1.2.4 Crowdsourced, Probe Vehicle, and Connected Vehicle Data

Smartphones, probe vehicles, and connected vehicles all rely on GPS and have generated a tremendous amount of data that can be used for TSMO purposes. Since these data sources are based on similar technologies, they are thus discussed under the same main category. However, there are some subtle but important differences among these data sources.

Smartphones users contribute their data either actively or passively. One example of active smartphone data contribution is the Waze App. Waze users report roadway conditions such as incidents, debris, and speed traps. Both incidents and debris are critical safety hazards and need to be cleared as soon as possible. Such crowdsourced data are important for DOTs to improve highway safety. Smartphone users in many cases also passively contribute their data. For example, when drivers are using navigation Apps, they often contribute their speed and location information every few seconds. The speed and location information from all drivers collectively can be used to predict travel time, estimate travel speed, detect incidents, and identify safety hazards (e.g., locations with frequent harsh brakes). These Apps share some of the derived data with data contributors, but not all.

Probe vehicles refer to vehicles equipped with GPS and wireless communications technology. Sometimes this is called Automated Vehicle Location (AVL). Many buses and commercial trucks (e.g., owned by UPS and Walmart) are equipped with AVL. With AVL and other onboard sensors, system operators can know in real time where drivers are, how many times a heavy truck backs up, whether turn signals are used when they should be, speed violations, etc. Drivers of Transportation Network Companies such as Uber and Lyft need to install an App to connect with customers/passengers. These Uber and Lyft vehicles essentially work as probe vehicles. Companies such as INRIX and TomTom then take probe vehicle data from different sources (The exact sources are not disclosed and they may not include Uber and Lyft), clean them, and sell them to customers such as state DOTs. The INRIX and TomTom data are aggregated. One can only know the average segment speed or travel time, not individual vehicle speeds and locations. Given the original probe vehicle data, theoretically it is possible for INRIX and TomTom to provide disaggregated data to customers. However, this requires customers to have the capability of analyzing very detailed and large datasets.

USDOT has three ongoing connected vehicles pilot studies in Wyoming, New York City, and Tampa. In these studies, thousands of connected vehicles have generated very detailed vehicle trajectory data. However, such data only cover the three pilot sites. On the other hand, many new vehicles on the market are now equipped with GPS and wireless communications capability. Car manufacturers collect vehicle location and speed information, engine and wiper status, etc. and sell them to companies such as Wejo and Otonomo. Both probe vehicle data and connected vehicle data are based on GPS and wireless communications. The vehicle location and speed data are usually transmitted from vehicles to data center every few seconds. The processed data are then shared with customers in about one minute, which is sufficient for many TSMO applications such as incident detection. Different from INRIX and TomTom, Wejo and Otonomo provide disaggregated data to customers. Although such detailed data can be very useful, analyzing them is difficult. So far, none of the six New England state DOTs have used either Wejo or Otonomo data.

In the rest of this section, the above data sources are discussed in more detail. In this report, these data sources are grouped into the following three sub-categories:

- ***User Reported Data***: This specifically refers to data contributed actively by travelers using cellphones. Waze is a main source of such data. Some navigation Apps such as Google Maps also allow users to report incidents and speed traps.
- ***Aggregated Probe Data***: This sub-category includes aggregated speed and travel time data such as those provided by INRIX, TomTom, etc. Also, Uber Movement provides zone to zone travel time and road segment travel speed data. Google sells travel time data. All these aggregated datasets are based on GPS coordinates generated by smartphone Apps, AVL, or connected vehicles.
- ***Trajectory Data***: This sub-category is for unprocessed GPS coordinates generated by smartphone Apps, AVL, or connected vehicles. Examples include Wejo and Otonomo.

2.1.2.4.1 User Reported Data

Almost every driver now has a cellphone. When a crash occurs, it usually will not take much time for the driver(s) involved or passing by drivers to call 911 and report it. Some state DOTs rely a lot on such information for AID. A limitation of driver incident reporting is that non-collision (e.g., road debris) and property-damage-only (PDO) incidents may be under-reported. Also, for drivers calling 911, they sometimes do not know their exact locations on the road (unless they are using a mobile App).

Most state DOTs have access to Waze data and are using Waze incident reports for AID. Issues with Waze incident reports include: (1) submitting Waze incident report while driving is dangerous; (2) usually there is a delay between when an incident is spotted and when it is reported, which makes it difficult to directly identify the exact incident location; and (3) sometimes there are inaccurate reports. For example, an incident has already been cleared, but it still shows up in Waze. In a quantitative comparison by Iowa DOT of various sources of incident detection, Waze was ranked the 4th (out of 8) largest contributing sources. While essentially free, Waze incident reports still must be validated by other means, and it captured only 43.2% of ATMS recorded incidents during the analysis period (although this has most likely increased as the number of users increases).

2.1.2.4.2 Aggregated Probe Data

Several companies offer aggregate probe data, including INRIX, HERE, TomTom, and Google. A significant advantage of probe data is that state DOTs do not need to invest in any data collection infrastructure, and do not need to worry about the maintenance of such infrastructure either. Although purchasing data can be expensive, clearly many DOTs think it is worthy given the trouble and cost associated with maintaining their own data collection infrastructure. In addition, probe data usually has a much larger coverage than traditional data sources such as inductive loop detectors, microwave detectors, and CCTV cameras. Its actual coverage depends on how many users are contributing their data. Usually, there are more data contributors in urban than suburban areas.

Most probe data vendors provide information aggregated by road segments, such as segment speed and travel time. State DOTs take what these vendors provide and do not know the details of how the data are aggregated. The length of each segment is also decided by vendors. Different vendors often have different standards/ways to divide roads into segments. When DOTs have data from multiple vendors, they often face the challenge to reconcile data aggregated using different segment definitions, which is not a trivial task. In addition, state DOTs lose the opportunity to extract more granular and useful information from the aggregated probe data.

Using Automated Incident Detection (AID) as one example, DOTs may want to have short segments in areas prone to incidents (ideally in all areas if computational power is not a constraint). With short segments, changes in individual vehicles' speeds and travel times can be quickly reflected in the corresponding segment measures. On the other hand, providing aggregated data and hiding the details to some extent is beneficial to DOTs, as they often do not have the human resources to handle the large volume of raw trajectory data and extract critical information out of them.

As discussed previously, the aggregated probe data originally come from detailed vehicle trajectories. Besides the aggregated road segment measures, some data vendors (e.g., INRIX) provide more detailed data at the lane level. They can also generate incident and dangerous slowdown alerts.

Overall, state DOTs are satisfied with probe data quality. CTDOT has validated HERE travel time data. DOT staff had driven some routes to verify the travel time estimated by HERE and found that they meet the DOT data quality standards (FHWA 23CFR511 quality standard for traffic data). CTDOT noted they only display the probe data when it is accurate, which at this point is mostly during daylight hours and some early evening hours on weekdays and weekends. NHDOT was able to detect crashes based on TomTom speed data even before they were notified by state police.

The City of Boston partnered with Waze to identify traffic signals that need improvements. They worked with MBTA to measure impacts of signal timing along the Silver Line [*3*]. Although it was not explicitly mentioned what data from Waze was used, it is likely the travel speed data (similar to the probe vehicle data) and the more detailed vehicle trajectories (see discussion in the next subsection) were used.

2.1.2.4.3 Trajectory Data

Smartphones, probe vehicles, and connected vehicles can also generate vehicle trajectories, which are much more detailed than the aggregated probe data described in the previous subsection. Two trajectory data vendors are Wejo and Otonomo. They provide similar vehicle trajectory datasets, which include data elements such as longitude, latitude, speed, heading, wipers change, seat belt change, autonomous emergency braking, etc. These data elements are collected from some commercial vehicles sold in recent years at short intervals (e.g., every 3 seconds) and are transmitted back to a data warehouse and made available to end users within 60 seconds. Currently, there are over 10 million vehicles contributing trajectory data, and this number is growing as more new vehicles are being sold.

Such vehicle trajectory data can be used to measure traffic operations performance and derive surrogate safety measures such as harsh-braking events. For example, researchers from Purdue University used Wejo data to correlate harsh-braking events with crash occurrences near highway work zones [*4*]. Vehicle trajectories are also useful for some real-time applications, such as detecting traffic incidents and traffic slowdowns and generating safety hazard alerts (e.g., black ice on road).

The Eastern Transportation Coalition (previously known as the I-95 Corridor Coalition) conducted a pilot study to estimate traffic volume using Wejo data in real time [*5*]. Their study utilized data from six states: Alabama, Florida, Georgia, North Carolina, Tennessee, and Virginia. They found that Wejo data covered about 3% of all vehicles on the road and the pilot study received data from each connected vehicle every 3 seconds. Wejo generated over 230B data points in the 3-month pilot study period. Their study concluded that using Wejo data to estimate traffic volume in real time is a viable solution, particularly given that the number of connected vehicles is continuously growing.

Many DOT vehicles are equipped with the AVL system, allowing DOTs to track their vehicles (e.g., plow trucks) in real time. These vehicles can provide valuable trajectory information especially under severe weather conditions. Other public agencies also have AVL in their fleets, such as state police and transit. Integrating all fleet data can generate very useful traffic information benefiting all participating agencies (e.g., first responders always want to have accurate traffic information to find the best routes). In Delaware, all state vehicles have GPS based tracking. As part of their ATCMTD AI-ITMS project, DelDOT is equipping some DOT vehicles to monitor vehicle data port and to integrate the data into their AI-ITMS.

2.1.2.4.4 Summary

Crowdsourced, probe vehicles, and connected vehicles data are playing a critical role in TSMO. For example, RIDOT is exploring INRIX data for AID. Currently, RIDOT relies on CCTV cameras and reviews the footage manually to detect incidents. RIDOT also has access to the radio of state police, which is another important source (i.e., user reported data) for incident information. Although user reported data via Waze are a little noisy (see discussion in the second paragraph of Section 2.1.2.4.1), they are very useful to DOTs due to its coverage.

Overall, state DOTS are satisfied with probe data given that they are maintenance free and cover a very large area. Issues with probe data include low sampling rate in rural areas, reliability (sampling rate for the same segment changes over time), and data conflation.

It is estimated that by 2023 90% of new vehicles in the United States will be shipped with embedded connectivity. The near real-time connected vehicle trajectory data provides a source with many great possibilities for improving the understanding of traffic flows and developing advanced traffic management strategies. State DOTs should carefully monitor the development of connected vehicles and their impacts on traffic data collection and TSMO.

2.1.2.5 Other Mobile Device Location Data

Since almost every driver now has a cellphone, being able to accurately locate cellphones can help estimate vehicle speeds. Cellphones can be used to generate user reported data (Section 2.1.2.4.1) and aggregate probe data (Section 2.1.2.4.2). Besides cellphones, there are many other mobile devices such as tablets and smart watches. It is mentioned in a 2019 news article [*6*] that "Every minute of every day, everywhere on the planet, dozens of companies — largely unregulated, little scrutinized — are logging the movements of tens of millions of people with mobile phones and storing the information in gigantic data files." It is estimated that there is a $12 billion market [*7*] for such data. There is a long list of companies that use mobile device location data for various applications, including AirSage, SkyHook, Cuebiq, and SafeGraph.

There are mainly three types of mobile device location data: cell tower triangulation data, mobile device GPS location data, and mobile carrier data. These data sources are further detailed below.

2.1.2.5.1 Cell Tower Triangulation

Each cellphone must be connected to at least one cell tower. The distance between the cellphone and cell tower can be estimated by measuring the strength of wireless signals transmitted between them. Since the cell tower location is fixed and known, the location of the phone can be narrowed down to a circle. If the phone is communicating with two cell towers, its location can be further narrowed down to two points. With three cell towers, theoretically the phone location can be uniquely determined. However, the accuracy of distance estimation based on wireless signal strength is not perfect. Even with three cell towers, a cellphone can usually be located within an area of ¾ square miles.

Mobile device location data obtained via cell tower triangulation usually is not very precise, and cannot be used for calculating speed, travel time, etc. However, it can be used to estimate time-dependent OD data. OD data is important for understanding travel demand. It can be used together with traffic simulation tools to answer questions such as what may happen if a road segment is shut down due to major accidents or construction.

2.1.2.5.2 GPS Location or Location Based Service (LBS) Data

Smartphone are all equipped with GPS, which provides more accurate location information than cell tower triangulation. Most mobile device location applications are based on GPS location data. GPS location data (e.g., obtained via navigation and other LBS Apps) can be used to derive travel time and speed (see Section 2.1.2.4.2). It also has many other important applications. For example, GPS location data can be used to derive trip generation rates and help businesses find optimal retail locations. Similar to cell tower triangulation data, GPS location data can be used to derive OD data. In addition, it can potentially be used to drive trip chain, travel mode, and route choice information, which is important for both transportation planning and TSMO.

A major issue with GPS location data is latency. Unlike the connected vehicle trajectory data in Section 2.1.2.4.3, GPS location data in many cases are not immediately available to end users. An exception is the GPS location data obtained via navigation Apps, which is aggregated and made available to App users in real time. It would be ideal if Google and Apple could share their real-time navigation App data (e.g., trajectory, travel time, incidents) with state DOTs to improve TSMO (e.g., incident detection). Even historical GPS location data can be useful for DOTs. They can be used to identify and prioritize bottlenecks, safety hazards, etc.

2.1.2.5.3 Mobile Carrier Data

Through either cell tower triangulation or smartphone GPS, mobile carriers can have the location information of their subscribers. This data source is likely to have a much higher sampling rate of users than other sources such as probe vehicles and generate more accurate measurements of traffic speed and travel time.

The wireless communications solution for future connected vehicles is not clear at this moment. It could be based on DSRC, 5G, or 6G. If 5G or 6G is used as the backbone for connected vehicles, mobile carriers will play a critical role and will have access to a vast amount of vehicle related data, including the trajectory data discussed in Section 2.1.2.4.3.

Some mobile carriers have also shown great interest in ITS and smart cities. Verizon partnered with some cities (e.g., Boston and Kansas City) to install sensors in the pavement and connect cameras to traffic lights for detecting traffic [*8*] and improving traffic signal operations. AT&T also has an "AT&T Smart Cities Structure Monitoring" program, which adds AT&T LTE-enabled sensors to the existing lighting infrastructure in some U.S. cities (e.g., Atlanta, Dallas) to monitor traffic and parking, and detect gunshots [*9*].

2.1.2.5.4 StreetLight Data

StreetLight is essentially based on mobile device GPS location data (Section 2.1.2.5.2). It applies AI algorithms to integrate mobility device location data provided by sources such as Cuebiq [*10*], data from DOT permanent traffic counting stations, etc. to estimate AADTs, bike and pedestrian volumes, OD, and so on. It is listed here in a separate subsection because it is being used by both MaineDOT and MassDOT.

2.1.2.6 Social Media

Some researchers proposed to use data from social media such as Twitter for TSMO purposes. They use the Natural Language Processing (NLP) method to extract traffic incident related information from social media feeds. For instance, after identifying an incident-related tweet, words related to "when", "where", and "how bad the incident is" will be extracted and analyzed if they exist. An issue with this data source is that incidents are not guaranteed to be posted in a timely manner and with sufficient details to accurately determine their nature and location information.

Most state DOTs (e.g., NHDOT and MassDOT) use Twitter and Facebook to share traffic information with the public, not to model data (e.g., major crashes, traffic disruptions due to snowstorms) from those social media platforms.

2.1.2.7 Autonomous Vehicles

Each autonomous vehicle is equipped with a suite of sensors, which generate a vast amount of data each day. It is not a secret that Tesla collects data from its vehicle owners [*11*] to improve their self-driving algorithms. Even human-driven vehicles are now collecting and sharing data (e.g., Wejo and Otonomo). Some autonomous vehicle companies such as Lyft and Waymo have made part of their collected data (e.g., LiDAR, vehicle trajectory, camera) available to the public.

Autonomous vehicles can sense the surrounding traffic and generate more detailed information than vehicle trajectories. They can detect damaged traffic signs and guardrails, potholes, distracted pedestrians, aggressive drivers, debris on the road, etc. However, they are not obligated to share anything with state DOTs. An interesting question is whether it is ethical for car manufacturers to collect data from drivers but do not share it with public agencies (e.g., state DOTs) for the benefit of drivers. The same question can be brought up to tech companies that collect mobile device location information.

2.1.2.8 Artificial Intelligence (AI)

AI technologies are well known for being data hungry. They often require a tremendous amount of data for model training and validation. On the other hand, AI is also an important tool for generating data. With cameras, drones, LiDAR, etc., transportation agencies have accumulated enormous images, video, and point cloud data that they sometimes cannot effectively utilize. Well-trained AI models can be used to turn such data into useful information. For example, traffic counts and assets can be derived from videos and LiDAR point cloud, respectively. AI algorithms are also widely used in autonomous driving to process camera, LiDAR, and microwave sensor data.

2.1.2.9 Summary

The data sources discussed are summarized in Table 2-1 below. The discussion above suggests that the landscape of traffic data collection has changed substantially in the past two decades given the advancements in sensors, wireless communications, the Internet of Things (IoT) and smart cities, GPS and mobile devices, connected vehicles, and automated driving. Among them, mobile devices and GPS probably have the most profound impacts on traffic data collection. They significantly expand the coverage of traditional sensors (e.g., loop detectors, cameras) and provide a maintenance free approach for transportation agencies to collect detailed data elements such as vehicle trajectory, wipers change, seat belt change, and autonomous emergency braking. Another important front is the wide applications of AI technologies in sensor data processing, generating valuable traffic measurements for data-driven decision making.

Table 2-1. Traditional and Emerging Data and Data Sources for TSMO

	Traditional Data & Data Source
Highway	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
Incidents and Crashes	Incident/crash records (e.g., location, time, duration), highway safety patrol records, 511 phone records
Arterial	Traffic signals, vehicle detectors, cameras, data from Automated Traffic Signal Performance Measures (ATSPM) system, queue length from drone.
Transit	GTFS, transit fare collection data (e.g., smart card, Mobile ticket), traffic cameras, APC data, ridership, etc.
Parking	Static (e.g., location and # of lots) and dynamic data (e.g., parking duration), parking fee data, Mobile parking app data
Assets	*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
Maintenance & Work Zone	*Maintenance*: real-time locations and speeds of plow trucks, National Weather Service Data *Work Zone*: smart work zone data, location, duration, configuration, etc.

New and Emerging Data and Data Source
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)

2.2 Potential AI Applications in Modeling TSMO Data

In the Sources Sought Notice 693JJ3-21-SS-0013 released by the FHWA [*12*] in July 2021, the following applications of AI in TSMO have been identified:

- Predict/detect traffic incidents efficiently and proactively using AI and multi-source/multi-sensor data and generate response plans,
- Predict multimodal delays in real-time using AI,
- Model urban network traffic as completely as possible using AI,
- Optimize signal timing plans offline to service all modes of transportation by predicting vehicle and pedestrian arrivals, queues, and delays,
- Optimize traffic signals in real-time using AI,
- Enhance ramp metering strategy to rapidly adapt to anticipated or predicted conditions,
- Use AI techniques to validate and verify datasets,
- Use AI techniques to detect work zone location, schematic, and hazards; alert construction crews; and disseminate traveler information,

- Detect and predict queues and shockwaves to harmonize speeds for reducing work zone crashes and delays,
- Predict road surface conditions before they become dangerous and respond accordingly,
- Proactively identify target speeds, lane assignments, ramp metering rates, etc. for improved traffic flow and throughput,
- Collaborate across agencies in real-time using Decision Support Systems (DSS) and Knowledge Based Expert Systems (KBES),
- Improve situational awareness by fusing data from multiple sources and multiple sensors across the region, and
- Test advanced traffic management and connected & automated vehicle technologies.

Based on the above ideas, the team further identified the following more specific potential applications of AI for the NETC project panel and the six New England state DOTs to consider. Two of these topics are considered for the case studies in Phase II of this project.

- Use AI for data modeling and data-driven decision making:
    - Some DOTs (e.g., VTrans, MassDOT) have deployed Automated Traffic Signal Performance Measures (ATSPM) systems. How to make the full use of ATSPM data is an interesting and timely topic.
    - Integrate data from the National Oceanic and Atmospheric Administration (NOAA), weather stations, and sensors installed on vehicles (e.g., plow trucks) for Winter Maintenance Decision Support Systems. One possible application is to determine the optimal amount of salt to be applied.
    - Queue/slow moving traffic detection using detailed trajectories generated by connected vehicles.
    - Use incident data to optimize safety patrol schedules. Currently, highway patrol schedules are decided based on human intelligence not algorithms.
    - Use location-based data to estimate traffic volumes, especially for low-volume roads and intersections.
    - ***Use AI to model data from multiple sources, including probe and/or connected vehicle, for safety applications.***

- Use AI for data processing and reduction:
    - ***Automatically detect incidents/risky events and collect traffic data using advanced sensors such as thermal traffic cameras.***
    - Develop strategies to conflate and integrate data from different sources.
    - Use AI to process LiDAR data and automatically extract asset information such as asset type and location.

2.3 Interviews with Experts

In addition to reviewing existing and emerging data sources, the team conducted interviews with staff from the six New England state DOTs. The interviews covered a range of topics, including their data needs, data analysis, archiving, sharing, security, and privacy protection practices, among others. The team also interviewed staff from the Federal Highway Administration, as well as from the DOTs of Texas, Oregon, Virginia, and Delaware, along with the Eastern Transportation Coalition (previously known as the I-95 Corridor Coalition). The questions posed during these interviews and the results are summarized in the rest of this section.

2.3.1 Data and Needs

2.3.1.1 Data Sources

Question: Are we missing any major data elements/sources in Table 2-1?

Most DOTs found the data elements in Table 2-1 to be very comprehensive. One possible way to further enhance Table 2-1 is by providing more detailed information about specific data products, data locations, and temporal limitations (e.g., real-time data, archived data).

2.3.1.2 Data Needs

Question: Any existing and future data needs for TSMO (e.g., estimating OD in addition to segment AADTs)?

The identified future data needs for TSMO include:

- Integration of data from different sensors (e.g., loop detectors, AVL, pavement sensors), at various rates, and stored in different databases,
- Connected vehicle data (e.g., detailed vehicle trajectories),
- Data related to public and private truck parking spaces and availability on major highway corridors,
- Better data sharing with travelers, such as broadcasting traffic signal timing information to drivers,
- Data for determining incident duration, clearance time, and secondary incidents, as well as separating incidents from recurring congestion. A reliable data source for identifying secondary crashes is currently unavailable, and such incidents are sometimes under-reported in police reports (e.g., noted as primary incidents). Recurring congestion makes it challenging to determine when an incident is cleared,

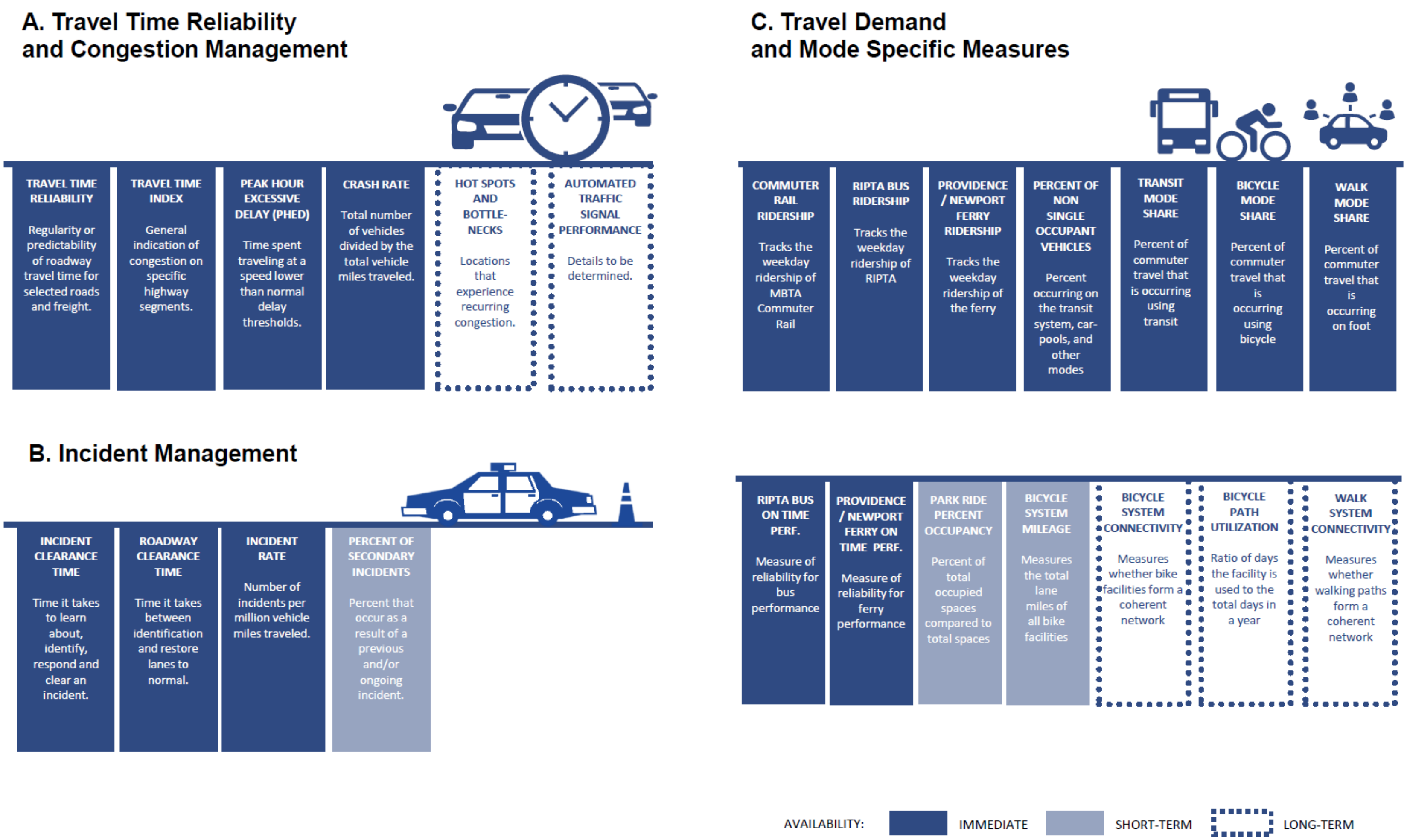


Figure 2-2. TSMO performance measure needs provided by RIDOT.

- Reliable queue length (or unexpected stops/slow-moving traffic on highways),
- Estimate highway and arterial traffic volume, density, and capacity from different locations in real time. Existing probe data only covers speed and travel time,
- OD data will significantly increase DOTs' ability to predict transportation system use and system response to demand changes. Reliable OD data, together with digital twins, AI, and simulation could substantially improve TSMO by quickly identifying changes, understanding their impacts, evaluating and recommending options, implementing those options, and continuously monitoring and adapting to the impacts of the changes,
- DOTs need more detailed and real-time condition information about ITS assets, and
- CAVs will generate a considerable amount of data that can be used for TSMO applications. On the other hand, CAVs will need precise data for making safe, efficient, and eco-friendly driving decisions. In the future, variable message signs may not be needed. Instead, DOTs need to provide traveler information in digital formats that can be unambiguously interpreted by CAVs.

Additionally, RIDOT provided a detailed list of performance measures needed for TSMO, presented in Figure 2-2. While this list is not specifically for data needs, the listed items will depend on data and will guide researchers and engineers to find appropriate data sources to support the development of such performance measures. One example is the percentage of secondary incidents. Such data does not currently exist in police reports and is difficult to obtain. However, it may potentially be obtained from new and emerging data sources, such as connected vehicle data.

In addition to data needs, the research team identified some needs for data analysis methods, including:

- Data conflation is a major issue faced by many DOTs. For probe data, such as INRIX, TomTom, and HERE, data vendors only provide aggregated information, such as segment speed and travel time. State DOTs adopt what these vendors provide and lack details on how the data are aggregated. The length of each segment is also decided by the vendors. Different vendors often employ different standards or methods to divide roads into segments. When DOTs possess data from multiple vendors, they encounter the challenge of reconciling data aggregated using different segment definitions (e.g., Linear Referencing System (LRS) used by DOTs to manage pavement conditions, bridges), which is not a trivial task. Additionally, state DOTs lose the opportunity to extract more granular and useful information from the aggregated probe vehicle and GPS data. Using automated incident detection (AID) as one example, DOTs may want to have short segments in areas prone to incidents (ideally in all areas if computational power is not a constraint). With short segments, changes in individual vehicles' speeds and travel times can be quickly reflected in the corresponding segment measures. On the other hand, providing aggregated data and obscuring the details to some extent is beneficial to DOTs, as they often do not have the human resources needed to handle the large volume of raw trajectory data and extract critical information from them.
- Data aggregation and mining are important. DOTs need a systematic way of integrating data from different sources (e.g., loop detectors, CCTV cameras, Waze, HERE, INRIX) and generating useful data for performance measurement, incident detection, traveler

information system, etc. Note that these data come in at varying rates, latencies, and accuracies. For example, Waze incident report data can be quite noisy with multiple reports for one incident, and the reported locations of the same incident may not match well. Another example is loop detector data, which can be used to complement and verify INRIX data. However, unlike INRIX data, loop detector data are often not streamed to the HOC in real time and are only for fixed locations instead of segments.

- Innovative data analysis methods and approaches are needed. The existing data analysis methods can be adapted and applied in different ways, depending on the available data. Using AID as one example, traditional AID methods are based on loop detector data and focus on identifying critical patterns/thresholds in terms of spot speed, volume, and occupancy measured at up- and downstream locations. With INRIX and HERE data, the same statistical or machine learning methods can be used but applied in a different way. The focus now is to compare the current and previous speeds of adjacent segments, taking segment length into consideration. If Wejo data (i.e., raw vehicle trajectories in real time) are available, the AID problem will become identifying changing points in time series (i.e., vehicle trajectories). Therefore, evolving data sources will necessitate new methods and innovative applications of existing methods.
- Data sharing and brainstorming: For example, LiDAR data can be used to derive accurate ramp geometry information, which can be combined with drone captured vehicle trajectories to identify safe entrance speed for speed advisory applications. The treasures hidden in various datasets require creative thinking and analysts with both a data science background and transportation engineering domain knowledge.
- With real-time data at more granular levels, we need to know how to process and store the data, and ensure that we do not overwhelm our communication and computing systems.

2.3.1.3 Data Quality

Question: Are you satisfied with the existing data quality and reliability? For example, many state DOTs have purchased Waze, INRIX, and StreetLight data. The quality of these datasets needs to be rigorously checked. Given their accuracies, they can then be used for appropriate applications.

DOTs in general are satisfied with probe data, which appears to be reliable on high-volume roadways (~8-10% penetration rate). For low-volume roadways, traffic data can be less reliable. The dependability of probe/crowdsourced data, such as Waze and INRIX, relies heavily on the number of users or data contributors. Therefore, such data for low-volume roads (e.g., rural roads) with few Waze or GPS users could pose challenges. Although averaging data over extended time periods can partially address this issue, the optimal solution is to increase the number of data contributors.

Probe and crowdsourced data are widely utilized by DOTs, with an awareness of their limitations but a recognition of the benefits derived from their extensive coverage and high temporal granularity. DOTs express interest in a systematic assessment of the quality of such data, considering factors like sampling rate, accuracy, and confidence levels. Additionally, there is a need to maintain and modernize traditional sensors, to monitor their conditions and stream data to highway operation center (HOC) in real time.

A gap exists between infrastructure assets management (e.g., bridges, pavement) and ITS/TSMO assets management. Maine, Vermont, and New Hampshire are actively working to bridge this gap, aiming to develop an asset inventory for their ITS and TSMO-related assets. This effort can be coupled with the modernization of traditional sensors. Unlike physical infrastructure like bridges or tunnels, ITS assets may appear normal but can stop functioning suddenly. Therefore, real-time status monitoring is crucial.

It is important to assess the quality of data from third-party vendors. Such evaluations ideally should be conducted at a regional level instead of by individual DOTs, as the cost of a thorough study can be high. At times, traditional data sources may be entirely absent. In such cases, considering nontraditional data sources is beneficial and likely better than having no data at all. DOTs typically lack a specific protocol for determining which data source(s) can be used for design and other purposes. They approve the use of nontraditional data source(s) on a case-by-case basis.

Several informal studies have been conducted to compare the traffic counts from StreetLight with field measurements. One concern was that StreetLight traffic counts were calibrated based on permanent counting stations on major highways. The accuracy of StreetLight traffic counts for local roads was not as high as that for major highways.

All DOTs have stationary sensors (e.g., loop detectors, microwave sensors) to collect traffic counts, speeds, etc., on major highways. Although this data could potentially be used to calibrate probe data, there are some issues: (1) these sensors are not connected to the highway operation center (HOC), and traffic operators do not have real-time information on the status of these sensors. Sometimes, DOTs lose months of data from a sensor before realizing it; and (2) most sensors are on interstate highways, making it challenging to validate probe data for other highways. Adding new sensors at strategically chosen locations on local highways for data validation would be beneficial. Additionally, being able to monitor sensor health status, similar to what the ATSPM system does, is important. Some DOTs have found that TomTom data matches their sensor data well, potentially due to TomTom's integration of information from diverse sources.

The Eastern Transportation Coalition has conducted many validation studies of the probe data primarily for highways. They plan to expand such efforts to cover arterials and local roads over the next few years.

2.3.2 Emerging Data Sources

2.3.2.1 Data Collection Methods

Question: Short- and long-term plans to meet agency's data needs while minimizing the life-cycle cost (e.g., relying on 3rd-party vendors vs. investing in data collection infrastructure) and maximizing the data collection system robustness and reliability (e.g., reliability of crowdsourced data depends on the number of data contributors).

DOTs do not have a consensus on the direction of future data collection methods. Currently, most DOTs utilize a combination of methods, including their own data collection infrastructure and third-party data vendors. It is likely that DOTs will continue with this dual approach in the near future.

Some DOTs prefer third-party data vendors over DOT's own data collection infrastructure due to considerations of life cycle and maintenance costs. Maintaining DOT's own infrastructure often requires setting up short-term work zones. Using data vendors such as HERE does not require significant infrastructure investment, and it reduces personnel risk during infrastructure maintenance. In some states, Bluetooth is primarily used in smart work zones for travel time data collection. Meanwhile, traffic detection is shifting to systems like radar and cameras, and in-pavement loop detectors are no longer being installed.

Traffic counters and weather stations are still being added in some states. The collected data can be compared with data products from third-party vendors, even though such data are limited to fixed locations. There is no clear consensus on the preferred direction (e.g., investing in their own data collection infrastructure vs. purchasing data from third-party vendors) for the future. DOTs acknowledge the benefits of leveraging data from private vendors, as it entails no initial investment and avoids the costs and challenges associated with maintenance.

Another challenge involved in making this decision is the constant evolution of technologies, which is difficult for public agencies to keep up with. Some technologies may become obsolete quickly, posing a risk to heavy investments. From this perspective, it makes sense to shift the risk to private companies and procure data products from them. In most cases, private companies can adapt to technological changes more swiftly than public agencies. The traditional model of design, build, and maintain may not be as effective in the future. It is likely more beneficial to adopt a data-as-a-service model.

The maintenance of ITS and data collection infrastructure is a significant challenge for DOTs. When DOTs own the infrastructure, they must invest in staff training, especially as new technologies are constantly introduced. Sometimes, hiring vendors becomes necessary for maintenance. Additionally, ensuring consistency in technology is critical. Otherwise, DOTs might end up with various types of devices on the road, requiring them to keep a broad range of spare parts and their maintenance staff to be familiar with diverse technologies. Simply letting vendors handle operations and maintenance is not a perfect solution either. This approach gives rise to data integration issues. For example, some DOTs encountered difficulties in obtaining data from toll road authorities. Moreover, different vendors use distinct data structures and formats, creating barriers for integrating data from various sources for in-depth data analysis.

DOTs are constantly facing the question of whether to repair and install additional traffic sensors or simply purchase data from private companies. A main reason for this dilemma is the maintenance cost associated with existing traffic sensors. If companies can provide accurate volume data, many DOTs are likely to favor third-party data vendors to avoid the hassle and costs of maintaining existing sensors. Another issue with buying data from vendors is their reliability, which is significantly affected by the sample size. In Northern New Hampshire, the sample size for probe data is lower than Southern New Hampshire, and understandably, the data quality is also lower.

In addition to building and investing in data collection infrastructure, maintenance is critical. Maintaining a state of good repair of ITS infrastructure is important but has not received enough attention. DOTs are not only concerned with the price of data products provided by vendors; they are also concerned with the life-cycle costs for maintaining their own data collection infrastructure.

2.3.2.2 Emerging Data Sources

Question: Ridesharing companies (e.g., Uber, Lyft), logistics companies (e.g., UPS), Cell Phone data (e.g., from Verizon), and Connected and Automated Vehicles are producing tons of data each day. Any plans to utilize data from emerging sources for TSMO?

Some of these datasets are also included in the probe data such as HERE and INRIX. In this sense, they have been widely utilized by DOTs. Overall, DOTs are aware that there are many emerging data sources beyond the probe data that could be useful. But DOTs have not yet extensively incorporated those. One reason is that DOTs have not seen convincing examples demonstrating how the emerging data could significantly benefit TSMO.

Although some mobile carriers have approached DOTs, specific applications have not been implemented yet. Cameras have been widely used by DOTs for generating traffic counts. DOTs acknowledge that, compared to probe vehicle data, this method requires equipment installation and is difficult to scale up.

Autonomous vehicle (AV) companies have collected a vast amount of data, including high-resolution vehicle trajectories of AVs and nearby vehicles, and roadway conditions derived from camera/LiDAR data. There is no clear legislative guidance on what data could and should be shared. AV companies are concerned about data privacy, which hinders data sharing. On the other hand, public agencies do not know what data is available, what to request from AV companies, and how to securely store and analyze the massive data. Data from such tech companies will become increasingly important, which may cause some of the existing ITS technologies to be obsolete (e.g., variable message signs). At the moment, DOTs lack a formal plan on how to prepare for this future.

2.3.2.4 Data Sharing

Question: Transit agencies (e.g., GTFS data), DOT's maintenance vehicles (e.g., plow trucks), and Automated Traffic Signal Performance Measures (ATSPM) systems can also be used to generate a lot of valuable data. Any plans to coordinate different DOT divisions?

Many DOTs are either using or planning to use AVL and their maintenance vehicles for data collection. Several DOTs have deployed the ATSPM system and are interested in exploring how such data can be used for improving traffic operations and safety.

There is a consensus that data sharing among different DOT divisions would be beneficial. However, there are not many data-sharing activities among different divisions of DOTs. Several reasons could contribute to this. First, there are not many urgent needs to share data. Second, individuals are not familiar with the types of data owned by other divisions own. Third, there was a lack of full understanding regarding the value of the data they possessed for other divisions.

2.3.3 Data Integration and Analysis

2.3.3.1 Artificial Intelligence (AI) and Data Analysis

Question: AI and edge computing technologies are making it possible to extract accurate traffic data using existing traffic cameras (or drone-mounted cameras) and turn cameras into smart sensors, better utilizing the existing data and data collection infrastructure. Is your agency investigating/interested in such technologies?

There are many potential AI and edge computing applications such as traffic camera data processing. Issues related to camera data processing, including software licensing, access, privacy, and recording, have limited DOTs' ability to apply them on a large scale. Besides cameras, AI has been used for modeling TSMO data, although this is usually outsourced rather than done in-house by DOTs. Another important application area for AI and edge computing is CAV. Edge processing could help with the structuring and application of CAV data, allowing them to be more effectively utilized for downstream analysis.

With the assistance of consultants, most DOTs have experience with using AI for various applications, including pedestrian detection in tunnels using existing cameras, processing drone videos, and counting traffic. Some DOTs do not have an immediate plan to use AI and edge computing technologies. They are concerned about the significant efforts needed to upgrade and expand the existing camera network. Overall, although most DOTs have not conducted in-house case studies using AI and edge computing, they express interest in innovative technologies that can enhance efficiency and reduce costs.

As part of DelDOT's AI-ITMS project, they are developing and testing machine vision capabilities. Their goal is to gradually replace in-pavement detection technologies with "non-intrusive" detection technology. The entire transition is expected to take many years, given that DelDOT is responsible for between 20,000 and 30,000 loop detectors. DelDOT is reviewing detector design requirements to fully leverage the potential of AI.

2.3.3.2 In-House Data Analysis or Outsourcing

Question: Should data analysis be done in-house or by consultants? How to integrate the data analysis efforts of different DOT divisions?

Most DOTs engage in data analysis work both in-house and through consultants, depending on the type of work, available resources, and their workload. Some DOTs have established a dedicated data analysis team within their agency. It is unclear to what extent the in-house analysis would cover. For the in-house data analysis, the priority areas are like to be crash data and traffic data analysis (e.g., probe data, ITS data).

Many DOT vendors store their data on cloud servers. Therefore, IT support is important for DOTs to integrate data from vendors and other sources, such as downloading those datasets to a local server and integrating them. Data vendors typically are profit-driven and may not always be motivated to integrate their data with DOT data. As a result, some of the data integration work may need to be performed by DOT staff. This issue is closely related to workforce development, DOT staff training, etc.

At times, DOTs prefer to conduct data analysis in-house if resources are available. While hiring consultants can be helpful, it is beneficial to have experts in DOTs knowing what consultants are doing and what DOTs are paying for, underscoring the importance of developing in-house data analytics capabilities.

2.3.3.3 Investing in Data Analytics

Question: Short- and long-term plans for investing in data analytics and workforce development.

Although DOTs may not have specific plans to invest in data analytics and workforce development, they all recognize the importance of such efforts. Several state DOTs have recently hired data analysts/scientists or created related positions/offices.

2.3.4 Data Archiving, Sharing, Security, and Privacy

2.3.4.1 Data Management Protocols

Question: Any protocols for how data should be reduced and how long should a dataset be archived (in the original form or reduced/processed form)?

How long a dataset should be retained generally depends on the nature of the data and agency data retention policies. DOTs have established policies governing data retention, privacy, and security. However, many DOTs face challenges in handling the growing volumes of data and extracting insights from the massive datasets. Data storage is a major challenge for the IT departments of state DOTs. It is critical to involve IT in data collection, storage, and sharing processes to ensure that data sharing align with agency privacy and security policies.

2.3.4.2 Measures for Protecting Privacy and Security

Question: What kind of measures is in place to protect privacy and security?

All DOTs recognize the importance of data privacy and security. Some have a position specifically for handling data and policy. Some have their IT department work closely with the TSMO division to handle data security issues.

2.3.4.3 Data Sharing

Question: Standards and protocols to guide practices such as: what kind of data should (or should not) be shared and how to share them (e.g., using Amazon Web Services or DOT owned servers)?

Most DOTs use both third-party cloud services and in-house servers for data storage. Each state has its own record retention policies that apply to the collected data. These policies specify how long records or data should be retained and when they can be destroyed. New system vendors often host data on the cloud, and they need to comply with state data security requirements while providing a data retention plan. In some cases, hosting data on the cloud is the only option.

Most data collected by public agencies is subject to the Public Information Act. Agencies typically avoid collecting data with personally identifiable information (PII) or business confidential information, making it easier to share the collected data.

2.3.5 Stakeholders and Workforce

2.3.5.1 Stakeholders

Question: List of stakeholders? We may want to identify stakeholders within DOT to support the investment in data collection and analytics.

Stakeholders identified include:

- Internal: IT, ITS, TMC, safety/traffic, GIS, maintenance, planning, performance measures, communications to the public, legal. Essentially, all divisions within a DOT should be considered key stakeholders for data collection and analytics because any decision-making process should be data-driven. It is important to involve IT, especially for system and data integration, as well as maintenance.
- External: 911, transit, emergency response/first responders, state police, MPO

2.3.5.2 Organizational Structure

Question: What kind of organizational structure changes are both feasible and necessary to better prepare DOTs for the future? For example, Iowa DOT has an Office of Analytics; Arizona DOT has a Data Analytics section responsible for reporting, maintaining, collecting, analyzing, and visualizing the data on roadways in Arizona; and Florida DOT has a Transportation Data and Analytics Office that is FDOT's central clearinghouse and the principal source for highway, traffic, travel time, multimodal, and freight and passenger data information.

There is no one-size-fits-all solution regarding organizational structure, as it should be based on the needs and available resources of the specific organization. Data analyst positions are being created within most DOTs. Sometimes, these positions are distributed across different divisions of a DOT, leading to coordination problems. Establishing a central data analytics division to coordinate efforts is critical. Creating a central data office/section will provide data analysts with a sense of belonging to a core group, making it easier for them to exchange ideas and learn from each other. In some cases, DOTs also assign the central data office with tasks such as innovative research and grant applications.

2.3.5.3 Workforce Development Needs

Question: Workforce development needs and current strategies?

Workforce development is crucial, and providing competitive salaries is essential to attract and retain skilled data analysts. DOTs should prioritize hiring data analysts with a background in both transportation and data science. Data scientists without a civil/transportation background may struggle to understand the nuances, leading to potential problems. Consider training civil students and encouraging them to take data analytics courses. This approach allows new hires to focus on transportation data analytics and commit to it.

Mainstreaming the importance of data and data analytics would be a great strategy. Additionally, training the existing workforce to acquire the necessary data skills is another viable approach. Some DOTs offer in-house training courses developed by consultants.

2.3.5.4 Additional Thoughts

Question: Additional Thoughts: Any thoughts you have related to data driven TSMO applications.

There is no boundary when it comes to collaboration in data-driven TSMO applications. Collaboration among divisions within DOT is crucial, and inter-agency collaboration in this area is equally important. It is also desirable to investigate the impact of Connected and Autonomous Vehicles (CAV) on TSMO.

3.0 Assessment of Data Needs, Emerging Data Sources, and Data Processing and Analytics

Building upon the review and interview results in Chapter 2, this chapter offers recommendations in four sections: data needs, emerging data sources, data processing and analytics, and others, respectively, to state DOTs in the New England region. Based on the recommendations and the potential AI application topics identified in Section 2.2, three case study topics are selected, which are

- Speed behavior on highway horizontal curves,
- Speed and lane-changing behavior prior to highway work zone, and
- Network-wide speeding activity analysis using probe vehicle data.

The first two case studies aim to demonstrate the advantages of using portable sensors to collect detailed vehicle trajectory data for studying driver behavior under different circumstances, which corresponds to recommendations #2 in Table 3-1 and #10 in Table 3-2. They also showcase the capability of using AI to analyze and reduce the collected trajectories (see recommendation #20 in Table 3-3), generating meaningful conclusions. The third case study is to show the power of combining location-based service (LBS) data and traditional road inventory data to study driver speeding activities at a network scale, which reflects recommendations #9 in Table 3-2 and #15 in Table 3-3.

3.1 Recommendations on Data Needs

Table 3-1. 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. The penetration rates of emerging connected vehicle data (e.g., Wejo, Otonomo) are continuously growing. DOTs should not invest in additional roadside sensors such as Radar and camera for incident detection, TIS, and travel time estimation purposes, unless it is for areas that are poorly covered by the above data sources, or these data sources are unreasonably expensive.
2	• Vehicle trajectories	Safety is an important aspect of TSMO. Safety analysis has been done reactively and based primarily on historical crash data. It is interesting to use vehicle trajectory data to proactively evaluate safety risk in the future. Vehicle trajectories from connected vehicles (e.g., Wejo, Otonomo) cover a large area but only a small sample of all vehicles.

ID	Data Needs	Recommendations
		Roadside sensors (e.g., high-resolution Radar, camera, LiDAR) cover a short road segment but can capture all passing vehicles. Both data sources are important for proactive safety risk analysis. DOTs are encouraged to investigate both data sources (i.e., connected vehicles and roadside sensors). When investing in new roadside sensors, DOTs are encouraged to consider sensors that can generate vehicle trajectories during both daytime and nighttime.
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. This may potentially be done for passenger vehicles and heavy trucks separately. Such OD information is not only important for planning purposes, but also will substantially increase DOTs' ability to understand driver behavior and predict transportation system use and response to disruptions. TSMO and planning divisions are encouraged to work together on deriving and evaluating OD information using LBS and vehicle ReID data.
4	• Traffic volume and capacity	Existing probe data only covers speed and travel time. Estimating traffic volume and capacity (e.g., under different weather conditions) can be very interesting. Such information can be used together with OD to predict when congestion (not caused by incidents) may occur and the corresponding queue growing and dissipating processes. Although some data vendors claim that they can provide traffic volume data such as segment AADTs and intersection turning movement counts, the accuracy of such data needs to be thoroughly evaluated, especially for rural areas where there are not many permanent traffic monitoring stations to provide calibration data. Existing traffic monitoring stations are mainly on major highways to satisfy the HPMS requirement. 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. For example, retrofitted loop detectors and camera + AI technologies can

ID	Data Needs	Recommendations
		differentiate among vehicles such as flatbeds, dry goods semitrailer, tankers, refrigerated trucks, and recreational vehicles. DOTs are encouraged to consider such technologies. DOTs are not encouraged to install new loop detectors due to their high installation and maintenance costs. However, retrofitting existing loop detectors can extend their service life and generate more useful information.
6	• Travel time	For areas without good probe data coverage, DOTs are encouraged to consider installing Bluetooth sniffers/readers to collect travel time data. DOTs can also install sensors to read E-ZPass transponders. For example, New York City has been using E-ZPass transponder data to track vehicles and measure travel time.
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. Tracking such data is important for ensuring system safety (e.g., a malfunctioning traffic signal can cause accidents) and developing preventative maintenance plans. It is strongly recommended that DOTs invest in this area. Some of the data does not need to be transmitted to the Traffic Management Center (TMC) in real time. For instance, the detector condition data may be reported every hour instead of minute to the TMC.

3.2 Recommendations on Emerging Data Sources

Table 3-2. Recommendations on Emerging Data Sources

ID	Emerging Data Sources	Recommendations
9	• Connected vehicles and travelers	It may take many years for automated vehicles to occupy the streets. However, connected vehicles are very close to us now. Many auto makers have already been collecting data using their new vehicle models. These datasets are packed and sold by companies such as Wejo and Otonomo. They

ID	Emerging Data Sources	Recommendations
		include vehicle trajectories as well as event logs such as wiper speed and activation/deactivation. Travelers nowadays depend heavily on mobile devices and various Apps, even knowing that their privacy is at risk. These mobile devices and Apps are contributing critical data (e.g., StreetLight) for understanding traveler behavior under different traffic conditions. 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. DOTs should explore the potential applications of such datasets and their impacts on traffic operations and safety. DOTs should also work with legislators to push technology companies such as Google to make such datasets available to public agencies. Such datasets are collected from the public and probably should be made available for free or at a reduced price to public agencies for the benefits of whoever contribute the data.
10	• Sensors powered by AI and edge computing: thermal and RGB cameras, loop detectors, LiDAR, Radar, E-ZPass transponder	Advanced sensors powered by AI and edge computing technologies will be another important data source. 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. New York City has been using E-ZPass transponder data to estimate travel time. DOTs are encouraged to explore the potential of traditional and new sensors mounted on portable platforms. These portable platforms can be moved to different locations to (1)

ID	Emerging Data Sources	Recommendations
		collect trajectory data for safety studies, and (2) collect speed and travel time data to complement the probe and connected vehicle data in rural areas.
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. There are already commercial products based on probe (e.g., INRIX) and connected vehicles (e.g., Wejo, Otonomo) data. It is anticipated that there will be commercial datasets available in the future that are collected by semi- or fully automated vehicles. DOTs should take this potential data source into consideration when making future data and data collection infrastructure decisions.

3.3 Recommendations on Data Processing and Analytics

Table 3-3. Recommendations on Data Processing and Analytics

ID	Data Processing and Analytics	Recommendations
12	• Data quality validation	DOTs should continuously 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	It would be interesting to integrate crash history, pavement condition, and probe vehicle data to find connections among them. However, these datasets are organized using different referencing systems. Crash data is often based on x and y coordinates; pavement condition data is typically stored using linear referencing systems; while probe data is organized by segments (e.g., INRIX uses XD segments). Data conflation is a major issue faced by many DOTs and should be given enough attention.
14	• More detailed incident data analysis	With probe data such as TomTom and INRIX, DOTs can derive more detailed incident information, including

ID	Data Processing and Analytics	Recommendations
		duration, queue length, clearance time, and effects on secondary incidents. Such information can be correlated with incident characteristics such as # of lanes closed, # of vehicles involved, and injury and casualty to establish models to predict future incident impacts. In addition, probe data can be used to separate recurring congestion from incidents and for queue detection and warning. The recurring congestion information in conjunction with OD and travel mode choice (e.g., from StreetLight) data can be used to develop comprehensive transportation network improvement solutions. DOTs are encouraged to explore this area and conduct more detailed analysis of probe data.
15	• Connected vehicle data analysis	USDOT has funded three connected vehicle pilot projects. These vehicles have generated a vast amount of exciting data. In the meantime, many auto makers have already been collecting data using their new cars. These datasets are packed and sold by companies such as Wejo and Otonomo. These datasets are not aggregated by segments (like what TomTom and INRIX do) and contain more details. DOTs are encouraged to investigate such datasets and explore their applications beyond incident detection and travel time estimation. They can potentially 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. For example, StreetLight data is mainly used for planning purposes. It can provide useful OD and mode/route choice information for developing contingency traffic management plans for special events, major construction projects, and accidents. 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. Waze data is not seamlessly integrated with INRIX or TomTom data for incident detection/verification. DOTs are encouraged to explore methods to integrate such data sources and automate the process of integrating them.
17	• ATSPM data analysis	Several New England State DOTs have implemented or are planning to implement the Automated Traffic Signal Performance Measure (ATSPM) system. ATSPM allows

ID	Data Processing and Analytics	Recommendations
		DOTs to detect traffic signal related hardware and control plan issues in real time and remotely from the Traffic Management Center (TMC), identify potential causes, and quickly dispatch staff as needed. It helps to minimize the impacts of traffic signal control malfunction and improve traffic safety at signalized intersections. 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, which has not been adequately investigated. ATSPM is getting increasingly popular. DOTs are encouraged to explore such datasets for both traffic operations and safety applications.
18	• Innovative data analysis methods	Emerging data sources such as probe vehicles, connected vehicles, and ATSPM require innovative data analysis methods. For example, previous incident detection methods based on loop detectors are not applicable to probe vehicle data. DOTs should investigate innovative analysis methods to get the most out of these new data sources.
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. With the shared data, DOTs may hold data analytics competition among college and high school students to identify interesting ideas and attract students into the transportation data analytics area.
20	• AI + Edge computing for data analysis and reduction	Most DOTs struggle with the growing data volumes and how to extract insights out of the massive data. With real-time data at more granular levels, DOTs need to investigate how to best process and store the data, and how not to overwhelm communication and computing systems. For example, DOTs are encouraged to explore AI and edge computing technologies to speed up the processing of images and videos. This will significantly reduce the amount of data that needs to be transferred and stored. DOTs are encouraged to work with universities on this topic.
21	• Road Weather Information System	Although all six New England state DOTs have invested a lot in stationary and mobile weather stations, more still needs

ID	Data Processing and Analytics	Recommendations
		to be done to analyze the collected data. For example, such data can be used to estimate the optimal amount of deicing materials to be applied.

3.4 Other Recommendations

Table 3-4. Other Recommendations

ID	Others	Recommendations
22	• Collaboration among DOTs	State DOTs in the New England region face many similar issues that are unique to this region (e.g., winter maintenance). It is strongly recommended that leaders from their TSMO divisions get together regularly to share best practices, experience, and issues encountered. For procurement decisions (e.g., which probe vehicle dataset to purchase), working together will give New England state DOTs more bargaining power.
23	• Organizational changes	Since we are increasingly relying on data to make decisions, DOTs should have a central office to handle data related issues. Instead of hosting data scientists/analysts in different DOT divisions, having a central office is beneficial for workforce training, recruiting, and retaining. Employees in this data office can easily help and learn from each other, which is helpful for data modeling. The data office will be similar to the IT department. Every DOT division can have some IT experts. However, it makes more sense to have a central IT department. Almost every DOT division depends on data and needs to collect, analyze, and store data. Having an office of data analytics will allow things to be done more efficiently and professionally (in terms of data safety, retention, sharing, etc.). With a holistic view of all the DOT data assets and how they are being utilized, it would be easier to develop data sharing, retention, privacy, and security policies. This central office can discuss the data retention needs and sharing policies with individual DOT divisions.

ID	Others	Recommendations
24	• Data storage and sharing among different DOT divisions	Most DOTs use both third-party cloud services and in-house servers for data storage. Most states have their own formal and informal record retention policies that apply to the collected data. DOTs are recommended to move their data to the cloud when applicable, which will make it easy to share data and help to ensure data safety, security, privacy, and integrity. More work needs to be done to promote and facilitate data sharing among different divisions of DOTs and different agencies (e.g., Transit vs. Highway; Turnpike vs. TSMO). Having a central Data Office may help to facilitate data sharing.
25	• Workforce	Many DOTs are creating data scientist/analyst positions, and they are encouraged to continue doing this as needed. Although DOTs can always outsource the data analytics work to private companies, it is important for DOTs to understand what is being done by private companies.
26	• Personalized TIS with more dynamic and precise traffic information	A major part of TSMO is TIS. In the future, personalized data sharing with travelers would be important (e.g., sharing traffic signal timing data with connected vehicles, Alexa type of system instead of 511 phone system, recommender system that provides personalized traffic information based on a traveler's location and trip history). Google maps to some extents are doing this. With detailed and comprehensive (e.g., Transit, work zone) information, DOTs should explore what roles public agencies can play in future TIS. For example, can DOTs develop an App to share information not readily available on Google maps (e.g., scheduled work zones) with travelers in this region? Such an App can also collect travelers' mobile device GPS information (when it is within the boundary of state highways) for estimating travel time and detecting incidents. Such information will not be used for any commercial purpose unlike Google maps. Connected and Automated Vehicles (CAV) will generate a lot of data that can be used for TSMO applications. On the other hand, CAV will need precise traffic data for making safe, efficient, and eco-friendly driving decisions. In the future, variable message signs most likely will be phased out. Instead, DOTs need to provide traffic information in digital formats that can be easily and precisely interpreted by CAV.

ID	Others	Recommendations
		The traffic information will be much more detailed than what is displayed on a variable message sign today and can include information such as which lane is closed, taper length, distance to lane closure point, average left-turn phase duration, average queue length, etc.
27	• Drone as a data collection platform	Drones have been widely used by many DOTs for infrastructure inspection and providing situational awareness. DOTs are encouraged to 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	Some DOTs are reluctant to invest in new roadside traffic sensors such as inductive loops, Radar and camera due to installation and maintenance costs. They are more willing to simply purchase probe data. DOTs should conduct studies to compare the life-cycle costs of relying on data vendors and their own data collection infrastructure. In the future, DOTs can invest in mobile/portable data collection units (similar to portable variable message signs) for areas that are not well covered by probe data. These portable data collection units can also be used to collect trajectory data for safety studies. Also, DOTs should invest in retrofitting existing traffic cameras and loop detectors using AI and edge computing technologies to expand the capacities of these traditional sensors. DOTs may work together and develop data and communication interface standards for vendors. In this way, DOTs can easily switch from one vendor to another to obtain the same data elements. This flexibility and independence may potentially increase the competition among vendors and reduce the sensor maintenance and replacement costs.

4.0 Case Study on Speed Behavior on Highway Horizontal Curves

It is estimated that over 25% of fatal crashes are on horizontal curves. It is important to have a clear understanding of how vehicles behave on those segments in response to different warning signs and pavement markings and under various weather and environmental conditions. This case study utilized advanced radar and thermal camera sensors to collect vehicle trajectories on horizontal curves. The collected data was analyzed and compared with the corresponding TomTom data.

The main purpose of this study was to demonstrate the advantages of using portable sensors to collect detailed vehicle trajectory data and use AI techniques to model the data for understanding driver behavior. Given the limited time and resources available, this study did not include altering existing traffic signs and pavement markings and investigating the impacts of such changes on driver behavior. However, the proposed portable sensors can be used for this purpose.

4.1 Site Identification and Data Collection

With the inputs from the NHDOT, the research team identified five high-risk horizontal curves. We initially planned to identify sites based on crash history. However, such data was unavailable. The selected sites are provided in Table 4-1 and listed in Figure 4-1 through Figure 4-5.

Table 4-1. Selected 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

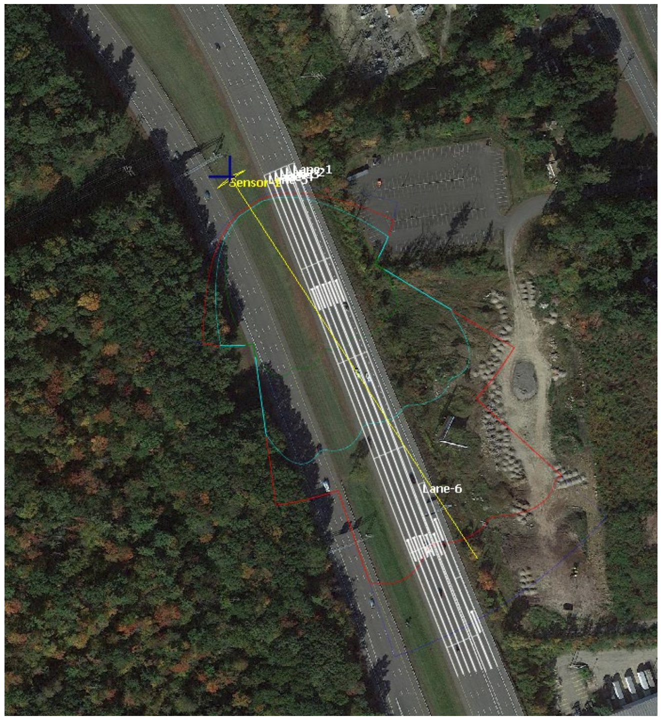


Figure 4-1. Overview of the Nashua Site

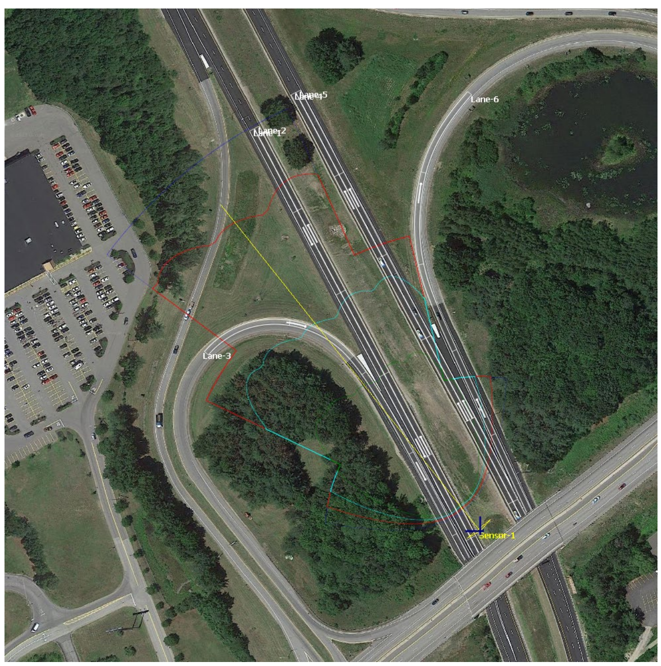


Figure 4-2. Overview of the Tilton North Site

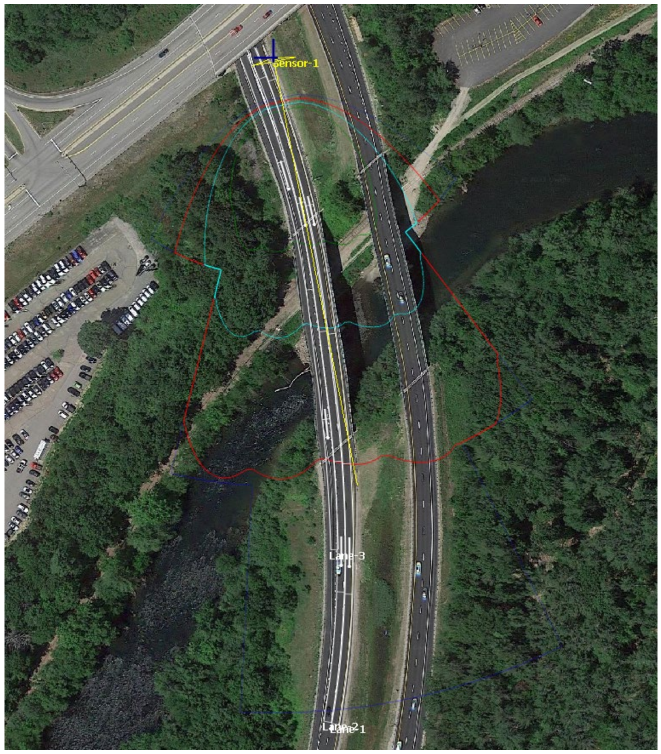


Figure 4-3. Overview of the Tilton South Site

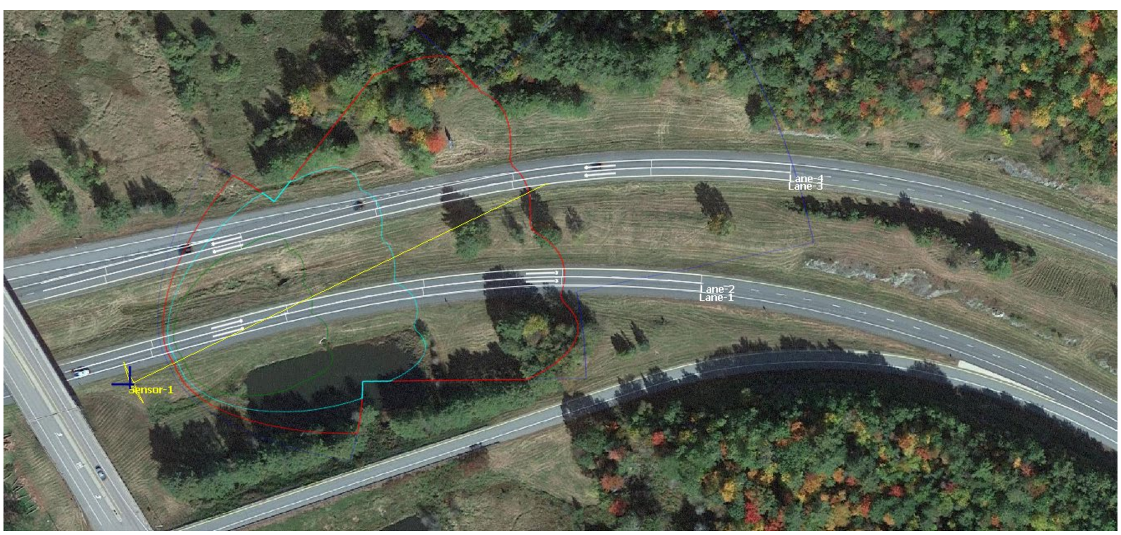


Figure 4-4. Overview of the Littleton North Site

Figure 4-5. Overview of the Littleton South Site

Ultra-high-definition radar and thermal cameras were used to collect traffic data at each site. The radar sensor used can detect and track individual vehicles up to 1,000 ft. Both the radar and camera sensors were powered by battery and solar panel. Some lessons learned regarding data collection through this case study include:

- The radar sensor should be mounted between 20 and 26 ft above the pavement.
- Ideally, both radar and camera sensors should be mounted on a stable structure (e.g., a sign gantry) directly above the traffic, although achieving this in many cases is difficult.
- Mounting radar and camera on the roadside makes it challenging to separate trajectories by lane. Therefore, if the research aims to study lane changes, mounting sensors on the roadside will make the subsequent data analysis very challenging.
- Both radar and camera were mounted on a trailer. The trailer's vibrations affected the quality of the collected data.
- The radar unit reliably detected and tracked small-sized vehicles. For heavy trucks, it sometimes generated phantom objects. This was likely caused by the radar's mounting position. Mounting the radar directly above the traffic may well address this problem.

4.2 Radar Data Analysis

The purpose of collecting the radar data was to understand how drivers behave when approaching and traversing a horizontal curve. Although interesting empirical findings were obtained from the collected radar data, the data was collected from only five sites and the sample size was not enough to show how roadway geometry such as lane width and curve radius affect driver behavior. Therefore, this study also utilized StreetLight data for highway horizontal curves and ramps in Maine and conducted a network-wide speeding analysis. The speeding analysis results are presented in Chapter 6 of this report.

The speed data generated by radar was analyzed from the following three perspectives, and the analysis results are presented in the remaining part of this section.

- How do average vehicle speeds change over time for different sites? This includes speed distributions at fixed locations and how average vehicle speeds vary across a road segment.
- Comparison of radar speed data and TomTom speed data.
- Detection of outliers in speed profiles using artificial intelligence (AI) algorithms.

4.2.1 Speed Distributions and Profiles

This subsection presents the following results:

- **Distribution of Segment Average Speeds**: The speeds of vehicles across a road segment were averaged, and histograms of the average speeds for small and large vehicles were then plotted. For each site, histograms were generated for four time periods to illustrate speed variations over time: 0:00-4:00 AM, 6:00-8:00 AM, 10:00 AM-2:00 PM, and 4:00-6:00 PM.
- **Average Speed Profiles across a Road Segment**: For each site, the entire road segment was divided into 30-ft sections. For each section, the average speed of all vehicles was calculated. This average speed varied across road sections during the above-mentioned four time periods and the variations were plotted.
- **Average Speed over Time at a Specific Location**: We selected a location on the road that is 350 feet away from the radar. The average speed of all vehicles passing this location was recorded. This average speed varied over a 24-hour period, and the variations were plotted for each of the five locations.

Given the large number of figures generated, they are included in Appendix A instead of this section. Some key findings from these figures are:

- Average vehicle speeds during 0:00-4:00AM were clearly lower than those during the other three time periods considered in this study.
- Although the average small-sized vehicle speeds for all sites were close to 70 mph, some small-sized vehicle speeds exceeded 80 mph.
- The average large-sized vehicle speeds were close to 65 mph, with few large-sized vehicles exceeding 70 mph.
- Vehicles did not change speeds significantly when approaching/traversing a horizontal curve. This is probably because the horizontal curves selected were all on interstate highway 93, which has a high design standard. Also, there were no speed limit signs prior to the five selected horizontal curves.

4.2.2 Comparison of Radar Data with TomTom Data

With the generous support of the NHDOT, we obtained TomTom speed data for the two sites in Tilton, NH. The TomTom dataset included speed data for both Tilton North (upstream) and Tilton South (downstream). The TomTom upstream data was for a single segment of about 1,600 feet. The downstream data was for ten shorter segments of varying lengths. To facilitate a meaningful comparison between TomTom and our radar data, we selected the upstream TomTom segment and three TomTom downstream segments with the longest lengths. These TomTom segments were matched with our radar data.

Table 4-2. Upstream speed comparison

Time	TomTom Mean (mph)	TomTom Sample Size	Mean (mph)	Radar Sample Size	Standard Deviation (mph)
00 – 01	75	6	69, 69, 69	32, 32, 32	7, 7, 6
01 – 02	69	7	66, 66, 66	16, 16, 16	7, 8, 7
02 – 03	75	9	74, 74, 74	25, 25, 25	6, 6, 6
03 – 04	74	7	67, 67, 67	21, 21, 21	15, 15, 14
04 – 05	74	13	72, 71, 71	34, 34, 34	8, 10, 9
05 – 06	74	17	73, 73, 73	75, 75, 75	7, 7, 7
06 – 07	76	45	75, 75, 74	147, 147, 147	5, 5, 5
07 – 08	74	95	73, 73, 73	306, 306, 306	6, 6, 6
08 – 09	76	158	74, 74, 74	443, 443, 443	6, 6, 5
09 – 10	74	206	73, 73, 73	566, 565, 566	5, 5, 5
10 – 11	75	232	73, 73, 73	677, 677, 677	6, 6, 6
11 – 12	75	244	74, 73, 73	626, 626, 625	5, 5, 5
12 – 13	74	248	73, 73, 73	617, 618, 618	6, 6, 6
13 – 14	72	247	72, 72, 71	555, 555, 555	6, 6, 6
14 – 15	75	243	74, 74, 74	543, 643, 643	6, 6, 6
15 – 16	75	269	75, 75, 75	590, 590, 590	6, 6, 6
16 – 17	77	242	75, 75, 75	409, 409, 409	6, 6, 6
17 – 18	76	243	76, 76, 76	144, 144, 144	5, 5, 5
18 – 19	76	244	76, 76, 76	207, 207, 207	6, 6, 6
19 – 20	76	184	75, 75, 75	205, 205, 205	6, 6, 6
20 – 21	76	136	75, 75, 75	51, 51, 51	5, 5, 5
21 – 22	74	159	77, 76, 76	51, 51, 51	7, 7, 7
22 – 23	75	83	73, 73, 73	161, 161, 161	6, 6, 6
23 – 24	74	37	73, 73, 73	104, 104, 104	6, 6, 6

The comparison results are presented in Appendix B (Figure 9-49 through Figure 9-72) with each figure for one hour. It can be seen from these figures that:

- For the upstream segment, TomTom data matched radar speed data better when the traffic volume was higher. Significant differences between the two sets of data were observed between 0:00-4:00 AM. This is likely due to the low sample rates during those periods, as shown in Table 4-2 and Table 4-3.
- For downstream segments, there were significant differences between the two sets of data throughout the day. A possible reason is that TomTom data only reported the speeds of vehicles from the interstate highway, not those from the on-ramp. Our radar captured vehicles from both the highway and the ramp, and ramp vehicles were slower than highway vehicles.
- For both upstream and downstream segments, TomTom average speeds appeared to be higher than radar speeds. This could be due to the different ways that TomTom and radar

average speeds were calculated. We do not know exactly how the TomTom speed data was calculated.
- Overall, the TomTom data at Tilton North (upstream) seemed accurate except for the early morning period when the sample sizes were relatively small as shown in Table 4-2 and Table 4-3.

Table 4-2 and Table 4-3 provide a comparison of data between TomTom and radar for both the upstream and downstream locations in Tilton, NH.

Table 4-3. Downstream speed comparison

Time	TomTom	TomTom		Radar	
	Mean (mph)	Sample Size	Mean (mph)	Sample Size	Standard Deviation (mph)
00 – 01	75, 75, 75	6	58, 59, 60	72, 72, 72	10, 9, 9
01 – 02	70, 71, 68	7	58, 58, 59	42, 42, 42	8, 8, 8
02 – 03	73, 73, 71	9	63, 64, 65	45, 45, 45	10, 9, 9
03 – 04	73, 73, 71	7	61, 62, 63	35, 35, 35	14, 13, 12
04 – 05	74, 74, 73	13	65, 66, 67	52, 52, 52	10, 9, 9
05 – 06	75, 74, 74	17	65, 66, 66	130, 130, 130	9, 8, 8
06 – 07	76, 75, 74	45	66, 67, 68	253, 253, 253	10, 9, 9
07 – 08	74, 74, 73	95	66, 67, 67	484, 484, 484	9, 8, 8
08 – 09	76, 75, 74	158	67, 68, 68	689, 689, 689	9, 8, 8
09 – 10	74, 74, 73	206	66, 67, 68	826, 825, 825	9, 8, 7
10 – 11	74, 74, 73	232	67, 68, 68	1012, 1011, 1009	9, 8, 7
11 – 12	74, 74, 73	244	67, 68, 69	1017, 1019, 1017	9, 8, 7
12 – 13	75, 75, 75	248	65, 66, 66	920, 922, 922	9, 8, 8
13 – 14	72, 71, 70	247	64, 65, 66	909, 909, 910	9, 8, 7
14 – 15	75, 74, 73	243	66, 67, 68	1011, 1011, 1010	9, 8, 8
15 – 16	74, 74, 74	269	68, 69, 69	1018, 1019, 1020	9, 8, 8
16 – 17	76, 76, 75	242	68, 69, 69	1019, 1019, 1019	9, 8, 8
17 – 18	76, 75, 74	243	68, 68, 68	987, 987, 987	9, 8, 8
18 – 19	76, 75, 75	244	69, 69, 69	974, 974, 975	9, 9, 8
19 – 20	76, 75, 74	184	67, 68, 68	746, 746, 757	9, 9, 8
20 – 21	75, 74, 73	136	67, 67, 67	573, 573, 573	9, 8, 8
21 – 22	73, 73, 72	159	66, 66, 66	586, 586, 585	9, 9, 8
22 – 23	73, 74, 73	83	64, 65, 65	377, 377, 377	10, 9, 9
23 – 24	74, 76, 73	37	60, 61, 62	331, 331, 331	9, 9, 8

4.2.3 Speed Profile Outlier Detection

The previous two subsections present the radar data in aggregated forms. One major advantage of radar data is its high level of detail. It provides the trajectories of individual vehicles. Artificial Intelligence (AI) algorithms and other heuristic methods can be developed to analyze radar trajectory data and automatically identify risky behavior and dangerous interactions among vehicles (i.e., developing surrogate safety measures such as speed variance). Without such algorithms, DOTs would have to review the vast amount (e.g., several weeks or months) of trajectory data manually. If before-and-after data are collected from the same site, the effectiveness of the safety treatments can be evaluated based on vehicle trajectories as well. In addition, surrogate safety measures derived from trajectories can be compared with historical crash data to see whether they demonstrate consistent trends. For example, they may both show high crash risks during certain time periods or in a specific section of horizontal curves.

As mentioned in Section 4.1, in this study the radar sensor was mounted on roadside. This setup was not ideal and limited the accuracy of lateral position measurements, making it difficult to differentiate vehicles in adjacent lanes. Therefore, the analysis in this subsection focuses on longitudinal trajectories.

One thing we noticed was that radar was good at reliably detecting and tracking small-sized vehicles. For large trucks, the radar sometimes generated phantom objects surrounding the true objects (i.e., heavy trucks). This was likely due to the suboptimal radar setup. Mounting the radar directly above traffic will give a top-down view, which can avoid the signals reflected by the two sides of a large truck.

We first cleaned the radar trajectory data by removing short and fragmented trajectories. After the data cleaning step, both AI and heuristic methods were used to detect outliers. For the AI method, we experimented with the Autoencoder (AE) [*13*] and Variational Autoencoder (VAE) [*14*] models.

The AE model is an unsupervised neural network model. Its main idea is to encode and reduce data dimensionality, generate intermediate data (referred to as Embedding), and then reconstruct an output that closely resembles the original data from this Embedding. The AE component responsible for encoding and dimensionality reduction is called the Encoder, while the Decoder is responsible for data reconstruction. As a variation of the AE model, the VAE model still consists of an Encoder and a Decoder. During the encoding process, VAE maps the data into a probability distribution beyond simply encoding it. Similarly, during the decoding process, it draws samples from the generated probability distribution. The introduction of this probability distribution allows VAE to better capture data diversity and enhance data generation richness.

Both AE and VAE models required users to define what a "speed profile outlier" is. We assumed that most vehicles follow a similar pattern of movement. Therefore, if a particular vehicle's behavior deviates significantly from this pattern, that vehicle is likely an outlier. Different types of deviations can be categorized as follows:

- A vehicle's speed consistently remains significantly higher or lower than the average speed for that period.
- A vehicle's speed continuously decreases or increases during its movement.
- A vehicle experiences a significant speed drop within a short time frame (e.g., harsh braking).

Although both AI and heuristic methods have been applied to outlier detection, we found the heuristic methods to be more intuitive, straightforward to implement, and robust for the problem under investigation. Therefore, only the results based on the heuristic methods are presented here. The AE and VAE algorithms need to be further investigated to improve their performance. Other AI models will also be explored in our future research.

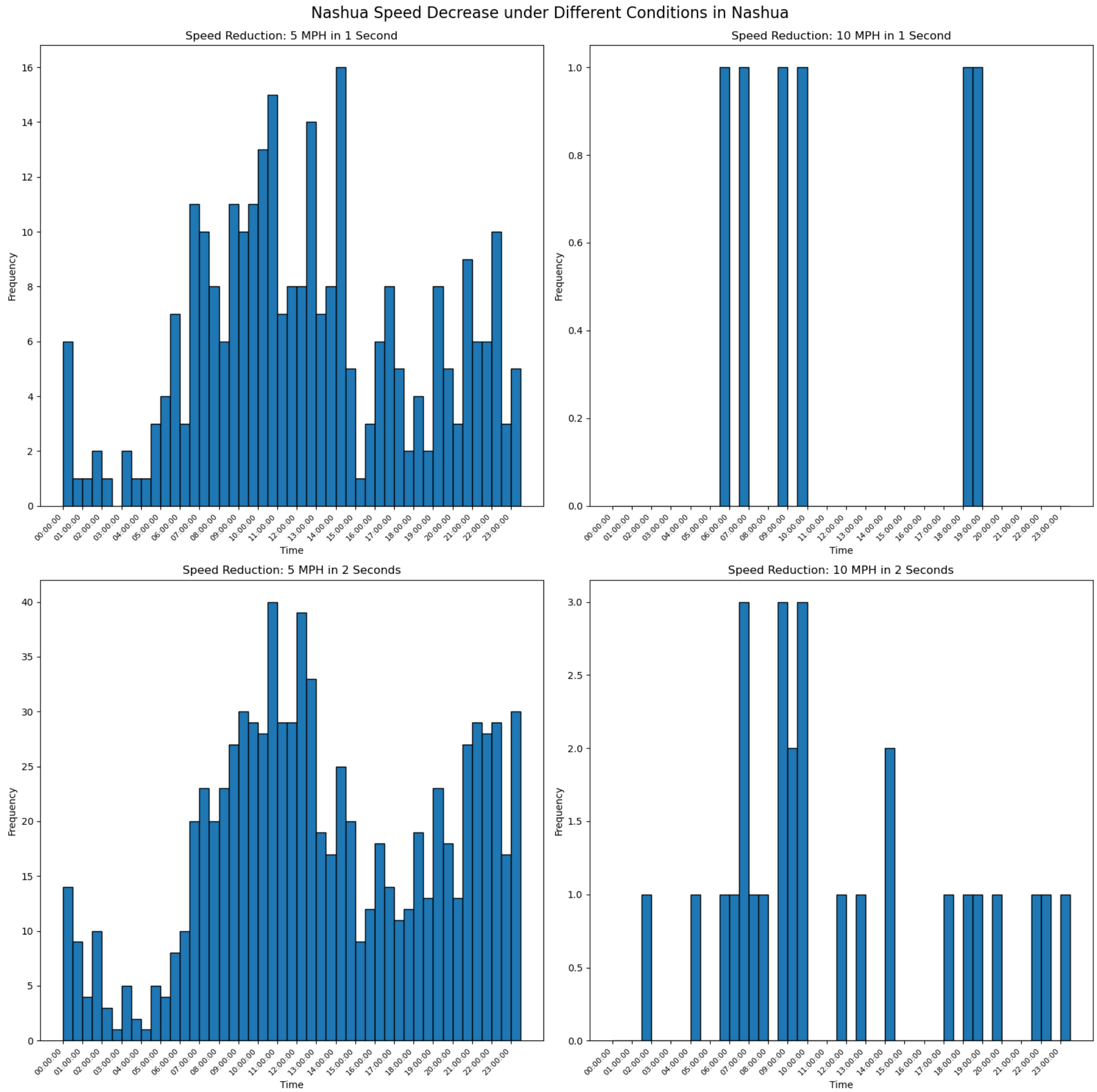


Figure 4-6. Speed Reduction Outliers in Nashua, NH

First, we focused on situations involving a sudden decrease in speed or harsh braking. Such situations were further categorized into four scenarios: a vehicle reduces its speed by 10 mph in 2 seconds, 5 mph in 2 seconds, 10 mph in 1 second, or 5 mph in 1 second. These four types of events can be easily identified based on the trajectories collected by the radar. We counted the number of such events and plotted the hourly results in Figure 4-6, which suggests that harsh braking events seem to occur mostly during the day around rush hours.

The radar data can show very detailed results. Vehicle "6570" decreased its speed by 10 mph in 1 second at 08:45 am on April 15, 2023 in Nashua. Its trajectory and the trajectories of surrounding vehicles are plotted in Figure 4-7, in which

- The red line represents the trajectory of the "risky vehicle" (ID: 6570).
- The green line represents other vehicles that were traveling at normal speeds within the same time window. These vehicles were in different lanes.
- Additionally, if there is another risky vehicle, occurring within a 3-second interval after vehicle marked as red one, we will mark it as blue. But in this study, we did not find 2 risky vehicles in the same time window.

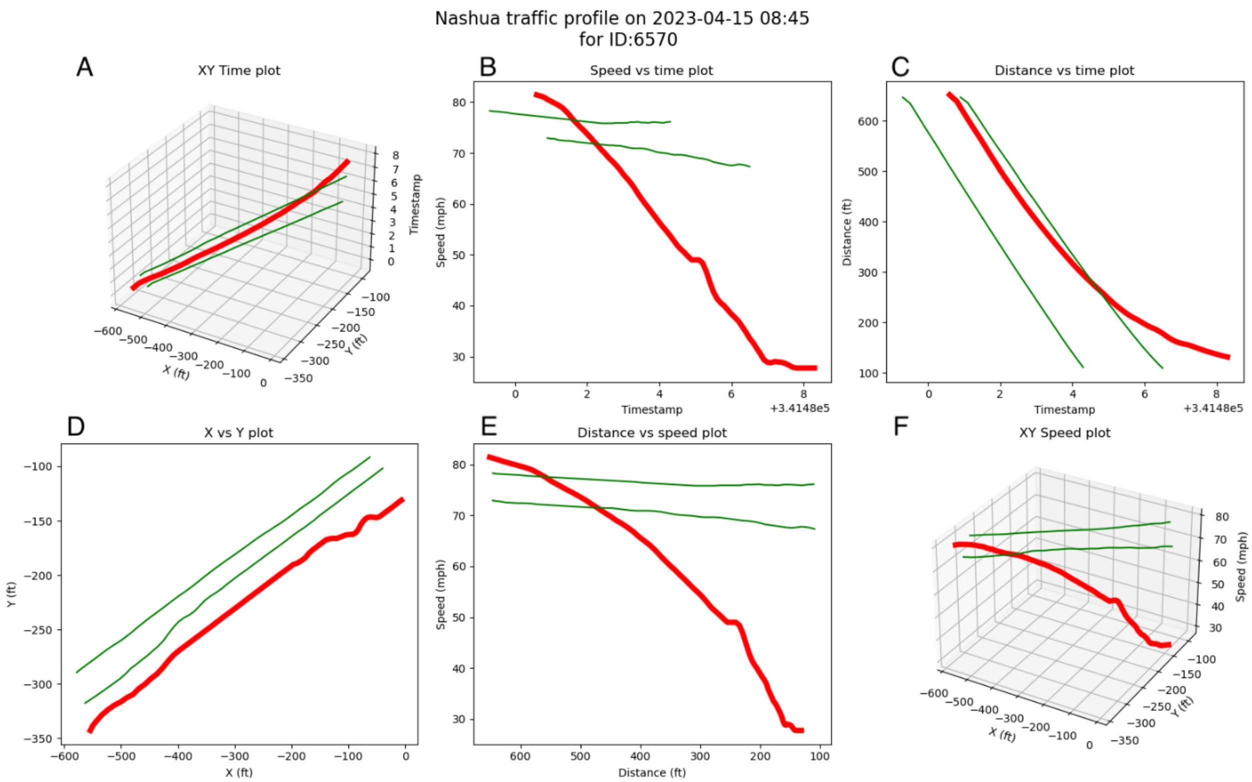


Figure 4-7. Trajectories of Vehicle 6570 and its surrounding vehicles

Out of the six subfigures in Figure 4-7,

- (a) shows a 3D visualization of the x and y coordinates of each vehicle at different timestamps. This subfigure probably is the most informative one.
- (b) shows the speeds of each vehicle at different timestamps. Since this subfigure misses the location (x and y coordinates) information, interpreting it is a little challenging. It shows the sharp deceleration in speed for the red vehicles.
- (c) displays vehicle distance to the radar sensor vs timestamp.
- (d) shows the x and y coordinates for all vehicles.
- (e) shows the speed vs distance relationship.
- (f) is a 3D visualization of the x and y coordinates of each vehicle and the corresponding speeds, showing that the red vehicles significantly decelerate as it approaches the radar sensor.

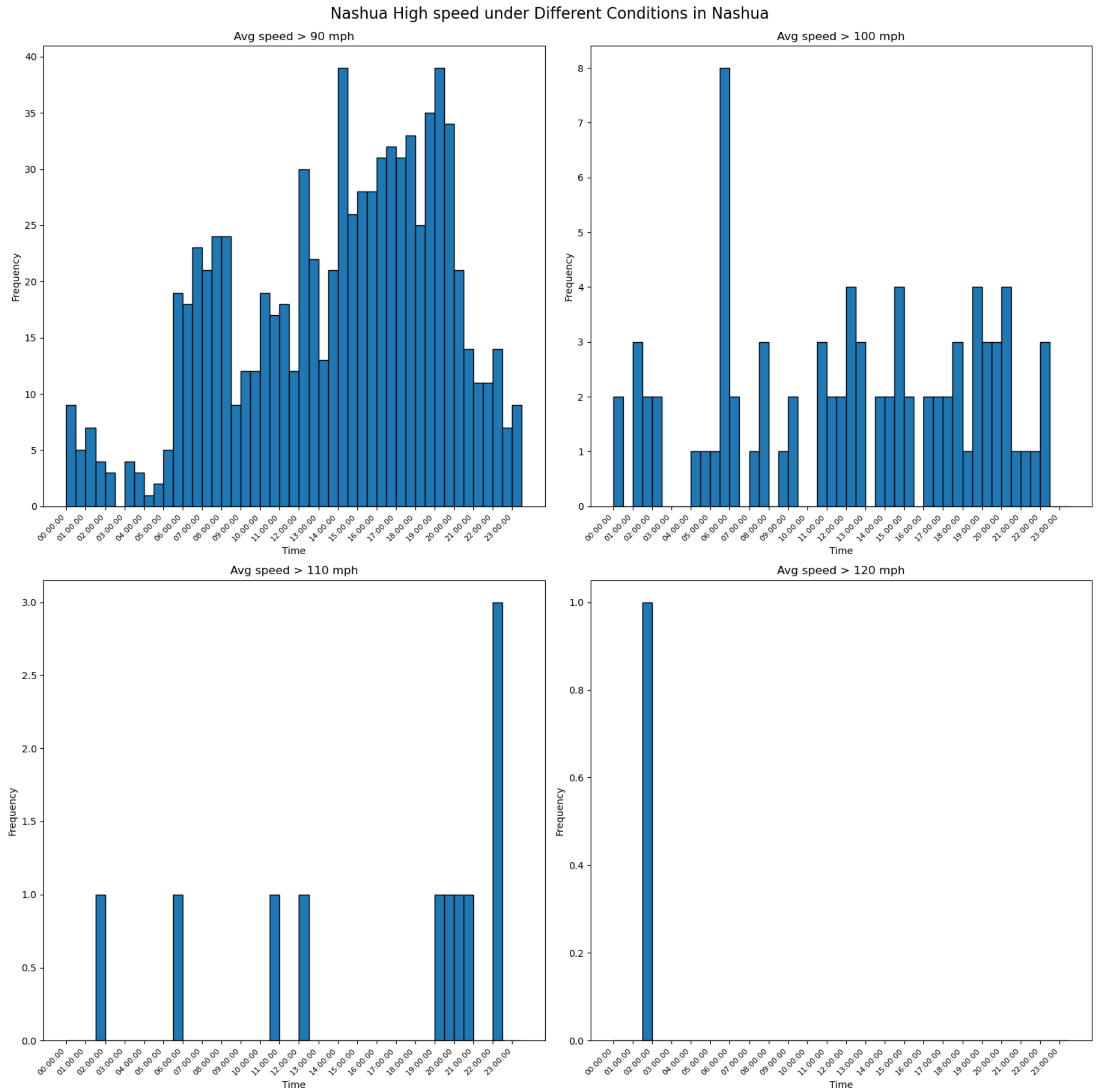


Figure 4-8. Distribution of Excessive Speeding Activities at the Nashua Site

In addition to harsh braking events, excessive speeding activities were also investigated. Vehicles exceeding 90, 100, 110, and 120 mph were counted, and their distributions are presented in Figure 4-8. For instance, several vehicles traversed the Nashua NH segment at a speed greater than 110 or 120 mph, one at 19:16 on April 11, 2023 and the other one at 23:03 on April 13, 2023. The corresponding thermal video footages for these two events were extracted and screenshots are provided in Figure 4-9 and Figure 4-10. Specifically, the 23:03 event was due to a speeding motorcycle.

Figure 4-9. Speeding at 19:16 on April 11, 2023

Figure 4-10. Speeding at 23:03 on April 13, 2023

4.3 Video Data Analysis

For video data analysis, we utilized the YOLOv8 L (large) model, specifically trained on our custom dataset of thermal images. YOLO, short for "You Only Look Once," is a widely recognized and efficient family of deep learning models designed for object detection. These models come in varying sizes, offering a tradeoff between speed and accuracy. In our case, the YOLOv8 L model was chosen, prioritizing the accuracy of vehicle detection over the speed at which detections occur.

One of the key advantages of using YOLOv8 and similar off-the-shelf deep-learning solutions lies in their ready-made architectures. By employing established models, we eliminate the need to invest resources in designing and testing new architectures. This not only saves time but also ensures that the model benefits from the collective knowledge of the broader deep learning community. Additionally, the use of off-the-shelf solutions facilitates faster deployment. With models like YOLOv8, the implementation process is streamlined, allowing us to quickly integrate the model with other modules. Furthermore, the replicability and portability of these models enhance their appeal. Once trained, they can be effortlessly replicated and applied to similar tasks, whether for remote analysis or in the field.

To derive vehicle trajectories and conduct related analyses, we employed the ByteTrack tracking algorithm to process the results obtained from YOLOv8. This integrated approach ensured a seamless and precise analysis process.

Our video data analysis comprised of the following four primary components:

- Camera view change detection
- Volume and time headway analysis
- Risky behavior detection
- Merging point distribution analysis

4.3.1 Camera View Change Detection

Upon analyzing the acquired video data, we observed frequent changes in the camera view across all locations. Figure 4-11 illustrates an instance of this occurrence. Using the trees on the left edge of the video frames as references, we can clearly observe the change in camera view from 11:59:13 AM to 11:59:14 AM on June 8, 2023, at the Tilton North site. This was because the camera and radar sensors were mounted on a trailer instead of a fixed structure. The trailer mast, being less stable, exhibited frequent movements, adversely impacting the quality of both camera and radar data. To address this challenge, we developed a computer vision algorithm specifically designed to detect camera view changes. This algorithm can be beneficial, especially when the camera is remotely connected to a control center. It can generate alerts in response to camera view changes, signaling instances such as strong winds or the movement of the trailer. Such alerts are valuable in maintaining situational awareness and responding promptly to potential issues.

(a) Normal camera view. (b) Camera view change detected.

Figure 4-11. Camera view change detection at Tilton North.

Video stabilization is crucial for accurate vehicle tracking and subsequent analysis, as it establishes a foundation for consistent and stable video footage. This stability is essential for precise object recognition and motion analysis, particularly in dynamic environments. Traditionally, achieving video stabilization involved meticulously tracking specific salient features, requiring prior knowledge of their positions in the initial video frame. However, our approach eliminates the necessity for such prior information, human input, and involvement, providing a more streamlined and efficient process.

Our method focused on the automatic detection of points of interest in a video frame. Subsequently, we mapped these points in successive frames and computed affine image transformations between adjacent frames, all accomplished without the use of deep learning techniques. Frames shown in Figure 4-12 are from the Nashua site, and they highlight the substantial view change in a one-second interval. The yellow lines in Figure 4-12(c) show the changes between two video frames. The distances between the points of interest across frames allow us to measure how much the camera moved during a given interval.

(a) Frame at time 9:01:41 AM: Points of interest (indicated by Os) identified in the initial frame.

(b) Frame at time 9:01:42 AM: Points of interest (marked with Xs) detected in the subsequent frame.

(c) Mapping Points of Interest: Yellow lines show the correspondence between Os and Xs.

Figure 4-12. Frame-to-frame mapping of points of interest to detect camera movement.

In certain locations, we annotated segmentation for roadways, merge lanes, and gore areas. In such instances, we chose to train a deep learning model. By utilizing the annotated roadway segmentation results, we could seamlessly integrate the road context into our system. This integration proved invaluable for making precise adjustments to vehicle positions based on the road layout. This approach not only streamlined the tracking process but also significantly improved its accuracy.

Illustrating this concept with data from the Tilton North site, Figure 4-13 demonstrates how we were able to identify and keep track of various regions even when the camera moves or shakes. This information was used to refine subsequent analysis, ensuring the integrity of our analytical procedures.

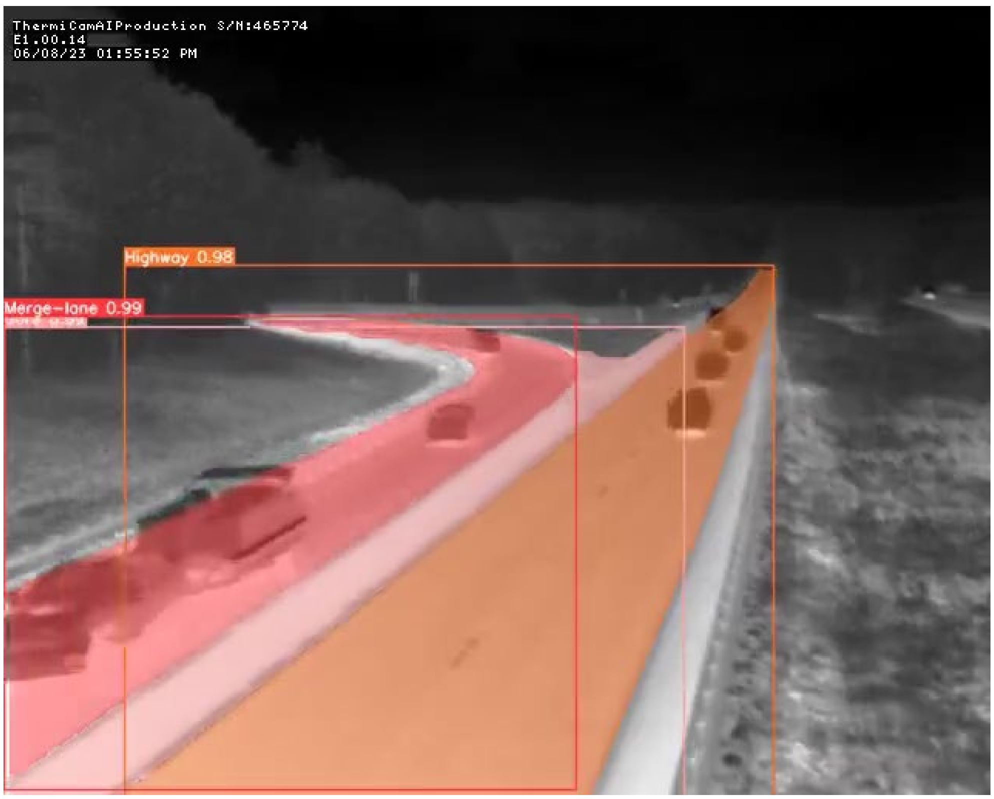


Figure 4-13. Segmentation mask illustrating detected roadway, gore, and merging lane areas.

4.3.2 Volume and Time Headway Analysis

Volume analysis examines the number of vehicles in specific sections, identifying peak activity periods. Time headway analysis explores temporal gaps between vehicles, providing insights into traffic safety and arrival patterns. Examining these metrics provides valuable insights into traffic behavior, guiding the development of effective traffic management strategies.

The incorporation of advanced technologies, notably deep learning for vehicle detection and tracking, transforms the methodology for volume and time headway analysis. Our system operates with minimal human intervention, significantly streamlining the data analysis process. By employing deep learning techniques, specifically YOLOv8 in our case, we achieved precise vehicle detection and tracking, eliminating the necessity for manual input and human supervision. This automated approach not only accelerates the analysis but also guarantees a high level of accuracy by minimizing the potential for human errors.

Figure 4-14 visually represents the hourly vehicle volumes at the Littleton South site. Additional visualization and charts for other locations are in Appendix D.

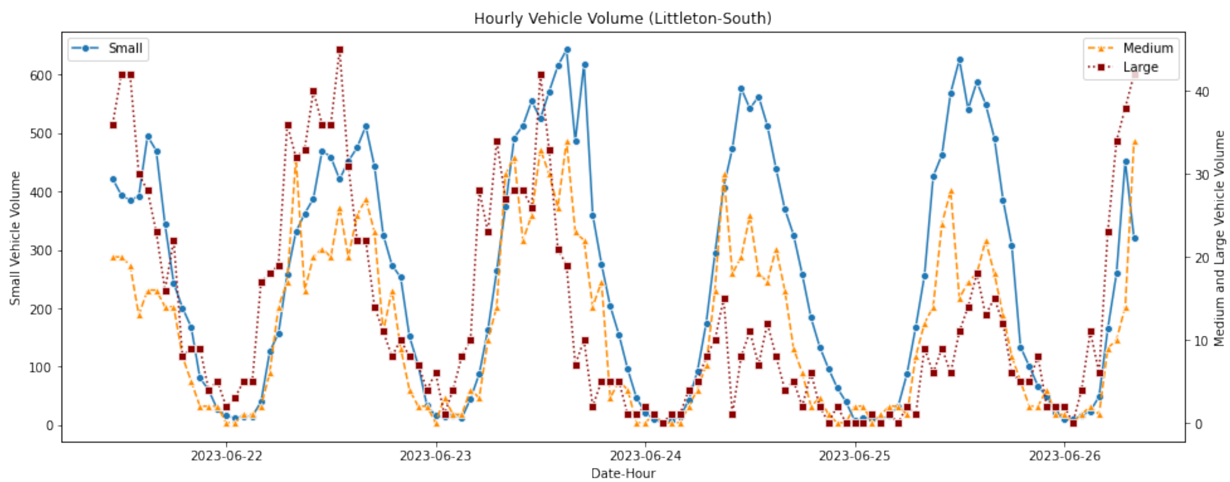


Figure 4-14 Hourly vehicle counts for Littleton South.

We conducted an analysis of hourly volume variations between weekends and weekdays. Figure 4-15 illustrates these variations for the Littleton South location, and for a more detailed understanding, hourly volumes are presented as a heatmap in Figure 4-16. This heatmap offers a comprehensive overview of traffic dynamics throughout different hours. This analysis was done for all other locations, and the remaining charts are presented in Appendix D. In some figures within Appendix D, such as Figure 9-89 an Figure 9-93, certain cells appear blank due to corruption in the corresponding thermal videos. Each thermal video has a duration of one hour.

In our time headway analysis, we define a specific line segment on the road surface. When a vehicle crosses this designated line, our algorithm records the exact time. The time headway is then computed by measuring the time gap between this vehicle and the one preceding it, which had previously crossed the same line. This method allows us to accurately assess the time intervals between vehicles in each lane, providing us with valuable insights into traffic dynamics and vehicle arrival patterns. The pseudocode for this approach is given below.

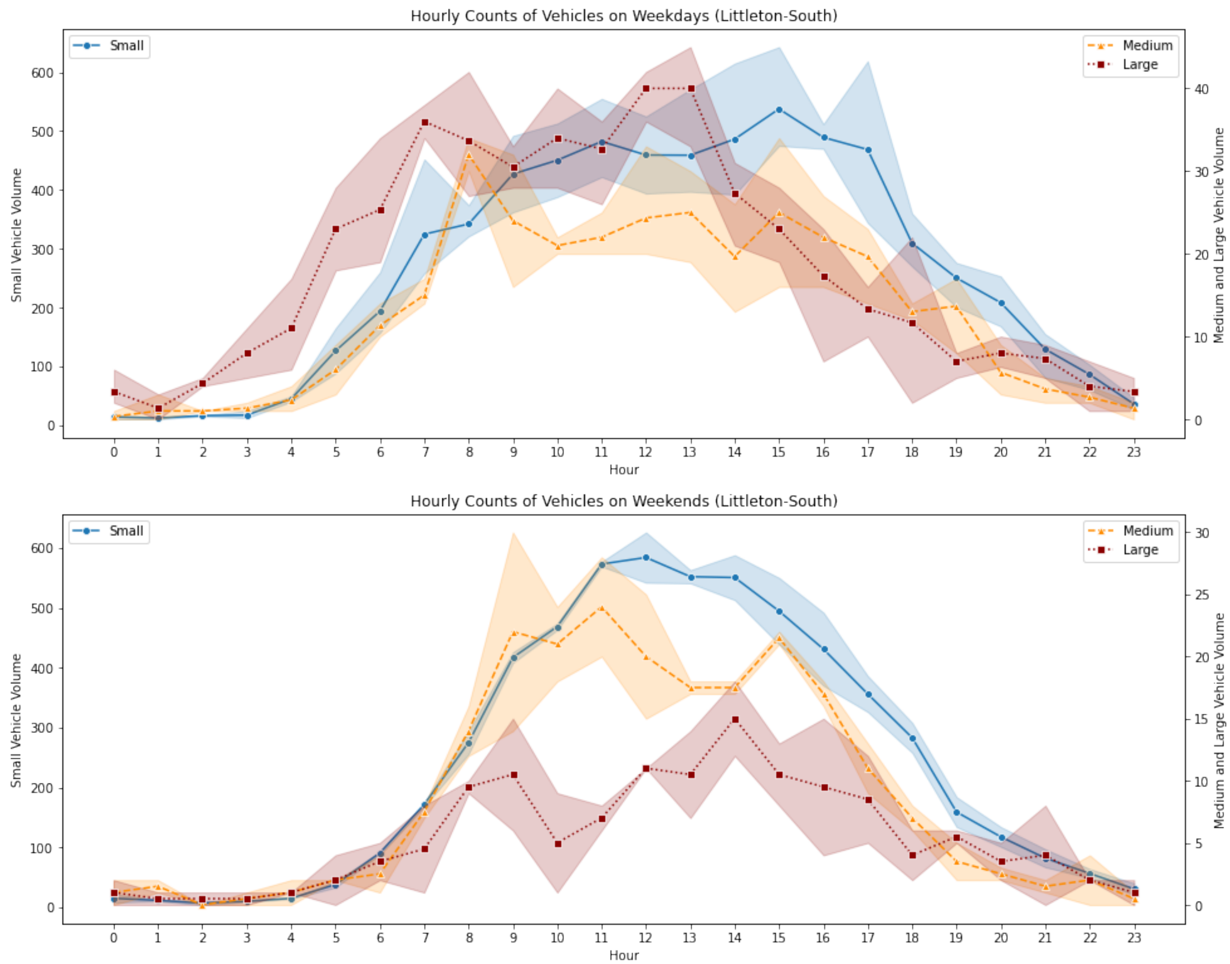


Figure 4-15. Hourly Vehicle Counts on Weekends and Weekdays at Littleton South.

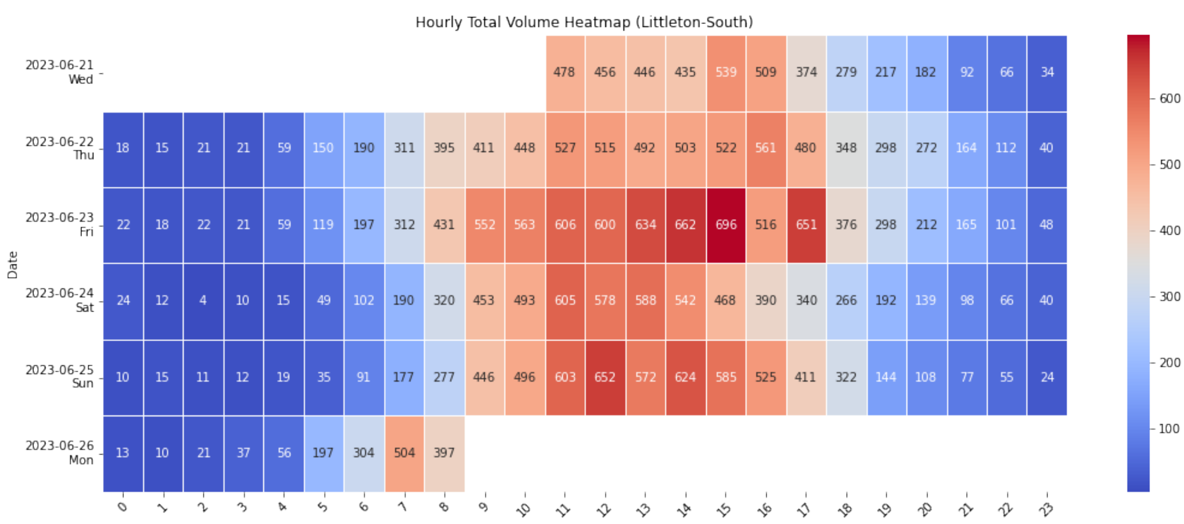


Figure 4-16. Heatmap depicting the hourly volume of vehicles at Littleton South.

```
1. Initialize empty dictionaries to store crossing times for each lane:
   - lane1_times = {}
   - ...
   - laneN_times = {}
 (keys of the dictionaries will be trackID of the vehicle.)

2. Define a line segments [L0, L1 ,..., LN]on the road surface for each
lane.

3. While monitoring the traffic:
    a. When a vehicle is detected in lane i:
        i. Record the time of crossing Li as crossing_time.
        ii. Check the dictionary for lane i:
            - If it's the first vehicle in this lane:
                - Store crossing_time as initialization time.
            - If it's not the first vehicle:
                - Calculate time headway as crossing_time -
lanei_times[last_vehicle_in_lanei].
                - Store the time headway data for analysis.
                - Update lanei_times dictionary with the current
vehicle's crossing_time.

4. End monitoring when video/live feed ends.
```

Figure 4-17 illustrates the results of the time headway analysis presented as a line chart. The analysis is for the Tilton South site on data recorded on June 8, 2023, from 6 PM to 7 PM. The chart provides a detailed breakdown of the time headway gap frequencies observed at this site, categorized by individual lanes. Note that we excluded time headway gaps longer than 30 seconds from this calculation. This detailed breakdown by lane aids in understanding the traffic behavior in specific areas of the roadway, providing essential data for traffic management, safety assessments, and future road infrastructure planning.

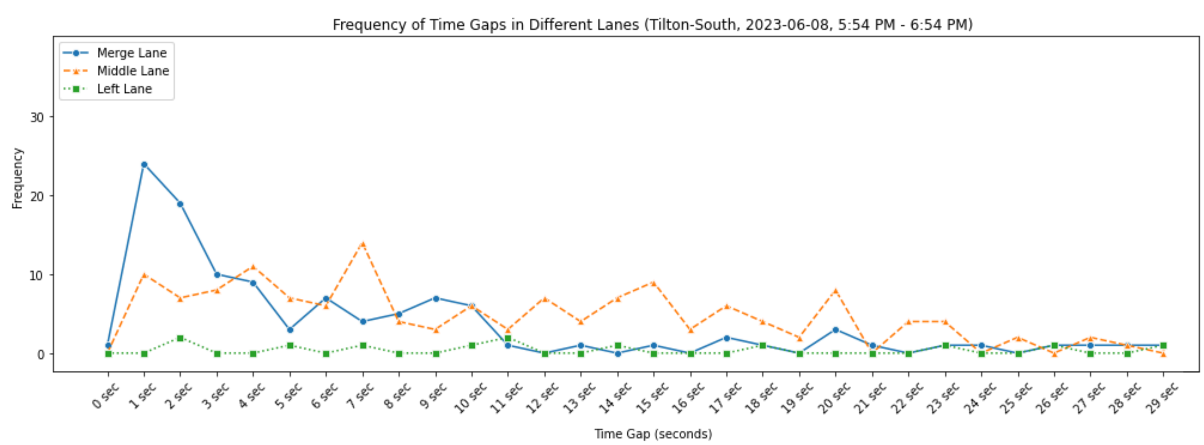


Figure 4-17. Time headway gap frequencies by lane and category for the Tilton South site.

4.3.3 Risky Behavior Detection Using AI Algorithms

The application of deep learning techniques has revolutionized our ability to detect and analyze risky behaviors among drivers. One notable instance is the detection of vehicles straying into restricted areas, such as gore regions during merging, a behavior that poses a significant threat to road safety. Employing cutting-edge deep learning models, we can accurately identify these violations and record crucial data for further analysis.

At the Tilton North location, we utilized an approach to monitor vehicles merging onto the highway. Employing a combination of vehicle and road segmentation results, we developed a pseudocode-based system to identify instances where vehicles crossed into the gore region, a prohibited area during merging. By assessing the overlap between vehicle segmentation masks and the designated gore region in each frame, we recorded the frame numbers, corresponding track IDs, and timestamps of these events. Such techniques allowed us to precisely pinpoint and analyze instances of risky behavior, enabling a deeper understanding of the factors contributing to hazardous driving practices in merging scenarios. Figure 4-18 shows two such instances. The pseudocode for this method is given below.

```
Initialize an empty list to store vehicles that drive over the gore region

For each frame in the video:
    # Extract segmentation masks for vehicles and road/gore/merging ramp
    vehicle_mask = get_vehicle_segmentation_mask(frame)  # Returns binary mask
of vehicles
    road_mask = get_road_segmentation_mask(frame)  # Returns binary mask of
road, gore, merging ramp

    # Use track IDs to identify individual vehicles in the vehicle mask
    vehicle_track_ids = extract_track_ids(vehicle_mask)  # Extract track IDs
from vehicle mask

    # Iterate through each identified vehicle
    For each vehicle_track_id in vehicle_track_ids:
        # Get the segmentation mask for the specific vehicle
        specific_vehicle_mask = extract_specific_vehicle_mask(vehicle_mask,
vehicle_track_id)

        # Check if the vehicle mask overlaps with the gore region in the road
mask
        if mask_overlap(specific_vehicle_mask, road_mask, gore_threshold):
            # Record the frame number, vehicle track ID, and timestamp
            record_event(frame_number, vehicle_track_id, timestamp)

# Function to check mask overlap using a threshold
Function mask_overlap(mask1, mask2, threshold):
    Intersection = Count overlapping pixels between mask1 and mask2
    Union = Count total pixels in mask1 + mask2 - Intersection
    OverlapRatio = Intersection / Union
    return OverlapRatio > threshold
```

Figure 4-18. Detecting Vehicles Going Over Gore Region in Tilton North Location.

Figure 4-19 and Figure 4-20 are the heatmaps visualizing the percentage and volume of vehicles travelling over the gore region. This gives us a better understanding of the pattern of occurrence of such events. The percentage is calculated with respect to the ramp volume in that hour. The data shows that travelling over the gore region happened quite frequently at the Tilton North site.

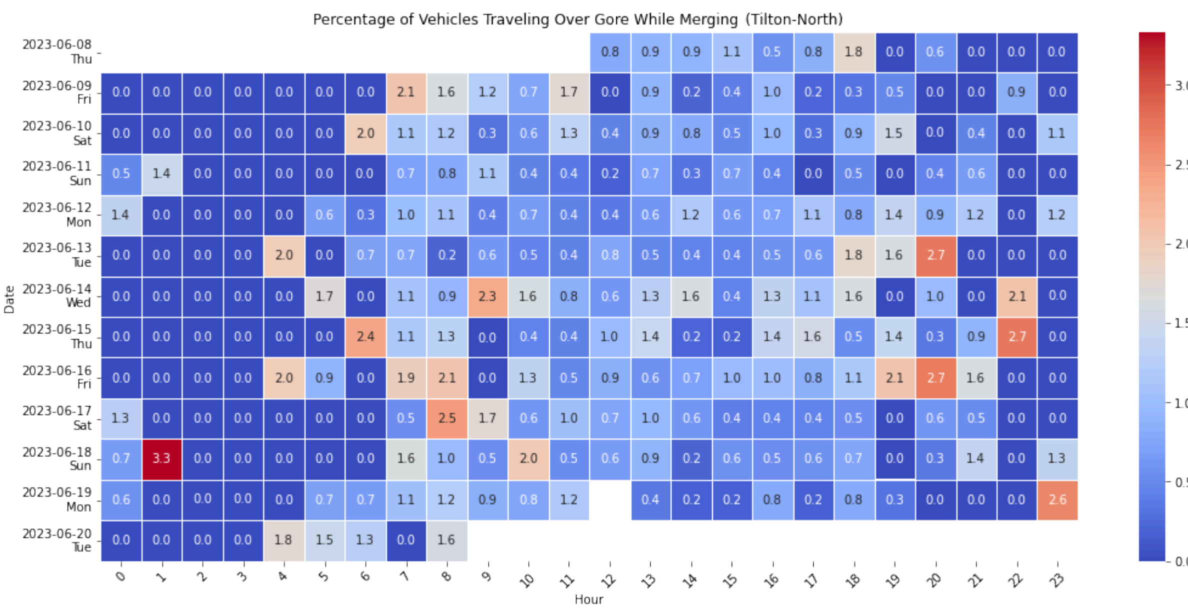


Figure 4-19. Heatmap showing percentage of vehicles going over the gore area.

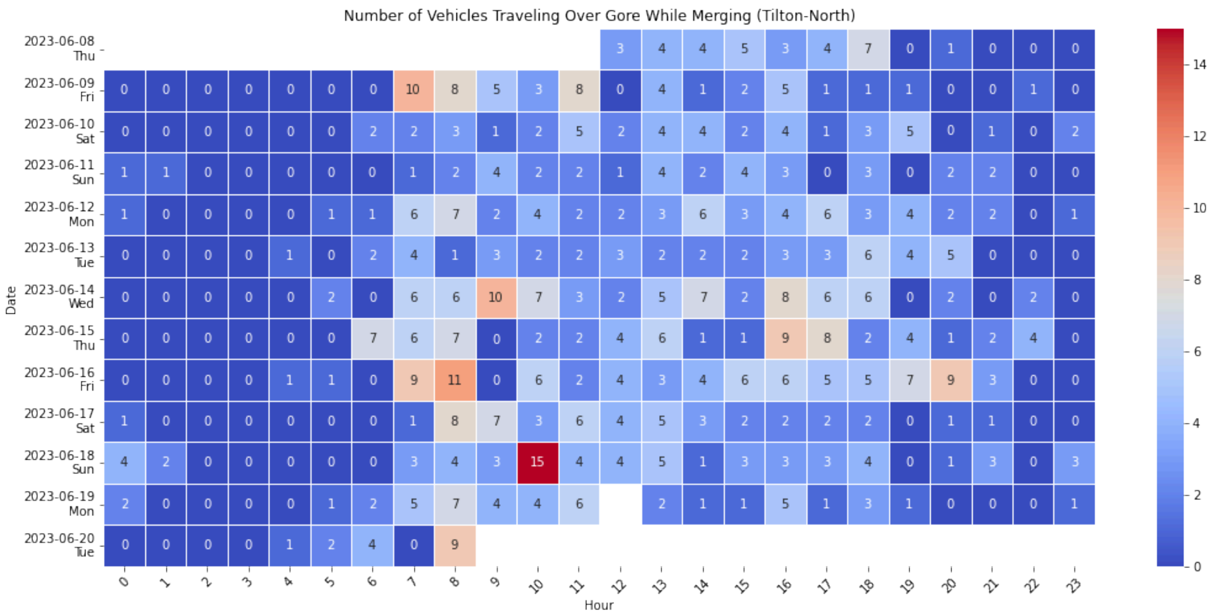


Figure 4-20. Heatmap showing number of vehicles going over the gore area.

4.3.4 Merging Point Analysis Using AI Algorithms

Merging point analysis is for investigating merging behavior and safety at highway ramps. Our analysis was based on 24 hours of videos captured on June 22, 2023 at Littleton South. It utilized object segmentation and tracking algorithms to gain insights into vehicle merging patterns.

We initiated our analysis by segmenting the merging lane, a process vital for isolating the specific area under investigation. This segmentation delineated the boundaries of the merging lane, providing a visual reference for our subsequent assessments. Within the segmented merging lane, we further categorized specific regions. As shown in Figure 4-21, our classification includes:

- Risky Area (Over the Gore) (RED): This zone is identified as particularly hazardous, representing instances where vehicles breach the gore area. Detection of vehicles in this area signifies risky behavior.
- Far Region (BLUE): The 'far' region refers to a section of the merging lane that is farther from the gore area, indicating a relatively safer position for vehicles merging onto the main highway.
- Farthest Region (GREEN): Positioned at the outermost edge of the merging lane, the 'farthest' region represents the safest area for vehicles attempting to merge.

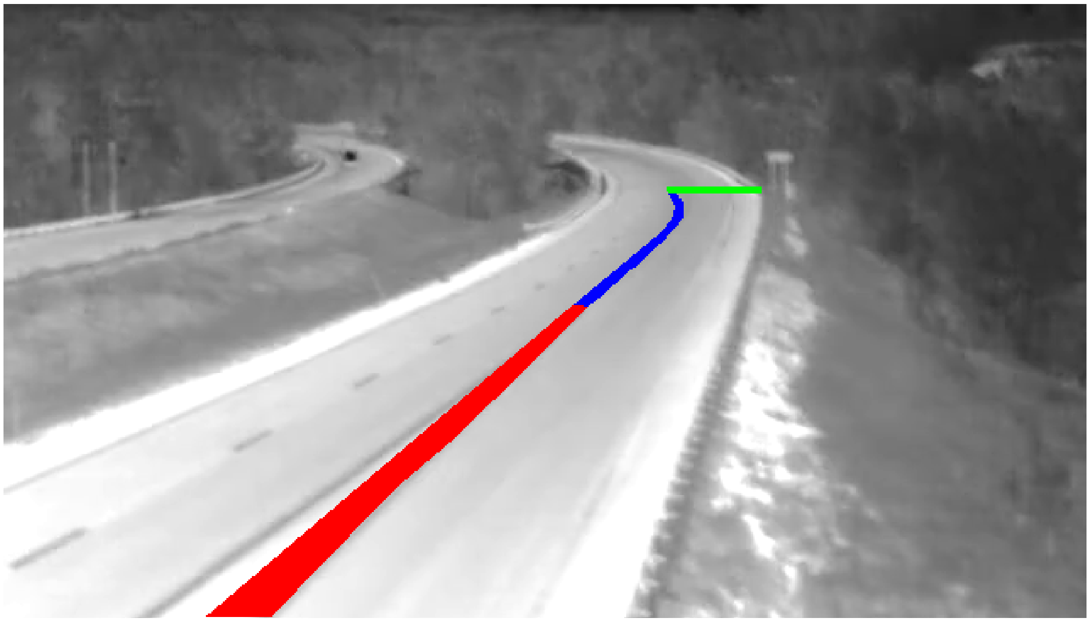


Figure 4-21. Merging region classification for merging point analysis at Littleton South.

In Figure 4-22, we present a visualization of the percentage of vehicles merging onto the main highway at various times of the day. Figure 4-23 compares the merging ramp volume and the number of vehicles merging in different regions, and Figure 4-24 compares freeway traffic volume and the merging pattern.

Given the limited time, we did not conduct a before-and-after study to investigate how different pavement markings may affect the merging pattern. Another interesting direction is to further look into the gaps between ramp merging vehicles and highway mainline vehicles. With the AI tools developed in this research, such studies can be made possible.

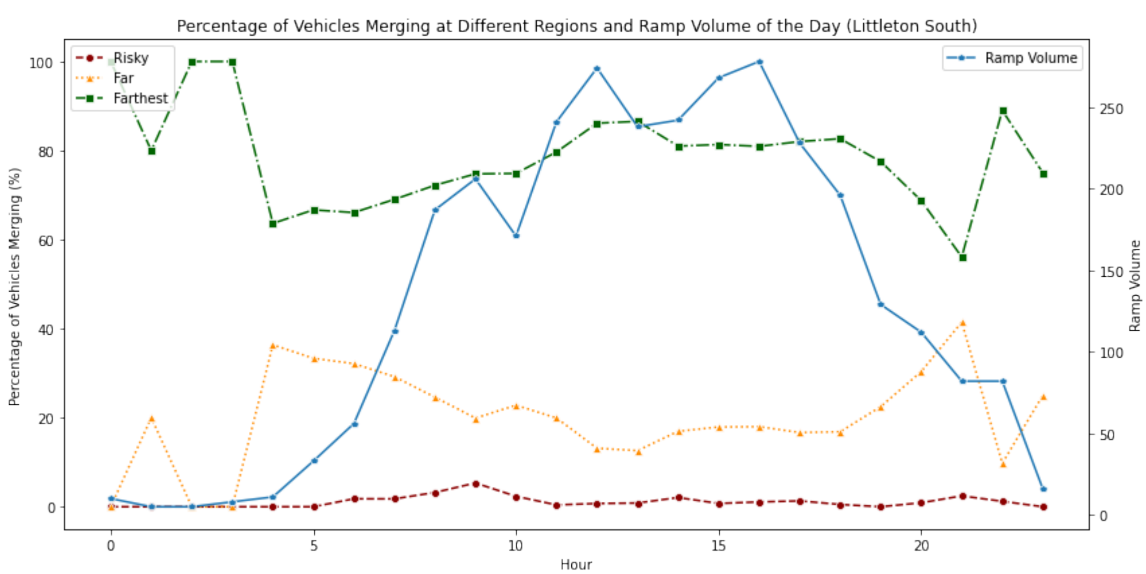


Figure 4-22. Percentage of vehicles merging at Littleton South location.

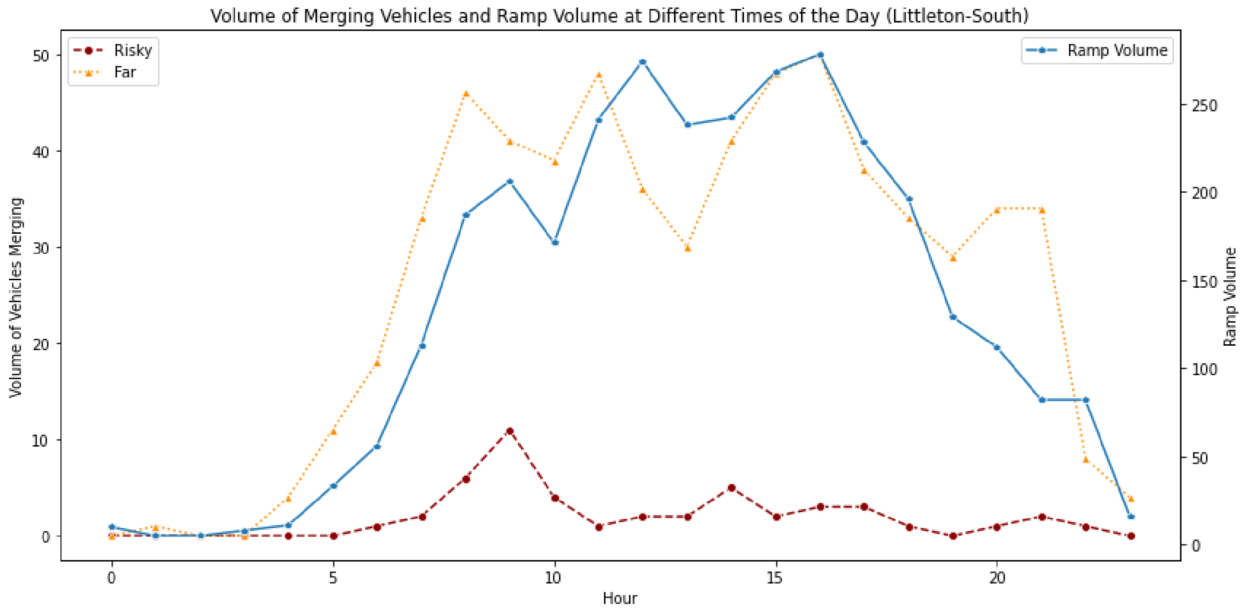


Figure 4-23. Comparison of merging ramp volume and merging pattern at Littleton South.

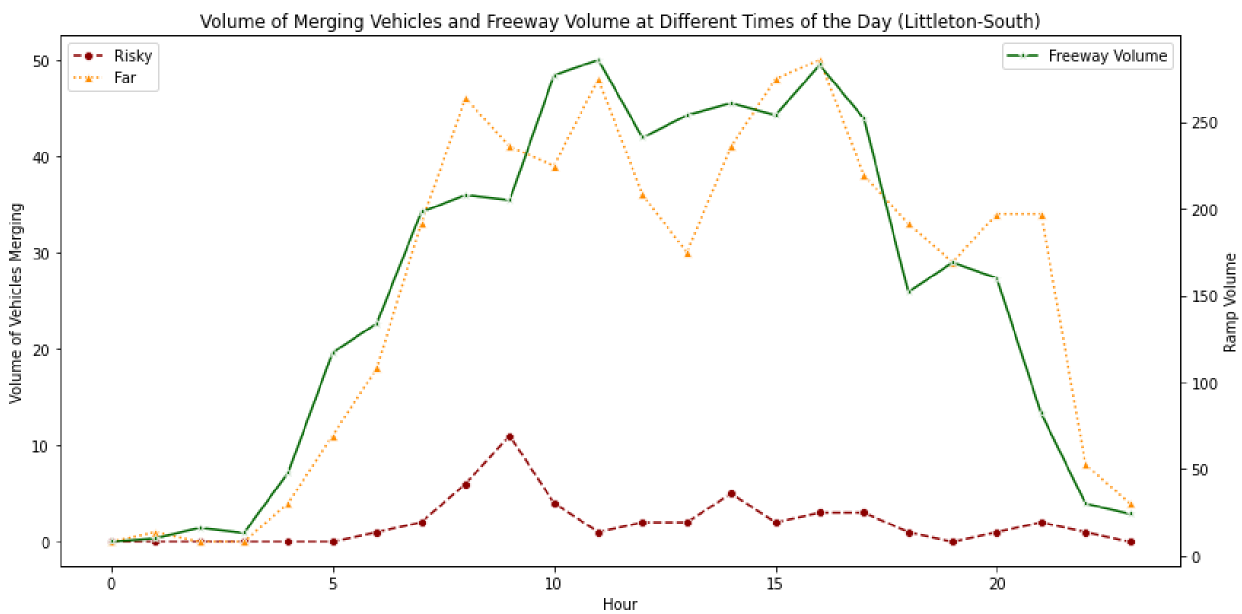


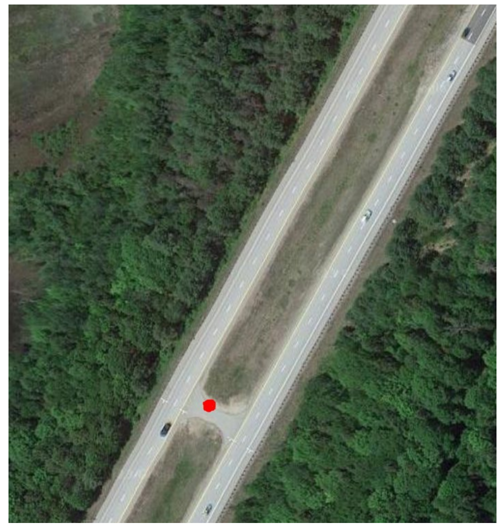
Figure 4-24. Comparison of freeway volume and merging pattern at Littleton South.

This page left blank intentionally.

5.0 Case Study on Speed and Lane Changing Behavior Prior to Highway Work Zone

Besides horizontal curves, highway work zones are another hot spot for crashes. Understanding how drivers behave and react to different control strategies when they approach highway work zones is important. This pilot study focused on a highway zone on I-93 Southbound in Campton, NH. The exact location of this work zone is shown in Figure 5-1.

(a) Trailer and field view



(b) Satellite site view (Red dot indicates trailer location)

Figure 5-1. Work Zone on I-93 in Campton, NH

For this southbound work zone, the left passing lane was closed from August 17 to August 31 in 2023. The Southbound I-93 left lane merge taper started at mile marker 86.2. Two flashing & flanking speed limit signs were located at mile marker 86.4 on the two sides of the highway. The right-side sign coordinate was 43.84554167, -71.64636944, and the left-side sign coordinate was 43.84548611, -71.64611667. A portable changeable message sign (PCMS) was located in the median cross-over at mile marker 88.2, and the coordinate was 43.86655556, -71.66287778. The PCMS sign was two miles upstream of where the left lane merge taper started. Table 5-1 shows how the flashing speed limit signs and the PCMS were controlled from 08/17/2023 to 08/31/2023.

At this work zone, we collected radar data from 08/17/2023 to 08/23/2023 and thermal video data from 08/17/2023 to 08/31/2023. The radar sensor stopped working on 08/23/2023, causing the radar data to be shorter than the thermal video data. The exact reason for the radar sensor to stop working is unknown. It was likely due to the power supply. Based on the historical weather data on wunderground.com, it was cloudy for most of the time from 08/17/2023 to 08/23/2023 in Campton, NH. This could have caused the power voltage to be unstable, and the radar sensor was sensitive to this issue. After the power voltage was restored to the normal level, the radar data logger couldn't reestablish connection to the radar sensor and retrieve the data.

Table 5-1. Work Zone Control Strategies

Date	Flashing Speed Limit Signs			PCMS Messages	
				LEFT LANE CLOSED	POSSIBLE SLOW OR STOPPED
	UP/ON	UP/OFF	DOWN/OFF	MM 86.4 MERGE EARLY	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

This pilot study aims to demonstrate how the radar and camera sensors can be used to investigate two important aspects of work zone traffic operations: (1) vehicle approaching speed, and (2) vehicle merging behavior. The first aspect is to find out how vehicles adjust their speeds when approaching a work zone. For the second one, the purpose is to count last-minute lane changes and compare them with upstream traffic control strategies.

5.1 Camera and Radar Data Collection

Both the radar and camera sensors were mounted on a trailer (see Figure 5-1(a)), which was different from the trailer used in Chapter 4. This new trailer also moved slightly during the data collection, which could be seen from the collected thermal videos and were also reflected in the collected radar data. The trailer was deployed in the median cross-over (see Figure 5-1(b)), not in the closed lane. This location caused some issues during radar and camera data analysis, making it difficult to differentiate between the two lanes. Although we managed to separate vehicles by lane during camera and radar data processing, the accuracy could be significantly improved if the radar and camera sensors had been mounted closer to the highway. Ideally, these sensors should be mounted directly above the traffic.

5.2 Approaching Speed Analysis Based on Radar Data

To address the issue of sensor mount vibrations, we separated vehicle trajectories into 30-ft segments and calculated the centerlines for the left and right lanes. Based on the centerlines, trajectories were categorized into the left and right lanes. The average speeds for the two lanes were also calculated. This process was repeated for each 30-ft segment and each hour. Doing this for different hours separately is to further reduce the impacts of sensor mount vibrations, with the assumption that the sensor mount did not move much within an hour.

Although we could conduct the approaching speed analysis for all 24 hours, that may generate too much information. Instead, we chose the following periods to perform the speed analysis, and the results are shown in Figure 9-73 through Figure 9-79.

- 4-5am
- 7-8am
- 10-11am
- 2-3pm
- 6-7pm
- 10-11pm

The approaching speed data should be compared with the work zone control strategies in Table 5-1 as well as the traffic volume data obtained from the thermal videos (see Section 5.3 below). Some observations are:

- Vehicles in both lanes clearly decelerated as they were approaching the work zone lane closure taper.
- When the traffic was congested, the approaching speed decreased to as low as 40 mph. In this case, the right lane had a lower speed than the left lane. A possible reason is that the left lane was about to be closed, and vehicles in the left lane tried to maintain a higher speed to facilitate merging into the right lane.
- The approaching speeds from 10-11PM were lower than those during other periods except for those congested ones (e.g., 10-11AM and 6-7PM on August 20, 2323)
- Except for some very congested periods, the average vehicle speeds at the beginning of the lane closure taper were around 65 mph before 7PM. From 10-11PM, the corresponding average speeds were lower and were close to 60 mph.

5.3 Merging Point Analysis Based on Video Data and AI

Another important and interesting aspect of work zone traffic operations is how vehicles merge. Thermal cameras can “see” well during the night, which makes it very convenient to monitor how vehicles merge prior to a work zone and how they may behave differently during daytime and nighttime. Although sensor vibrations made it challenging to analyze the collected thermal videos, we developed a clustering algorithm to categorize the extracted vehicle trajectories into four groups as shown in Figure 5-2. The boundary between the red and yellow zones is at the first traffic barrel. The four types of trajectories illustrated in Figure 5-2 are defined as:

- If a vehicle stays consistently in the right lane prior reaching the green detection zone, its trajectory is in the green group.
- If a vehicle starts in the left lane and shifts into the right lane during the yellow zone, its trajectory is in the yellow group.
- If a vehicle starts in the left lane and shifts into the right lane during the red zone, its trajectory is in the Red 1 group.
- If a vehicle starts in the left lane and remains in the left lane by the time it exits the red zone, its trajectory is in the Red 2 group.

Given the resolution of the thermal camera used in this study, we cannot recognize objects beyond the beginning of the green zone. Higher-resolution thermal cameras can be helpful. However, they are much more expensive. Another alternative is to utilize regular cameras if capturing video data during nighttime work zone operations is not a priority.

Figure 5-2. Work Zone Vehicle Trajectory Classification

Utilizing trajectory data to analyze driver behavior is a nuanced and intricate process. By plotting these trajectories, we gained valuable insights into the subtleties of driver behavior, allowing us to discern patterns, anticipate responses, and uncover the underlying dynamics of traffic flow. Additionally, we computed the Kernel Density Estimation (KDE) which provided us with distribution and density of traffic. Figure 5-3 shows the vehicle trajectories from 08/17/2023, 5 PM to 6 PM along with the KDE of trajectory points.

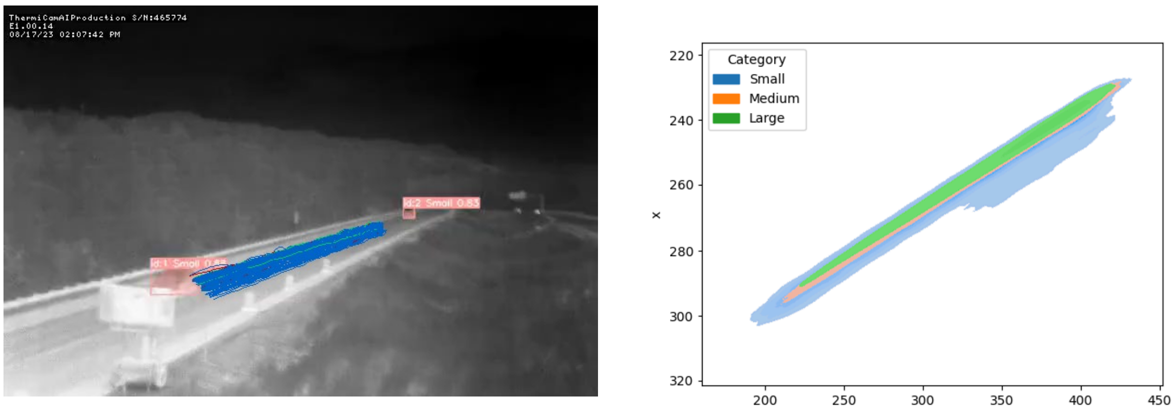


Figure 5-3. Vehicle trajectories and KDE from video recorded at Campton.

Figure 5-4 visualizes the reconstructed trajectories of vehicles. Vehicle 527 in the left subfigure merged earlier in the Yellow region, whereas vehicle 3117 in the right subfigure did not merge even when it was in the Red zone.

Figure 5-4. Samples of trajectory analysis showing vehicles merging in different regions.

Figure 5-5 shows the aggregated hourly trajectory analysis results for the entire 14 days at the Campton site in NH. A few sample results are included in this section. Among them, Figure 5-6 shows the detected safe and risky events on August 23, 2023. Figure 5-7 and Figure 5-8 present the hourly traffic volumes for the two weeks from August 17 to August 31, 2023. The analysis of data from the Campton site has resulted in numerous charts and figures, all of which are presented in Appendix D. The heatmaps depicted in Figure 9-104, Figure 9-105, and Figure 9-106 contain two blank cells each, resulting from corruption in two thermal videos. Additionally, line charts in Figure 9-115, Figure 9-116, Figure 9-119, and Figure 9-120 exhibit gaps in the lines, which are due to factors such as excessive camera movement caused by high winds, adverse weather conditions, or other unavoidable issues.

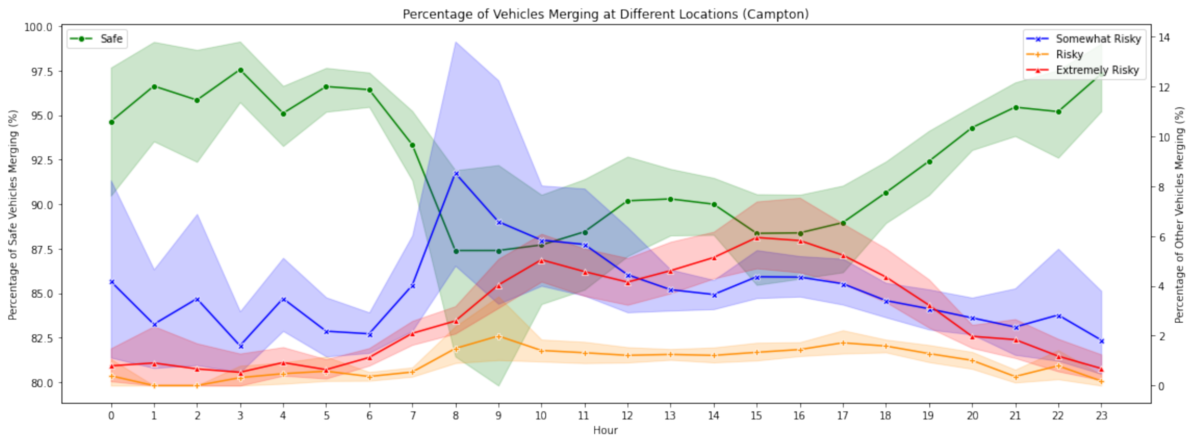


Figure 5-5. Work Zone Vehicle Trajectory Classification Results

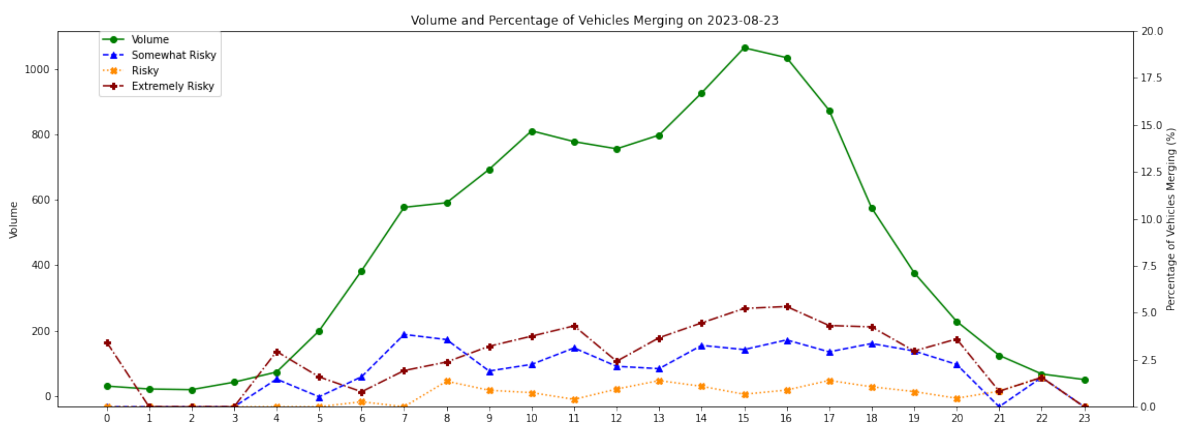


Figure 5-6. Visualization of merging pattern at Campton on Aug 23, 2023.

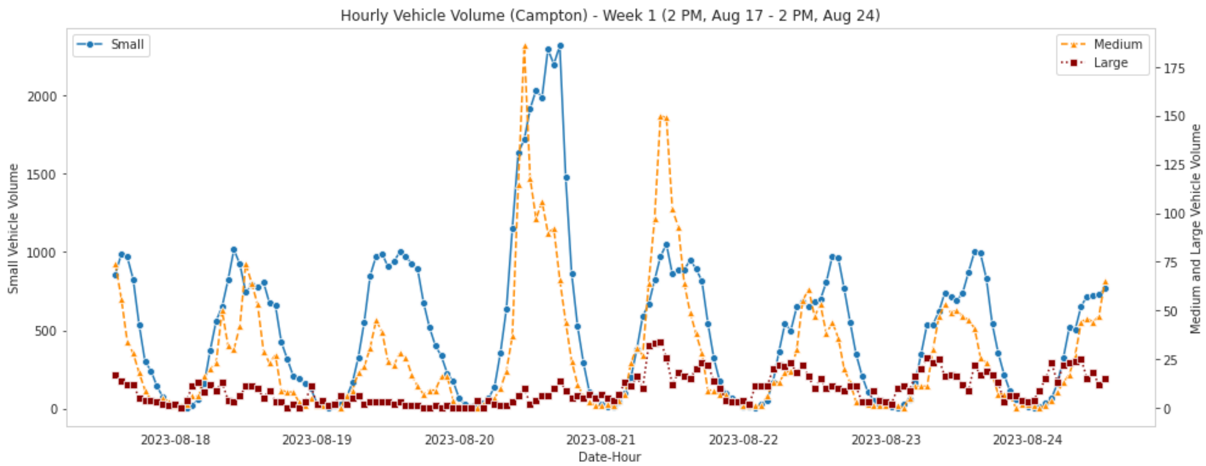


Figure 5-7. Hourly volume at Campton for the first week of data collection.

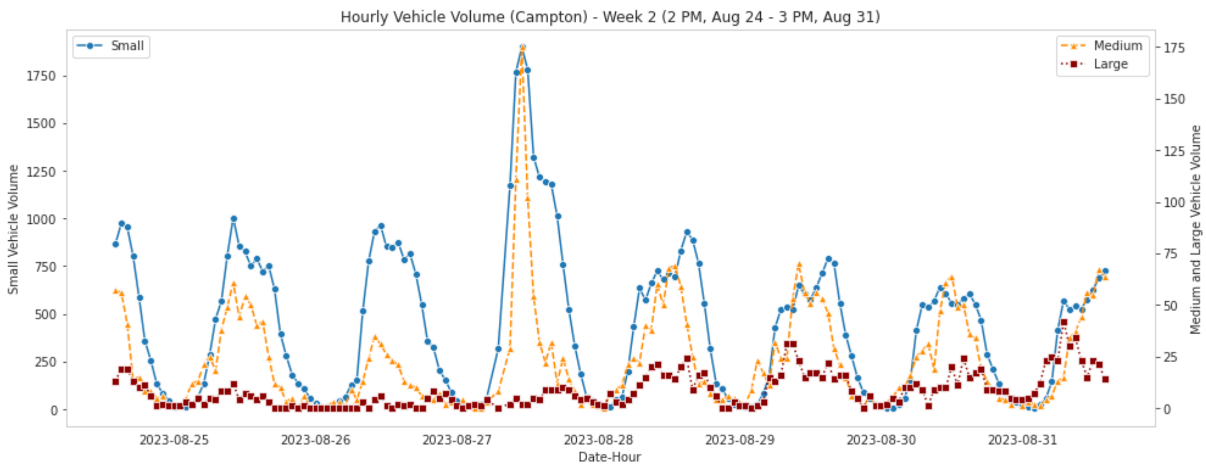


Figure 5-8. volume at Campton for the second week of data collection.

5.4 Summary of Work Zone Pilot Study

Overall, the speed and merging point analysis results suggest that:

- The flashing speed limit signs are helpful in reducing the speeds of vehicles approaching the work zone.
- The flashing speed limit signs and the PCMS seem to be helpful in prompting drivers in the left lane to merge into the right lane earlier.
- Quite a few vehicles traverse the red zone without merging into the right lane. This generate significant safety hazards to both vehicles and workers.

This page left blank intentionally.

6.0 Case Study on Network-Wide Speeding Activity Analysis Using Probe Vehicle Data

Lane departure collisions account for many roadway fatalities across the United States. Many of these crashes occur on horizontal curves or ramps and are due to speeding. This research investigates factors that impact speeding on Interstate horizontal curves and ramps. We collected and combined two unique sources of data. The first database involves comprehensive curve and ramp characteristics collected by an automatic road analyzer (ARAN) vehicle; the second database includes volume, average speed, and speed distribution gathered from probe data provided by StreetLight Insight®. We evaluated the impacts of level of service (LOS), which reflects traffic density or level of congestion, time of the day (morning, evening, and off-peak hours), time of the week (weekdays and weekends), and month of the year (Jan-Dec), and various geometric characteristics such as curve radius, arc angle, and superelevation on speeding. The results show that the odds of speeding increase at horizontal curves with improved levels of service, as well as at those with larger radii and superelevation. The odds of speeding decrease on curves with larger arc angles, as well as during the winter months of the year. Similar results were also observed in models developed for ramps, except for the ramp radius, which was found to be an insignificant factor. The results show the importance of using speed enforcement and other countermeasures to reduce speeding on curves and ramps with low traffic volumes, high speed limits, and large radius and superelevation, especially those located in rural areas. The results could be used to prioritize locations for installation of speed countermeasures or signage such as advisory speed signs, as well as dispatching enforcement resources to high-priority locations and times.

6.1 Background

Lane departure crashes constitute the majority of severe and fatal collisions in the United States (U.S.) [*15*,*16*,*17*]. A significant portion of these crashes occur on horizontal curves and ramps. This is particularly due to the impact of centrifugal force and challenges in negotiating these roadway elements [*18*]. Several other factors, however, also impact the disproportionate rate of severe and fatal crashes on these road elements, including but not limited to weather and environmental factors, driver behavior, and roadway familiarity [***Error! Bookmark not defined.****,19,20*]. Speeding stands as an additional factor that influences lane departure crashes on horizontal curves and ramps. While several research studies investigated the influence of driver behavior on lane departure crashes on horizontal curves or ramps [***Error! Bookmark not defined.***,*21*,*22*,*23*], limited research has been devoted to understanding the factors that influence speeding on these roadway elements. This research addresses this gap by exploring the relationship between different factors, such as (1) time of the day (morning, evening or off-peak hours), time of the week (weekdays and weekends), and month (Jan to Dec), (2) traffic density (or level of service), (3) area type (urban and rural) and (4) various curve geometric characteristics (e.g., arc angle, superelevation, curve radius, shoulder width, lane width, and curvature) and speeding on Interstate horizontal curves and ramps.

Located on the east coast of the United States, Maine is a state with a population of approximately 1.39 million. In summer, Maine attracts a significant number of tourists and visitors, especially in areas such as the Acadia National Park, as well as other recreational areas and landmarks. Maine has the highest road fatality rates across the New England region. Lane departure crashes account for over 70% of fatal crashes in the state; most of these crashes occur on horizontal curves [***Error! Bookmark not defined.***,*24*]. Several factors are associated with this disproportionate rate of crashes or fatalities, including but not limited to adverse and long winter seasons, aging infrastructure, and older population [*25*]. Speeding is another factor that influences the crashes in the state [*26*,*27*]. Given the severity of crashes on horizontal curves, and ramps coupled with the impact of speeding on these crashes, it is important to explore what factors impact the odds of speeding on horizontal curves and ramps to better design countermeasures, plan interventions, or dispatch enforcement to reduce speeding on these road elements. Furthermore, Maine is a rural state. Research studies showed increased lane departure crashes on horizontal curves located in rural areas compared to urban areas [***Error! Bookmark not defined.***,*28*]. We will investigate if speeding also occurs at higher rates in rural areas compared to urban regions.

Naturalistic Driving Study (NDS) data have been used as a major data source to analyze driving behavior on horizontal curves [***Error! Bookmark not defined.,Error! Bookmark not defined.,Error! Bookmark not defined.***,**Error! Bookmark not defined.**,29,*30*,*31*,*32*,*33*,*34*]. Limited research, however, has been devoted to the application and use of probe or crowdsourced data to better understand driver behavior (e.g., speeding) on horizontal curves or ramps. StreetLight InSight®[1] and other probe data providers use technologies like cell phones to extract positioning information of vehicles to compute speed and volume in the roadway network [*35*]. In this study, we leverage probe data provided by StreetLight InSight® to gather information about traffic volume, average speed, and speed distribution [*36*]. While the NDS data were useful to understand driver behavior on select curves where data are available, the probe data sources provide networkwide information to understand the impact of speeding or other driver behaviors and factors on all curves or ramps across the network. The availability of probe data sources on roadway segments also allows for the calculation of traffic density and computing the level-of-service (LOS) of roadway segments and understand the impact of congestion on speeding [***Error! Bookmark not defined.Error! Bookmark not defined.***].

In summary, this research contributes to the existing literature in multiple ways. First, as noted above, speeding is a major factor in lane departure crashes on curves and ramps, especially in Maine. We develop models to understand contributing factors on speeding at horizontal curves and ramps. Second, we will investigate the impact of congestion (reflected in LOS) on the odds of speeding for these road elements. Third, we demonstrate the application of probe data as a new data source to analyze speeding at horizontal curves and ramps. Fourth, due to the availability of probe data, and the complete curve and ramp database in Maine, we consider the entire inventory of Interstate ramps and horizontal curves in the state in our analysis. Fifth, we

[1] StreetLight applies proprietary big data processing resources and machine-learning algorithms to measure travel patterns of vehicles, bicycles, and pedestrians, and makes them available on-demand via its SaaS platform, StreetLight InSight®.

compare the speeding occurrence in rural and urban areas. Finally, we provide recommendations to reduce speeding on these road elements.

6.2 Data Description

Two major data sources were collected and combined to create a uniform dataset for analysis. The first dataset includes the characteristics of Interstate horizontal curves and ramps in Maine. MaineDOT has spent significant time and resources collecting information about horizontal curves and ramps across the Maine network using the automatic road analyzer (ARAN) vehicle over the past few years. The database includes a comprehensive inventory for characteristics of the horizontal curves and ramps across the state. The variables in the database include but are not limited to speed limit, curve radius (R), superelevation (SUP), arc angle (ARC), curvature, lane width, left and right shoulder widths, and the area type (i.e., urban, or rural). Table 6-1 shows the summary statistics of variables collected on Interstate horizontal curves and ramps. The speed limit for Interstate horizontal curves varies from 50 mph to 70 mph, and for Interstate ramps from 55 mph to 65 mph in 5 mph increments. However, there is only one Interstate ramp with a speed limit of 60 mph in the database. Therefore, this ramp was excluded from the analysis.

Table 6-1. Geometric Characteristics for the Selected Interstate Horizontal Curves and Ramps.

Variable	Horizontal Curves	Horizontal Curves	Horizontal Curves	Horizontal Curves	Ramps	Ramps	Ramps	Ramps
	Mean	**S.D.**	**Min**	**Max**	**Mean**	**S.D.**	**Min**	**Max**
Speed Limit (mph)	63.61	5.39	50	70	59.05	4.91	55	65
Curve/Ramp Radius (R) (ft.)	5061.04	1877.09	508.74	9441.88	2686.45	2408.18	116.03	9317.96
Superelevation (SUP) (%)	-0.34	3.12	-8.36	6.88	1.19	3.11	-9.95	8.42
Arc Angle (ARC) (degree)	19.29	11.39	2.07	70.62	23.28	26.76	2.08	255.36
Curvature (degree)	1.39	0.82	0.61	11.26	6.13	7.27	0.61	49.38
Lane Width (LW) (ft.)	12.00	0.00	12	12	14.01	1.97	8	16
Right Shoulder Width (SW) (ft.)	9.94	0.77	2	16	7.00	2.87	0	11
Left Shoulder Width (SW) (ft.)	4.00	0.09	4	6	5.15	2.90	0	11

The geometrics characteristics data was combined with probe volume (V), average operational speed (or space mean speed) ($\bar{S}$), and speed distribution data collected from the StreetLight InSight® platform. For this purpose, GIS maps are generated for Interstate roadway segments with homogenous characteristics. Then, the horizontal curve and ramp segments were separated from tangents and used for data collection. The information was collected for every hour in three years from 2017 to 2019. We excluded data from 2020 and beyond to avoid the effects of the COVID-19 pandemic. Table 6-2 and Table 6-3 show the summary statistics of the collected volume and operational speed data for different months of the year after removing LOS F data. To increase data accuracy, we ensured there were at least 15 observed vehicles in every hour of data collection.

LOS is used instead of traffic volume. LOS is determined based on density, which is a more accurate measure than volume for characterizing traffic conditions. For example, under free-flow and heavily congested conditions, the traffic volumes are similarly low, but their densities/LOS are very different. Using the information from traffic volume and operational speed (i.e., segment average speed), the traffic density (K) (vehicle/mile/lane) was estimated as shown in Eq. (1).

Table 6-2. Summary Statistics of Volume-Per-Lane on Interstate Horizontal Curves and Ramps.

Month	Speed Limit						
	Horizontal Curves	Horizontal Curves	Horizontal Curves	Horizontal Curves	Horizontal Curves	Ramps	Ramps
	50 mph	55 mph	60 mph	65 mph	70 mph	55 mph	65 mph
January	723.5 (396.1)	616.9 (328.8)	492.4 (252.4)	495.2 (262.1)	381.6 (146.4)	759.8 (453.5)	737.8 (433.9)
February	713.8 (391.8)	620.6 (330.9)	507.2 (268.0)	498.6 (261.8)	396.0 (149.4)	774.4 (457.7)	750.1 (437.7)
March	734.7 (397.9)	642.3 (339.3)	537.1 (267.6)	517.5 (263.0)	416.8 (162.1)	810.0 (463.5)	792.6 (448.9)
April	762.8 (402.8)	651.0 (326.9)	538.5 (272.9)	522.3 (262.9)	413.6 (156.8)	790.5 (462.9)	771.0 (448.8)
May	856.1 (429.2)	723.8 (356.0)	584.5 (282.2)	604.2 (311.6)	468.0 (176.9)	833.9 (477.0)	838.9 (463.0)
June	865.6 (447.7)	723.6 (359.6)	570.7 (278.6)	601.2 (317.5)	466.6 (183.2)	787.4 (466.5)	782.5 (450.0)
July	906.2 (454.1)	755.4 (352.2)	572.6 (274.5)	650.5 (332.9)	509.3 (205.7)	784.5 (472.1)	782.0 (451.3)
August	941.6 (460.1)	790.2 (366.6)	600.0 (290.5)	686.3 (350.0)	548.4 (223.1)	811.7 (475.4)	800.7 (446.4)
September	881.1 (450.6)	714.1 (355.1)	588.7 (299.1)	598.1 (305.2)	467.4 (183.0)	809.5 (476.4)	775.1 (454.3)
October	829.2 (423.4)	695.5 (339.2)	596.9 (288.3)	596.4 (294.0)	491.2 (187.5)	861.8 (483.1)	830.8 (457.5)
November	785.9 (407.5)	669.0 (333.6)	579.9 (287.0)	564.8 (283.1)	472.6 (183.3)	858.9 (478.5)	835.3 (457.8)
December	760.0 (402.6)	643.9 (326.4)	559.2 (292.7)	530.0 (270.6)	442.3 (175.2)	824.7 (469.7)	777.5 (440.8)

Table 6-3. Summary Statistics of Operational Speed on Interstate Horizontal Curves and Ramps.

Month	Speed Limit						
	Horizontal Curves	Horizontal Curves	Horizontal Curves	Horizontal Curves	Horizontal Curves	Ramps	Ramps
	50 mph	55 mph	60 mph	65 mph	70 mph	55 mph	65 mph
January	50.21 (6.637)	55.41 (5.903)	50.17 (6.021)	62.52 (6.716)	63.93 (6.848)	44.03 (10.22)	52.82 (11.45)
February	50.35 (6.513)	55.67 (5.903)	50.38 (5.926)	62.89 (6.422)	64.79 (6.378)	44.24 (10.36)	53.42 (11.48)
March	50.81 (5.975)	56.28 (5.407)	51.56 (5.514)	63.89 (6.202)	65.93 (5.931)	44.80 (10.47)	53.97 (11.79)

April	50.76 (5.826)	55.36 (5.210)	51.69 (5.562)	63.54 (6.251)	65.94 (5.990)	44.44 (10.44)	53.25 (11.78)
May	50.54 (5.807)	55.74 (5.386)	52.12 (5.407)	63.89 (6.209)	66.53 (5.436)	44.91 (10.74)	53.71 (12.05)
June	50.63 (6.426)	55.91 (5.739)	52.59 (5.347)	64.27 (6.040)	67.08 (5.298)	44.59 (10.60)	52.93 (12.08)
July	51.37 (6.536)	56.46 (6.142)	53.49 (5.295)	65.11 (6.186)	68.31 (5.075)	44.78 (10.75)	52.33 (12.73)
August	51.208 (6.744)	56.59 (6.065)	53.81 (5.287)	65.22 (6.310)	68.64 (5.005)	44.93 (10.80)	52.14 (13.00)
September	51.12 (6.147)	56.49 (5.361)	53.39 (5.300)	64.73 (5.991)	67.72 (5.108)	45.06 (10.47)	52.37 (12.34)
October	51.06 (6.282)	56.58 (5.126)	53.52 (5.295)	64.65 (5.954)	67.83 (5.244)	45.51 (10.75)	53.58 (12.40)
November	51.37 (6.695)	56.35 (5.886)	52.84 (5.714)	64.23 (6.153)	66.82 (6.064)	44.88 (10.77)	53.58 (12.16)
December	50.52 (7.170)	55.22 (6.487)	52.22 (6.132)	62.96 (6.857)	66.42 (6.255)	43.98 (10.58)	51.81 (12.33)

$$K = V/(n \times \bar{S}) \tag{1}$$

where,

V=Traffic volume (vehicle/hour).
$\bar{S}$= segment average speed (mph).
n= number of lanes.

Then, using the information from traffic density, the LOS of the horizontal curves and ramps was computed using the boundaries described in the Highway Capacity Manual (HCM) as follows [***Error! Bookmark not defined.***]:

LOS A: 0 < K ≤11 vehicles/mile/lane
LOS B: 11 < K ≤ 18 vehicles/mile/lane
LOS C: 18 < K ≤ 26 vehicles/mile/lane
LOS D: 26 < K ≤ 35 vehicles/mile/lane
LOS E: 35 < K ≤ 45 vehicles/mile/lane
LOS F, K > 45 vehicles/mile/lane

As noted earlier, data that represents LOS F was removed from the analysis as they represent a forced driving condition in which speeding rarely if at all occurs. Next, using the speed distribution, and speed limit, the percentage of the vehicles that drive by more than 10, 15, and 20 mph above the speed limit was computed. Data from these two data sources (i.e., ARAN data, and Streetlight Insight), along with computed traffic density, level of service, and speeding information were combined with dummy variables denoting the time of the day (e.g., morning, evening, and off-peak hours), time of the week (weekdays and weekends), and month of the year (January to December). Correlations and multicollinearity among variables were investigated. It was determined that arc angle and curvature are correlated. We also found that left and right shoulders, and lane width are also correlated. After careful consideration and testing various models, we found that models with arc angle, and lane width variables generally produce better results. Therefore, these variables were used in the final model.

To facilitate the modeling and interpretation of results, dummy variables were also created for various variables such as speed limit, curve radius, superelevation, arc angle, and lane widths.

Table 6-4 shows the final dummy variables and data used for modeling. The area type variable includes two alternatives, urban and rural. Rural was considered as a base (or reference) dummy variable. The congestion (or LOS) indicators include LOS A, B, C, D, and E. The LOS E was considered a base dummy variable. The time-of-the-day variable includes morning, evening, and off-peak hours. The off-peak hours were considered as a base dummy variable. Two alternatives (i.e., weekdays, and weekends) were considered for the time-of-the-week variable. The weekdays were considered as a base dummy variable. Dummy variables were created for each two months of the year from January to December. The months of July and August (summer months) were considered as a base variable. Speed limit dummy variables were created for speed limits of 50 to 70 mph, with speed limits of 50 and 55 being considered as a base dummy variable. Curve radius (R) was divided into four categories of $100 < R \leq 2500$, $2500 < R \leq 5000$, $5000 < R \leq 7500$, and $7500 < R$, with the range of $100 < R \leq 2500$ being a base dummy variable.

Table 6-4. Data Description.

Variables		**Classes**	**Definition of Dummy Variables**
Area type	**Rural/Urban**	Rural (=0)	Dummy variable denoting rural areas (base)
Area type	**Rural/Urban**	Urban	Dummy variable denoting urban areas
LOS Indicators	**Traffic Density (K) (or LOS)**	$0 < K \leq 11$ (LOS=A)	Density of 0 - 11 vehicle/mile/lane (LOS=A)
LOS Indicators	**Traffic Density (K) (or LOS)**	$11 < K \leq 18$ (LOS=B)	Density of 11 -18 vehicle/mile/lane (LOS=B)
LOS Indicators	**Traffic Density (K) (or LOS)**	$18 < K \leq 26$ (LOS=C)	Density of 18 -26 vehicle/mile/lane (LOS=C)
LOS Indicators	**Traffic Density (K) (or LOS)**	$26 < K \leq 35$ (LOS=D)	Density of 26 - 35 vehicle/mile/lane (LOS=D)
LOS Indicators	**Traffic Density (K) (or LOS)**	$35 < K \leq 45$ (LOS=E) (=0)	Density of 35 -45 vehicle/mile/lane (LOS=E) (base)
Time Indicators	**Time of the Day**	Off Peak (=0)	Off peak hours (10 am to 3 pm and 7 pm to 6 am) (base)
Time Indicators	**Time of the Day**	Morning Peak Period	Morning Peak hours (6 am to 10 am)
Time Indicators	**Time of the Day**	Evening Peak Period	Evening peak hours (3 pm to 7 pm)
Time Indicators	**Time of the Week**	Weekday (=0)	Weekdays (Monday to Friday) (base)
Time Indicators	**Time of the week**	Weekend	Weekends (Saturday and Sunday)
Time Indicators	**Time of the Year (Month)**	Jan. – Feb.	Months of January and February
Time Indicators	**Time of the Year (Month)**	Mar. – Apr.	Months of March and April
Time Indicators	**Time of the Year (Month)**	May – Jun.	Months of May and June
Time Indicators	**Time of the Year (Month)**	Jul. – Aug. (=0)	Months of July and August (base)
Time Indicators	**Time of the Year (Month)**	Sep. – Oct.	Months of September and October
Time Indicators	**Time of the Year (Month)**	Nov. – Dec.	Months of November and December

Curve or Ramp Variables[1]	**Speed Limit (mph)**	Speed Limit =50/55 (=0)	Segments with speed limit less than or equal to 50 or 55 mph (base)
Curve or Ramp Variables[1]	**Speed Limit (mph**	Speed Limit = 60 mph	Segments with a speed limit of 60 mph
Curve or Ramp Variables[1]	**Speed Limit (mph**	Speed Limit = 65 mph	Segments with a speed limit of 65 mph
Curve or Ramp Variables[1]	**Speed Limit (mph**	Speed Limit = 70 mph	Segments with a speed limit of 70 mph
Curve or Ramp Variables[1]	**Radius (ft.) (R)**	$100 < R \leq 2500$ (=0)	Radius of the curve/ramp is from 100 to 2500 feet (base)
Curve or Ramp Variables[1]	**Radius (ft.) (R)**	$2500 < R \leq 5000$	Radius of the curve/ramp is from 2500 to 5000 feet
Curve or Ramp Variables[1]	**Radius (ft.) (R)**	$5000 < R \leq 7500$	Radius of the curve/ramp is from 5000 to 7500 feet
Curve or Ramp Variables[1]	**Radius (ft.) (R)**	$7500 < R$	Radius of the curve/ramp is above 7500 feet
Curve or Ramp Variables[1]	**Superelevation (%) (SUP)**	$SUP \leq -6$	Superelevation of the curve is less than -6 %
Curve or Ramp Variables[1]	**Superelevation (%) (SUP)**	$-6 < SUP \leq -3$	Superelevation of the curve is from -6 to -3 %
Curve or Ramp Variables[1]	**Superelevation (%) (SUP)**	$-3 < SUP \leq 3$ (=0)	Superelevation of the curve is from -3 to 3 % (base)
Curve or Ramp Variables[1]	**Superelevation (%) (SUP)**	$3 < SUP \leq 6$	Superelevation of the curve is from 3 to 6 %
Curve or Ramp Variables[1]	**Superelevation (%) (SUP)**	$6 < SUP$	Superelevation of the curve is above 6 %
Curve or Ramp Variables[1]	**Arc Angle (ARC) (degree) (For Curves)[1]**	$0 < ARC \leq 15$ (=0)	Arc angle of the curve is below 15 degrees (base)
Curve or Ramp Variables[1]	**Arc Angle (ARC) (degree) (For Curves)[1]**	$15 < ARC \leq 30$	Arc angle of the curve is from 15 to 30 degrees
Curve or Ramp Variables[1]	**Arc Angle (ARC) (degree) (For Curves)[1]**	$30 < ARC$	Arc angle of the curve is above 30 degrees
Curve or Ramp Variables[1]	**Arc Angle (degree) (ARC) (For Ramps)[1]**	$0 < ARC \leq 10$ (=0)	Arc angle of the ramp is below 10 degrees (base)

Curve or Ramp Variables[1]	**Arc Angle (degree) (ARC) (For Ramps)**[1]	10 < ARC ≤ 20	Arc angle of the ramp is from 10 to 20 degrees
Curve or Ramp Variables[1]	**Arc Angle (degree) (ARC) (For Ramps)**[1]	20 < ARC ≤ 30	Arc angle of the ramp is from 20 to 30 degrees
Curve or Ramp Variables[1]	**Arc Angle (degree) (ARC) (For Ramps)**[1]	30 < ARC	Arc angle of the ramp is above 30 degrees
Curve or Ramp Variables[1]	**Lane Width (ft.) (SW) (for Ramps)**[2]	Wide Lane	Lane width ≥ 12 feet (base)
Curve or Ramp Variables[1]	**Lane Width (ft.) (SW) (for Ramps)**[2]	Narrow lane width	Lane width < 12 feet

[1]Arc angle and curvature were correlated; since models with the arc angle variable provided a better fit., only the arc angle variable (for horizontal curves and ramps) included in the model and reported in the table.
[2]For horizontal curves, the left shoulder, right shoulder, and lane width variables were insignificant. Therefore, they were not reported in this table. For ramps, these variables were correlated, but the models with the lane width variable provided a better fit. Therefore, only the information about the lane width dummy was included in this table.

Superelevation (SUP) was divided into four categories of -6 < SUP ≤ -3, -3 < SUP ≤ 3, 3 < SUP ≤ 6, and 6 < SUP, with the common superelevation of -3 < SUP ≤ 3 considered as a base dummy variable. For horizontal curves, the arc angle (ARC) variable was divided into three groups of 0 < ARC ≤ 15, 15 < ARC ≤ 30, and 30 < ARC, with 0 < ARC ≤ 15 being the base dummy variable. For Interstate ramps, the arc angle variable was divided into four groups of 0 < ARC ≤ 10, 10 < ARC ≤ 20, 20 < ARC ≤ 30, and 30 < ARC, with 0 < ARC ≤ 10 being the base dummy variable. The lane width indicator (for ramps) represents two categories: a narrow lane width of less than 12ft., and a wide lane width of 12 ft. and above. The variable indicating the wide lane width of 12 ft. and above was considered a base dummy variable.

6.3 Methodology

A mixed effect binomial model was used to model speeding [***Error! Bookmark not defined.**,**Error! Bookmark not defined.***,*37*]. The mixed effect model was used to account for the location heterogeneity, and repeated measures for each segment. In this model, it is assumed that in every hour, V_{ij} vehicles pass the i-th curve or ramp segment. From those, with a probability of p_i, y_{ij} out of V_{ij} vehicles speed, and (V_{ij}-y_{ij}) out of V_{ij} vehicles do not speed. Therefore, for each j-th time step, the binomial model can be written as follows [*38*].

$$y_{ij} \sim \text{Binomial}(P_{ij}, V_{ij}) \equiv \binom{V_{ij}}{y_{ij}} P_{ij}^{\,y_{ij}} (1-P_{ij})^{V_{ij}-y_{ij}} \quad (2)$$

Using a mixed effect logit link function, we correlated the log odds of speeding, i.e., $\text{Ln}(P_{ij}/(1-P_{ij}))$, with a set of dummy variables as follows:

$$\text{Logit}(P_{ij}) = \text{Ln}(P_{ij}/(1-P_{ij})) \sim \pi + K_{ij} + D_j + W_j + M_j + Area_i, R_i + Sup_i + Arc_i + SL_i + LW_i + \varepsilon_i \quad (3)$$

where,
π: intercept (constant).

K_ij: dummy variable denoting density (level of service) on the i-th element and the j-th time.
D_j: dummy variable denoting time of the day for the j-th time (morning, evening, or off-peak).
W_j: dummy variable denoting time of the week for the j-th time (weekdays, or weekends).
M_j: dummy variable denoting the month of the year for the j-th time (Jan.-Feb. to Nov.-Dec.).
Area_i: dummy variable denoting the area of i-th element (urban=1, and rural=0).
R_i: dummy variable denoting the radius range for the i-th element.
Sup_i: dummy variable denoting the superelevation range for the i-th element.
Arc_i: dummy variable denoting the arc angle range for the i-th element.
SL_i: dummy variable denoting the speed limit on the i-th element.
LW_i: dummy variable denoting a narrow lane width (< 12 ft.) for the i-th element.
ε_i: the random error component for the i-th element.

6.4 Modeling Results

Table 6-5. Modeling Results for Interstate Horizontal Curves.

Category		Variable	+10 mph Speeding		+15 mph Speeding		+20 mph Speeding	
			Mean (S.E.)	Odds Ratio	Mean (S.E.)	Odds Ratio	Mean (S.E.)	Odds Ratio
Constant	Constant	Intercept	-3.114 (0.01788)	-	-4.108 (0.10750)	-	-5.875 (0.19393)	-
Area	Area Type[1]	Urban	-0.448 (0.02503)	0.64	-0.454 (0.07355)	0.63	-0.255 (0.10192)	0.77
Time Variables	Time of Year (Month)[2]	Jan. – Feb.	-0.304 (0.00015)	0.74	-0.244 (0.00018)	0.78	-0.199 (0.00029)	0.82
Time Variables	Time of Year (Month)[2]	Mar. – Apr.	-0.136 (0.00014)	0.87	-0.079 (0.00017)	0.92	-0.049 (0.00026)	0.95
Time Variables	Time of Year (Month)[2]	May – Jun.	-0.204 (0.00013)	0.82	-0.135 (0.00015)	0.87	-0.072 (0.00024)	0.93
Time Variables	Time of Year (Month)[2]	Sep. – Oct.	-0.093 (0.00013)	0.91	-0.067 (0.00015)	0.93	-0.052 (0.00023)	0.95
Time Variables	Time of Year (Month)[2]	Nov. – Dec.	-0.234 (0.00014)	0.79	-0.184 (0.00017)	0.83	-0.159 (0.00026)	0.85
Time Variables	Time of the Day[3]	Morning Peak	0.070 (0.00011)	1.07	0.115 (0.00013)	1.12	0.125 (0.00020)	1.13
Time Variables	Time of the Day[3]	Evening Peak	0.062 (0.00010)	1.06	0.108 (0.00012)	1.11	0.118 (0.00018)	1.12
Time Variables	Time of the Week[4]	Weekend	0.315 (0.00010)	1.37	0.294 (0.00012)	1.34	0.283 (0.00019)	1.33
LOS Variables	Level of Service[5]	LOS=A	1.060 (0.00080)	2.89	0.940 (0.00116)	2.56	0.723 (0.00184)	2.06
LOS Variables	Level of Service[5]	LOS=B	0.999 (0.00079)	2.72	0.883 (0.00116)	2.42	0.641 (0.00184)	1.90
LOS Variables	Level of Service[5]	LOS=C	0.742 (0.00079)	2.10	0.596 (0.00116)	1.82	0.355 (0.00184)	1.43
LOS Variables	Level of Service[5]	LOS=D	0.379 (0.00083)	1.46	0.314 (0.00121)	1.37	0.221 (0.00192)	1.25
Curve Characteristics	Speed Limit (mph)[6]	Speed Limit = 60	Insig.	-	Insig.	-	0.812 (0.18996)	2.25
Curve Characteristics	Speed Limit (mph)[6]	Speed Limit = 65	1.522 (0.02508)	4.58	1.416 (0.08284)	4.12	1.812 (0.17296)	6.12
Curve Characteristics	Speed Limit (mph)[6]	Speed Limit = 70	1.951 (0.03005)	7.04	1.896 (0.10560)	6.66	2.454 (0.19799)	11.63
Curve Characteristics	Curve Radii (ft) (R)[7]	2500 < R ≤ 5000	0.255 (0.02670)	1.29	0.322 (0.08957)	1.38	0.500 (0.12568)	1.65

Curve Characteristics	**Curve Radii (ft) (R)[7]**	**5000 < R ≤ 7500**	0.596 (0.02226)	1.81	0.657 (0.08802)	1.93	0.673 (0.12228)	1.96
Curve Characteristics	**Curve Radii (ft) (R)[7]**	**7500 < R**	Insig.	-	Insig.	-	Insig.	-
Curve Characteristics	**Superelevation (%) (SUP)[8]**	**SUP < - 3**	Insig.	-	Insig.	-	Insig.	-
Curve Characteristics	**Superelevation (%) (SUP)[8]**	**3 < SUP ≤ 6**	0.235 (0.03532)	1.27	0.256 (0.10960)	1.29	Insig.	-
Curve Characteristics	**Superelevation (%) (SUP)[8]**	**6 < SUP**	0.851 (0.08228)	2.34	Insig	-	Insig.	-
Curve Characteristics	**Arc Angle (degree) (ARC)[9]**	**30 < ARC**	-0.248 (0.03123)	0.78	-0.265 (0.09182)	0.77	-0.225 (0.12778)	0.80

[1]rural area was considered as a base dummy variable.
[2]months of July and August were considered as a base variable.
[3]off-peak hours were considered as a base variable.
[4]weekdays was considered as a base variable.
[5]LOS of E was considered as a base dummy variable.
[6]speed limits of 50 and 55 mph were considered as a base dummy variable.
[7] curve radii of 100 < R ≤ 2500 was considered as a base dummy variable.
[8]supperelavation of -3 to 3, i.e., -3 < SUP ≤ 3, was considered as a base dummy variable.
[9]arc angle between 0 and 15 (ARC ≤ 15), was considered as a base dummy variable.

This section presents the modeling results and is divided into two parts. In the first part, the modeling results for Interstate horizontal curves are documented. In the second part, the modeling results for Interstate ramps are presented. Metrics such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) are used to find the best models. Significant variables at the 90% confidence interval are reported in the tables.

6.4.1 Horizontal Curve Models

Table 6-5 shows the modeling results for Interstate horizontal curves; the modeling results show that speeding in rural areas occurs at higher rates compared to urban areas. This is presumably due to factors such as reduced speed enforcement in rural areas, as well as encountering unexpected or unfamiliar curves when driving on rural compared to urban roads. Previous studies have shown that crash rates are higher on rural horizontal curves compared to urban curves [*39*,*40*]. The higher rate of speeding in rural areas could be one of the reasons for observing higher crash rates in rural areas. Our results particularly show that the odds of speeding by 10, 15, and 20 mph decreases by 36%, 37%, and 23%, respectively, on urban Interstate horizontal curves compared to those on rural roads.

As noted earlier, Maine experiences long winter seasons, often starting in October and ending in April. To investigate the impact of the month of the year on speeding behavior on Interstate horizontal curves, we created dummy variables for every two months, starting with January and February (Jan. – Feb.), and ending with November and December (Nov.- Dec.) and included them in the model. The two months of July-August (Jul.-Aug.) were considered as the base dummy variable. The modeling results show decreased odds of speeding in all month periods compared to Jul.-Aug., with the most significant effects observed in Jan.- Feb., and then Nov.-Dec. with frequent snowfall and cold winters. Particularly, the odds of speeding by more than 10, 15, 20 mph decreases by around 26%, 22%, 18% in Jan.- Feb., and by 21%, 17%, and 15% in Nov.-Dec., respectively, compared to Jul.-Aug. Modeling results show increased odds of speeding during morning and evening peak hours and during the weekends on Interstate horizontal curves.

To account for congestion on horizontal curves, we calculated the level of service (LOS) and used it to represent the level of congestion or a surrogate measure of traffic density. LOS E was considered as the baseline in the model. As shown in Table 6-2, LOS impacts speeding on horizontal curves. LOSs C and D are commonly observed on Interstates including curves and ramps. The models show increased odds of 46%, 37%, and 25% for speeding by more than 10, 15, and 20 mph for LOS D compared to LOS E. Similarly for LOS C, the odds of speeding by more than 10, 15, and 20 mph increases by 110%, 82%, and 43% respectively. The odds of speeding increases significantly for LOSs A and B. Particularly, the odds of speeding by more than 10 mph are 2.89 and 2.72 times more when the roadway operates at LOSs A and B compared to LOS E. These results show the importance of implementing appropriate countermeasures or installing signage such as advisory speed signs on low-volume horizontal curves.

The results show significant increases in odds of speeding on horizontal curves with higher speed limits. This finding is interesting but is different from patterns identified in relevant previous studies [***Error! Bookmark not defined.,Error! Bookmark not defined.Error! Bookmark not defined.Error! Bookmark not defined.**,*41*]. Particularly, compared to speed limits of 50 and 55 mph, the odds of speeding by more than 10, 15, and 20 mph increases by 4.58, 4.12 and 6.12 times when the speed limit is 65 mph, and by 7.04, 6.66, and 11.63 times when the speed limit is 70 mph. While these results could partially be due to higher design standards on high-speed horizontal curves, the modeling results show the importance of speed enforcement and proper countermeasures to reduce speeding on horizontal curves with higher speed limits.

The modeling results show that as the curve radius increases, the odds of speeding on Interstate horizontal curves increase as well. Particularly, the odds of speeding by more than 10, 15, and 20 mph increase by 29%, 38%, and 65% on Interstate horizontal curves with a radius of $2,500 < R \leq 5,000$ and by 81%, 93% and 96% on curves with a radius of $5,000 < R \leq 7,500$ compared to the lower radius of $0 < R \leq 2,500$. These results are interesting from the fact that on curves with smaller radii, lane departure is more likely to happen because of increased centrifugal forces. Therefore, curves with larger radii are often considered safer in terms of lane departure risk; however, our results show that curves with larger radii are more prone to speeding. As noted above, speeding is a surrogate measure of safety, and often results in severe or fatal lane departure crashes on horizontal curves. These results show that speeding activities probably are associated with drivers' perceived risk. Curves with large radii give drivers the impression of being safe. Therefore, they are more likely to speed on such curves than on sharper curves with smaller radii.

Table 6-5 shows that the odds of speeding do not change for curves with superelevation less than -3% compared to the baseline superelevation of -3% to 3%. Speeding, however, could increase on horizontal curves with higher superelevation. Particularly, the odds of speeding by more than 10 mph increases by 27% when superelevation is between 3% and 6% and by 134% when superelevation is greater than 3% compared to the baseline superelevation of -3% to 3%. With appropriate superelevation, drivers on sharp curves can feel more comfortable and are able to better overcome the impact of the centrifugal force. Therefore, speeding could occur at higher rates compared to when the superelevation is insufficient. The impact of larger superelevation on speeding by more than 20 mph was insignificant. The odds of speeding decrease when the arc angle is above 30 degrees. These results are expected, since shaper curves have larger arc angles. Finally, the impact of lane width on speeding was found to be insignificant for Interstate horizontal curves given that they were all 12ft lanes.

6.4.2 Ramp Models

Table 6-6 shows the speeding models for the Interstate ramps. Like horizontal curves, the odds of speeding decrease on urban ramps compared to rural ramps. This observation, presumably, is for the same reason as the one noted for Interstate horizontal curves. As noted above, speed enforcement as well as roadway familiarity (especially curve expectation) are often lower in rural areas compared to urban areas. Likewise, the odds of speeding on ramps decrease significantly during the months of Nov. through Feb. This is mainly due to the peak of the winter season during these months of the year in Maine, which often involves frequent snowfall, and icy and frozen roads. Specifically, the odds of speeding by more than 10, 15, and 20 mph decrease by 22%, 21%, and 22% during the Nov.- Dec. duration and by 20%, 18% and 15% during the Jan.- Feb. duration compared to the Jul.-Aug. duration.

Table 6-6. Modeling Results for Interstate Ramps.

Category		Variable	+10 mph Speeding		+15 mph Speeding		+20 mph Speeding	
			Mean (S.E.)	Odds Ratio	Mean (S.E.)	Odds Ratio	Mean (S.E.)	Odds Ratio
Constant	**Constant**	**Intercept**	-2.941 (0.06299)	-	-4.616 (0.58484)	-	-8.356 (0.07796)	-
Area	**Area Type**[1]	**Urban**	-0.842 (0.06781)	0.43	-1.035 (0.46075)	0.36	Insig.	-
Time Variables	**Time of Year (Month)**[2]	**Jan. Feb.**	-0.219 (0.00050)	0.80	-0.200 (0.00061)	0.82	-0.168 (0.00084)	0.85
Time Variables	**Time of Year (Month)**[2]	**Mar. – Apr.**	-0.100 (0.00047)	0.90	-0.091 (0.00057)	0.91	-0.072 (0.00078)	0.93
Time Variables	**Time of Year (Month)**[2]	**May – Jun.**	-0.099 (0.00044)	0.91	-0.086 (0.00053)	0.92	-0.063 (0.00073)	0.94
Time Variables	**Time of Year (Month)**[2]	**Sep. – Oct.**	-0.078 (0.00043)	0.93	-0.082 (0.00052)	0.92	-0.091 (0.00072)	0.91
Time Variables	**Time of Year (Month)**[2]	**Nov. - Dec.**	-0.251 (0.00047)	0.78	-0.242 (0.00057)	0.79	-0.247 (0.00080)	0.78
Time Variables	**Time of the Day**[3]	**Morning Peak**	0.135 (0.00037)	1.14	0.135 (0.00045)	1.14	0.117 (0.00063)	1.12
Time Variables	**Time of the Day**[3]	**Evening Peak**	0.040 (0.00034)	1.04	0.072 (0.00041)	1.07	0.080 (0.00057)	1.08
Time Variables	**Time of the Week**[4]	**Weekend**	0.241 (0.00037)	1.27	0.235 (0.00044)	1.26	0.205 (0.00061)	1.23
LOS Variables	**Level of Service**[5]	**LOS=A**	0.387 (0.00077)	1.47	0.410 (0.00100)	1.51	0.471 (0.00147)	1.60
LOS Variables	**Level of Service**[5]	**LOS=B**	0.285 (0.00067)	1.33	0.322 (0.00090)	1.38	0.362 (0.00135)	1.44
LOS Variables	**Level of Service**[5]	**LOS=C**	0.266 (0.00063)	1.31	0.310 (0.00086)	1.36	0.322 (0.00130)	1.38
LOS Variables	**Level of Service**[5]	**LOS=D**	0.202 (0.00062)	1.22	0.238 (0.00085)	1.27	0.239 (0.00129)	1.27
Ramp Characteristics	**Speed Limit (mph)**[6]	**Speed Limit = 65**	0.916 (0.09732)	2.50	1.297 (0.41143)	3.66	1.506 (0.13674)	4.51

Ramp Characteristics	**Curve Radii (ft) (R)**[7]	**2500 < R ≤ 5000**	Insig.	-	Insig.	-	Insig	-
Ramp Characteristics	**Curve Radii (ft) (R)**[7]	**5000 < R ≤ 7500**	Insig.	-	Insig.	-	Insig	-
Ramp Characteristics	**Curve Radii (ft) (R)**[7]	**7500 < R**	Insig.	-	Insig	-	Insig	-
Ramp Characteristics	**Superelevation (%) (SUP)**[8]	**SUP < - 3**	Insig.	-	Insig	-	Insig	-
Ramp Characteristics	**Superelevation (%) (SUP)**[8]	**3 < SUP ≤ 6**	0.959 (0.11643)	2.61	1.750 (0.41834)	5.75	1.630 (0.13608)	5.10
Ramp Characteristics	**Superelevation (%) (SUP)**[8]	**6 < SUP**	1.614 (0.75383)	5.02	1.631 (0.85125)	5.11	Insig	-
Ramp Characteristics	**Arc Angle (degree)(ARC)**[9]	**10 < ARC ≤ 20**	-1.735 (0.12069)	0.18	-2.411 (0.57432)	0.09	-2.232 (0.16803)	0.11
Ramp Characteristics	**Arc Angle (degree)(ARC)**[9]	**20 < ARC ≤ 30**	-1.820 (0.25356)	0.16	-3.003 (0.58499)	0.05	-3.363 (0.24915)	0.03
Ramp Characteristics	**Arc Angle (degree)(ARC)**[9]	**30 < ARC**	-1.138 (0.09288)	0.32	-1.523 (0.48606)	0.22	Insig.	-
Ramp Characteristics	**Narrow Lane (ft.) (LW)**[10]	**LW < 12 ft.**	-1.742 (0.14519)	0.18	-1.801 (0.75017)	0.17	-3.192 (0.25396)	0.04

[1]rural area was considered as a base dummy variable.
[2]months of July and August were considered as a base variable.
[3]off-peak hours were considered as a base variable.
[4]weekdays were considered as a base variable.
[5]LOS of E was considered as a base dummy variable.
[6]speed limits of 50 and 55 mph were considered as a base dummy variable.
[7] curve radii of 100 < R ≤ 2500 was considered as a base dummy variable.
[8]supperelavation of -3 to 3, i.e., -3 < SUP ≤ 3, was considered as a base dummy variable.
[9]arc angle between 0 and 10 (ARC ≤ 10), was considered as a base dummy variable.
[10]wider lane width (≥12) was considered as a base dummy variable.

Regarding LOS, as expected, the odds of speeding increase as the LOS improves. However, the increase in odds of speeding is not as high as the rates observed in models developed for Interstate horizontal curves. This observation could presumably be due to the driver's discomfort and challenges to negotiate Interstate ramps regardless of the congestion range. Specifically, the odds of speeding by more than 10 15, and 20 mph, at LOS D increases by 22%, 27%, and 27%, and at LOS C by 31%, 36%, and 38% compared to LOS E. Like horizontal curve results, models show the importance of speed enforcement, or use of advisory speed signage on low-volume ramps. The odds of speeding on ramps increase during morning and evening peak hours compared to off-peak hours, and during the weekends compared to weekdays.

Like horizontal curves, the odds of speeding increase at Interstate ramps with a higher speed limit of 65 mph. The odds of speeding by more than 10, 15, and 20 mph increases by 2.50, 3.66, and 4.51 times compared to when the speed limit is 55 mph. The impact of ramp radius is insignificant for all Interstate ramp models. This could again be due to the driver's discomfort and challenges to drive on ramps even when the radius is large. The impact of superelevation was insignificant for superelevation of smaller than -3%; however, the impact of larger superelevation was significant. Particularly the odds of speeding by more than 10, 15 and 20 mph increase by 2.61, 5.75 and 5.1 times when superelevation is between 3% and 6% compared to the baseline superelevation of -3% to 3%. Similarly, the odds of speeding by 10, and 15 mph increases by 5.02 and 5.11 times when superelevation is greater than 6% compared to the baseline superelevation of -3% to 3%.

The arc angle variable was significant for most Interstate ramp models. The results show a significant decrease in odds of speeding as the arc angle variable increases. These results are expected as speeding on sharp ramps could be challenging. This particularly showed in speeding at ramps with an arc angle of 20 to 30 degrees compared to those with arc angle of less than 10 degrees. As indicated in Table 6-6, the odds of speeding by more than 10, 15, and 20 mph decreases by 87%, 96%, and 95% at these ramps. Finally, the effect of narrow lane width was found to be significant and negatively associated with speeding on Interstate ramps potentially due to the inherent difficulties and the driver's discomfort to drive or speed on ramps with smaller lane width.

6.5 Summary and Conclusions

Lane departure crashes account for over 70% of roadway fatalities in Maine. Speeding is a main contributing factor in many of these collisions, especially on horizontal curves. This research examined the impact of various curve and ramp geometric characteristics, level of service, area type, and time factors on speeding at Interstate horizontal curves and ramps. Mixed-effect binomial regression models were developed, and the odds of speeding were computed and discussed. The results show increased odds of speeding on curves and ramps located in rural compared to urban areas presumably due to reduced speed enforcement and decreased roadway familiarity (or curve expectation) on rural roads. The odds of speeding on these road elements increase during the morning and evening peak hours, and weekends. In contrast, as expected, speeding decreases significantly during the peak of winter seasons with frequent snowfall, cold weather, and icy and frozen roads. Regarding congestion, speeding occurs at higher rates on less congested Interstate horizontal curves and ramps. These results suggest the importance of using enforcement or advisory speed signs (e.g., dynamic signs) on low-volume horizontal curves or ramps. For both horizontal curves and ramps, superelevation greater than 3% results in increased odds of speeding compared to the common superelevation of -3 to 3%. Curve radius was also found to be a significant factor. Higher odds of speeding observed as curve radius increases. The odds of speeding decrease at both horizontal curves and ramps when the arc angle variable increases. These results provide useful information to allocate speed enforcement resources to critical curve and ramps segments during times when speeding occurs at higher rates. It could also help to develop and prioritize safety countermeasures to reduce speeding.

6.6 Practical Applications

Speeding is an important contributing factor in many lane departure collisions on horizontal curves and ramps. It is crucial to allocate limited resources, such as funding or law enforcement, to high-priority locations and time windows prone to speeding activities. This research investigated locations and time windows that experience high rates of speeding activities on Interstate horizontal curves and ramps. It was found that the odds of speeding are higher in rural compared to urban areas, and at curves with larger radii and superelevation, and smaller arc angles. For ramps, the odds of speeding are higher on sites with larger superelevation and lane widths and smaller arc angles. These results show the importance of implementing advisory speed signs or other speed management countermeasures at locations with the above geometric characteristics. The results also show a significant increase in odds of speeding for less congested horizontal curves and ramps. This suggests the importance of speed enforcement on low-volume horizontal curves and ramps or during off-peak periods. Finally, the odds of speeding increase during the morning and evening peak hours and the weekends, indicating the need for more intensive enforcement during these time windows.

This page left blank intentionally.

7.0 Summary and Conclusions

Data is becoming increasingly important to state DOTs for strategic and day-to-day decision-making. This project aims to (1) provide DOTs with a clear and comprehensive picture of their data assets, needs, data analytics, and other data practices related to Transportation Systems Management and Operations (TSMO); (2) offer strategic and practical recommendations to prepare DOTs for future transportation data analytics; and (3) demonstrate the potential of Artificial Intelligence (AI) techniques and emerging data sources through three case studies to improve TSMO.

This research begins with a comprehensive review of data and data sources, presented in Section 2.1. The discussed data sources are further summarized in Table 2-1, indicating a substantial change in the landscape of traffic data collection over the past two decades. Mobile devices and emerging connected vehicles have significantly expanded the coverage of traditional sensors, such as loop detectors and cameras, providing a maintenance-free approach for transportation agencies to collect detailed data elements, such as vehicle trajectories. Another significant aspect is the widespread application of AI technologies in sensor data processing and modeling, generating valuable traffic measurements for data-driven decision-making.

The review is followed by interviews with domain experts. Overall, respondents recognize the need to

- Share data across divisions,
- Integrate data from different sources,
- Protect data privacy and security,
- Invest in data analytics using advanced tools and workforce development, and
- Introduce and develop innovative data analysis methods.

In general, they express satisfaction with the quality of probe and crowdsourced data. However, they acknowledge that such data may be less reliable in rural areas or during off-peak periods due to low sample size issues. All respondents rely on both private data vendors and their own infrastructure for data collection. Maintenance is a critical factor in making data infrastructure investment decisions. DOTs often find it challenging to keep up with technological evolution, posing challenges to maintaining and upgrading data collection infrastructure on their own. This may also lead to some technologies becoming obsolete quickly. Therefore, DOTs sometimes prefer to let private companies bear this risk and purchase data products from them. All respondents express interest in AI applications in TSMO, although most DOTs hire consultants for their existing AI-related work.

Based on the review and interviews, Chapter 3 provides recommendations to DOTs regarding transportation data analytics. DOTs are encouraged to continue exploring data from sources such as mobile devices, probe vehicles, and connected vehicles, which offer extensive coverage. Additionally, they should consider implementing AI and edge computing-powered roadside sensors at strategic locations or on a portable platform. While these sensors may have limited coverage compared to data generated by mobile devices or crowdsourcing, they capture rich information from all vehicles rather than a small sample, providing ground truth data. DOTs are also advised to invest in workforce development focused on data analytics and to restructure their organizations to facilitate both data sharing and synergistic data analysis.

Three case studies were conducted to demonstrate how AI and data from advanced radar and thermal camera sensors, along with emerging sources, can help DOTs understand driver speed and lane-changing behavior on horizontal curves and prior to a highway work zone. Specifically, the first case study focused on speed behavior on highway horizontal curves. Ultra-high-definition radar and thermal cameras were used to collect traffic data at five sites along I-93 in New Hampshire. AI models were developed to analyze the data, generating vehicle counts, headways, and speed distributions and profiles. AI models also performed camera view change detection, risky behavior detection, and vehicle merging point analysis. The data suggested that drivers do not change speed significantly when approaching a horizontal curve, probably because of the high geometric design standards of I-93. The camera view change detection algorithm can be useful for future AI + edge computing deployments for sensor self-calibration or generating sensor recalibration warnings.

The second case study investigated how drivers adjust speed and where they change lanes when approaching a work zone, which was equipped with two flashing speed limit signs on the two sides of the highway and a portable changeable message sign (PCMS). The speed data from radar and the data on vehicle lane-changing behavior, generated by AI models, suggested that both the flashing speed limit signs and the PCMS were effective in prompting vehicles to reduce speed and change lanes.

The final case study integrated probe data and road inventory data to model speeding activities on horizontal curves and ramps at a network level. The results showed that speeding occurs more frequently (1) in rural compared to urban areas, (2) at curves with larger radii and superelevation, smaller arc angles, (3) on less congested curves and ramps, and (4) during the morning and evening peak hours and on weekends.

The review and interview results, along with the recommendations, are directly usable by state DOTs for making informed decisions related to transportation data analytics. These case studies illustrate the benefits of utilizing detailed vehicle trajectories collected by a portable platform and how datasets from various sources can complement each other, providing a comprehensive view of driver behavior to improve highway traffic operations and safety. The results of the three case studies can assist DOTs in developing improved work zone temporary traffic control plans and strategies to address speeding on curves and ramps.

8.0 References

1 The Eastern Transportation Coalition. (2019). Optimal Traffic Monitoring in a New Data Age.

2 Tanaka, A., Schroeder, B., Trask, L., & Chase, T. (2019). NCDOT Guide on Automated Traffic Signal Performance Measures. North Carolina Department of Transportation.

3 Enwemeka, Z. (2015). Boston Tapping Waze Data to Help Improve Traffic Signals, Ease Congestion. Available online at https://www.wbur.org/news/2015/02/13/waze-boston-data-partnership, accessed on 10/26/2021.

4 Desai, J., Li, H., Mathew, J. K., Cheng, Y. T., Habib, A., & Bullock, D. M. (2021). Correlating hard-braking activity with crash occurrences on interstate construction projects in Indiana. *Journal of Big Data Analytics in Transportation*, *3*(1), 27-41.

5 The Eastern Transportation Coalition. (2020). Hurricane Proof of Concept Results States' Experience with Real-Time Connected Vehicle Data Coalition.

6 Thompson, S.A. & Warzel, C. (2019). Twelve Million Phones, One Dataset, Zero Privacy. Available online at https://www.nytimes.com/interactive/2019/12/19/opinion/location-tracking-cell-phone.html, accessed on 10/26/2021.

7 Keegan, J. & Ng, A. (2021). There's a Multibillion-Dollar Market for Your Phone's Location Data. Available online at https://themarkup.org/privacy/2021/09/30/theres-a-multibillion-dollar-market-for-your-phones-location-data, accessed on 10/26/2021.

8 Verizon. (2017). How the Internet of Things can create congestion-free cities - Verizon's Intelligent Traffic Management solution is already changing the way cities plan their streets. Available online at https://www.verizon.com/about/news/how-internet-things-can-create-congestion-free-cities, accessed on 10/26/2021.

9 AT&T. (2018). AT&T Expands Smart Cities Offerings with New Structure Monitoring Solution for U.S. Railways and Roadways. Available online at https://about.att.com/story/structure_monitoring_solution_for_railways_and_roadways.html, accessed on 10/26/2021.

10 Sawers, P. (2020). StreetLight Data raises $15 million to bring big data analytics to city transport Available online at https://venturebeat.com/2020/08/06/streetlight-data-raises-15-million-to-bring-big-data-analytics-to-city-transport/, accessed on 10/26/2021.

11 Lambert, F. (2020). Tesla is collecting insane amount of data from its Full Self-Driving test fleet. Available online at https://electrek.co/2020/10/24/tesla-collecting-insane-amount-data-full-self-driving-test-fleet/, accessed on 10/26/2021.

12 FHWA (2021). Sources Sought Notice 693JJ3-21-SS-0013 AI for ITS. Available online at https://govtribe.com/file/government-file/693jj3-21-ss-0013-ai-for-its-dot-pdf, accessed on 10/26/2021.

13 Bank, Dor, Noam Koenigstein and Raja Giryes. "Autoencoders." ArXiv abs/2003.05991 (2020): n. pag.

14 Kingma, Diederik P. and Max Welling. "An Introduction to Variational Autoencoders." ArXiv abs/1906.02691 (2019): n. pag.

15 Wu, L., Dadashova, B., Geedipally, S., Pratt, M. P., & Shirazi, M. (2021). Using naturalistic driving study data to explore the association between horizontal curve safety and operation on rural two-lane highways. Journal of Transportation Safety & Security, 13(8), 896–913. https://doi.org/10.1080/19439962.2019.1697776.

16 Al-Bdairi, N. S. S., & Behnood, A. (2021). Assessment of temporal stability in risk factors of crashes at horizontal curves on rural two-lane undivided highways. Journal of safety research, 76, 205-217.

17 Islam, M., Hosseini, P., & Jalayer, M. (2022). An analysis of single-vehicle truck crashes on rural curved segments accounting for unobserved heterogeneity. Journal of safety research, 80, 148-159.

18 Mannering, F. L., & Washburn, S. S. (2020). Principles of highway engineering and traffic analysis (Seventh edition). Wiley.

19 Pratt, M. P., Geedipally, S. R., Dadashova, B., Wu, L., & Shirazi, M. (2019). Familiar versus Unfamiliar Drivers on Curves: Naturalistic Data Study. Transportation Research Record: Journal of the Transportation Research Board, 2673(6), 225–235. https://doi.org/10.1177/0361198119846481.

20 Geedipally, S. R., Pratt, M. P., Dadashova, B., Wu, L., & Shirazi, M. (2020). Examining the Feasibility of Using Naturalistic Driving Study Data for Validating Speed Prediction Models. Transportation Research Procedia, 48, 1084–1094. https://doi.org/10.1016/j.trpro.2020.08.135.

21 Kanellaidis, G. (1995). Factors affecting drivers' choice of speed on roadway curves. Journal of Safety Research, 26(1), 49-56.

22 Hallmark, S. L., Tyner, S., Oneyear, N., Carney, C., & McGehee, D. (2015). Evaluation of driving behavior on rural 2-lane curves using the SHRP 2 naturalistic driving study data. Journal of Safety Research, 54, 17.e1-27. https://doi.org/10.1016/j.jsr.2015.06.017.

23 Wang, B., Hallmark, S., Savolainen, P., & Dong, J. (2018). Examining vehicle operating speeds on rural two-lane curves using naturalistic driving data. Accident Analysis & Prevention, 118, 236–243. https://doi.org/10.1016/j.aap.2018.03.017.

24 Sawtelle, A., Shirazi, M., Garder, P. E., & Rubin, J. (2023a). Exploring the impact of seasonal weather factors on frequency of lane-departure crashes in Maine. Journal of Transportation Safety & Security, 15(5), 445–466. https://doi.org/10.1080/19439962.2022.2086952.

25 Sawtelle, A., Shirazi, M., Garder, P. E., & Rubin, J. (2023b). Driver, roadway, and weather factors on severity of lane departure crashes in Maine. Journal of Safety Research, 84, 306–315. https://doi.org/10.1016/j.jsr.2022.11.006.

26 Shahlaee, A., Shirazi, M., Marshall, E., & Ivan, J. N. (2022). Modeling the impact of the COVID-19 pandemic on speeding at rural roadway facilities in Maine using short-term speed and traffic count data. Accident Analysis & Prevention, 177, 106828. https://doi.org/10.1016/j.aap.2022.106828.

27 Marshall, E., Shirazi, M., Shahlaee, A., & Ivan, J. N. (2023). Leveraging probe data to model speeding on urban limited access highway segments: Examining the impact of

operational performance, roadway characteristics, and COVID-19 pandemic. Accident Analysis & Prevention, 187, 107038. https://doi.org/10.1016/j.aap.2023.107038.

28 Wang, B., Hallmark, S., Savolainen, P., & Dong, J. (2018). Examining vehicle operating speeds on rural two-lane curves using naturalistic driving data. Accident Analysis & Prevention, 118, 236–243. https://doi.org/10.1016/j.aap.2018.03.017.

29 Ahmed, M. M., Khan, M. N., Das, A., & Dadvar, S. E. (2022). Global lessons learned from naturalistic driving studies to advance traffic safety and operation research: A systematic review. Accident Analysis & Prevention, 167, 106568. https://doi.org/10.1016/j.aap.2022.106568.

30 Hamzeie, R., Savolainen, P. T., & Gates, T. J. (2017). Driver speed selection and crash risk: Insights from the naturalistic driving study. Journal of Safety Research, 63, 187–194. https://doi.org/10.1016/j.jsr.2017.10.007.

31 Wu, J., & Xu, H. (2018). Driver behavior analysis on rural 2-lane, 2-way highways using SHRP 2 NDS data. Traffic Injury Prevention, 19(8), 838–843. https://doi.org/10.1080/15389588.2018.1524142.

32 Brewer, M. A., & Stibbe, J. (2019). Investigation of design speed characteristics on freeway ramps using shrp2 naturalistic driving data. Transportation research record, 2673(3), 247-258.

33 Hao, H., Li, Y. (Eric), Medina, A., Gibbons, R. B., & Wang, L. (2020). Understanding crashes involving roadway objects with SHRP 2 naturalistic driving study data. Journal of Safety Research, 73, 199–209. https://doi.org/10.1016/j.jsr.2020.03.005.

34 Richard, C. M., Lee, J., Atkins, R., & Brown, J. L. (2020). Using SHRP2 naturalistic driving data to examine driver speeding behavior. Journal of safety research, 73, 271-281.

35 StreetLight. (2022). StreetLight Data Sources and Methodology White Paper. StreetLight Data.

36 StreetLight. (2023). Speed Methodology and Validation White Paper. StreetLight Data.

37 Lee, W., & Grimm, K. J. (2018). Generalized Linear Mixed-Effects Modeling Programs in R for Binary Outcomes. Structural Equation Modeling: A Multidisciplinary Journal, 25(5), 824–828. https://doi.org/10.1080/10705511.2018.1500141.

38 Hilbe, J. M. (2014). Modeling count data. Cambridge University Press.

39 Fitzsimmons, E. J., Kvam, V., Souleyrette, R. R., Nambisan, S. S., & Bonett, D. G. (2013). Determining Vehicle Operating Speed and Lateral Position Along Horizontal Curves Using Linear Mixed-Effects Models. Traffic Injury Prevention, 14(3), 309–321. https://doi.org/10.1080/15389588.2012.701356.

40 Wang, B., Hallmark, S., Savolainen, P., & Dong, J. (2017). Crashes and near-crashes on horizontal curves along rural two-lane highways: Analysis of naturalistic driving data. Journal of Safety Research, 63, 163–169. https://doi.org/10.1016/j.jsr.2017.10.001.

41 Afghari, A. P., Haque, Md. M., & Washington, S. (2018). Applying fractional split model to examine the effects of roadway geometric and traffic characteristics on speeding behavior. Traffic Injury Prevention, 19(8), 860–866. https://doi.org/10.1080/15389588.2018.1509208.

9.0 Appendices

9.1 Appendix A. Speed Distributions and Profiles for the Five Horizontal Curves

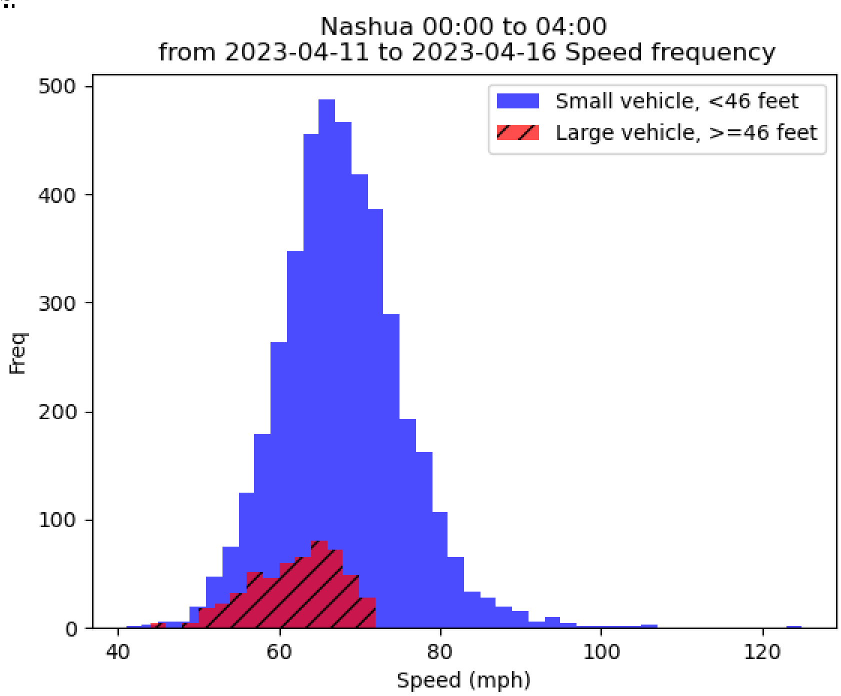


Figure 9-1. Speed Distribution for Nashua between 0 AM and 4 AM

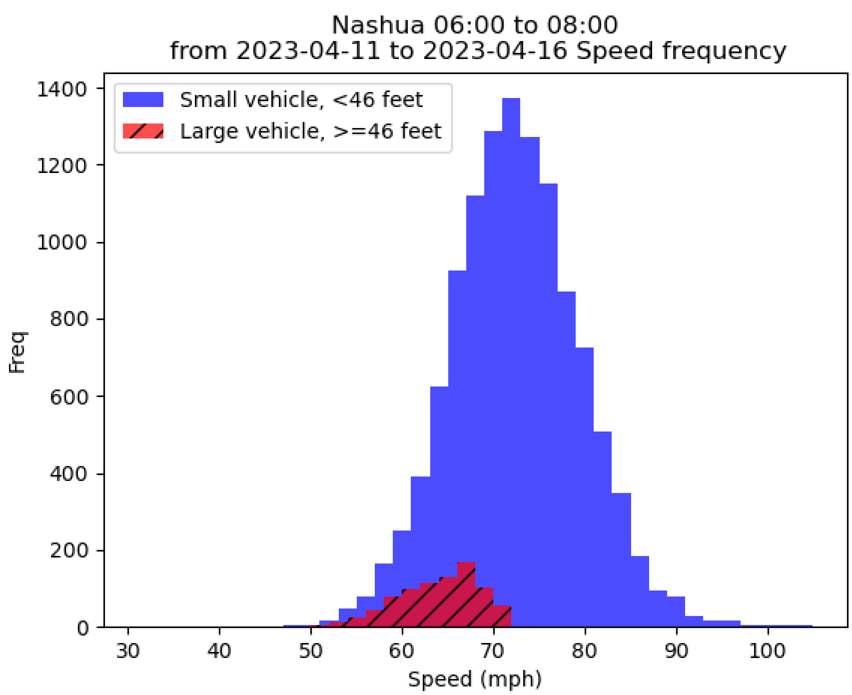


Figure 9-2. Speed Distribution for Nashua between 6 AM and 8 AM

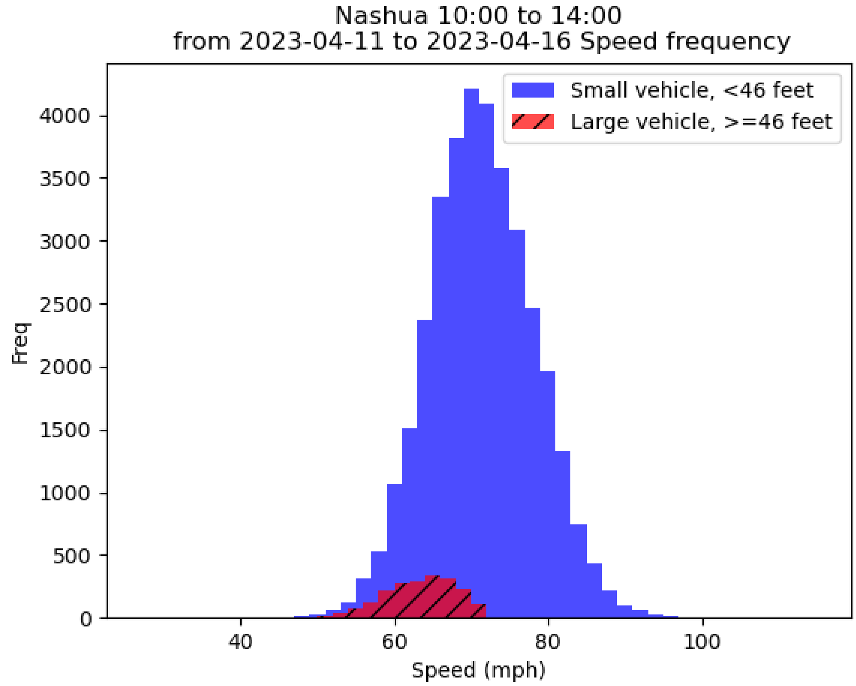


Figure 9-3. Speed Distribution for Nashua between 10 AM and 2 PM

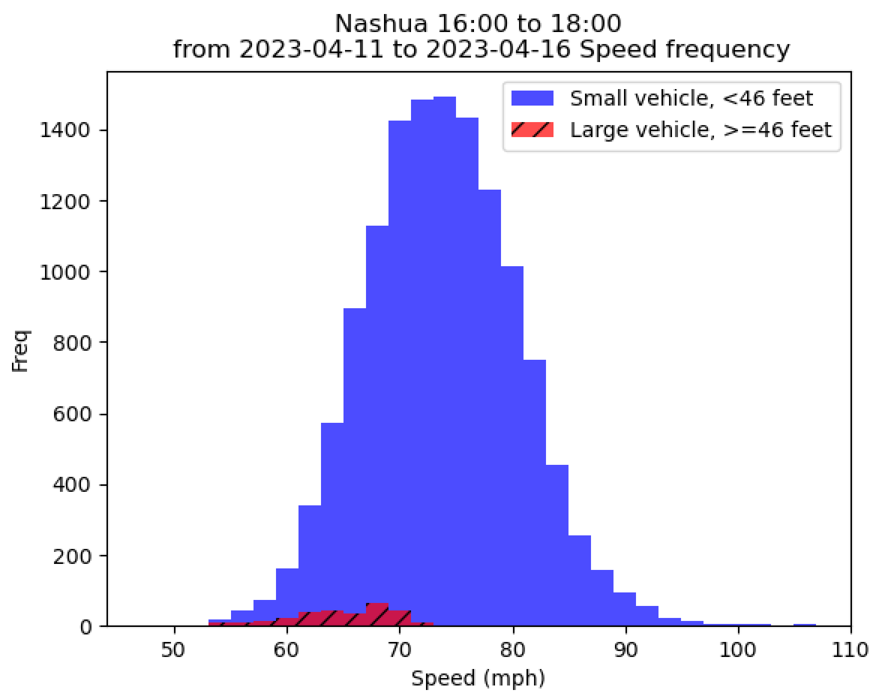


Figure 9-4. Speed Distribution for Nashua between 4 PM and 6 PM

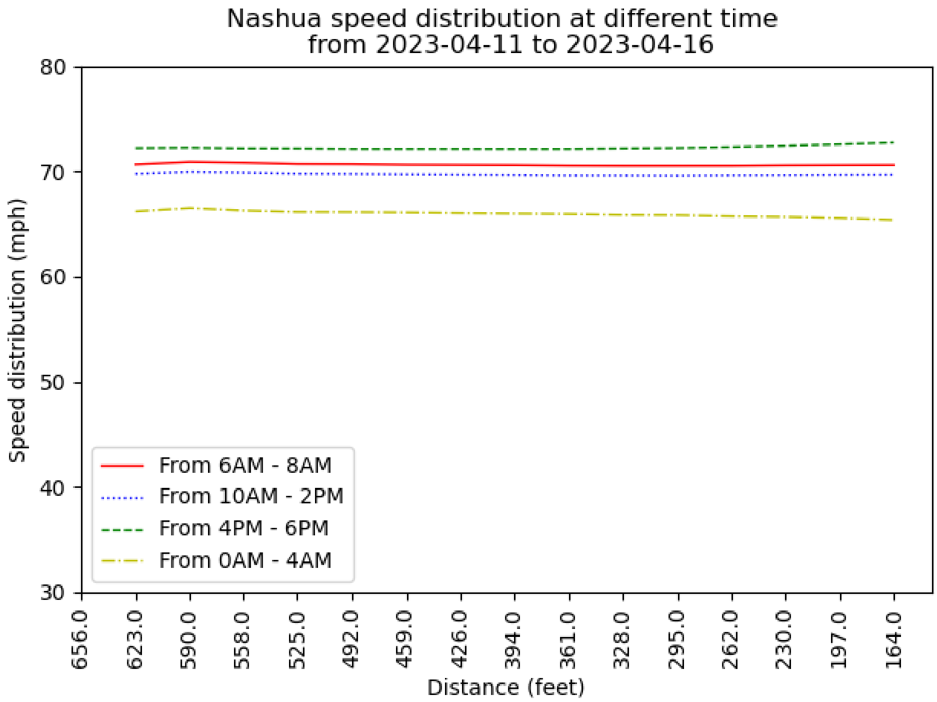


Figure 9-5. Average Segment Speed Profile at Nashua

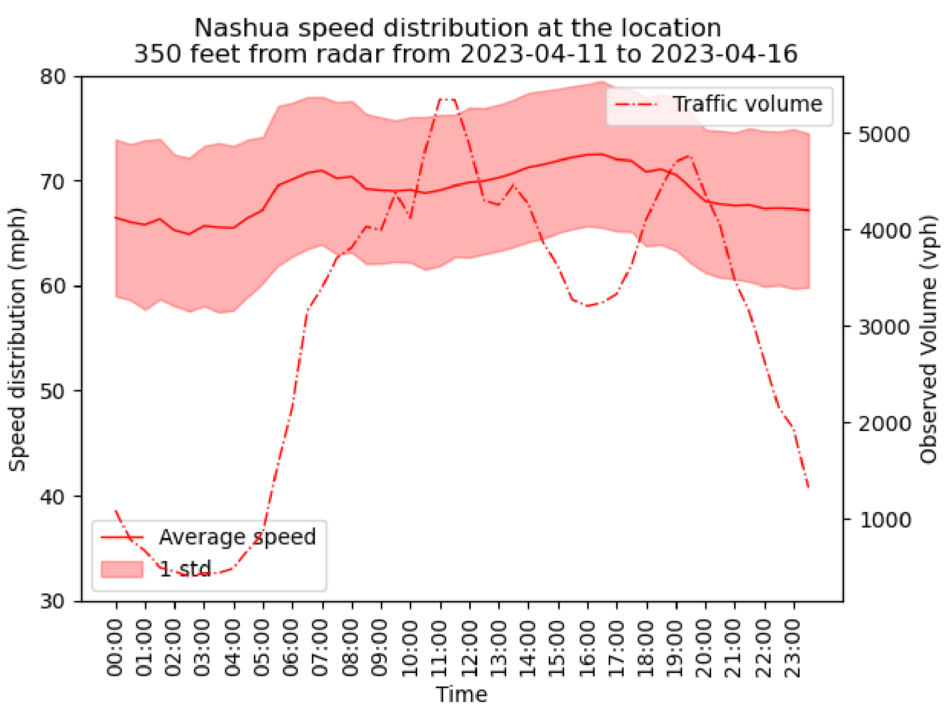


Figure 9-6. Average Spot Speed Profile at Nashua

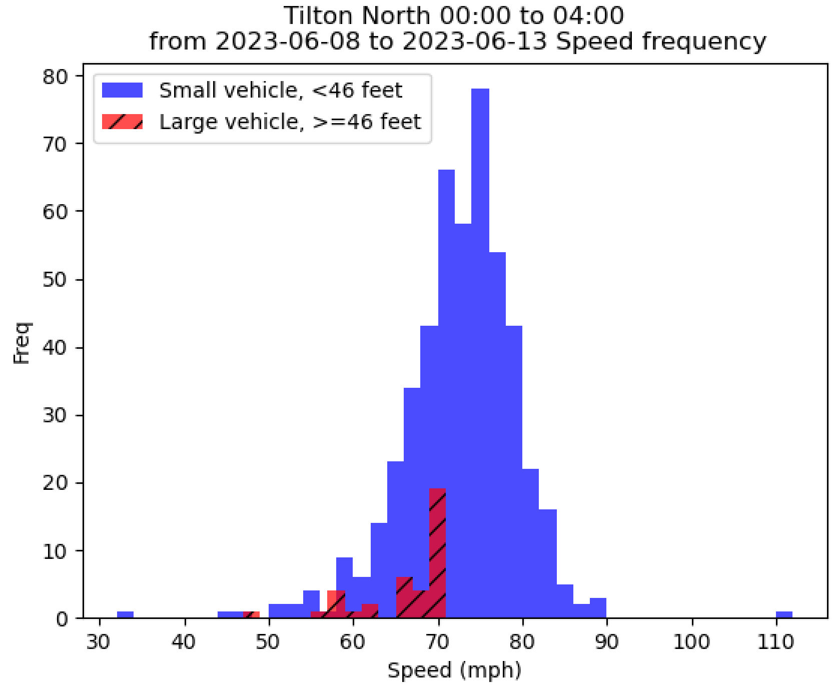


Figure 9-7. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 0 AM to 4 AM

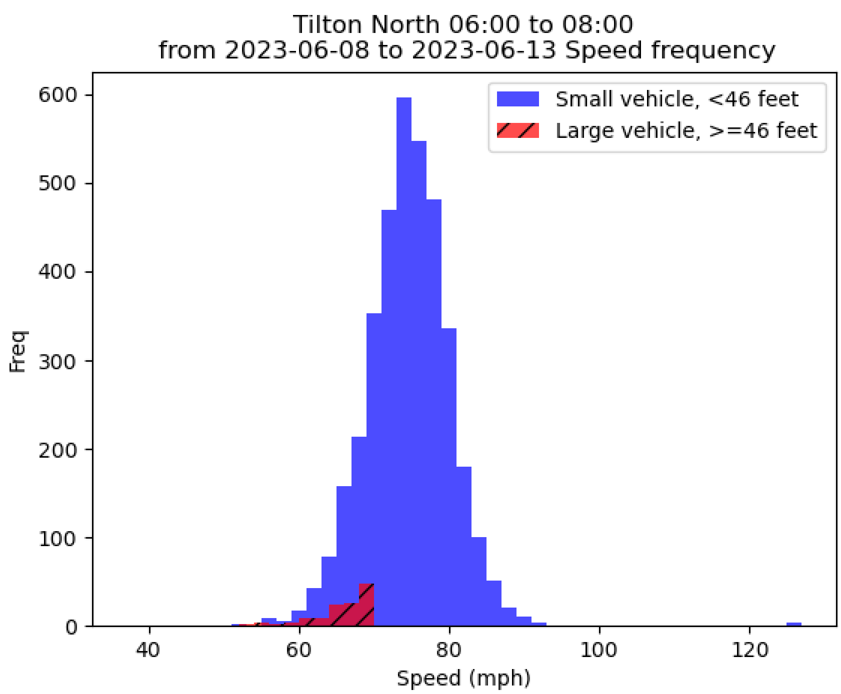


Figure 9-8. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 6 AM to 8 AM

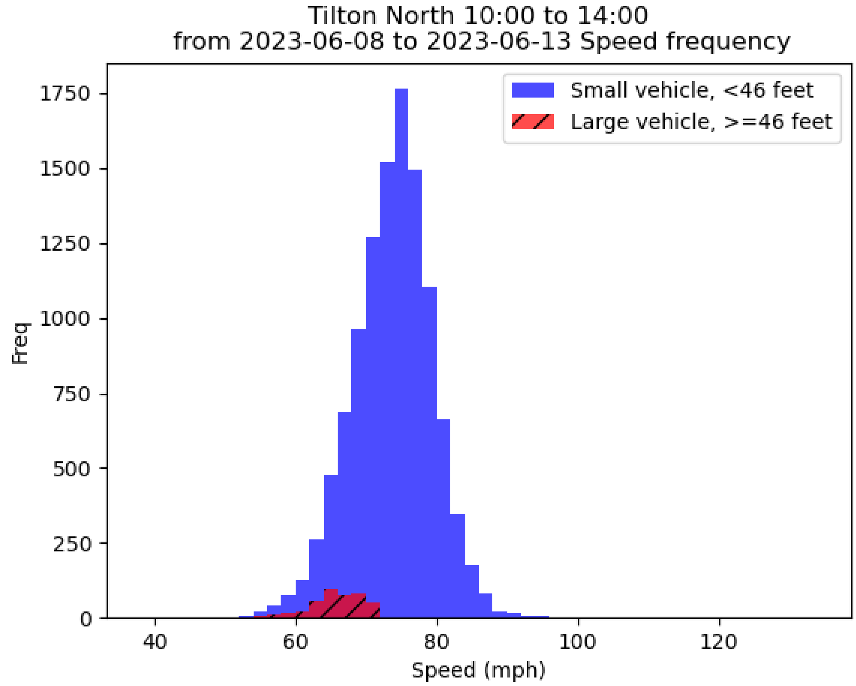


Figure 9-9. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 10 AM to 2 PM

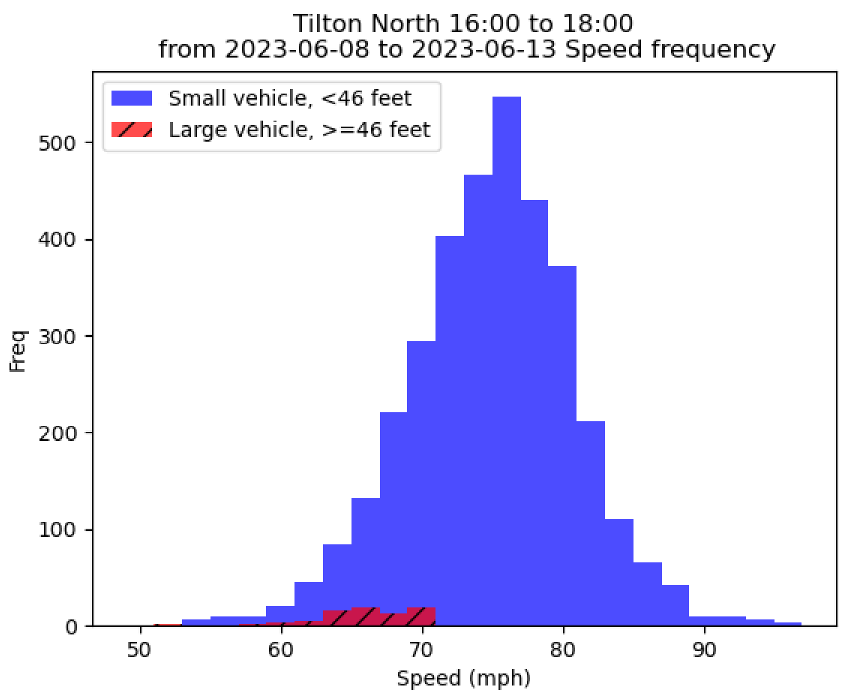


Figure 9-10. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 4 PM to 6 PM

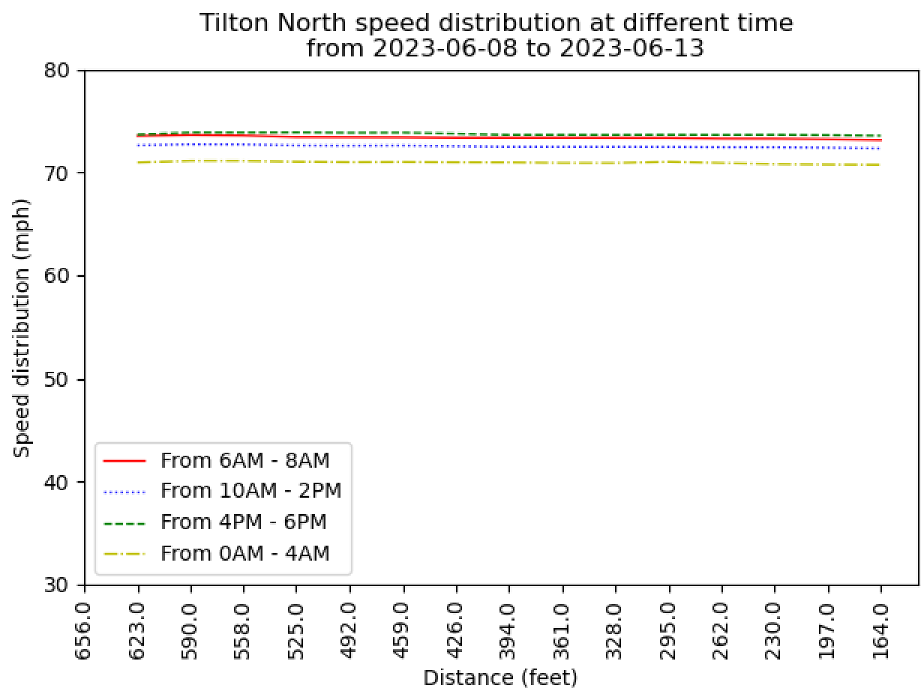


Figure 9-11. Average Segment Speed Profile at Tilton North Main Road

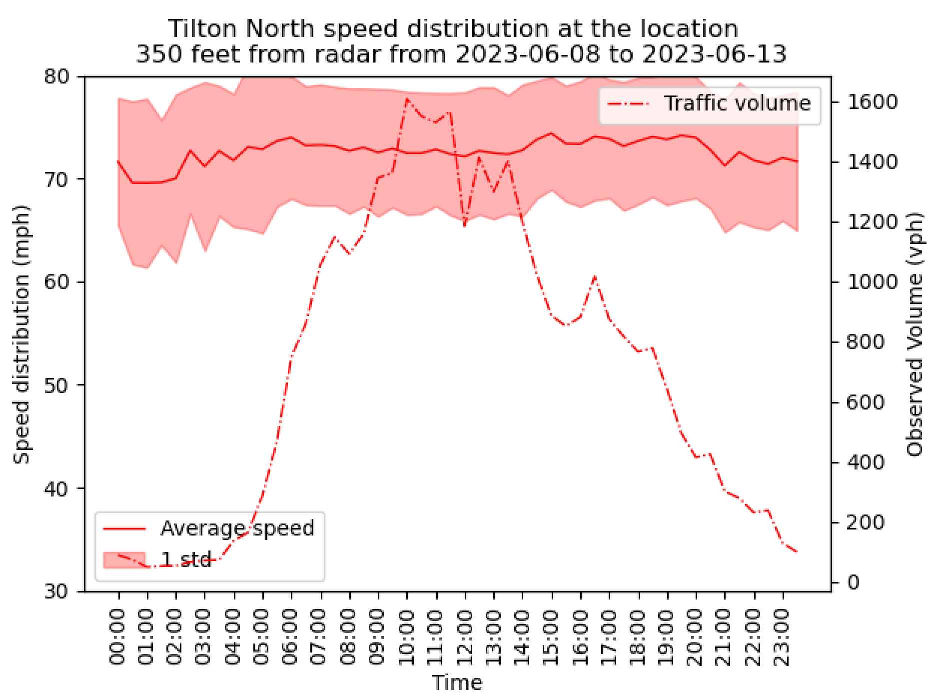


Figure 9-12. Time vs. Average Speed at Tilton North main road

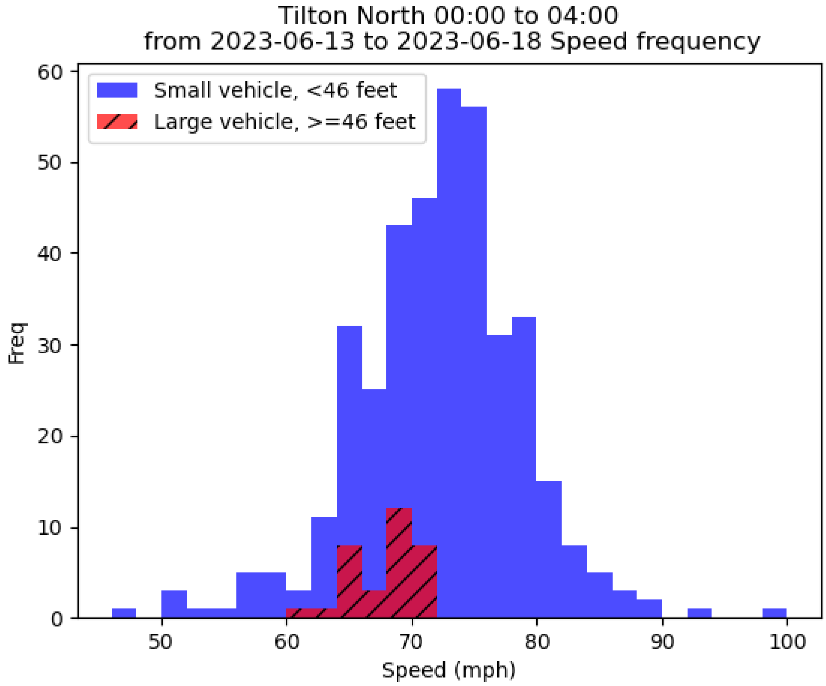


Figure 9-13. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 0 AM to 4 AM

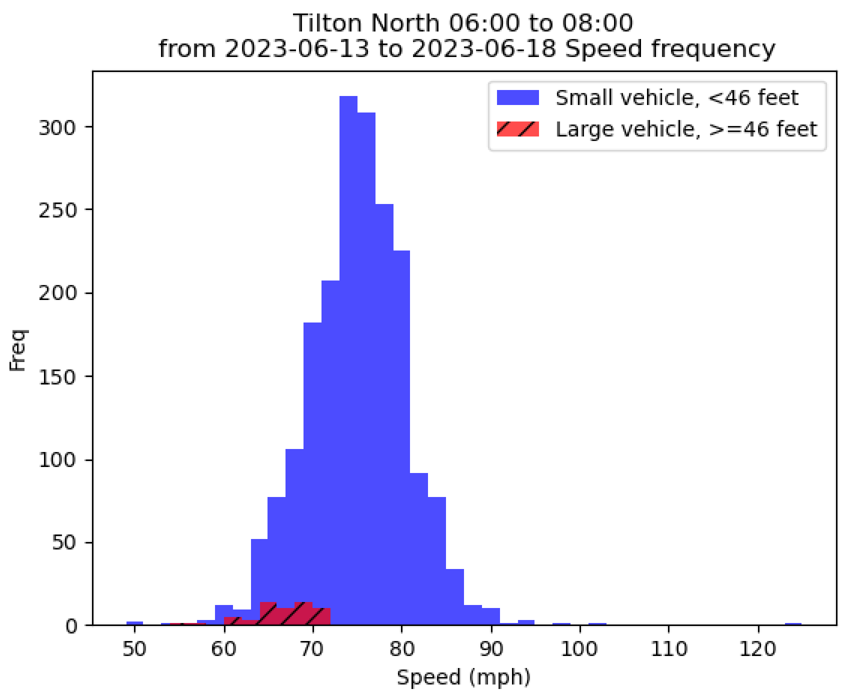


Figure 9-14. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 6 AM to 8 AM

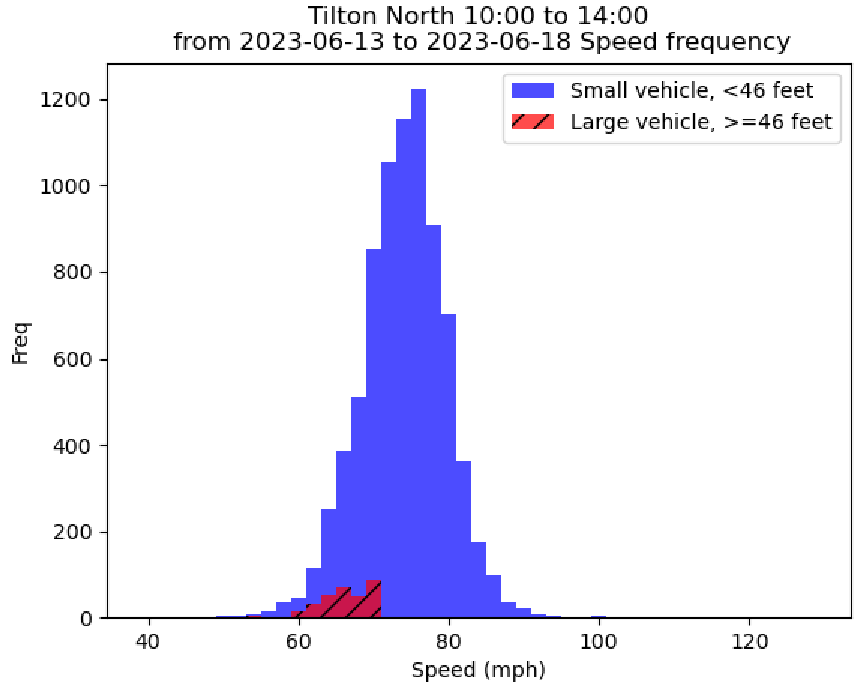


Figure 9-15. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 10 AM to 2 PM

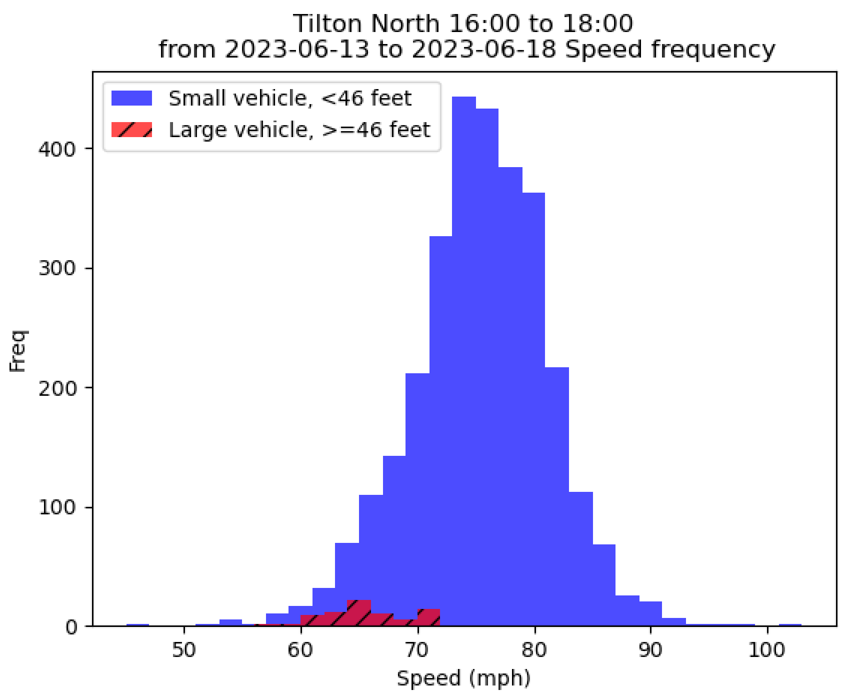


Figure 9-16. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton North main road 4 PM to 6 PM

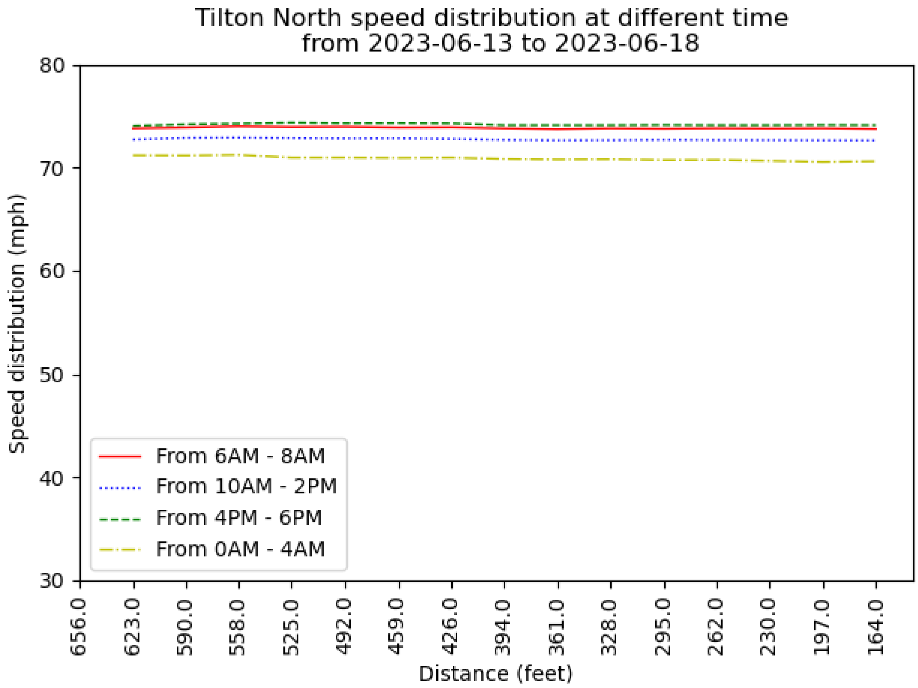


Figure 9-17. Average Segment Speed Profile at Tilton North main road

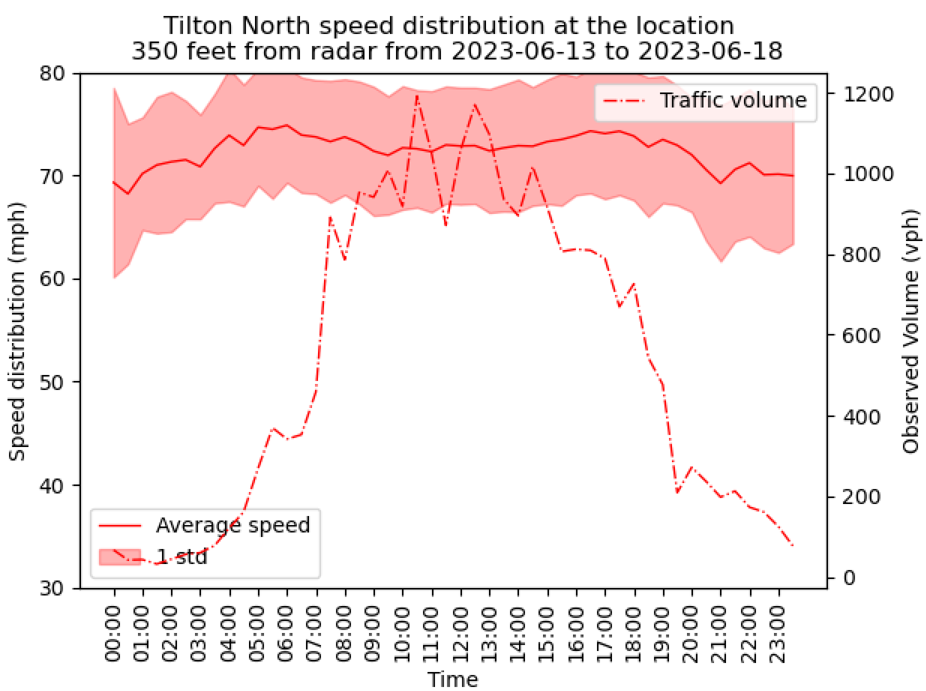


Figure 9-18. Time vs. Average Speed at Tilton North main road

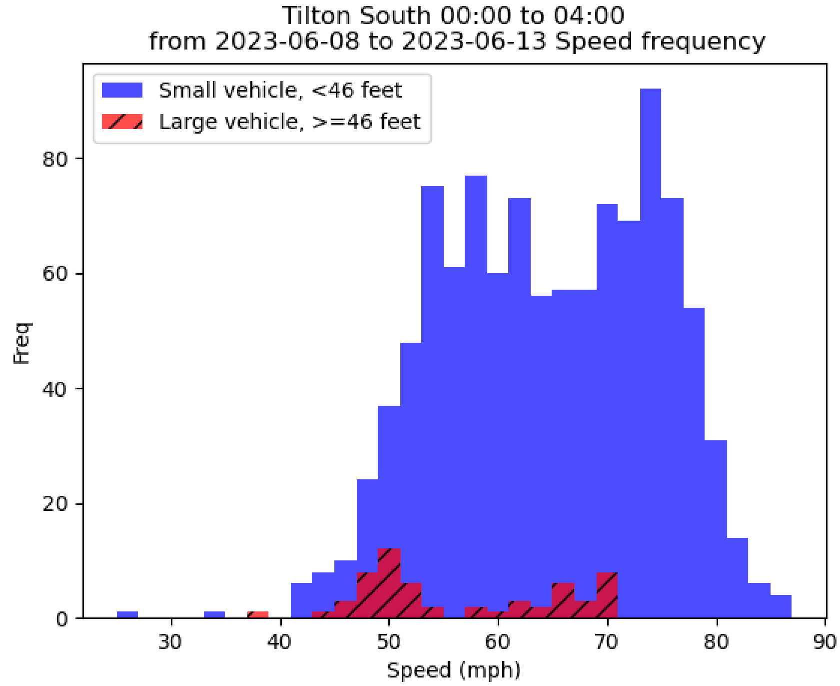


Figure 9-19. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 0 AM to 4 AM

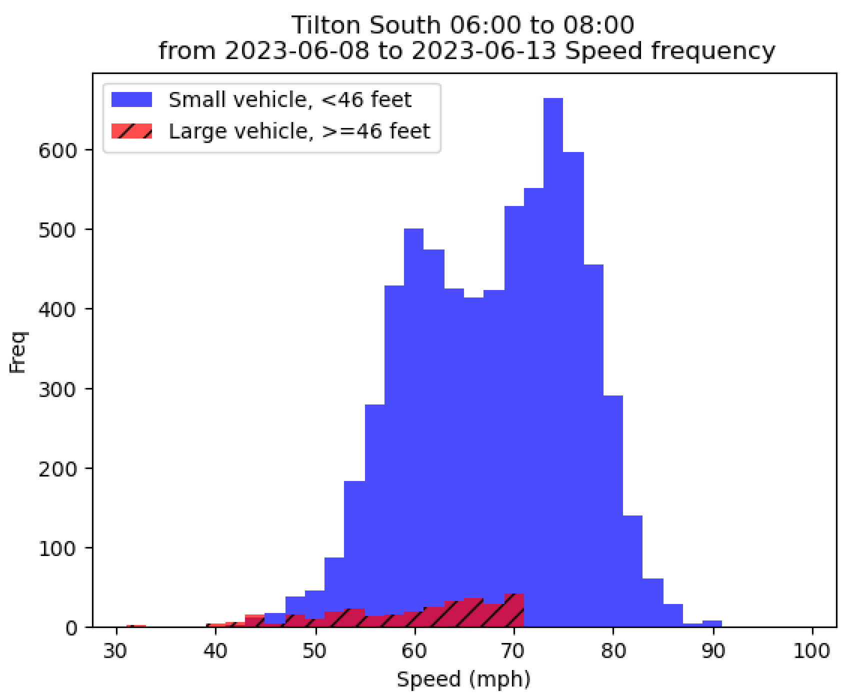


Figure 9-20. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 6 AM to 8 AM

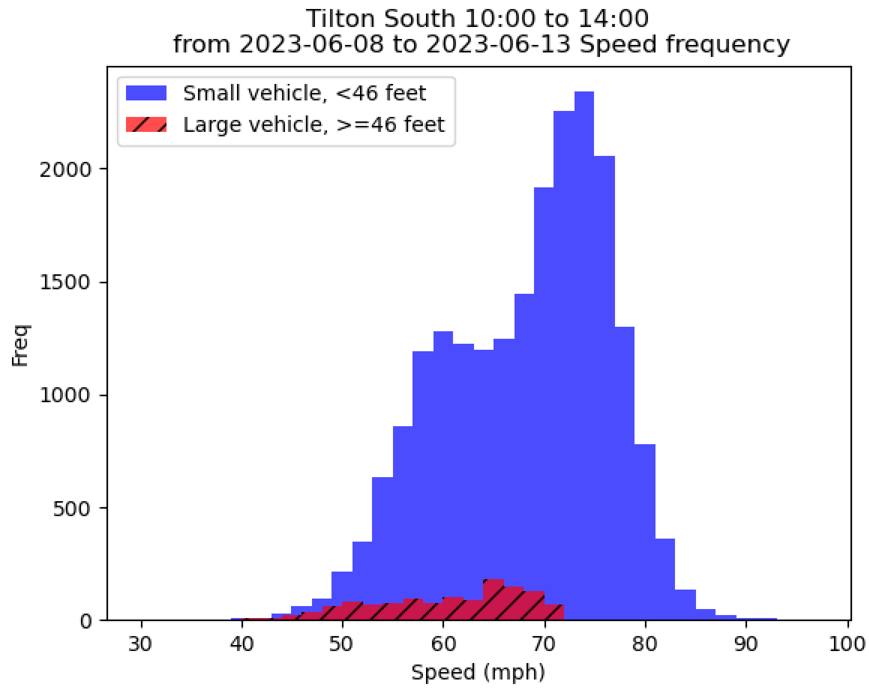


Figure 9-21. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 10 AM to 2 PM

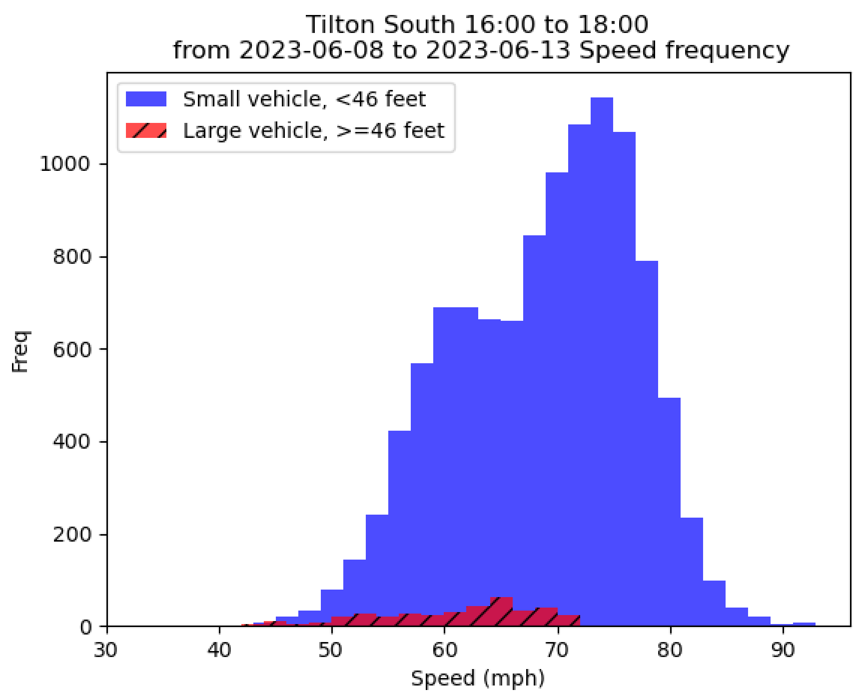


Figure 9-22. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 4 PM to 6 PM

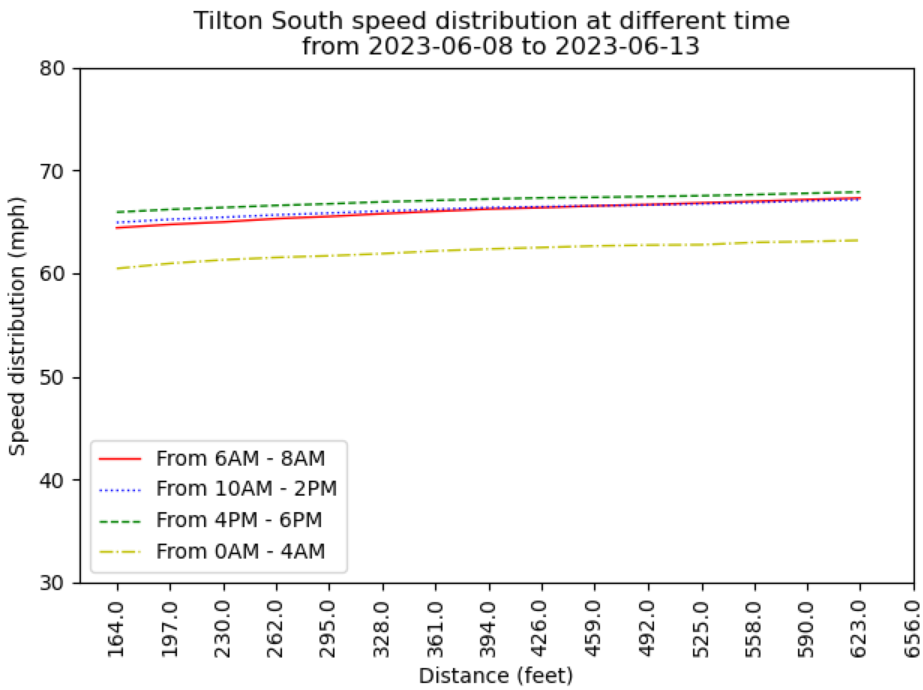


Figure 9-23. Average Segment Speed Profile at Tilton South

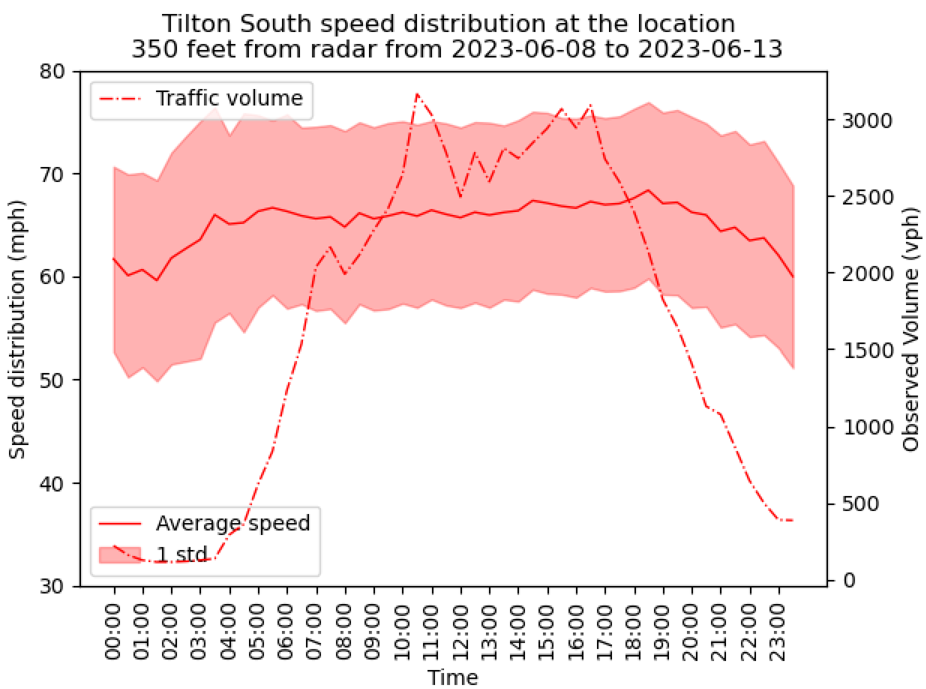


Figure 9-24. Time vs. Average Speed at Tilton South

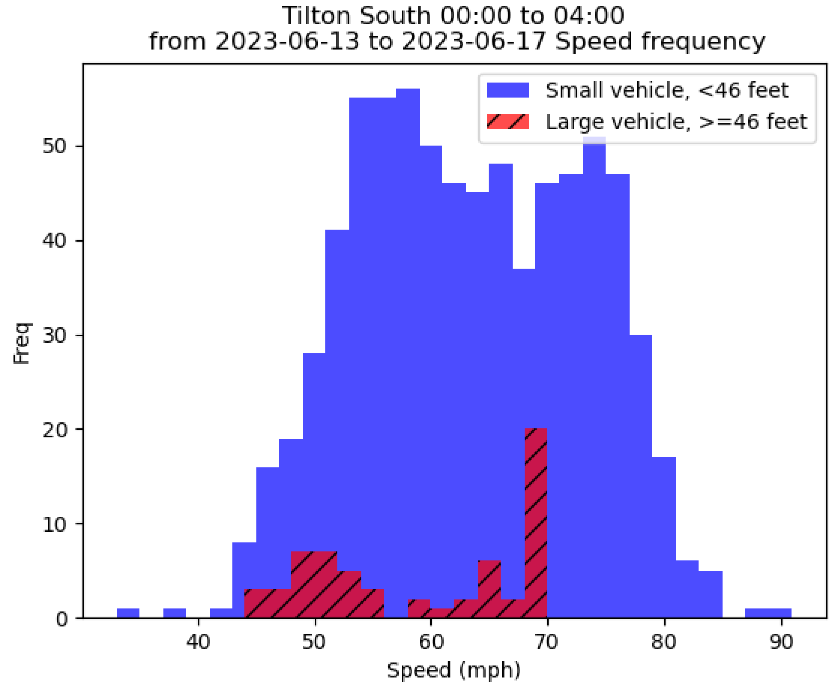


Figure 9-25. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 0 AM to 4 AM

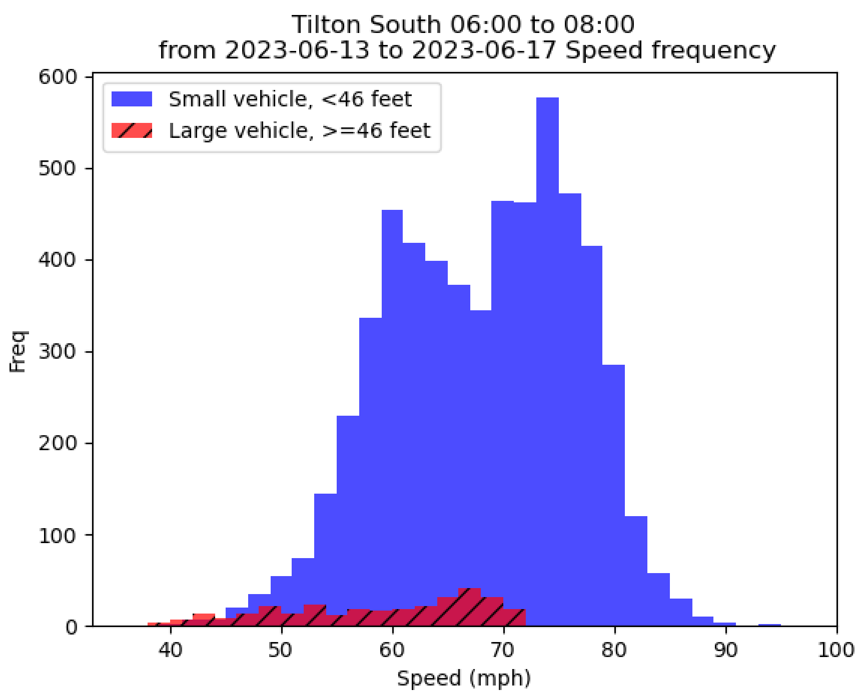


Figure 9-26. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 6 AM to 8 AM

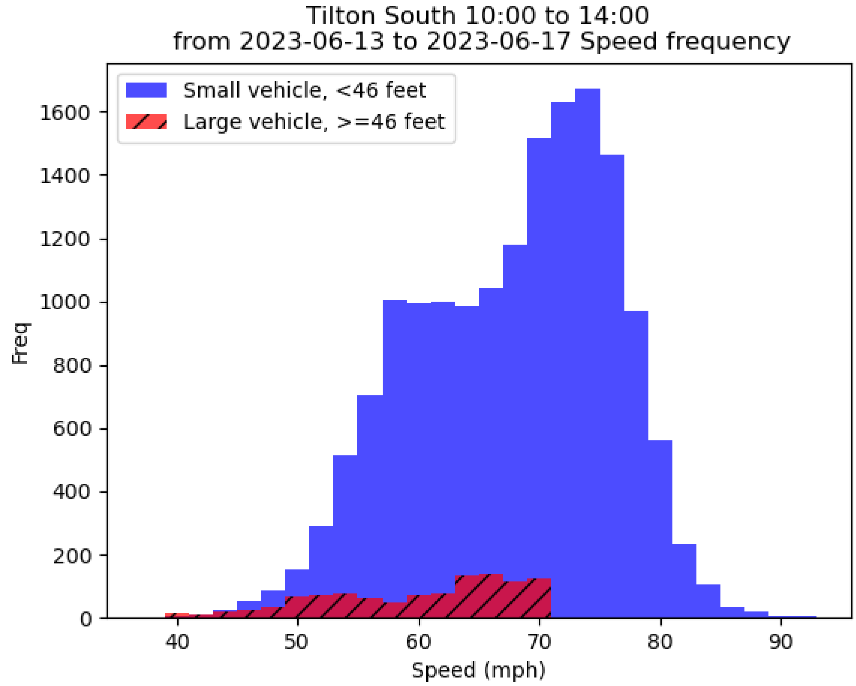


Figure 9-27. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 10 AM to 2 PM

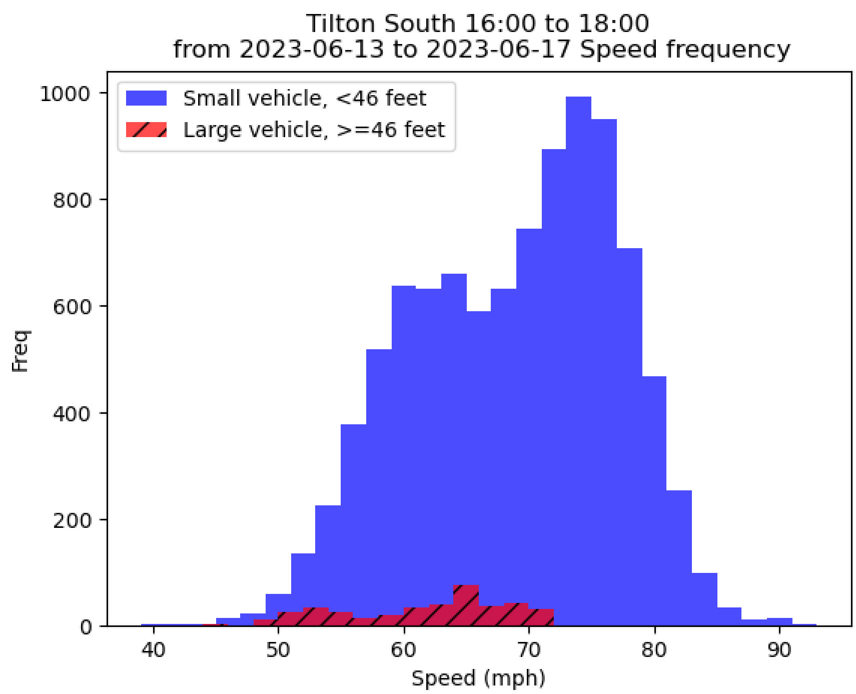


Figure 9-28. Average Speed vs. Frequency: Small vs. Large Vehicles at Tilton South 4 PM to 6 PM

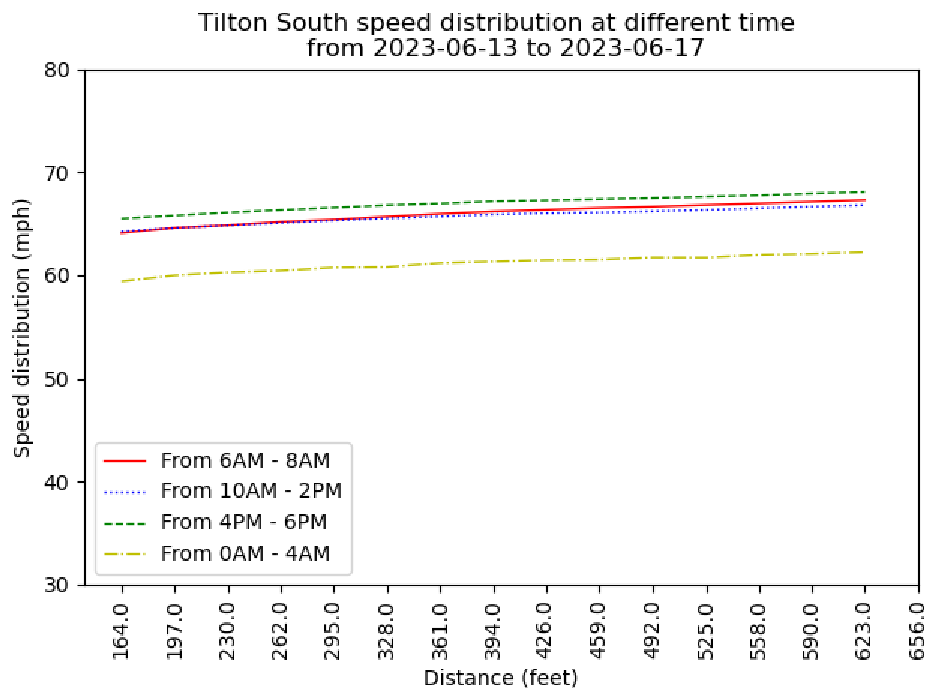


Figure 9-29. Observed Distance vs. Average Speed at Tilton South

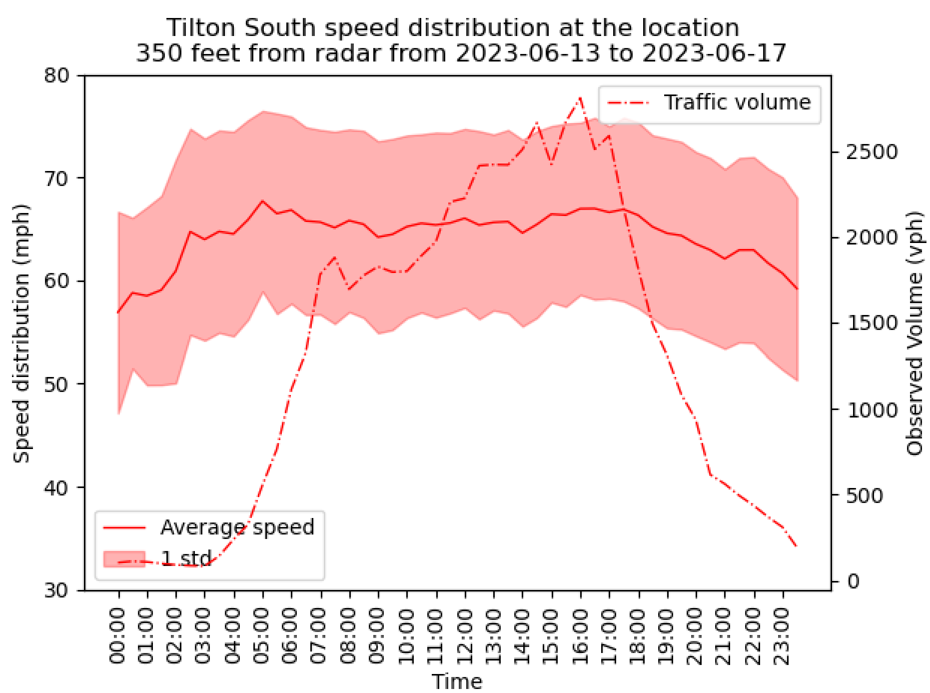


Figure 9-30. Time vs. Average Speed at Tilton South

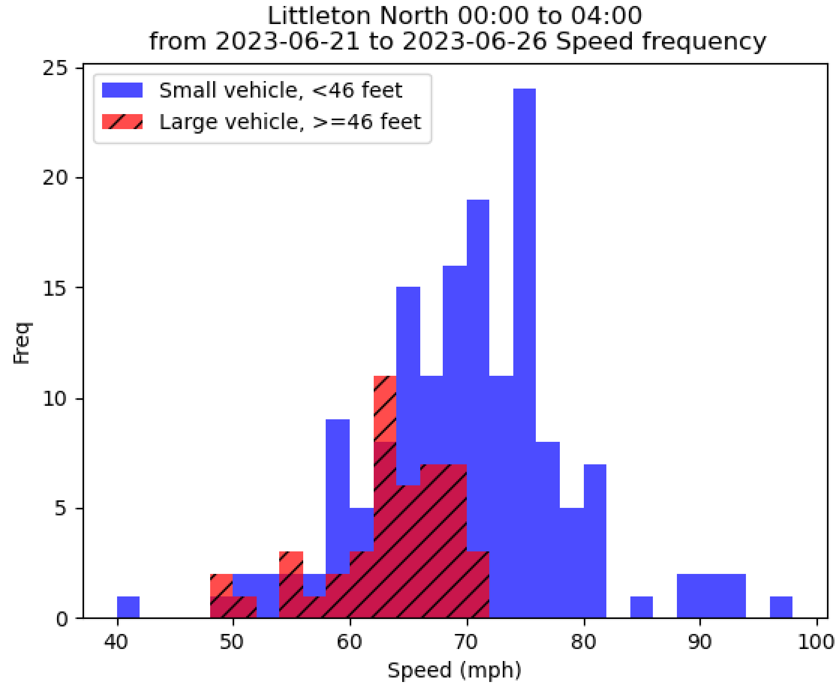


Figure 9-31. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton South 0 AM to 4 AM

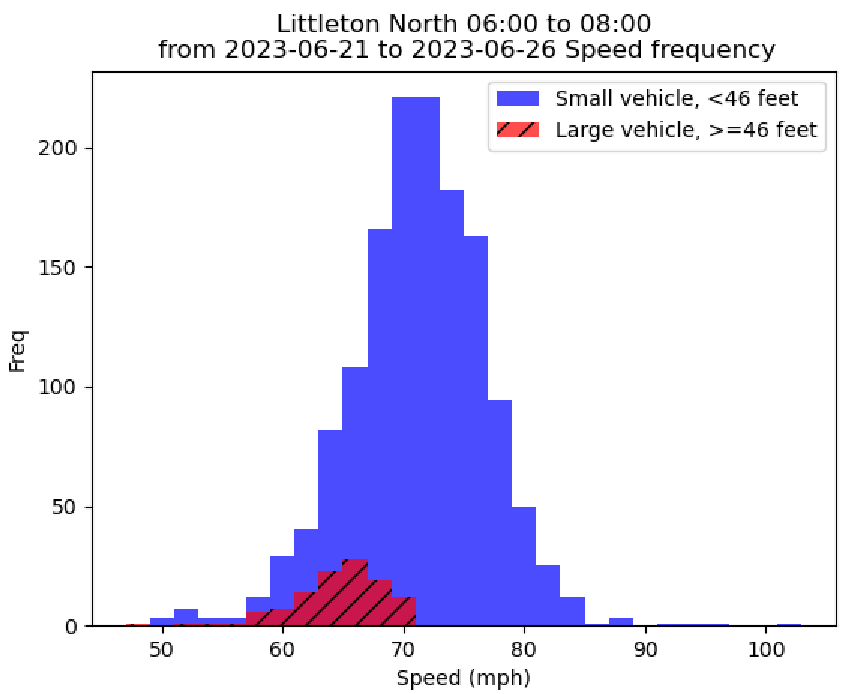


Figure 9-32. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton North 6 AM to 8 AM

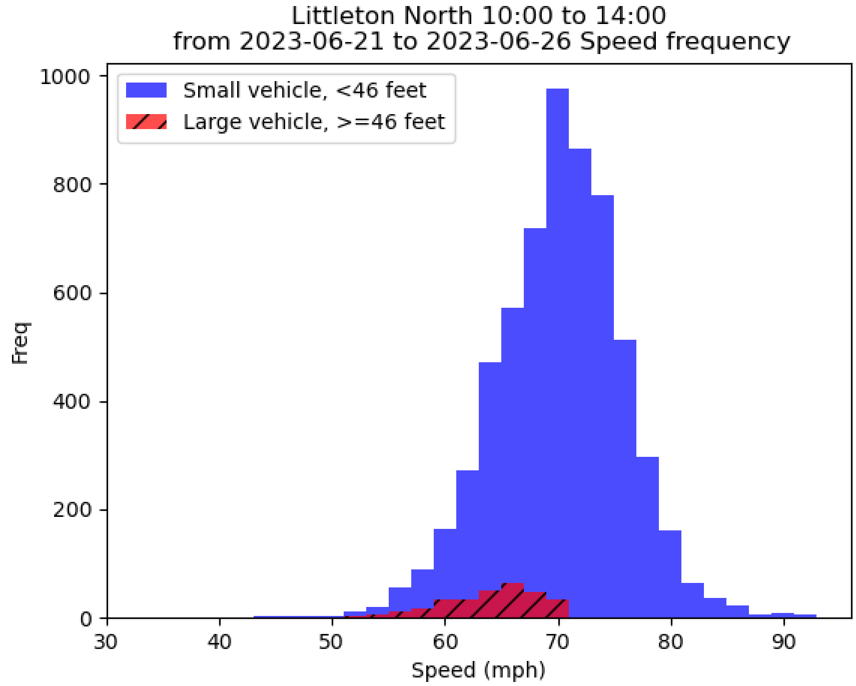


Figure 9-33. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton North 10 AM to 2 AM

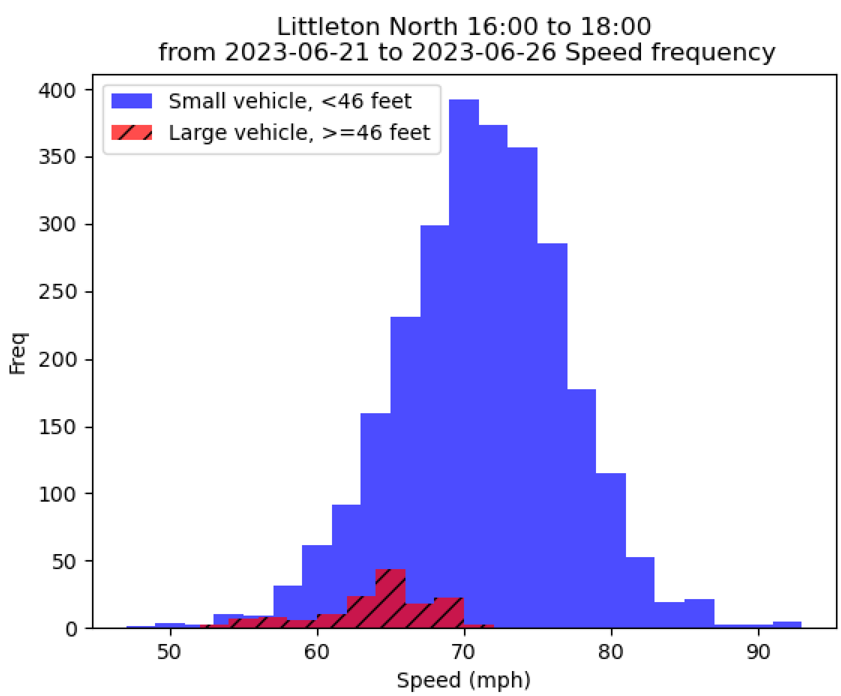


Figure 9-34. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton North 4 PM to 6 PM

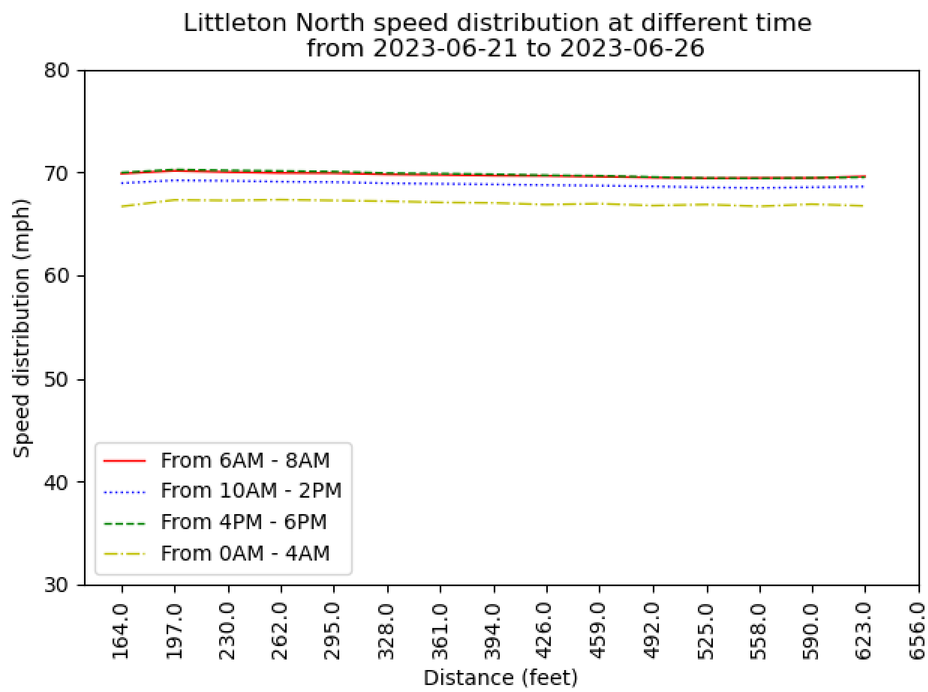


Figure 9-35. Observed Distance vs. Average Speed at Littleton North

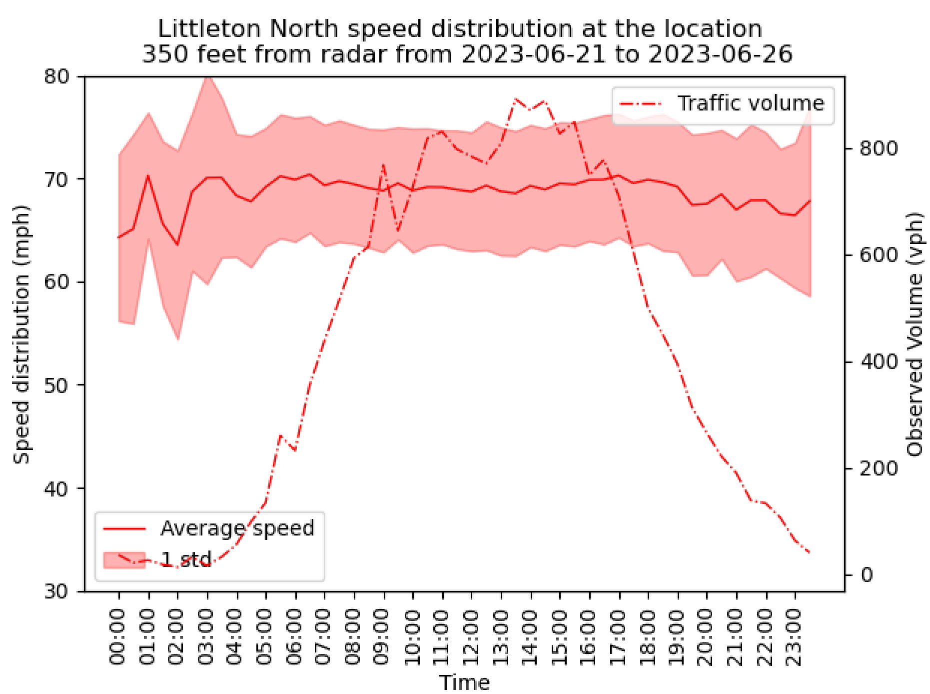


Figure 9-36. Time vs. Average Speed at Littleton North

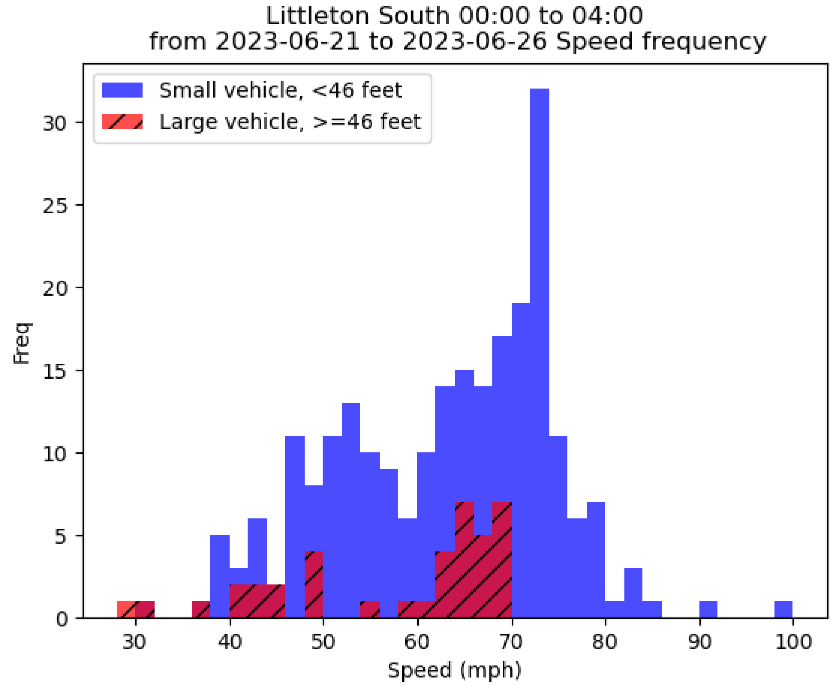


Figure 9-37. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton South 0 AM to 4 AM

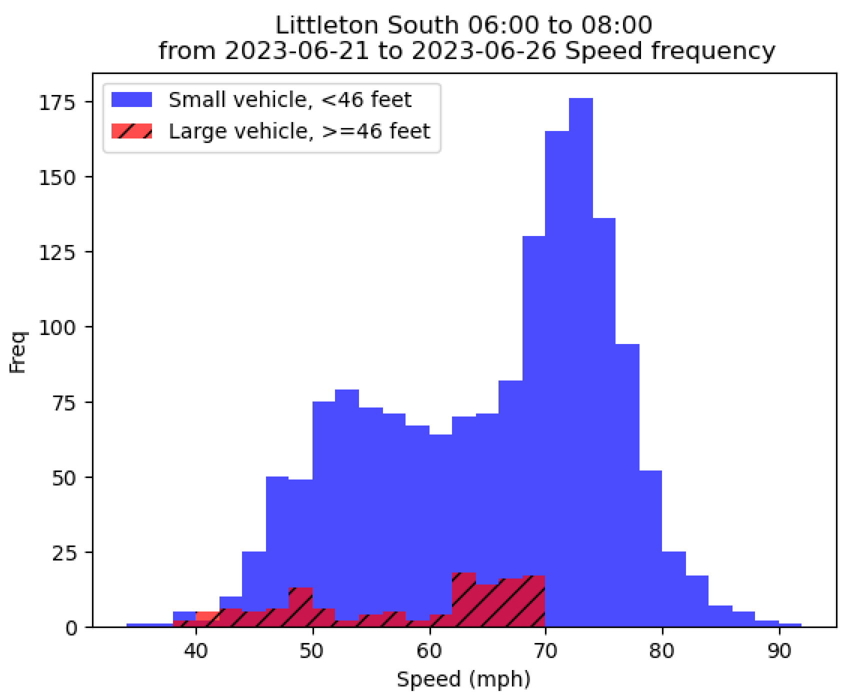


Figure 9-38. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton South 6 AM to 8 AM

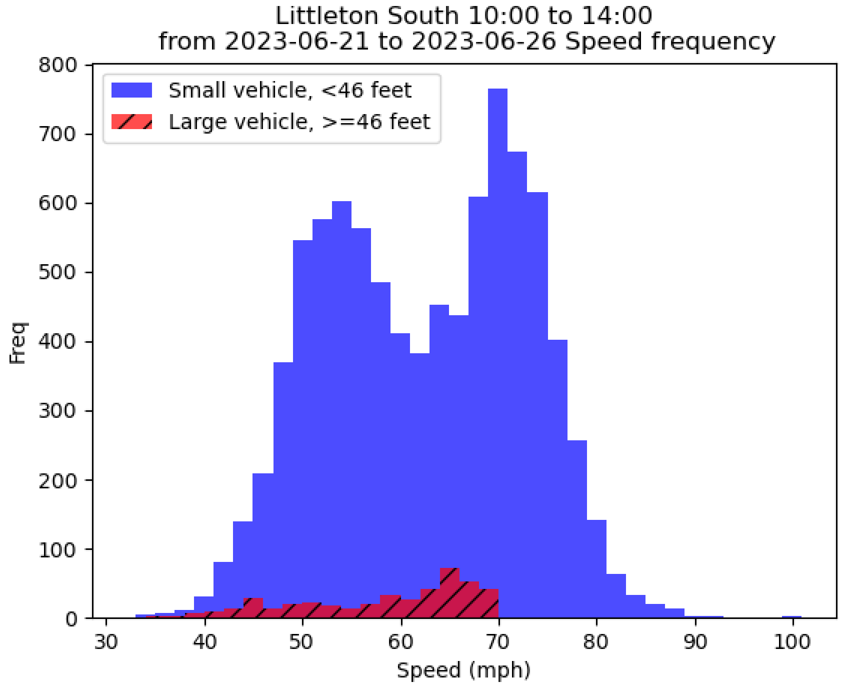


Figure 9-39. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton South 10 AM to 2 PM

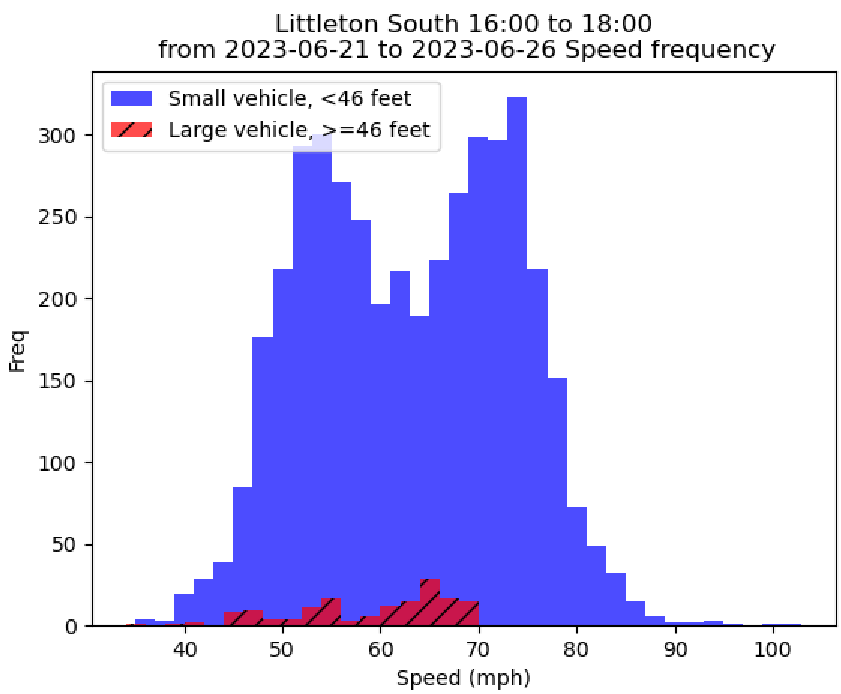


Figure 9-40. Average Speed vs. Frequency: Small vs. Large Vehicles at Littleton South 4 PM to 6 PM

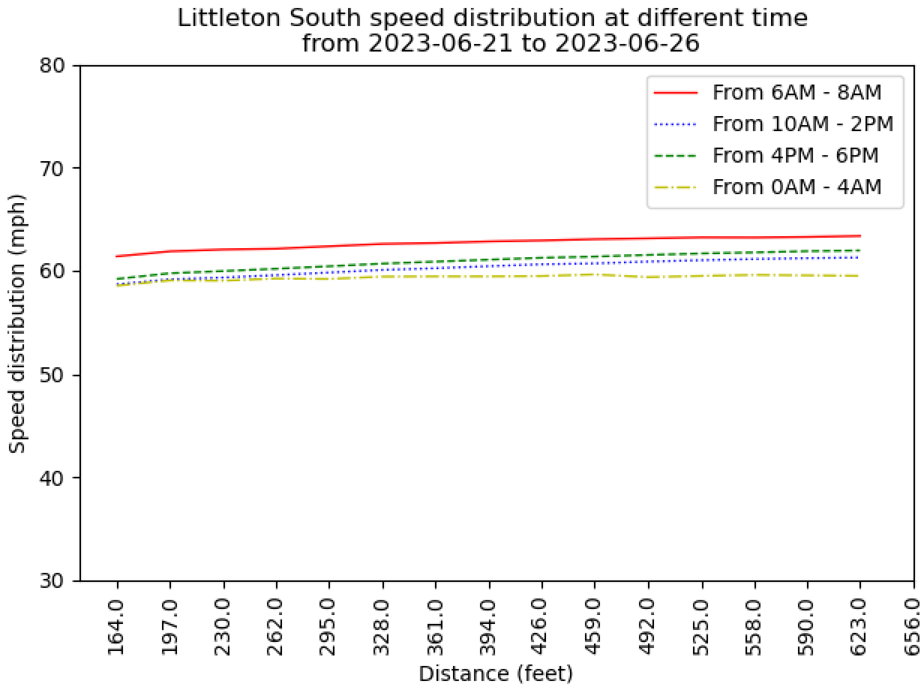


Figure 9-41. Observed Distance vs. Average Speed at Littleton South

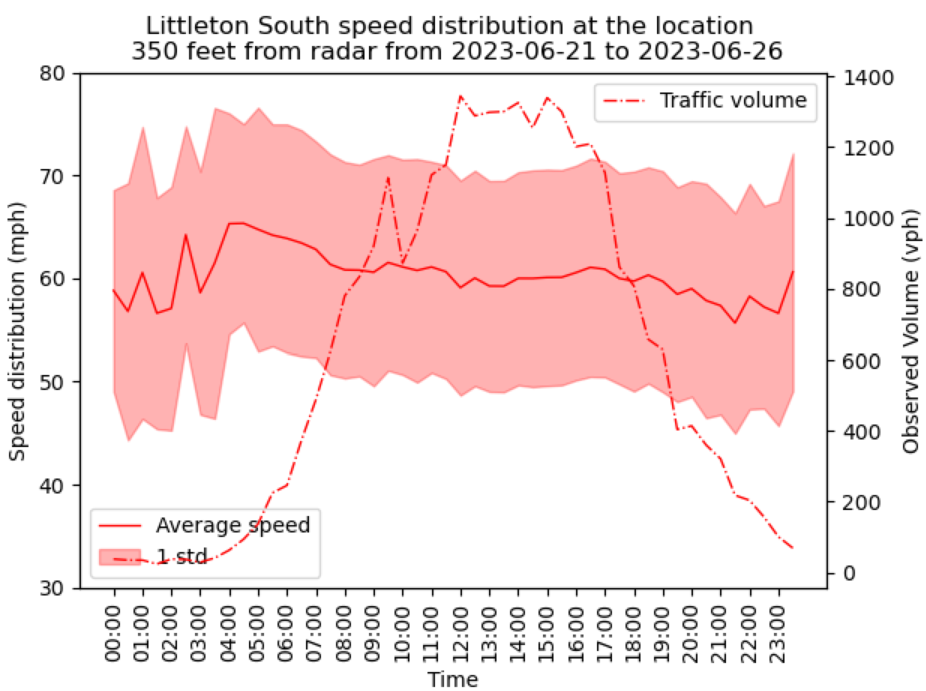


Figure 9-42. Time vs. Average Speed at Littleton South

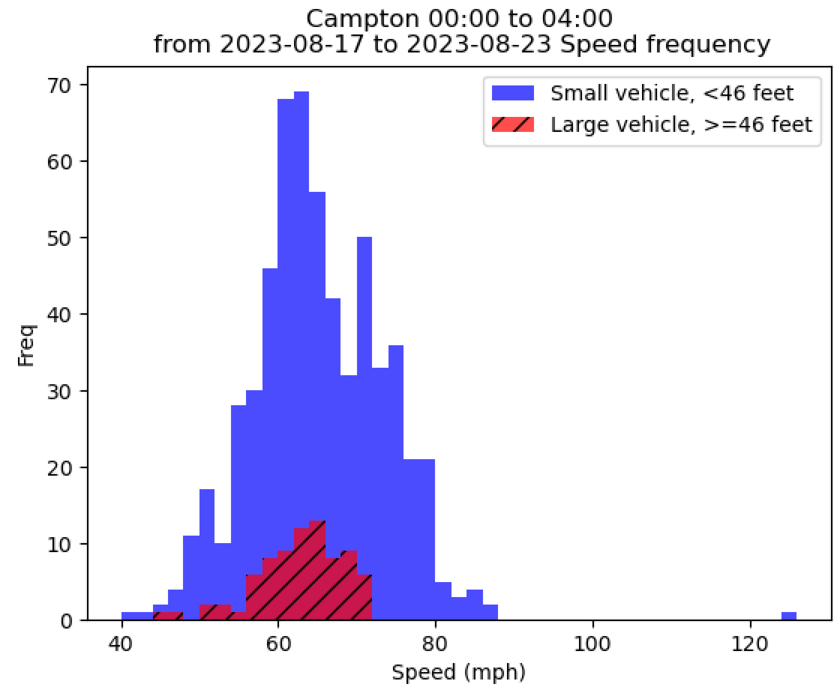


Figure 9-43. Average Speed vs. Frequency: Small vs. Large Vehicles at Campton 0 AM to 4 AM

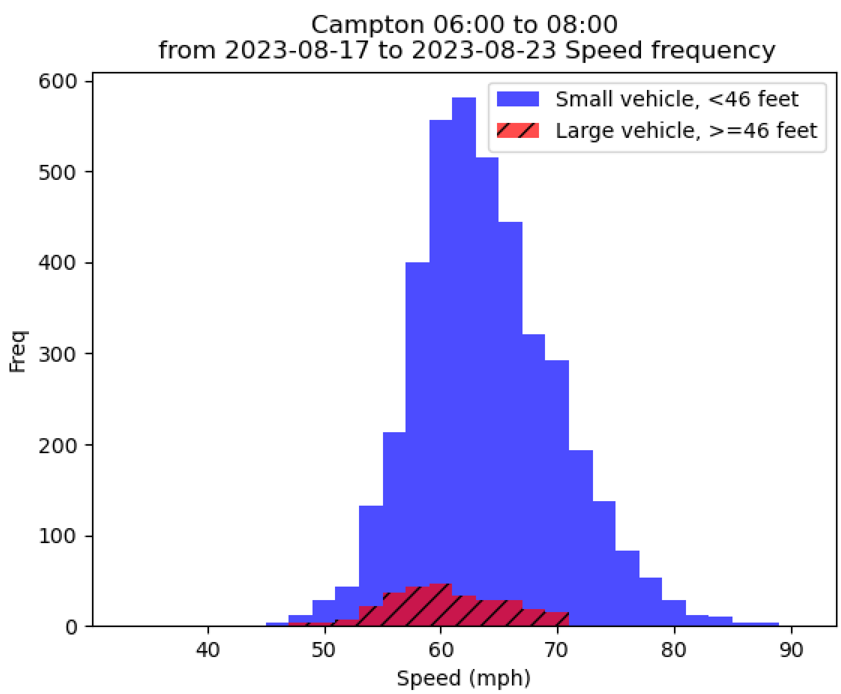


Figure 9-44. Average Speed vs. Frequency: Small vs. Large Vehicles at Campton 6 AM to 8 AM

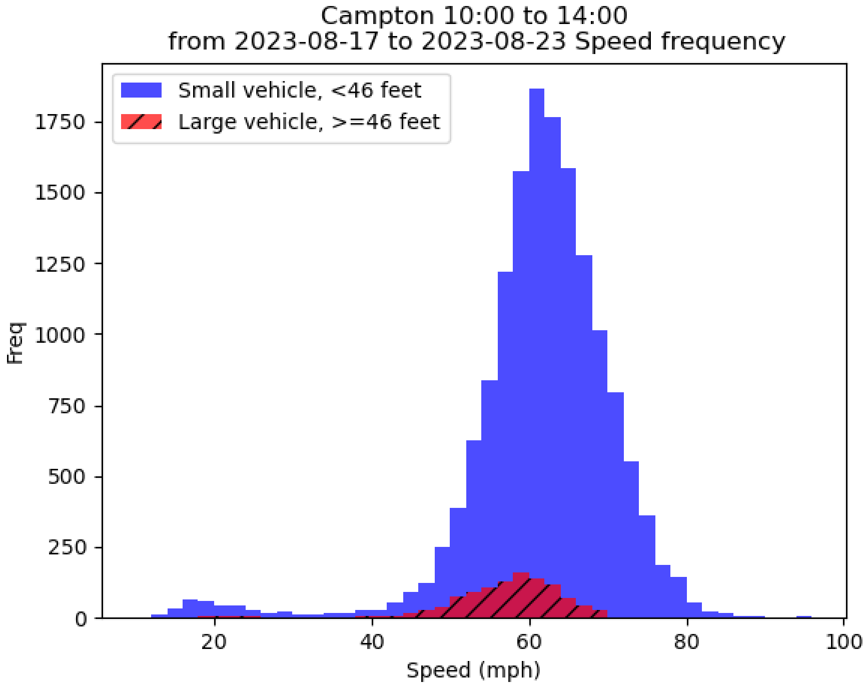


Figure 9-45. Average Speed vs. Frequency: Small vs. Large Vehicles at Campton 10 AM to 2 PM

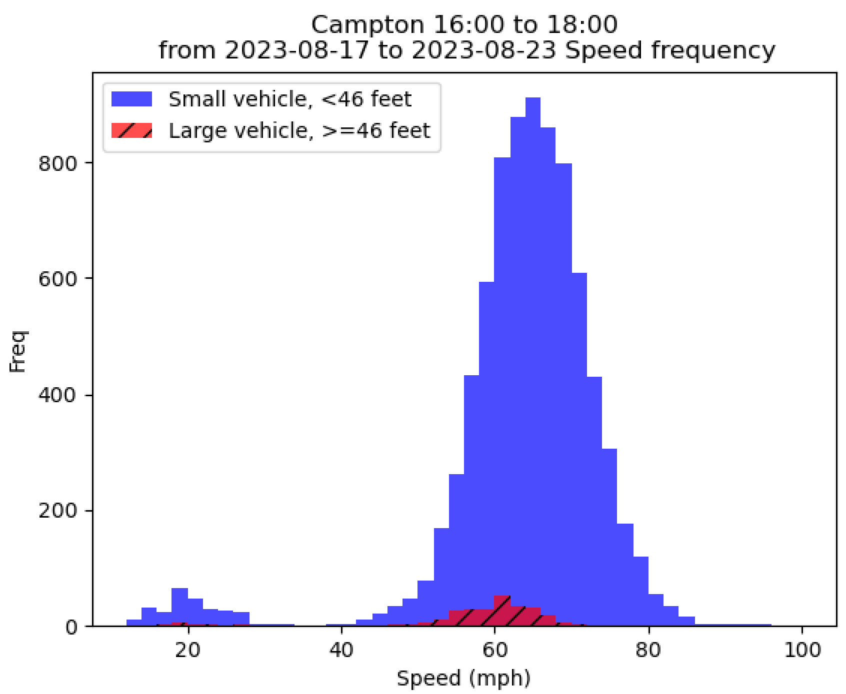


Figure 9-46. Average Speed vs. Frequency: Small vs. Large Vehicles at Campton 4 PM to 6 PM

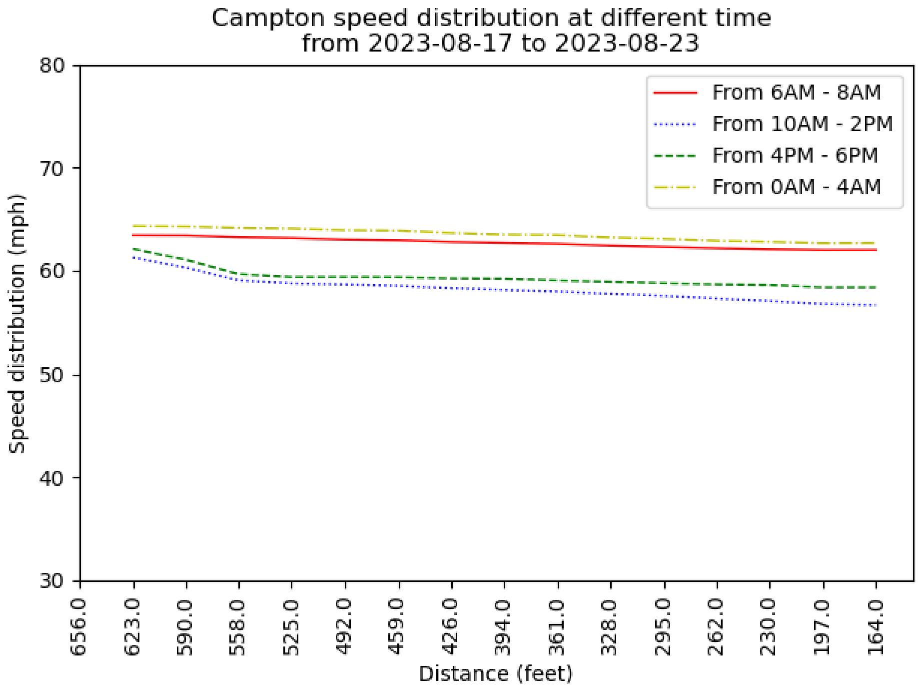


Figure 9-47. Observed Distance vs. Average Speed at Campton

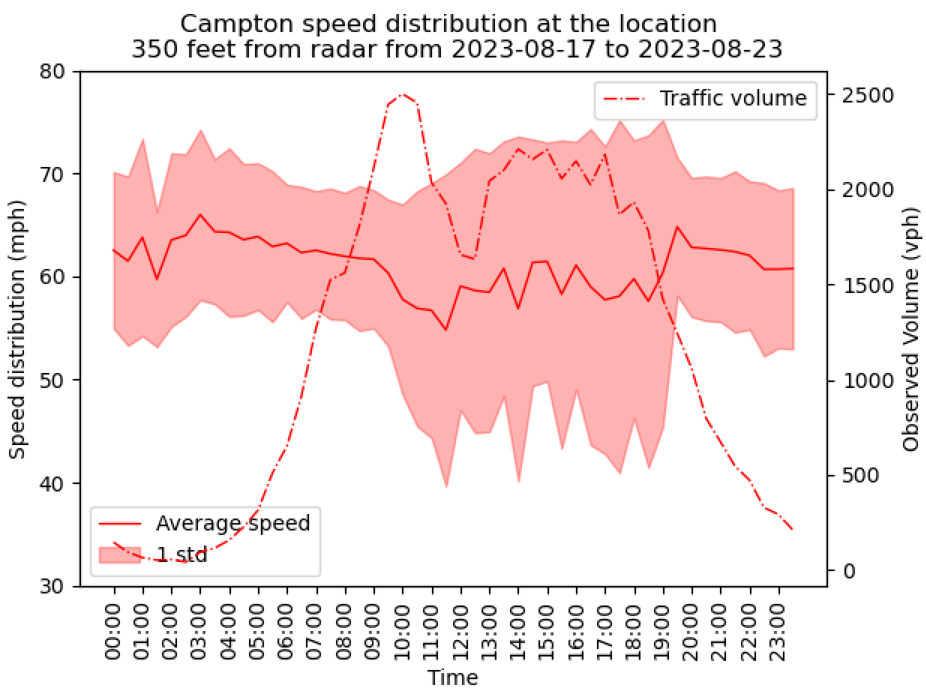


Figure 9-48. Time vs. Average Speed at Campton

9.2 Appendix B. Comparison of Collected Radar Speed Data and TomTom Speed Data

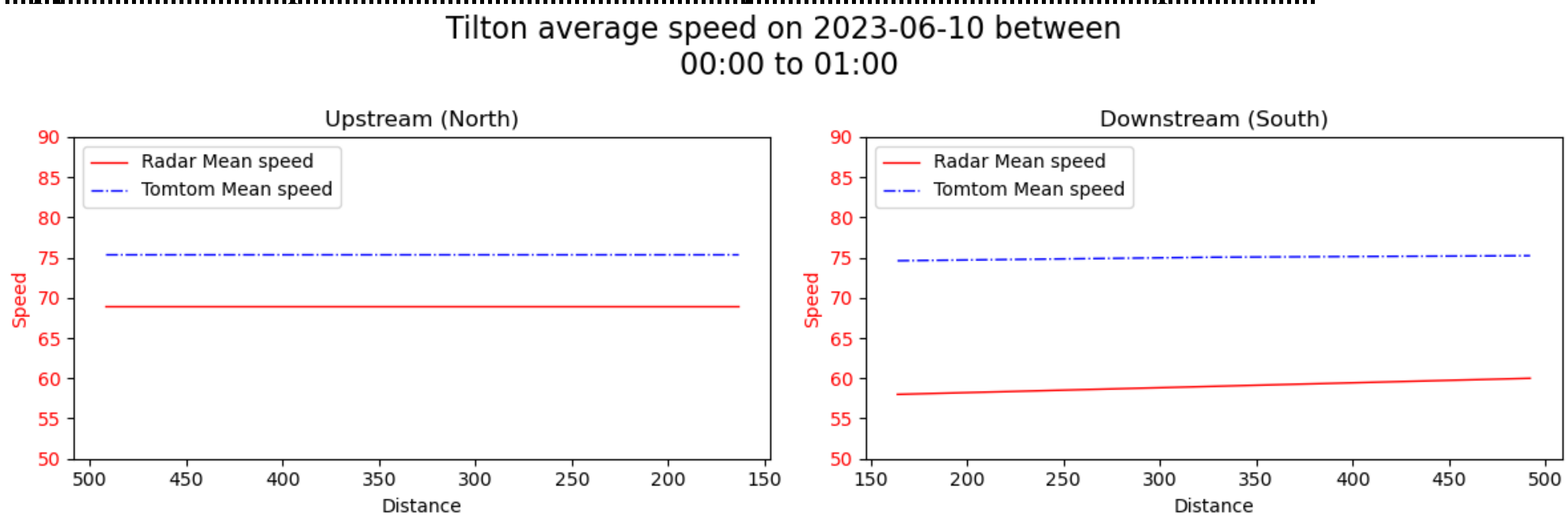


Figure 9-49. TomTom vs. Radar: Average Speed Comparison at Tilton (00:00-01:00)

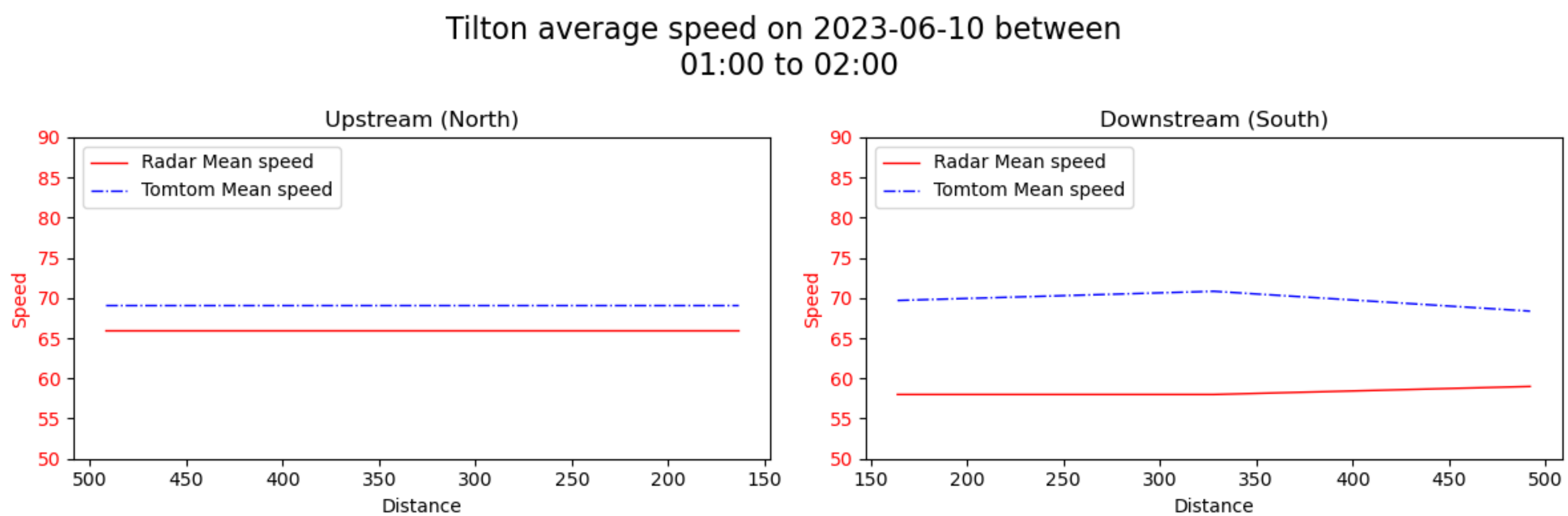


Figure 9-50. TomTom vs. Radar: Average Speed Comparison at Tilton (01:00-02:00)

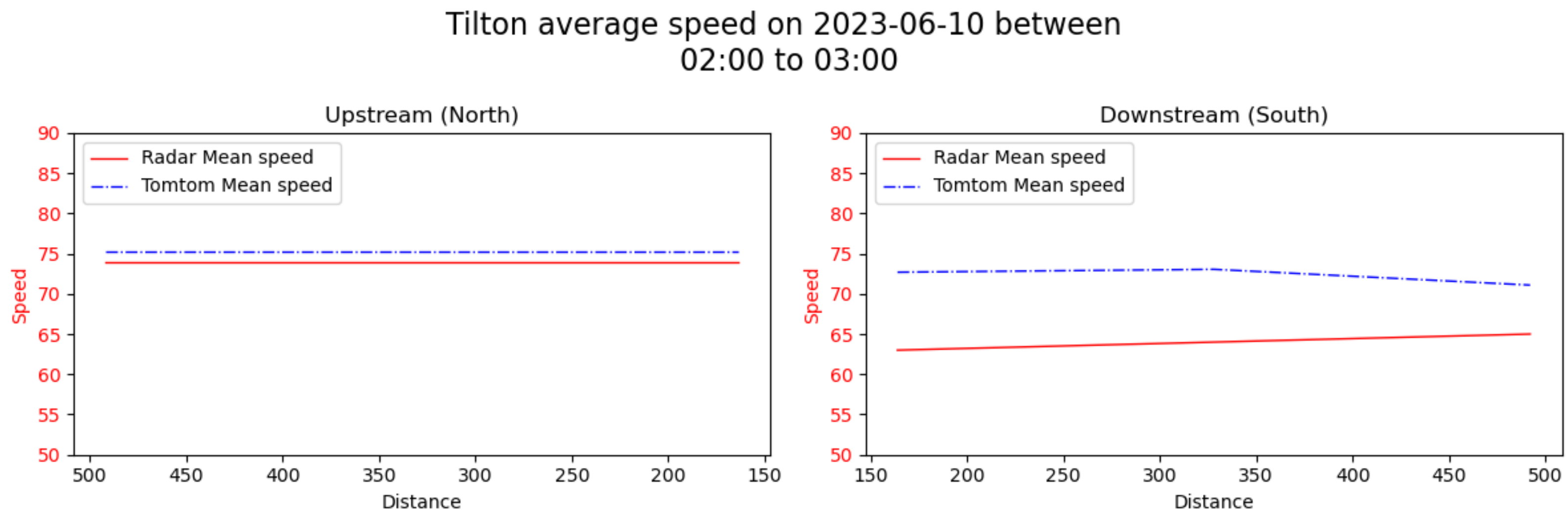


Figure 9-51. TomTom vs. Radar: Average Speed Comparison at Tilton (02:00-03:00)

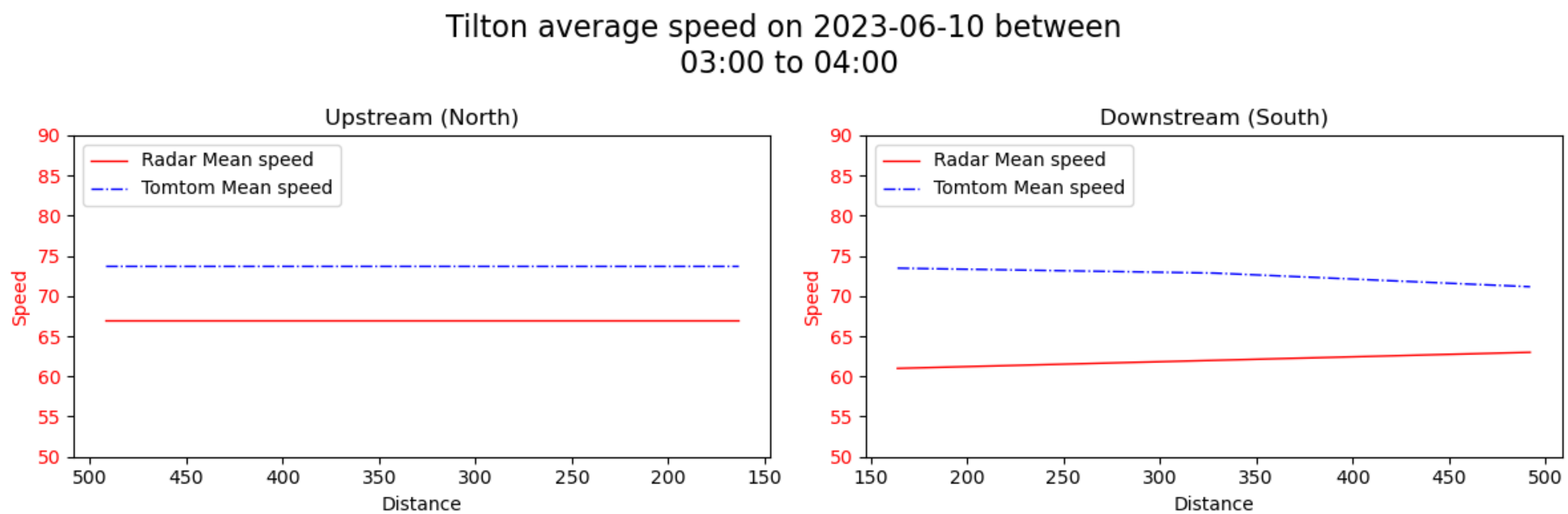


Figure 9-52. TomTom vs. Radar: Average Speed Comparison at Tilton (03:00-04:00)

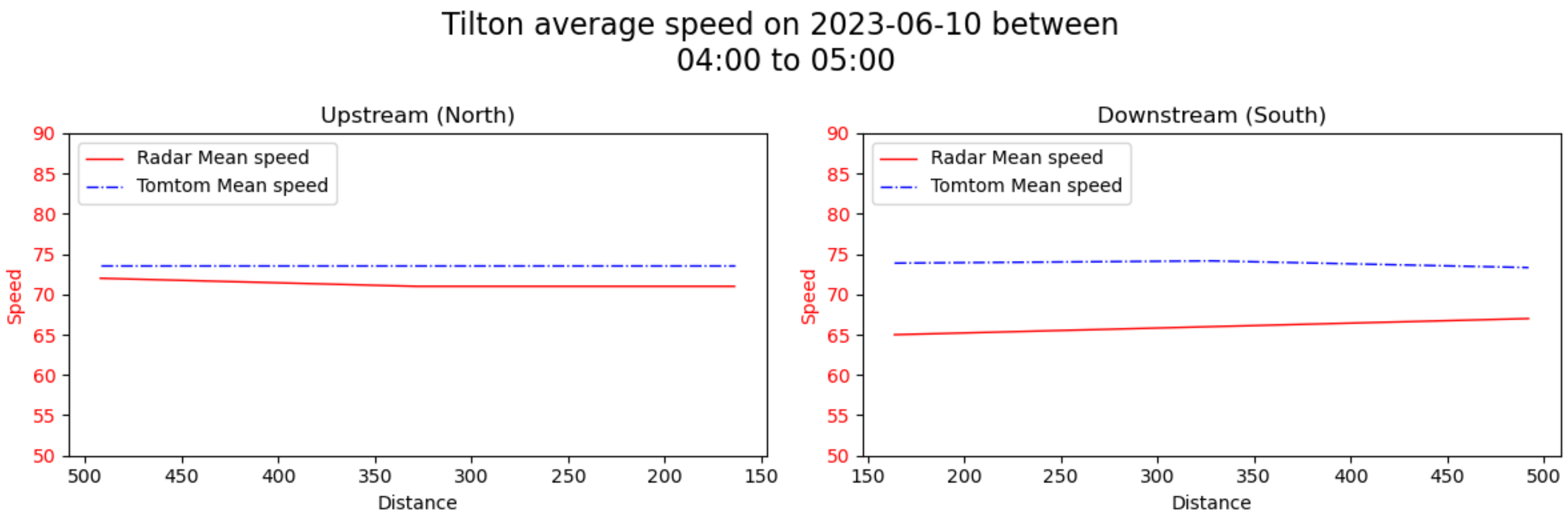


Figure 9-53. TomTom vs. Radar: Average Speed Comparison at Tilton (04:00-05:00)

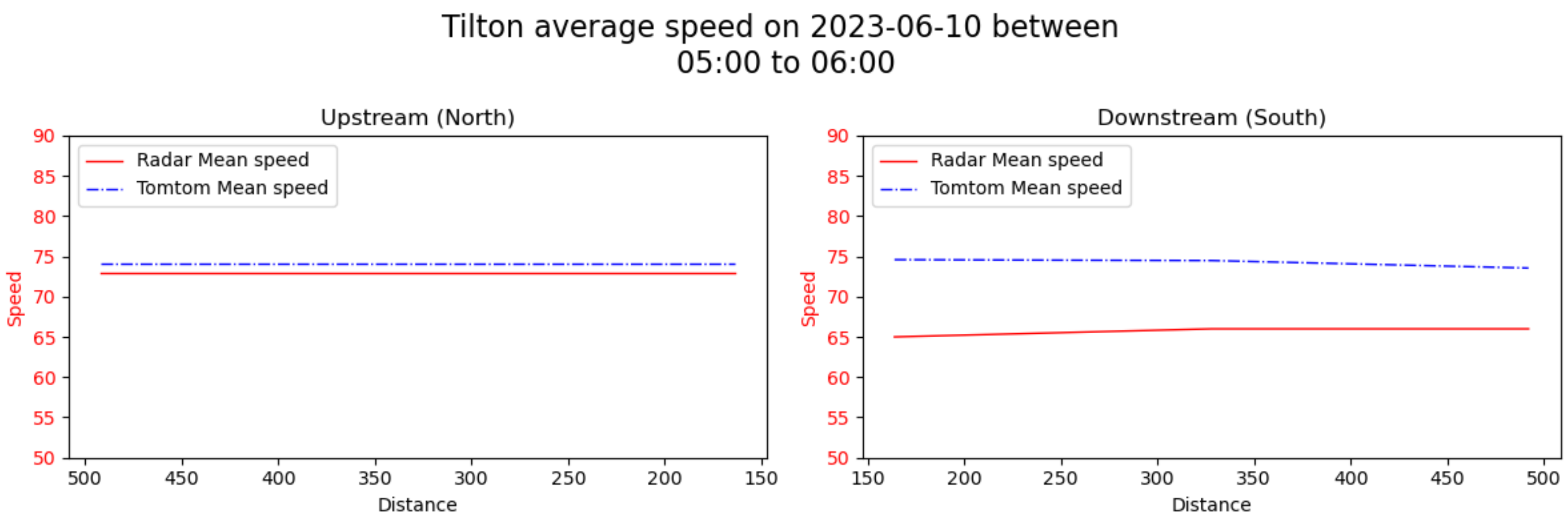


Figure 9-54. TomTom vs. Radar: Average Speed Comparison at Tilton (05:00-06:00)

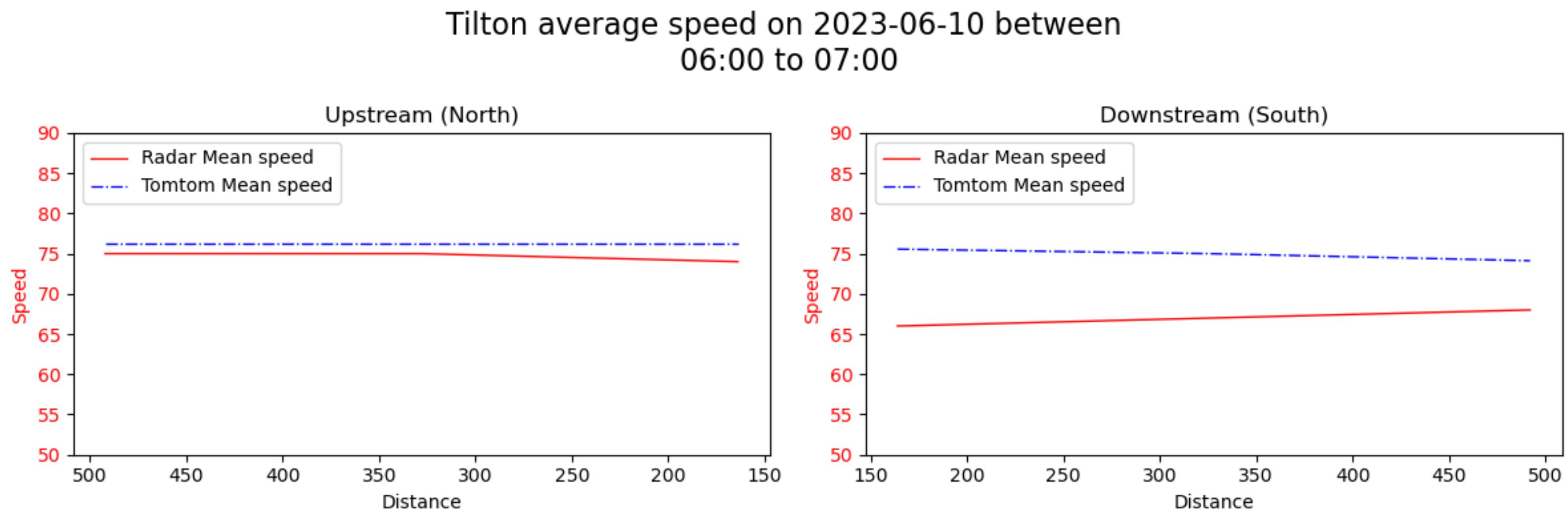


Figure 9-55. TomTom vs. Radar: Average Speed Comparison at Tilton (06:00-07:00)

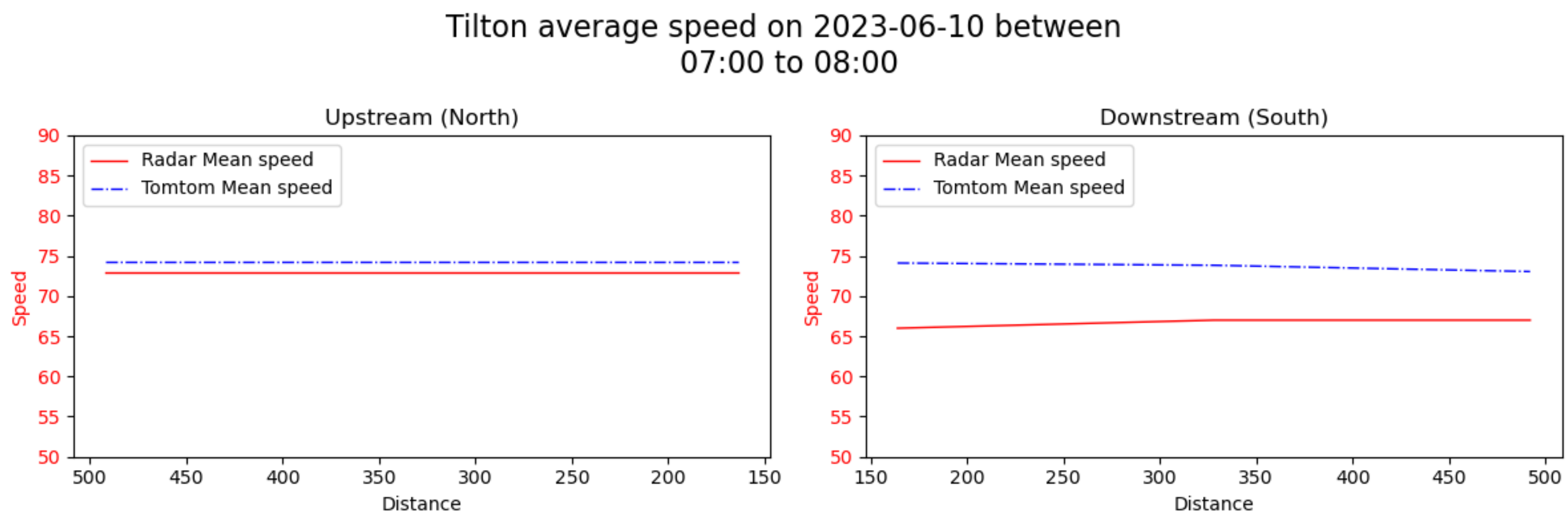


Figure 9-56. TomTom vs. Radar: Average Speed Comparison at Tilton (07:00-08:00)

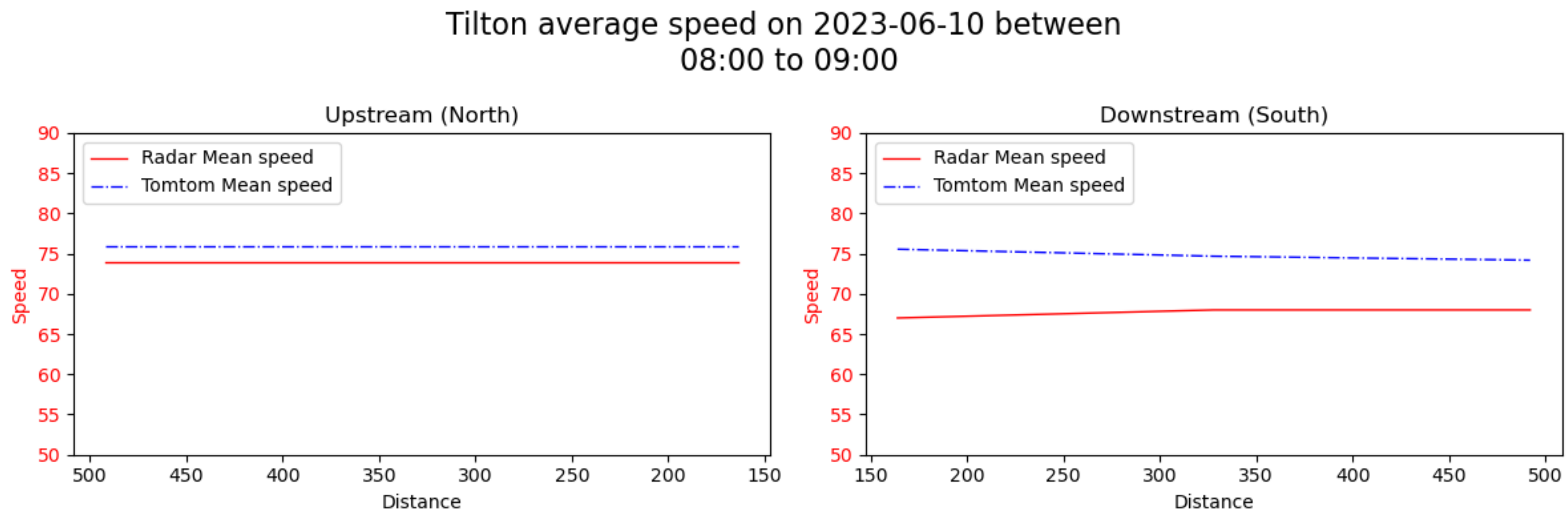


Figure 9-57. TomTom vs. Radar: Average Speed Comparison at Tilton (08:00-09:00)

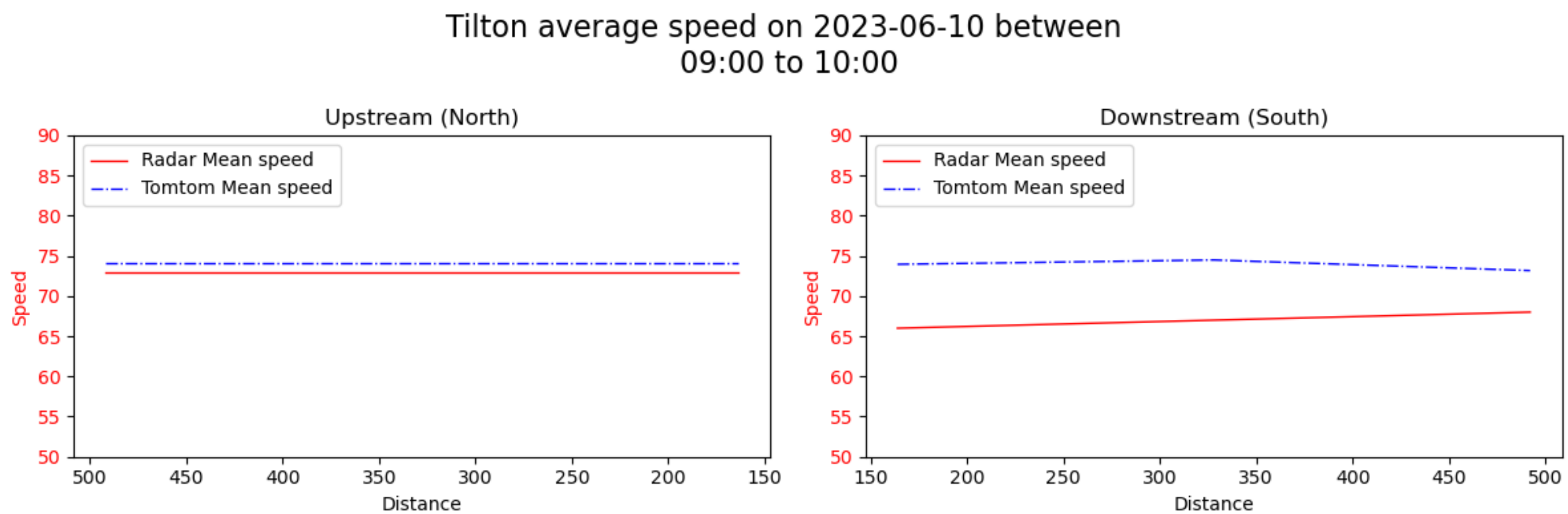


Figure 9-58. TomTom vs. Radar: Average Speed Comparison at Tilton (09:00-10:00)

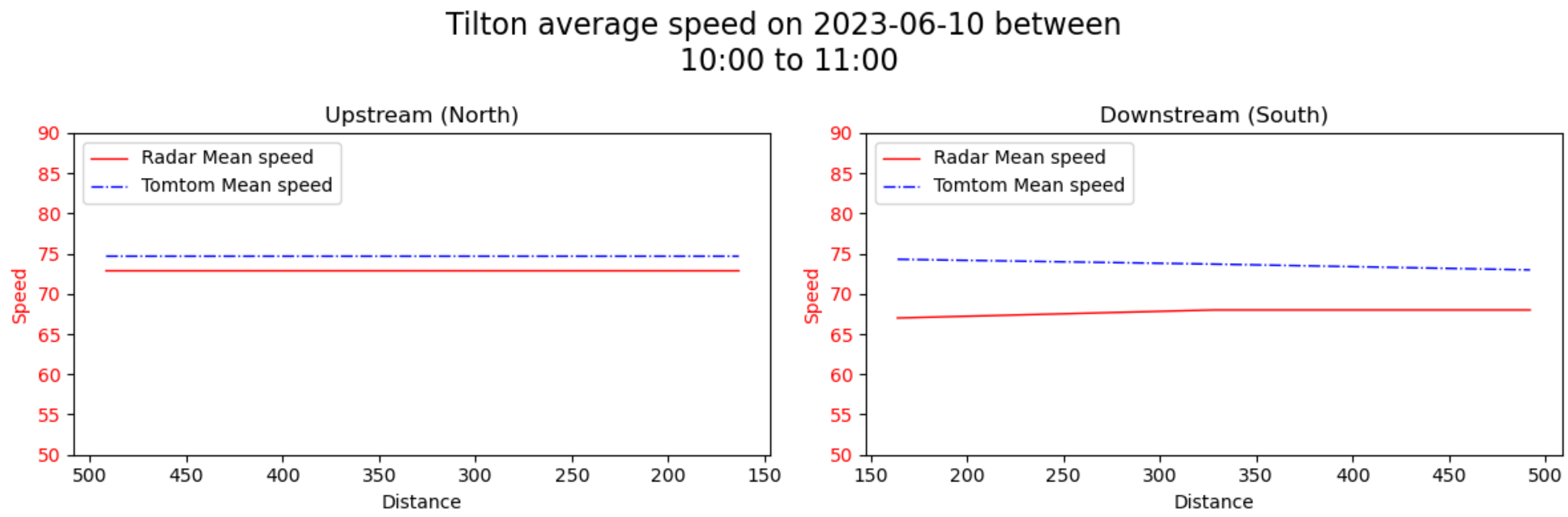


Figure 9-59. TomTom vs. Radar: Average Speed Comparison at Tilton (10:00-11:00)

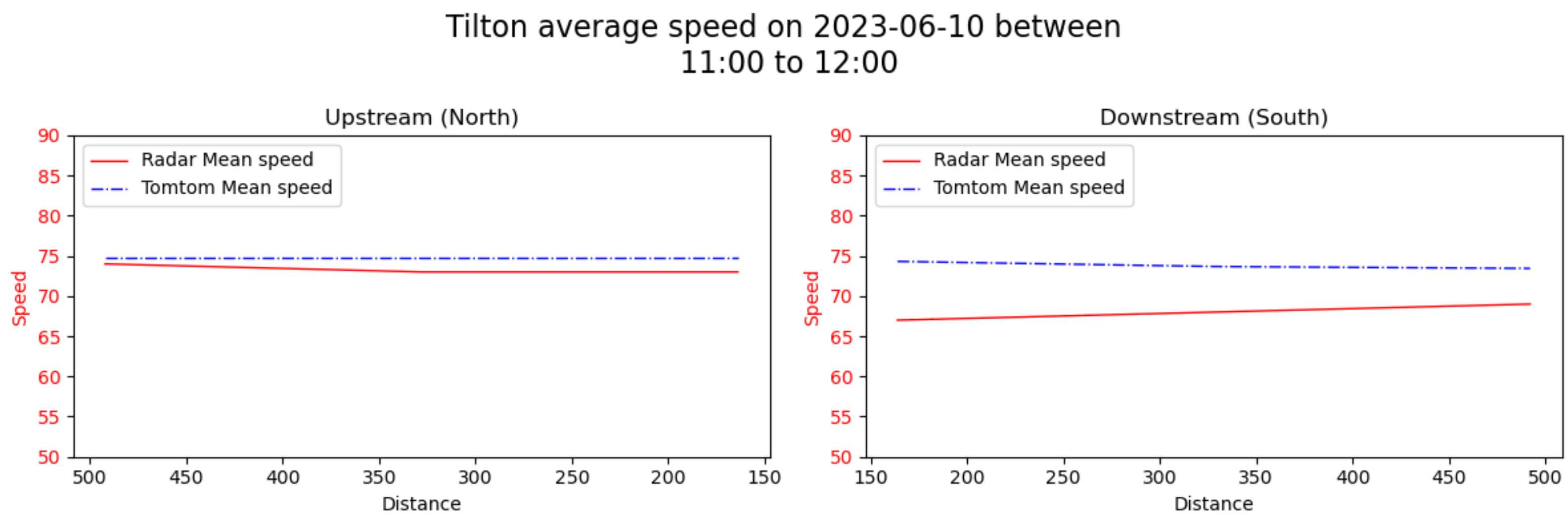


Figure 9-60. TomTom vs. Radar: Average Speed Comparison at Tilton (11:00-12:00)

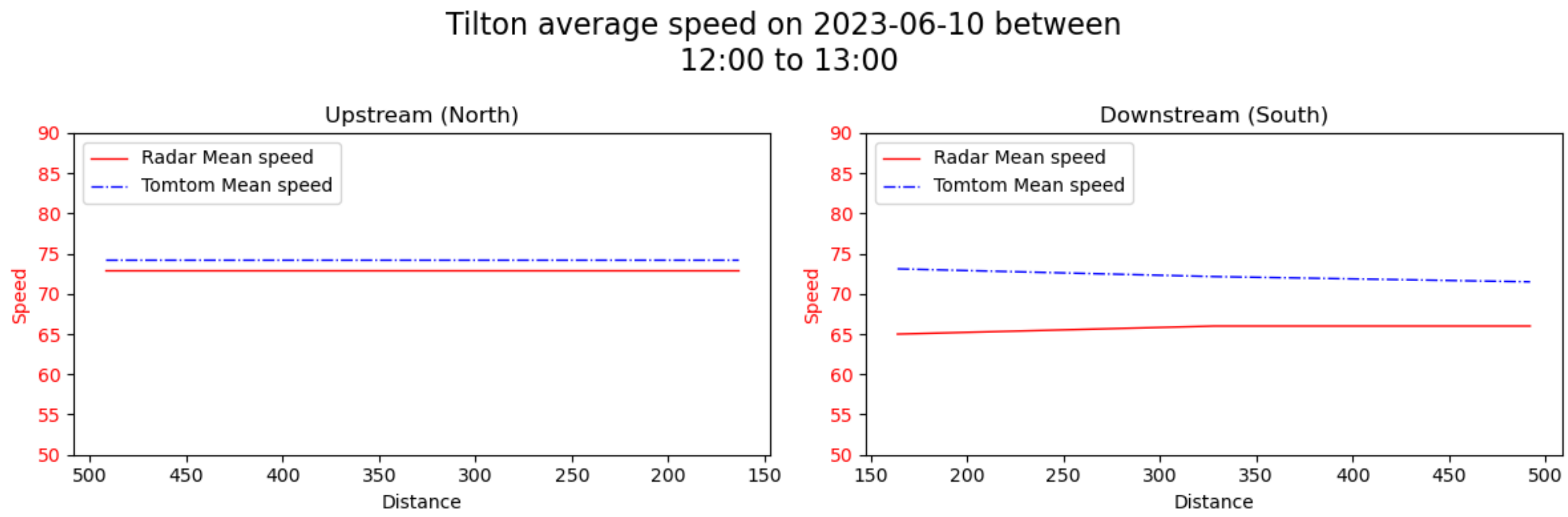


Figure 9-61. TomTom vs. Radar: Average Speed Comparison at Tilton (12:00-13:00)

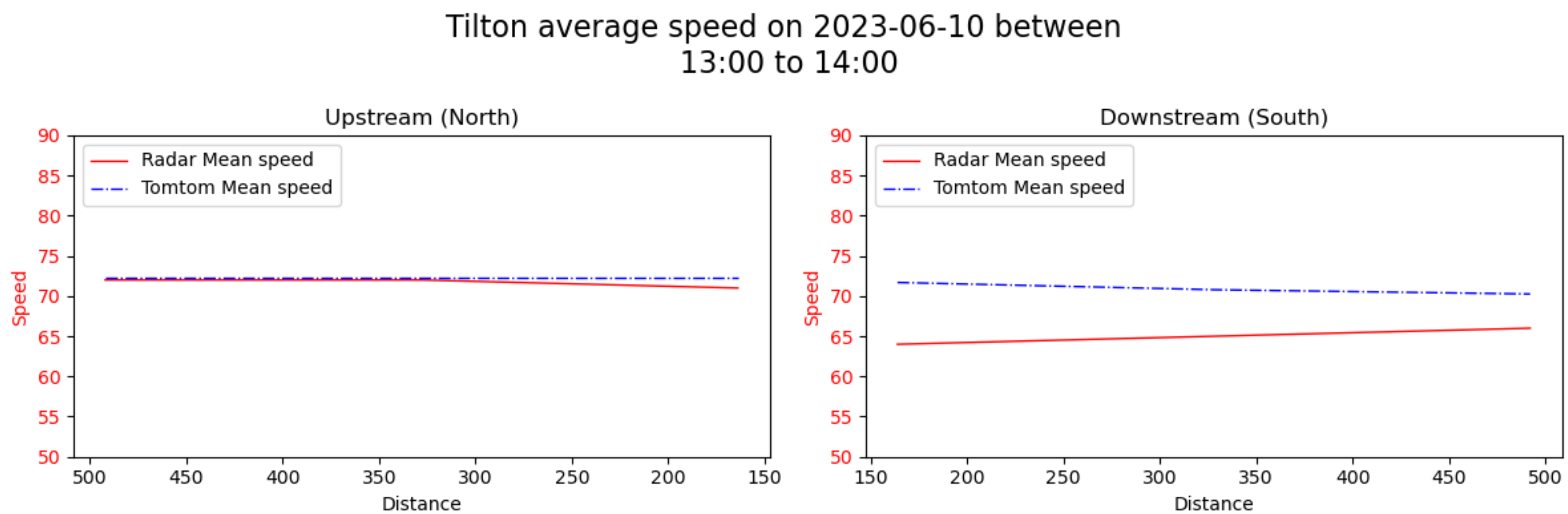


Figure 9-62. TomTom vs. Radar: Average Speed Comparison at Tilton (13:00-14:00)

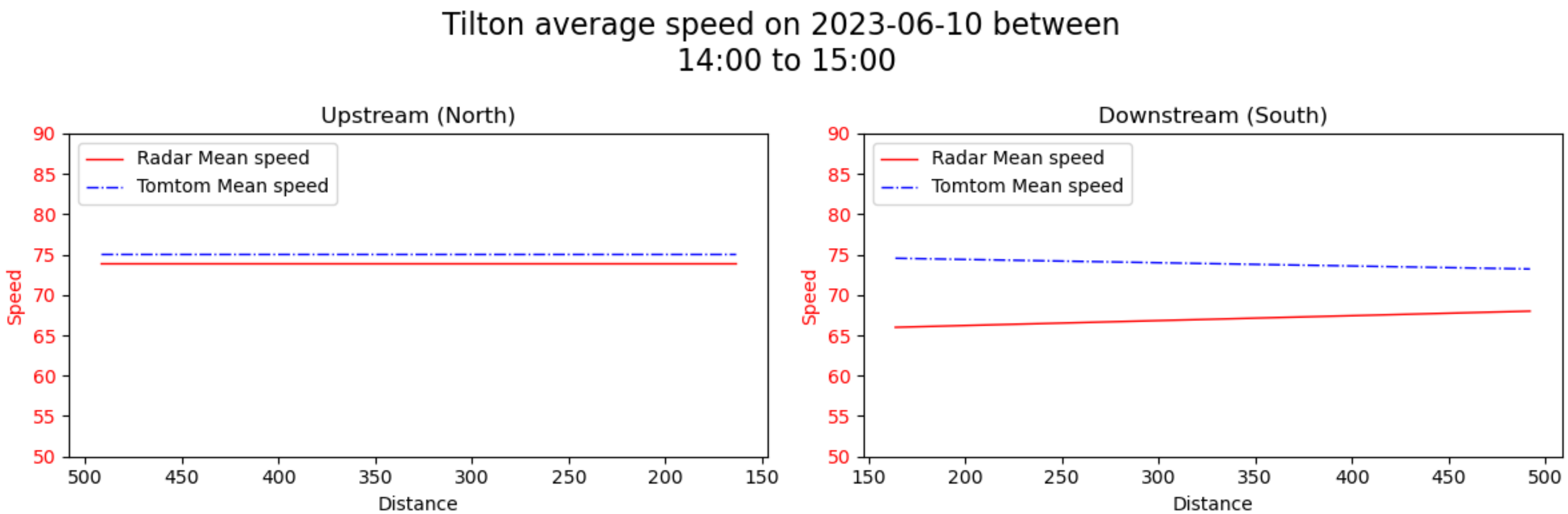


Figure 9-63. TomTom vs. Radar: Average Speed Comparison at Tilton (14:00-15:00)

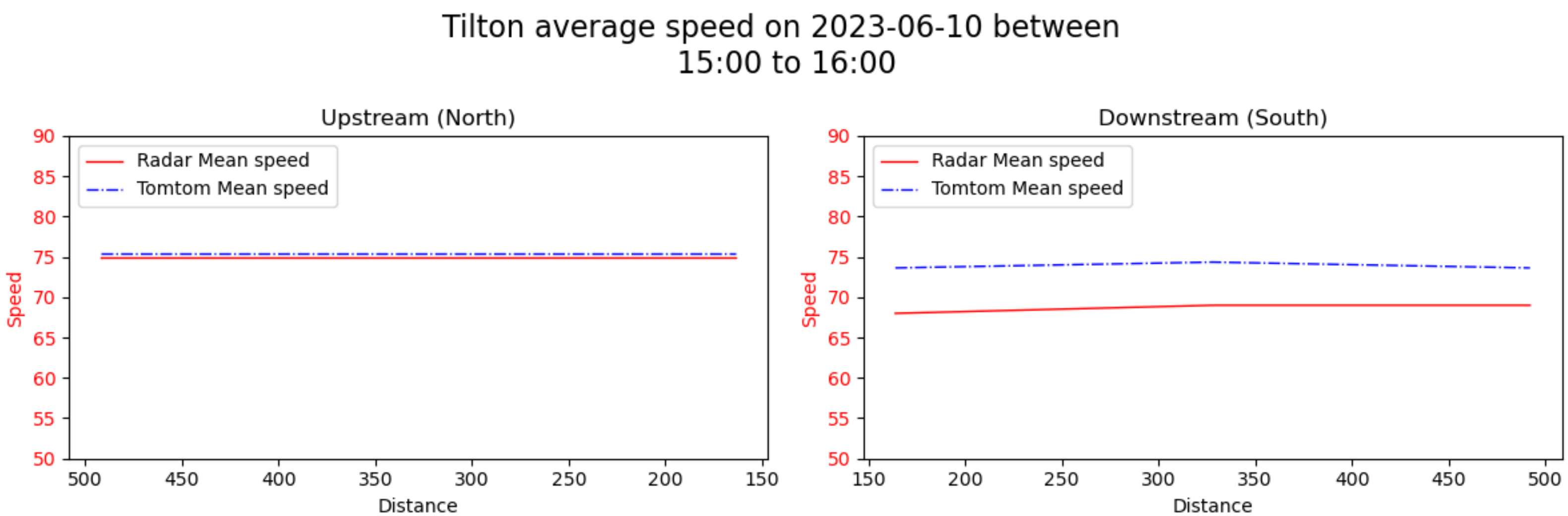


Figure 9-64. TomTom vs. Radar: Average Speed Comparison at Tilton (15:00-16:00)

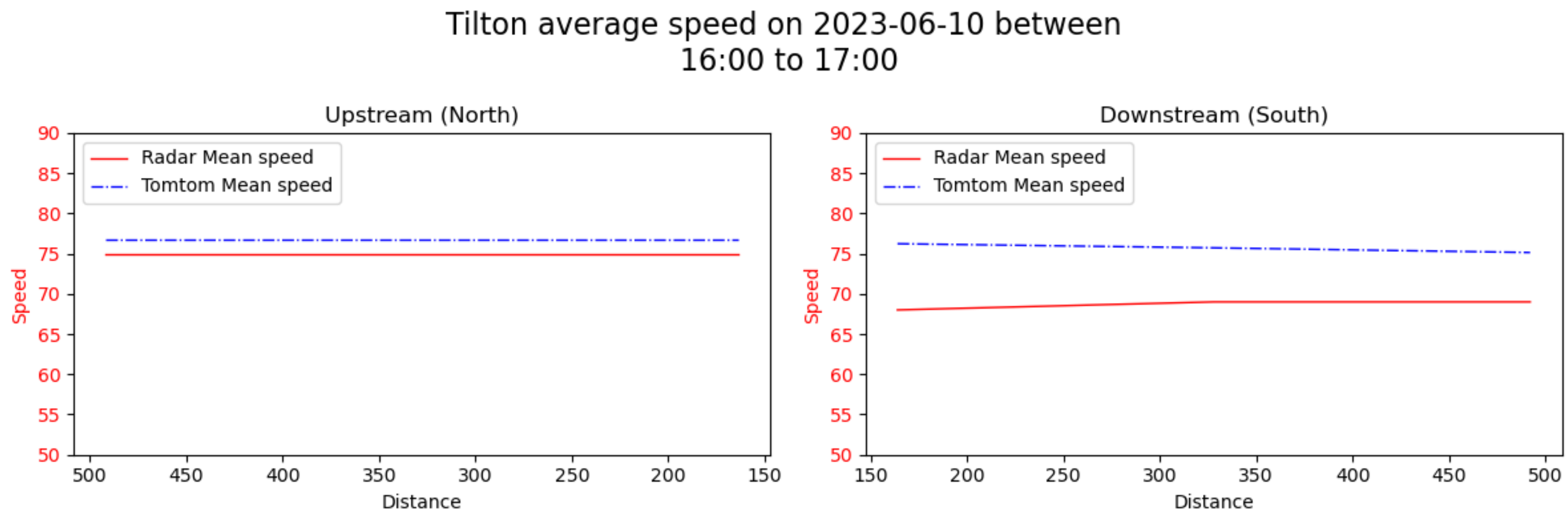


Figure 9-65. TomTom vs. Radar: Average Speed Comparison at Tilton (16:00-17:00)

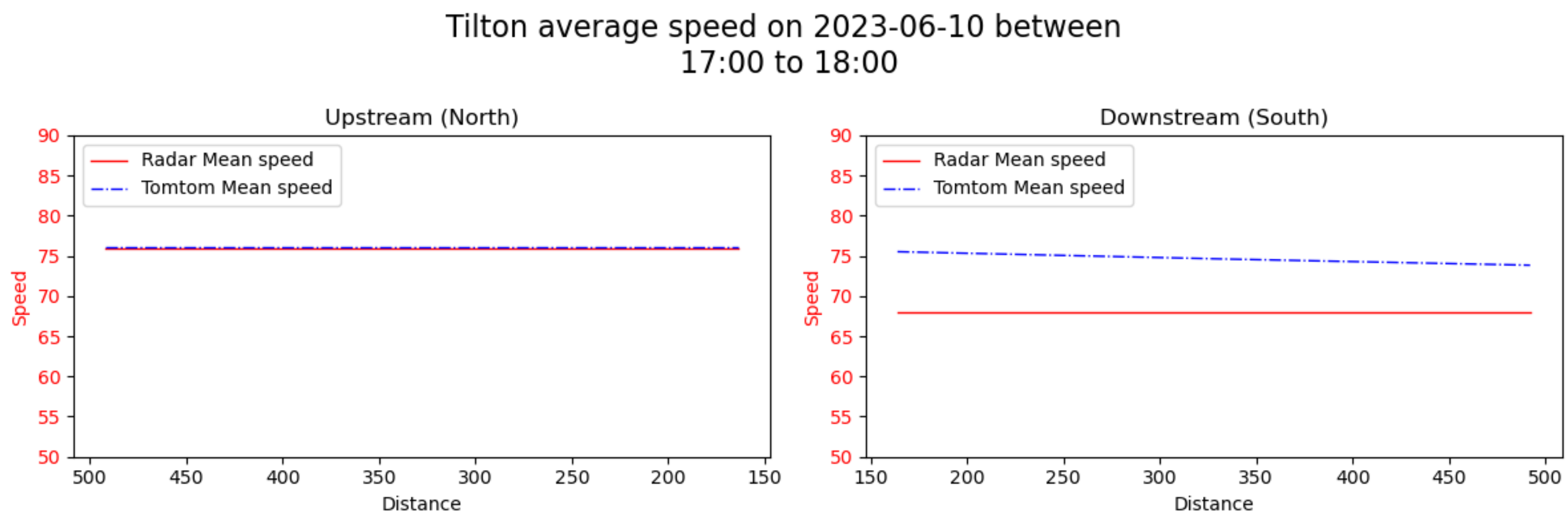


Figure 9-66. TomTom vs. Radar: Average Speed Comparison at Tilton (17:00-18:00)

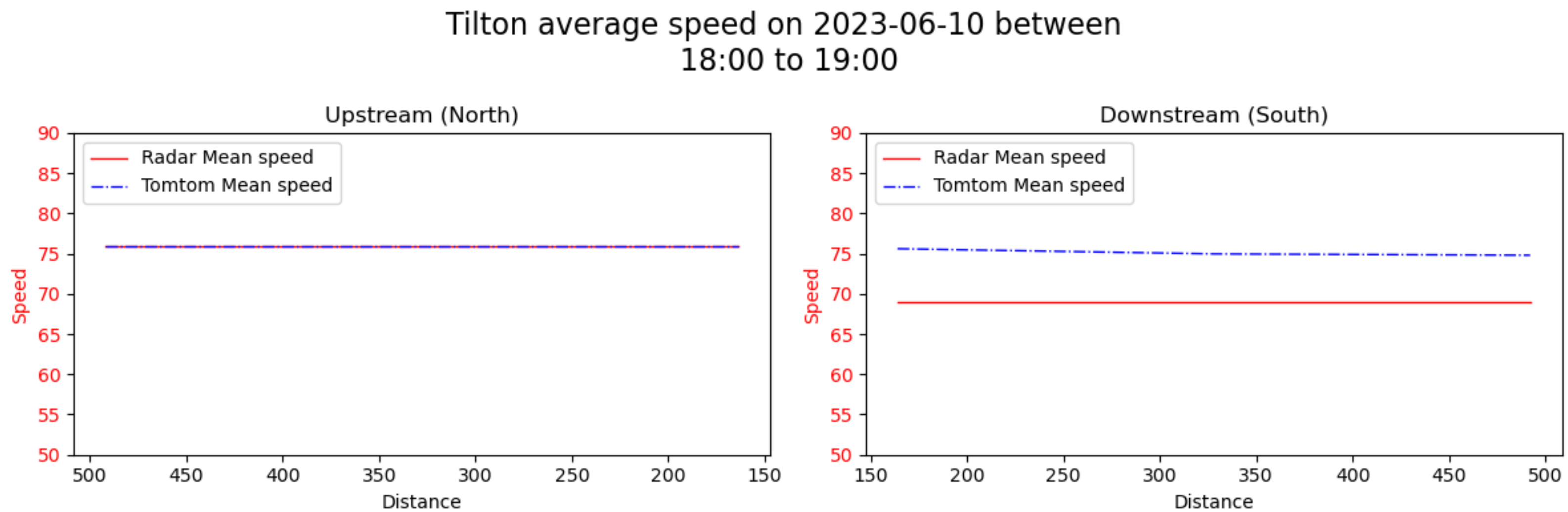


Figure 9-67. TomTom vs. Radar: Average Speed Comparison at Tilton (18:00-19:00)

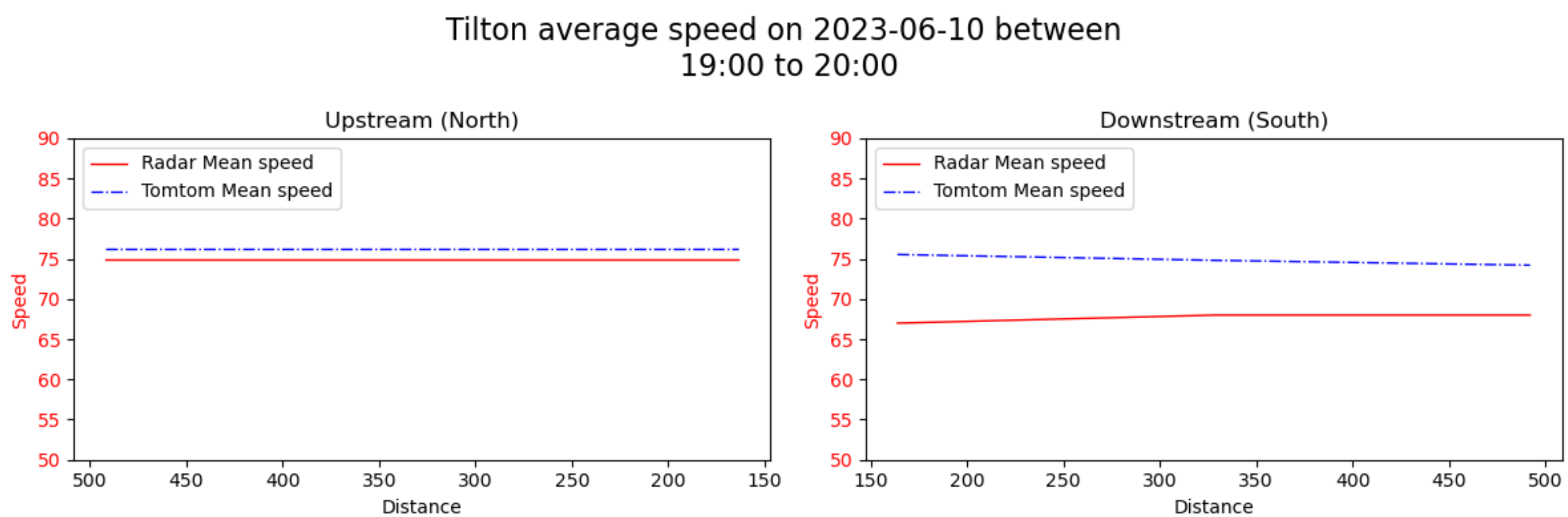


Figure 9-68. TomTom vs. Radar: Average Speed Comparison at Tilton (19:00-20:00)

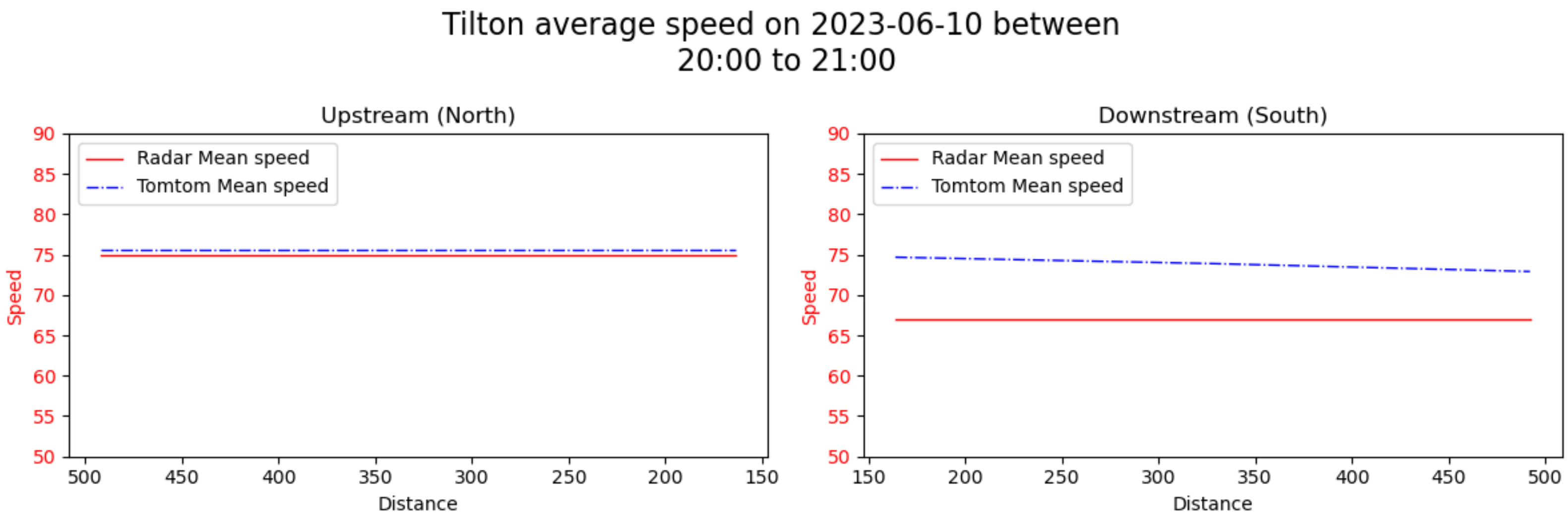


Figure 9-69. TomTom vs. Radar: Average Speed Comparison at Tilton (20:00-21:00)

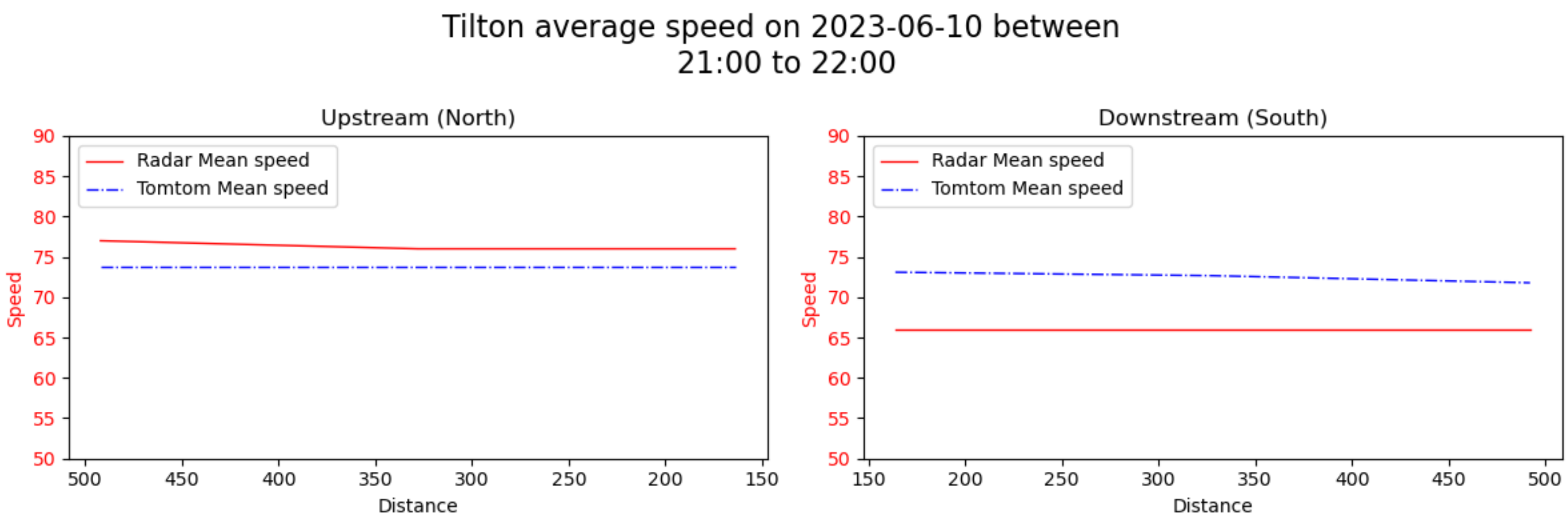


Figure 9-70. TomTom vs. Radar: Average Speed Comparison at Tilton (21:00-22:00)

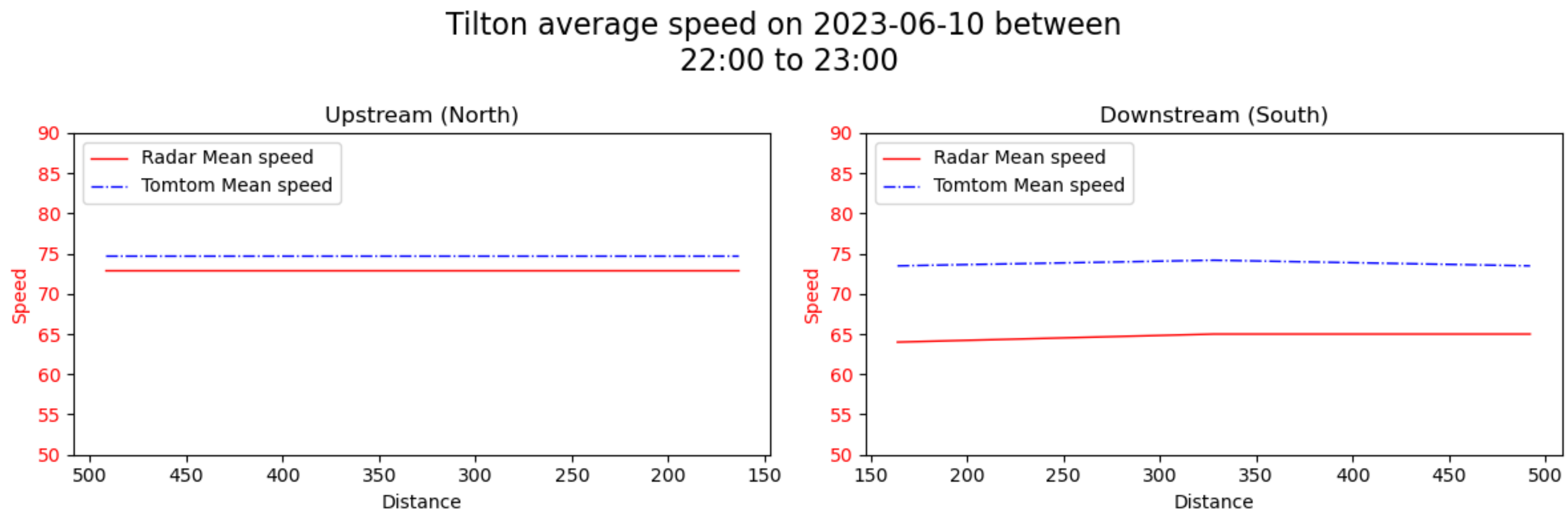


Figure 9-71. TomTom vs. Radar: Average Speed Comparison at Tilton (22:00-23:00)

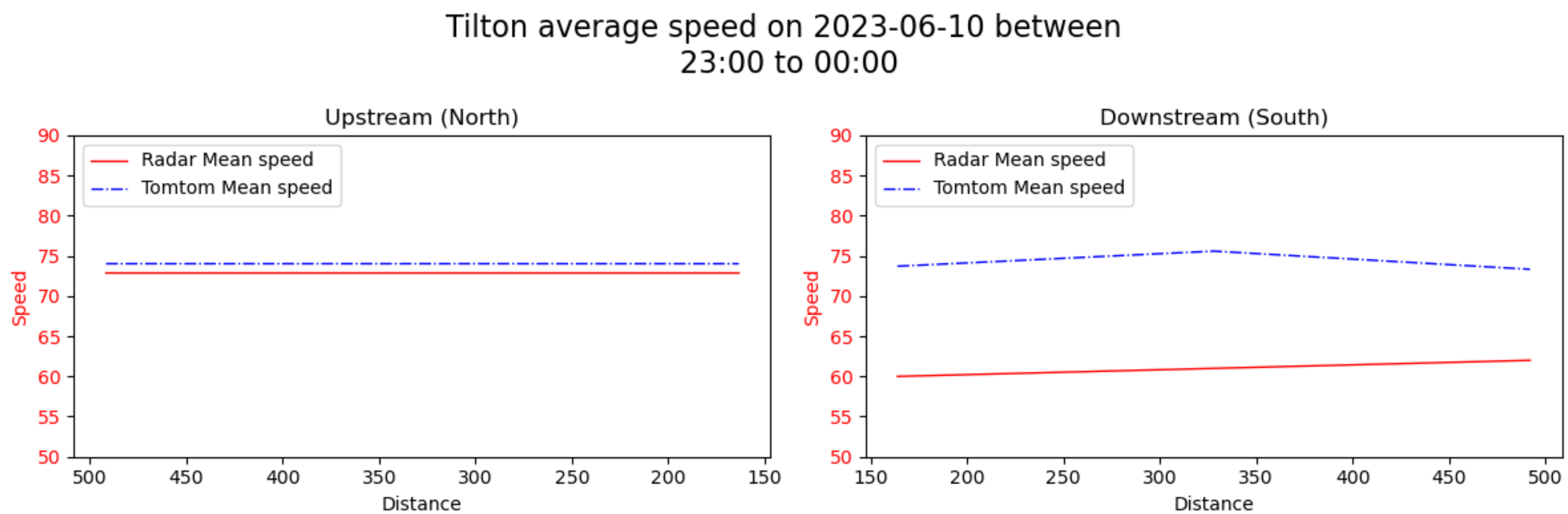


Figure 9-72. TomTom vs. Radar: Average Speed Comparison at Tilton (23:00-24:00)

9.3 Appendix C. Campton Speed Profiles

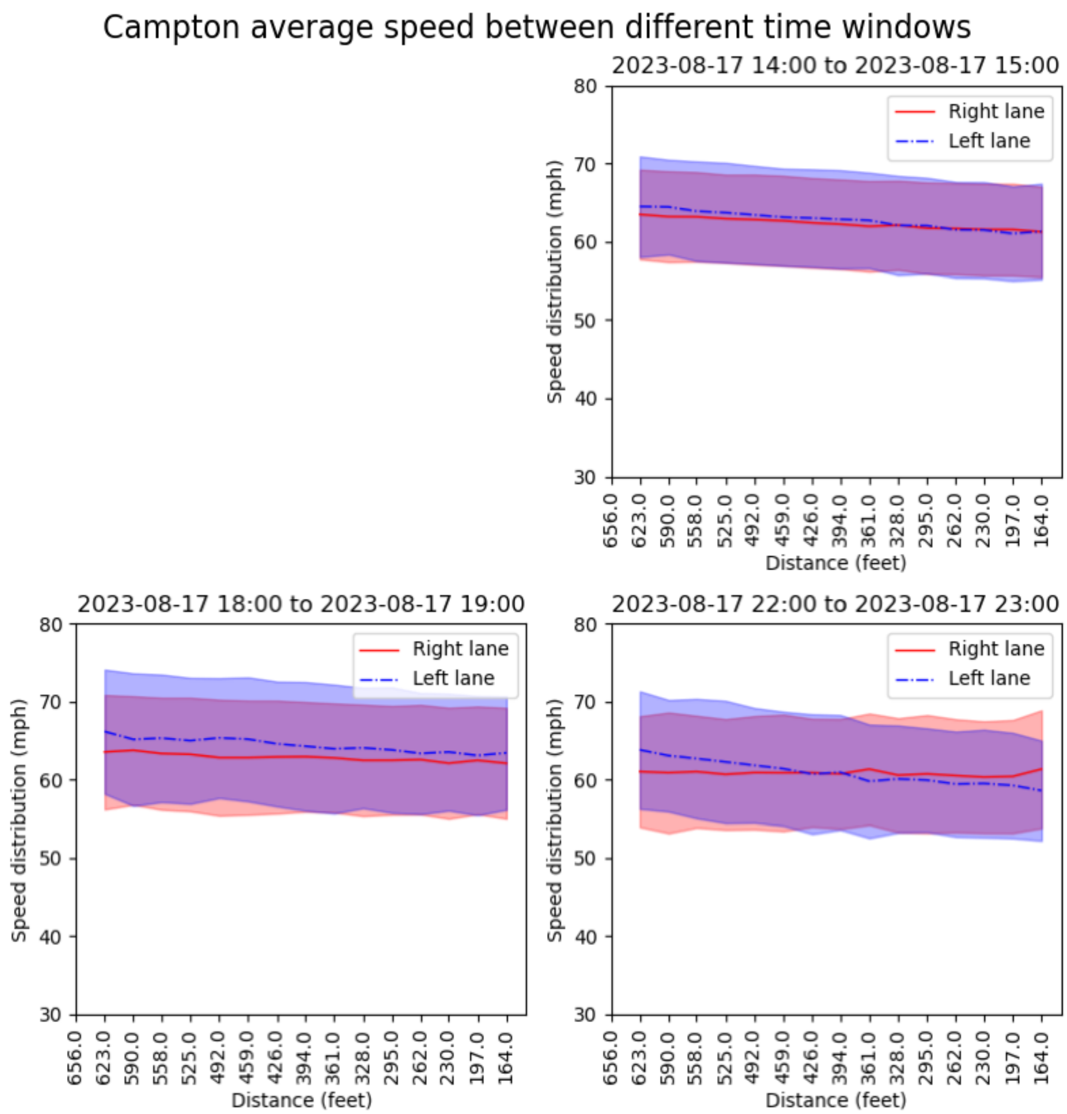


Figure 9-73. Average Speed Comparison between Left and Right Lane across Dates

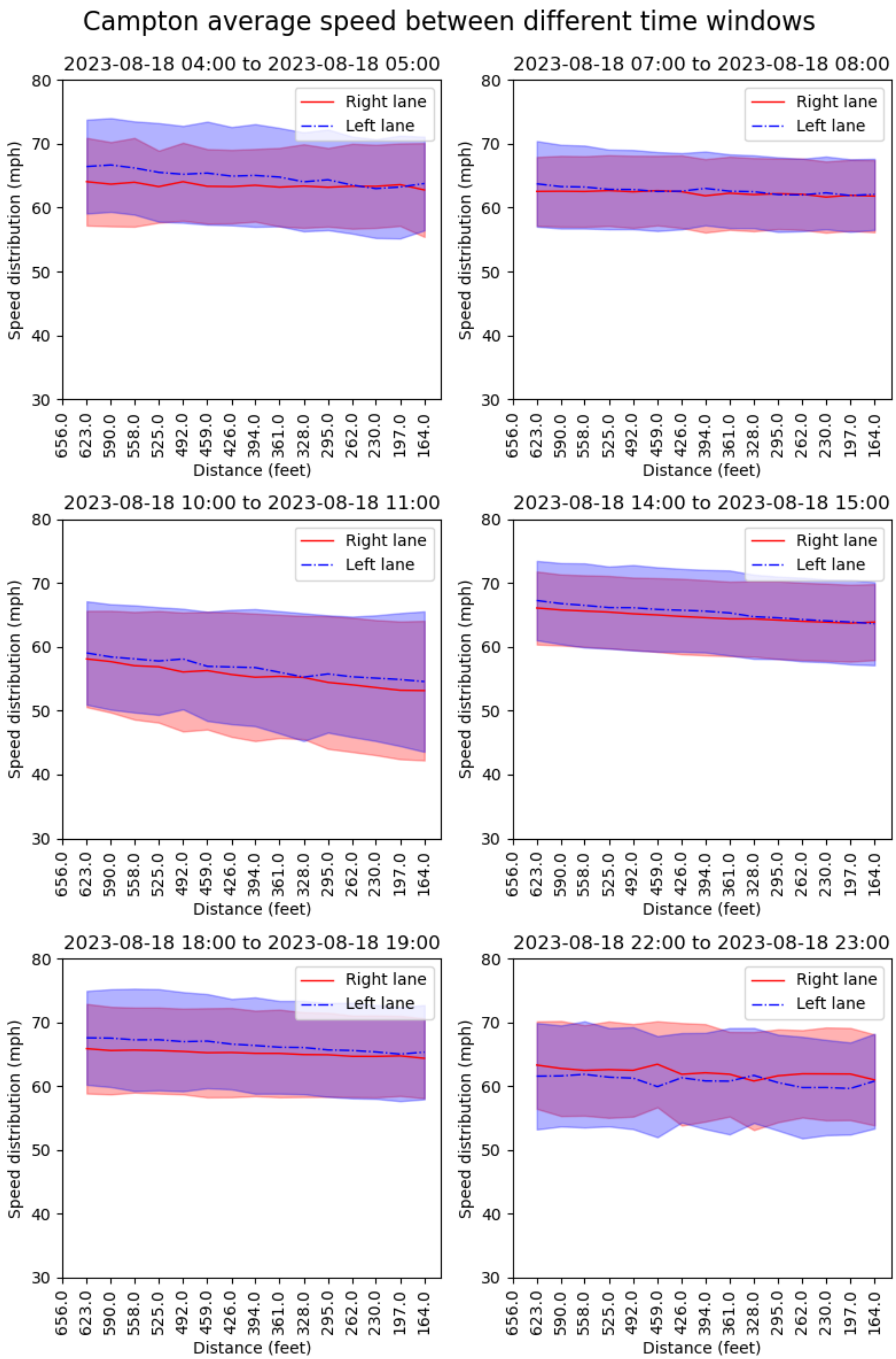


Figure 9-74. Average Speed Comparison between Left and Right Lane across Dates

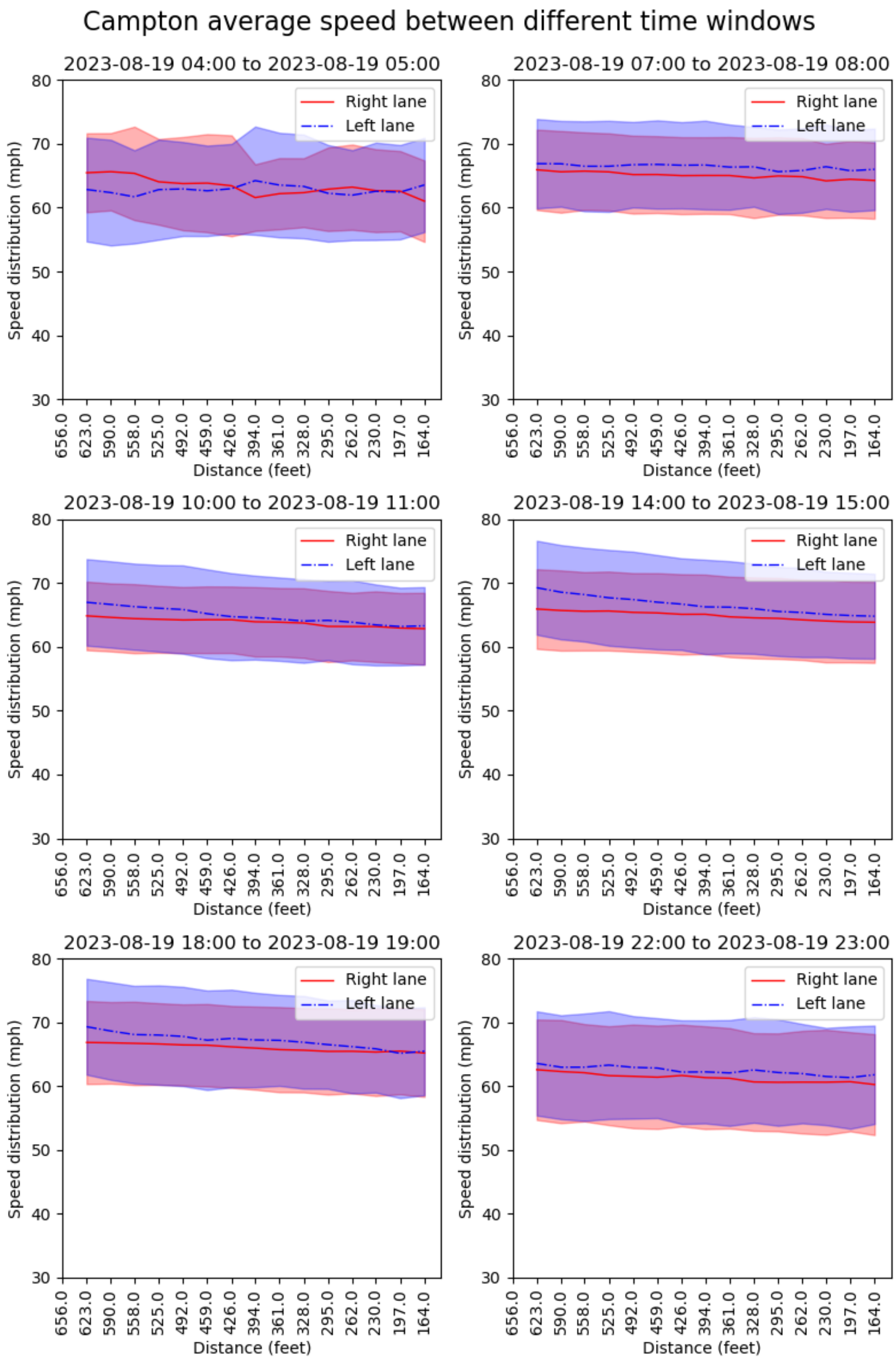


Figure 9-75. Average Speed Comparison between Left and Right Lane across Dates

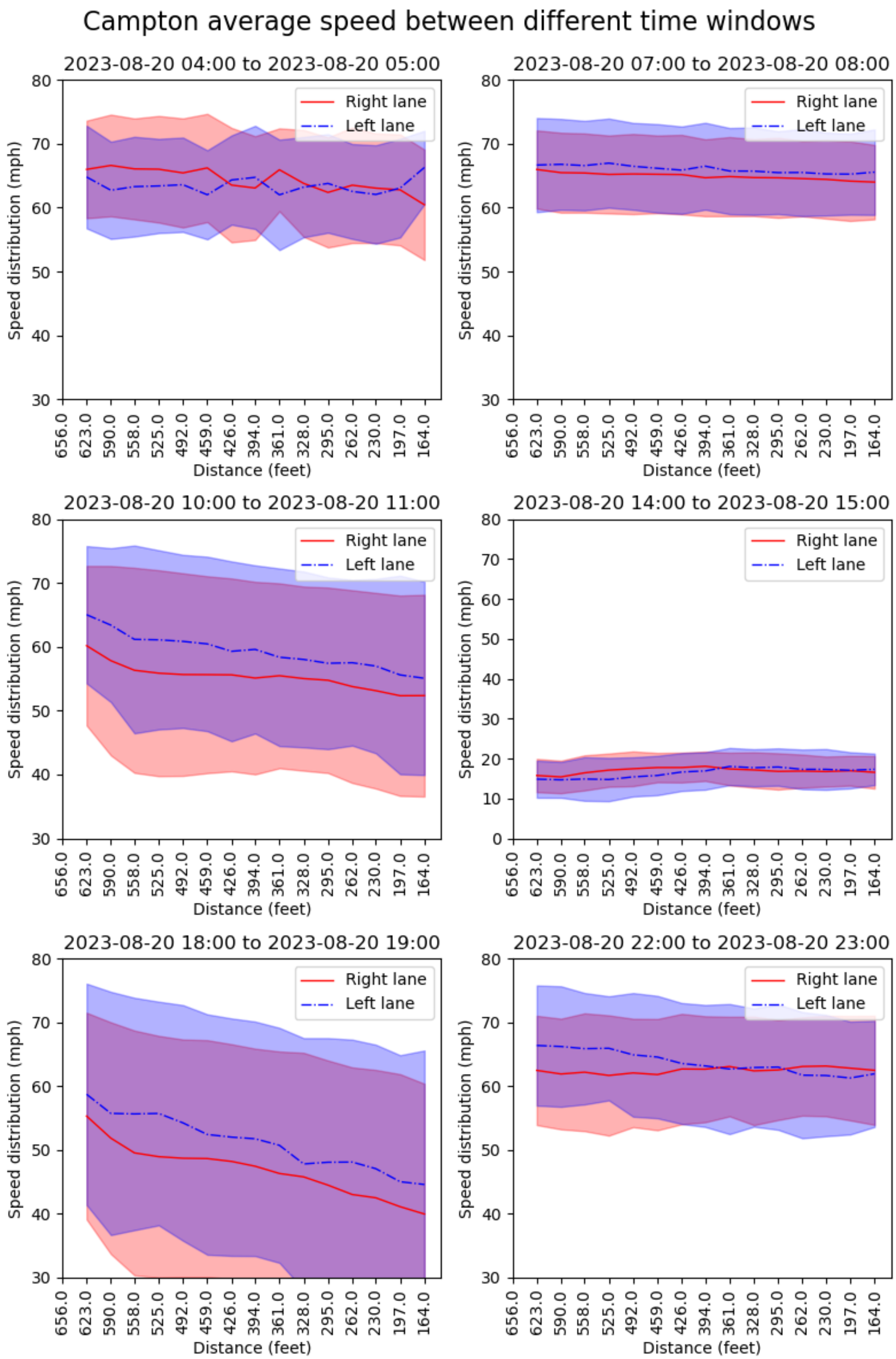


Figure 9-76. Average Speed Comparison between Left and Right Lane across Dates

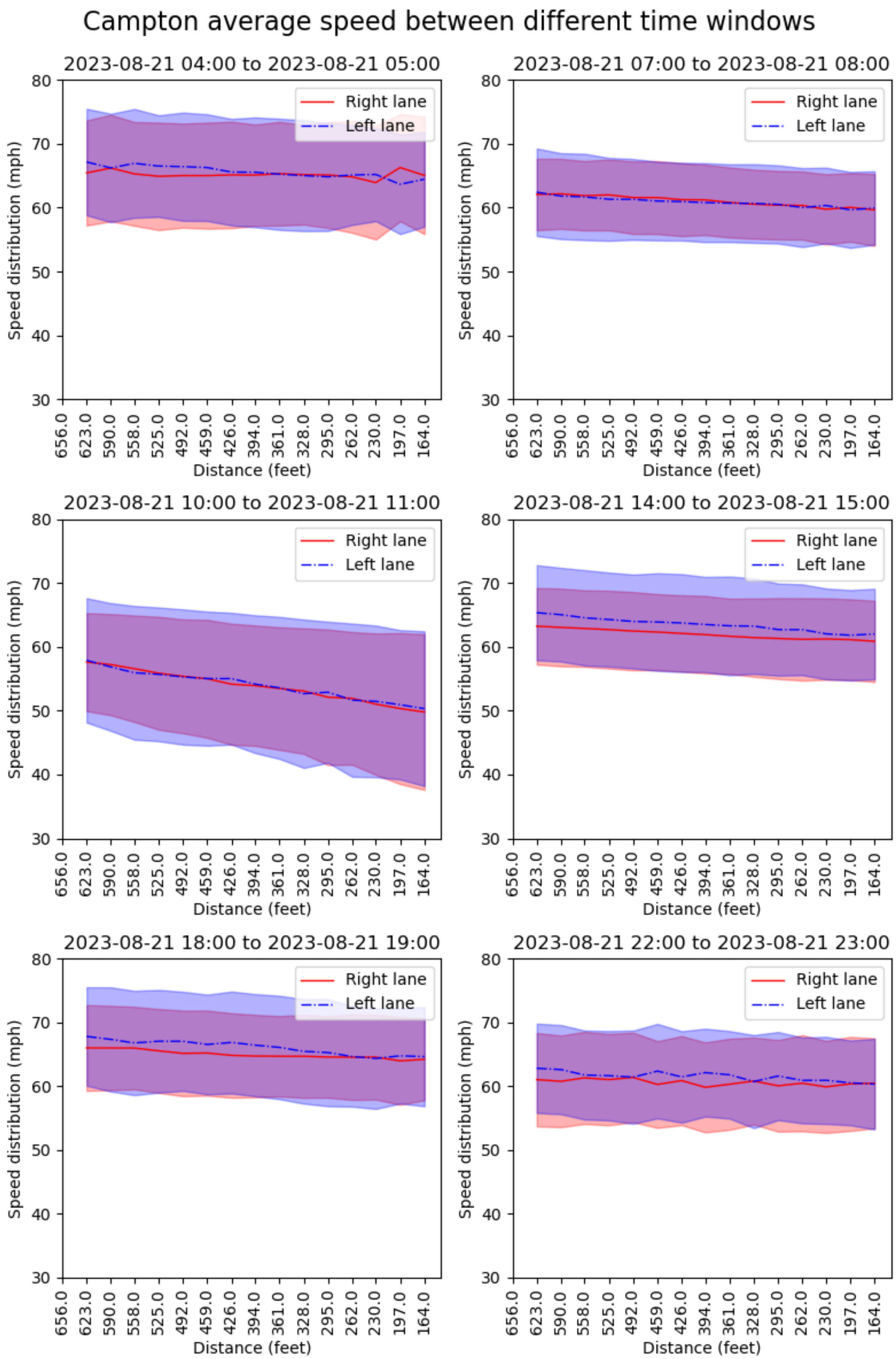


Figure 9-77. Average Speed Comparison between Left and Right Lane across Dates

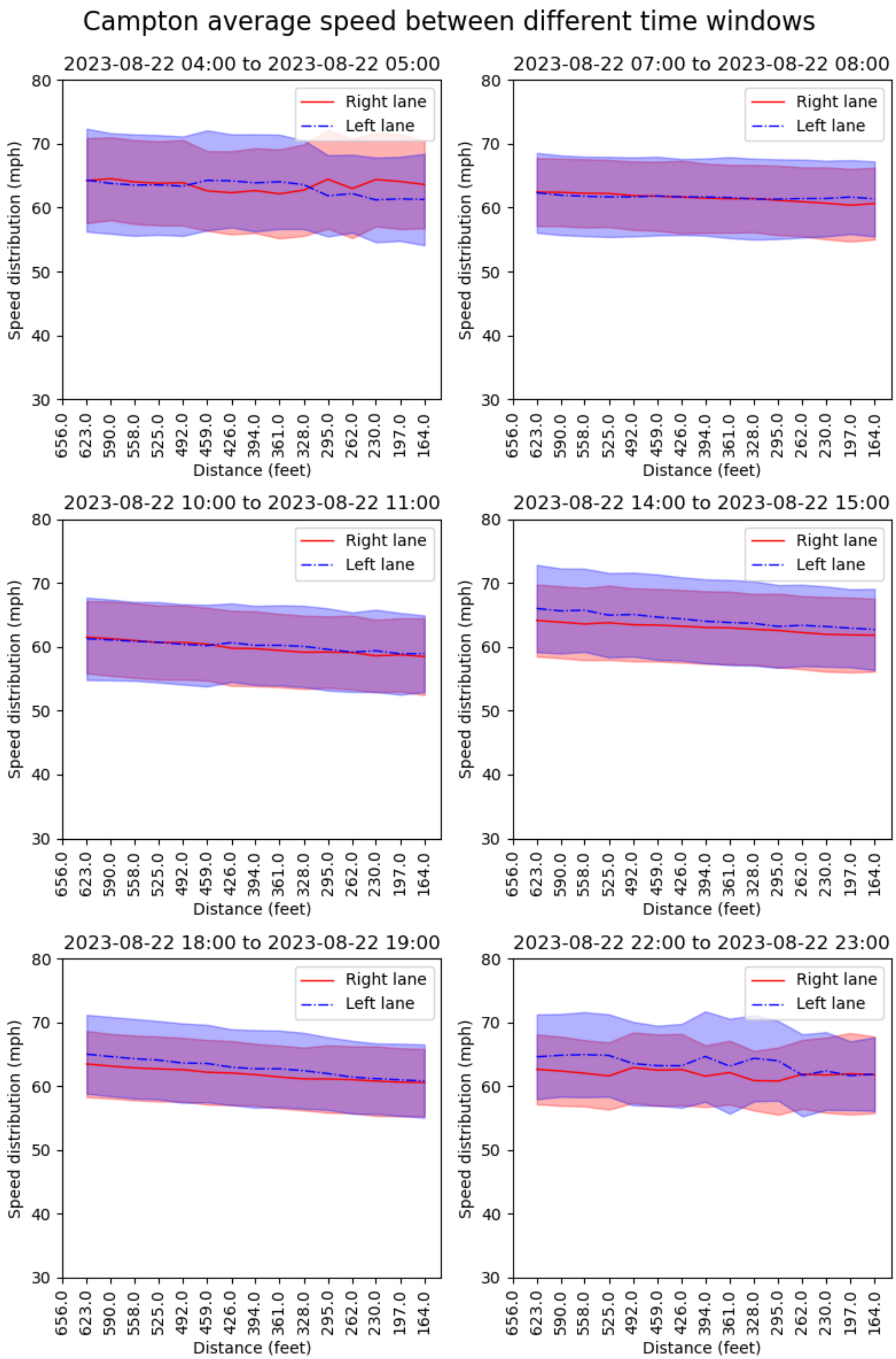


Figure 9-78. Average Speed Comparison between Left and Right Lane across Dates

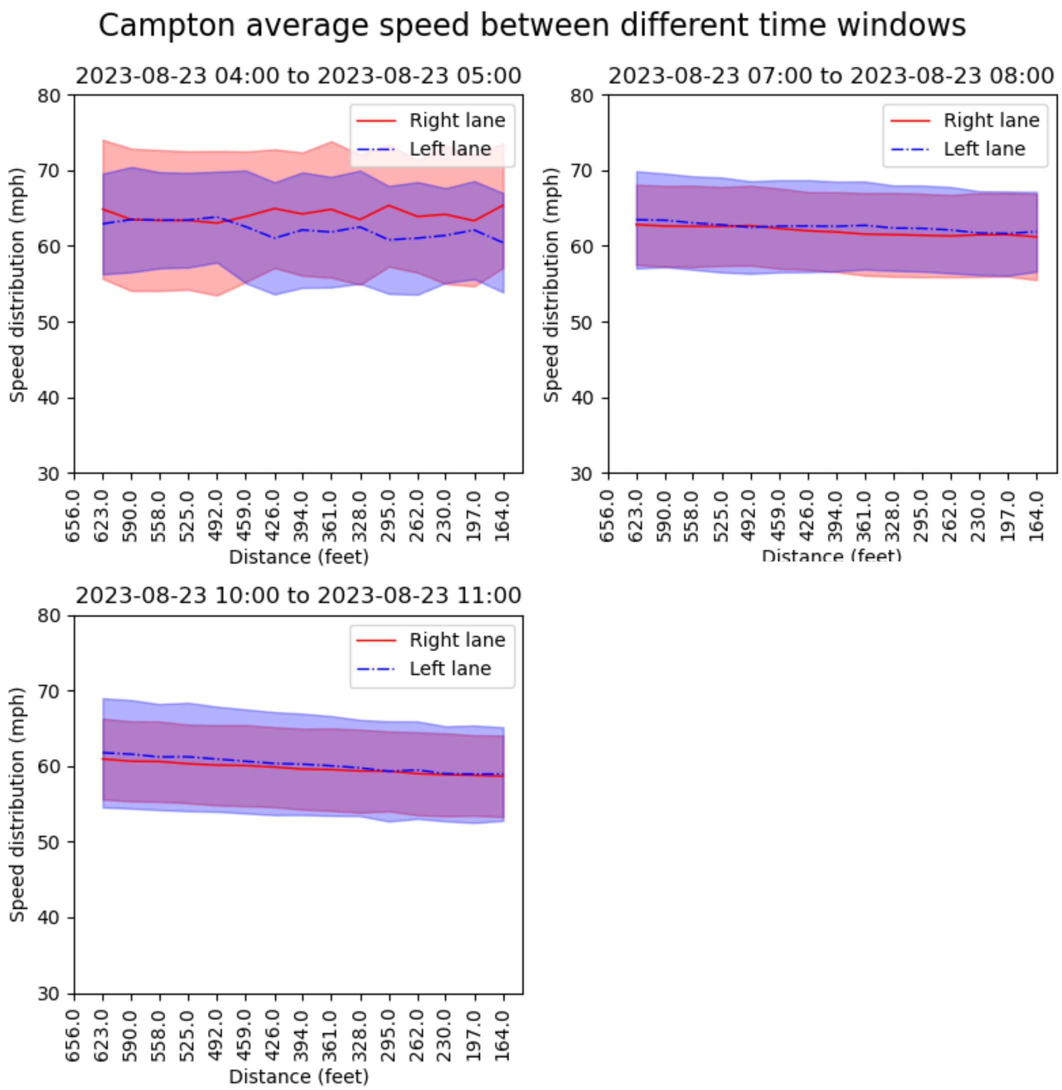


Figure 9-79. Average Speed Comparison between Left and Right Lane across Dates

9.4 Appendix D. Traffic Volume Data

Littleton North

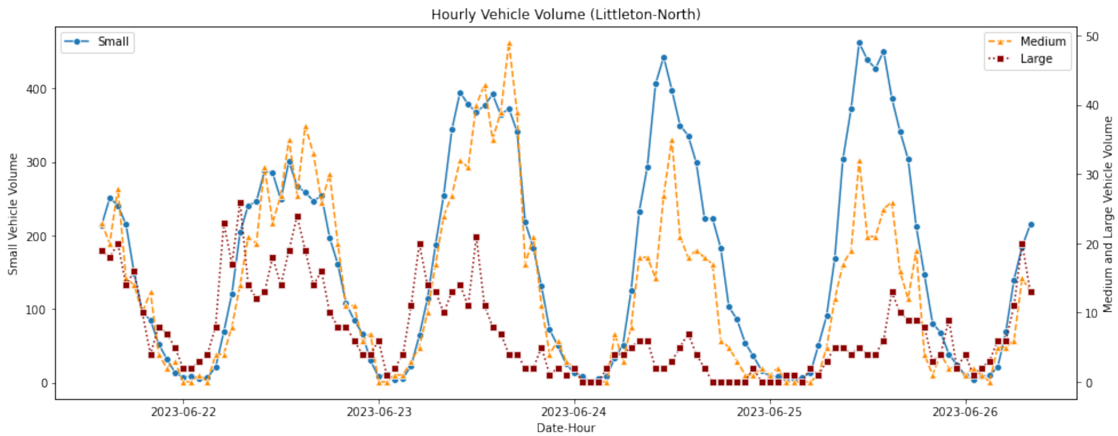


Figure 9-80. Hourly vehicle volume at Littleton North for all days.

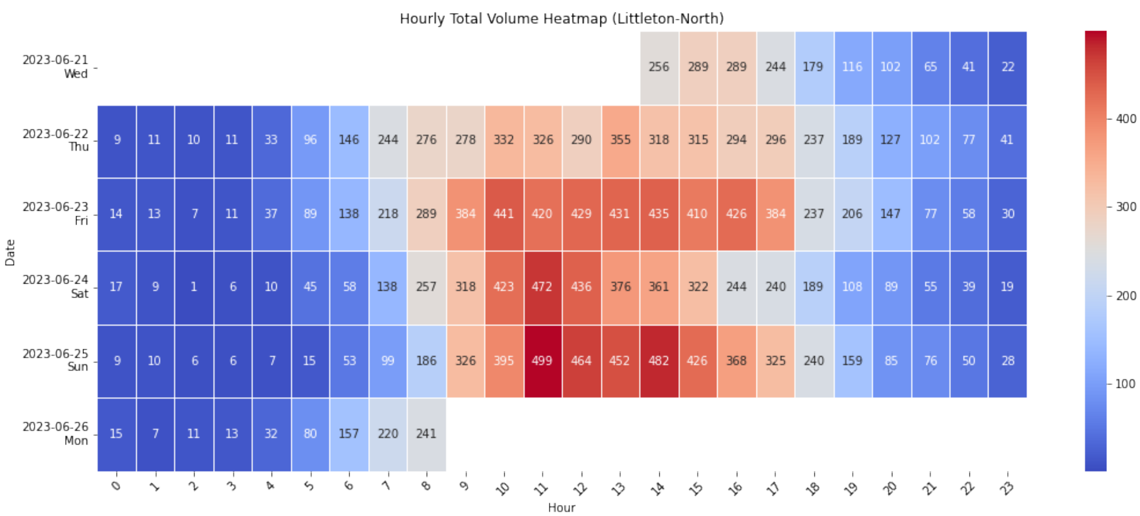


Figure 9-81. Heatmap illustrating hourly vehicle volume at Littleton North.

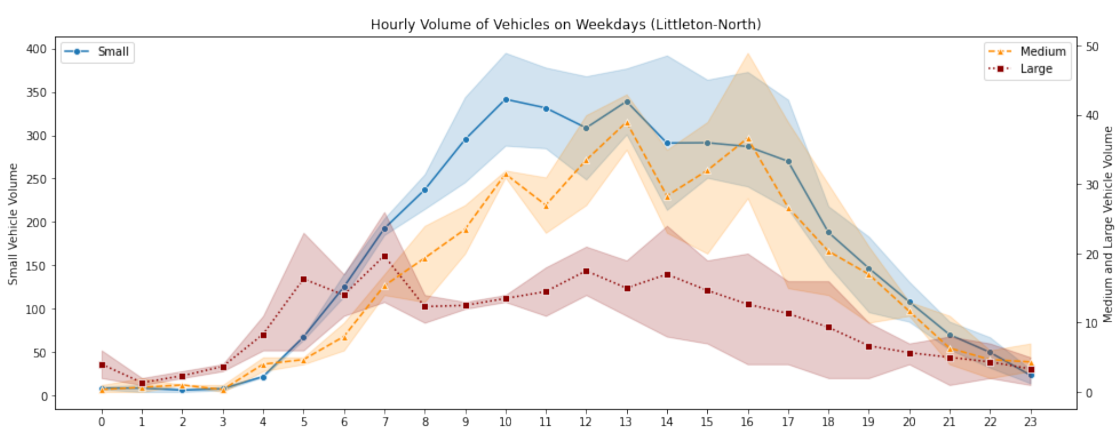


Figure 9-82. Average hourly volume of vehicles on weekdays at Littleton North.

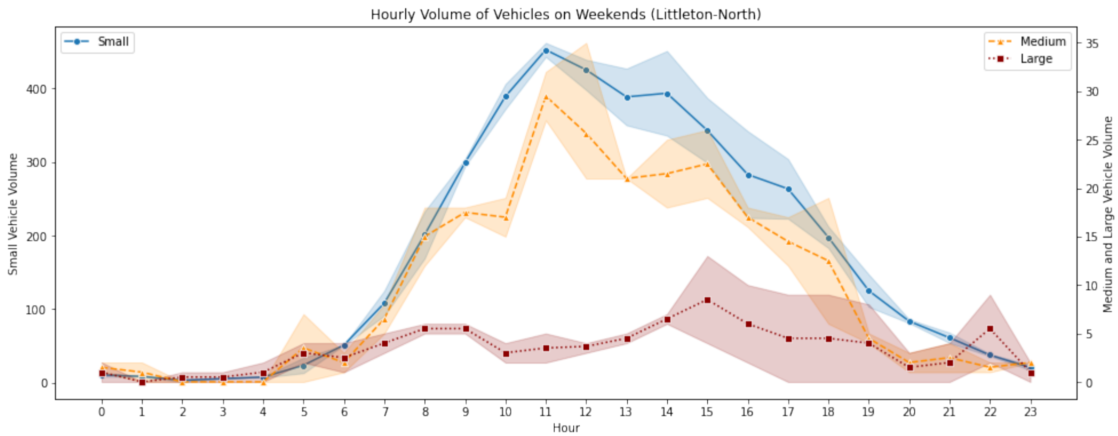


Figure 9-83. Average hourly volume of vehicles on weekends at Littleton North

Littleton South

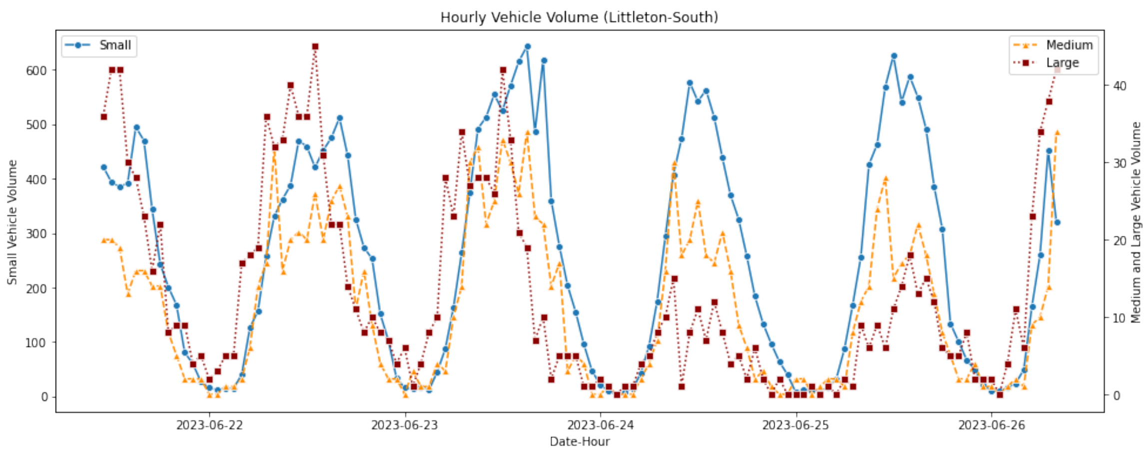


Figure 9-84. Hourly vehicle volume at Littleton South for all days.

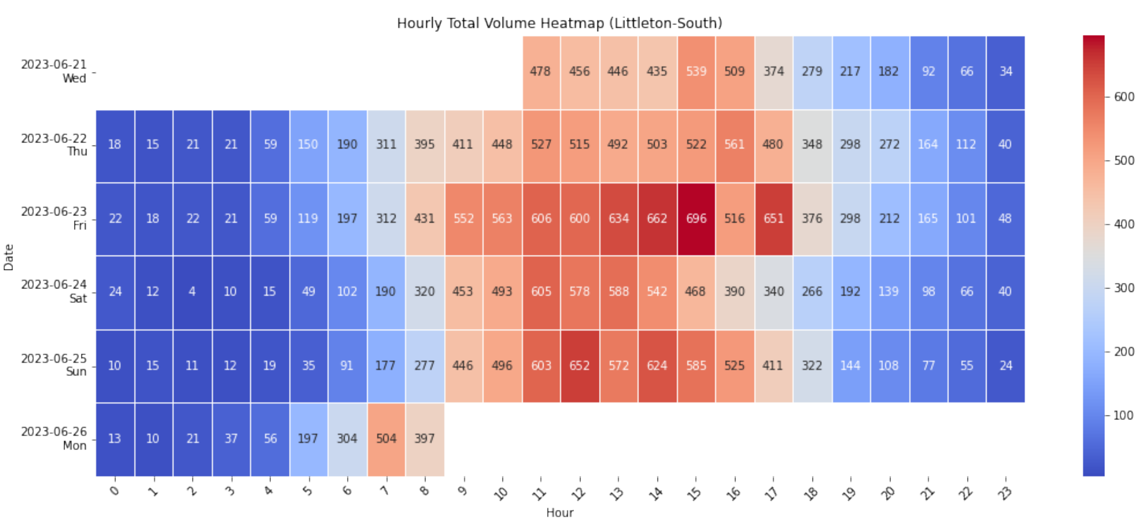


Figure 9-85. Heatmap illustrating hourly vehicle volume at Littleton South.

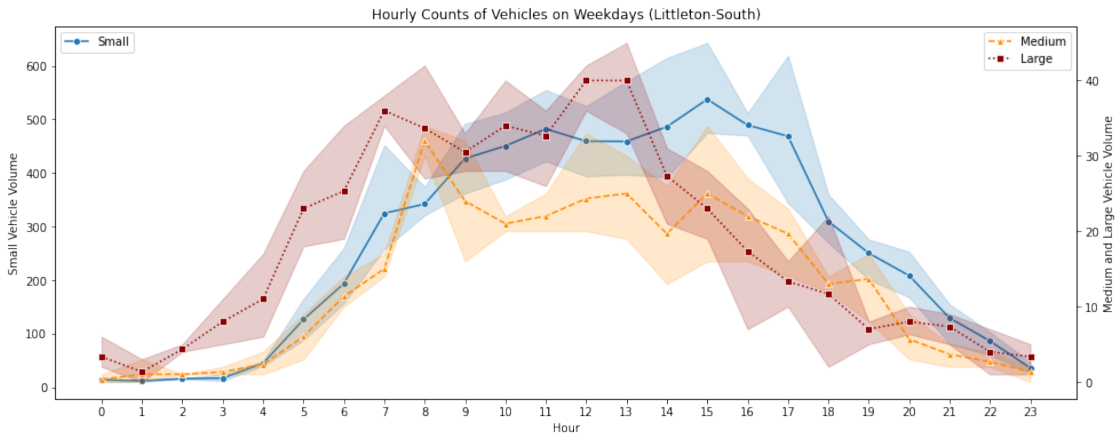


Figure 9-86. Average hourly volume of vehicles on weekdays at Littleton South.

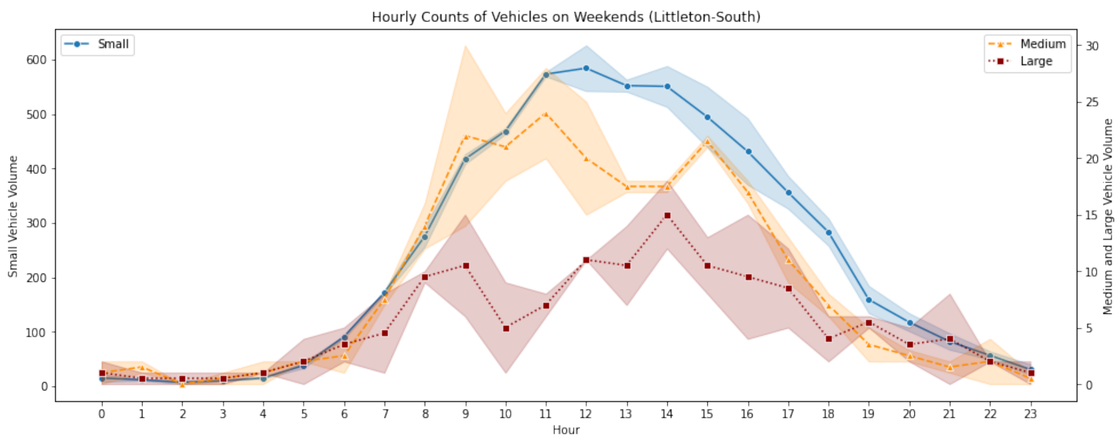


Figure 9-87. Average hourly volume of vehicles on weekends at Littleton South.

Tilton North

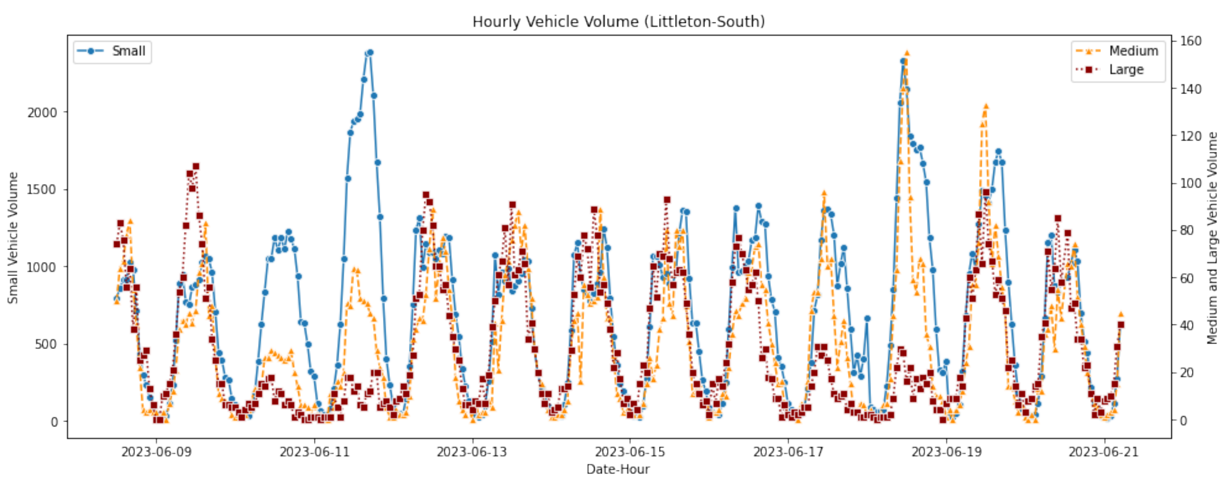


Figure 9-88. Hourly vehicle volume at Tilton North for all days.

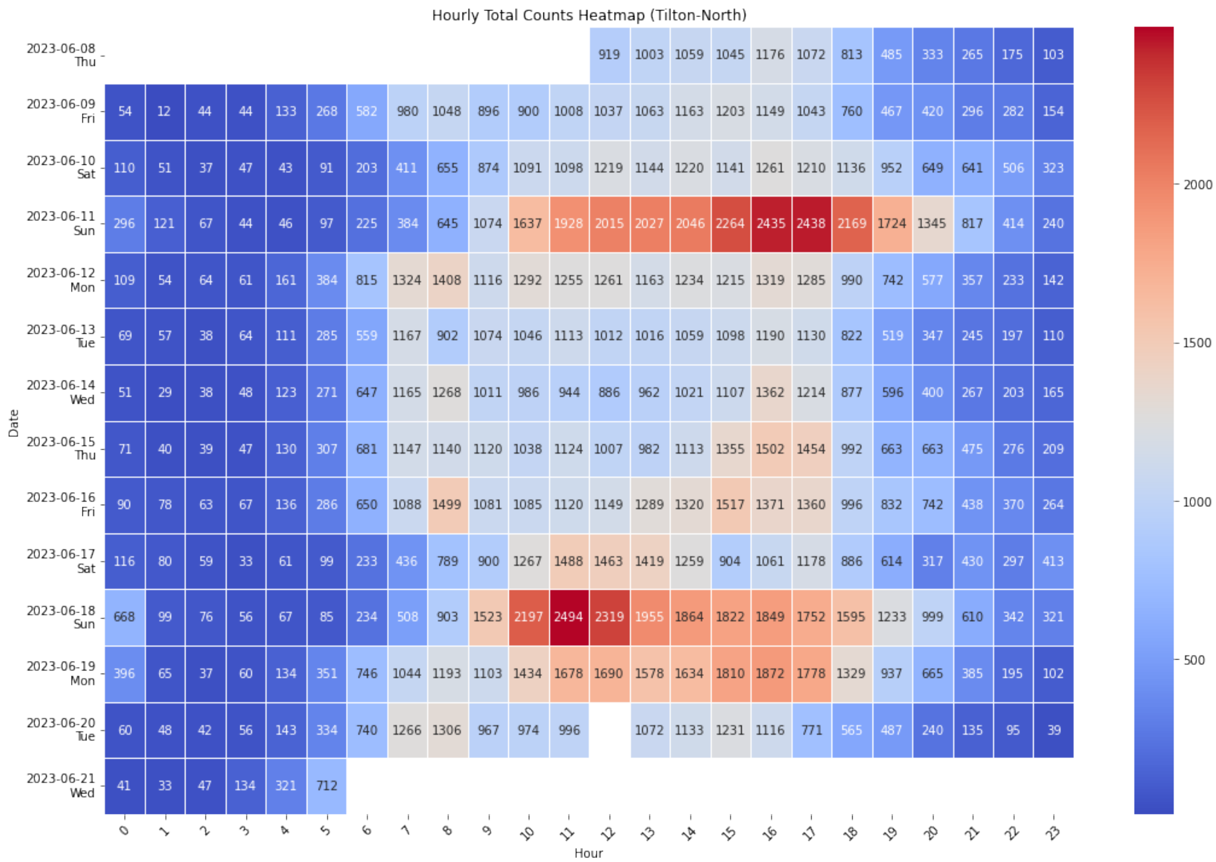


Figure 9-89. Heatmap illustrating hourly vehicle volume at Tilton North.

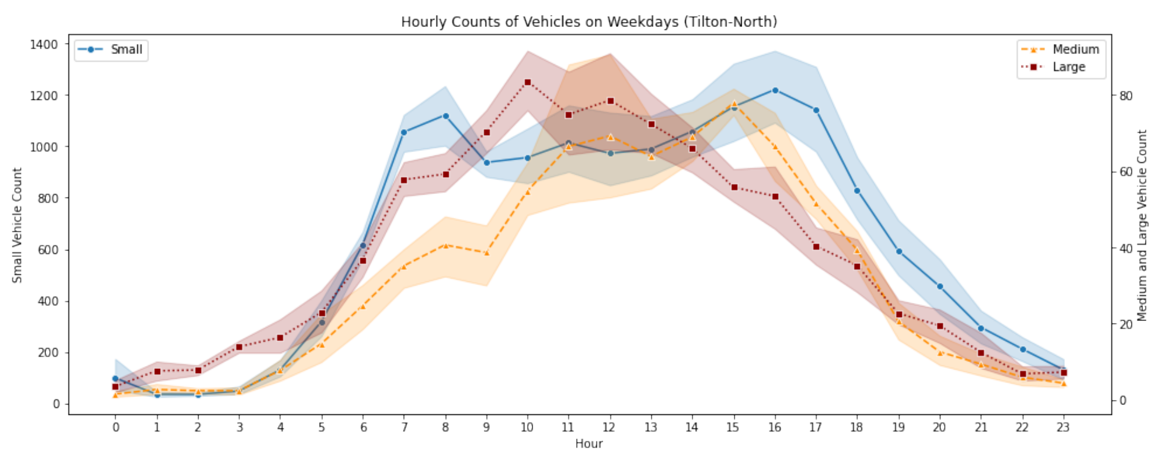


Figure 9-90. Average hourly volume of vehicles on weekdays at Tilton North.

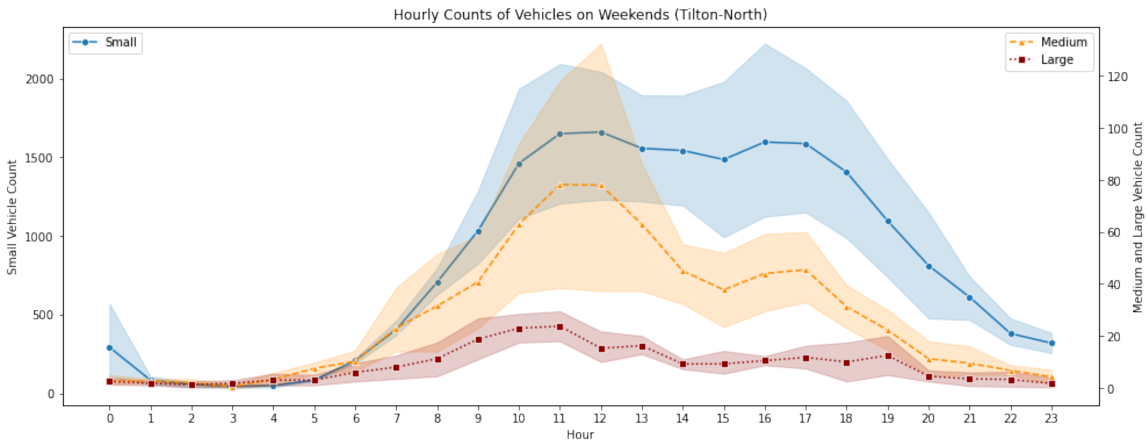


Figure 9-91. Average hourly volume of vehicles on weekends at Tilton North.

Tilton South

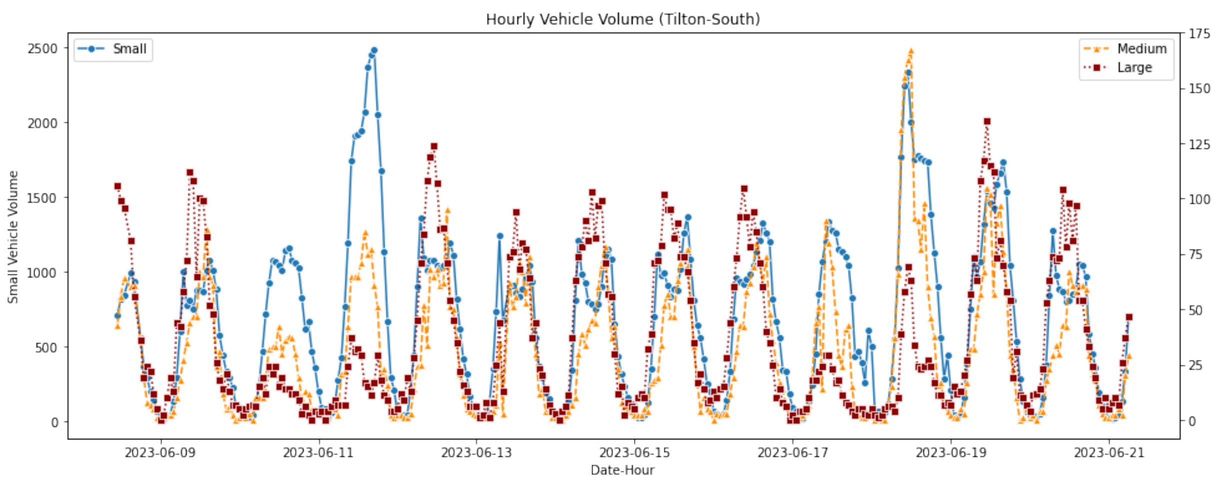


Figure 9-92. Hourly vehicle volume at Tilton South for all days.

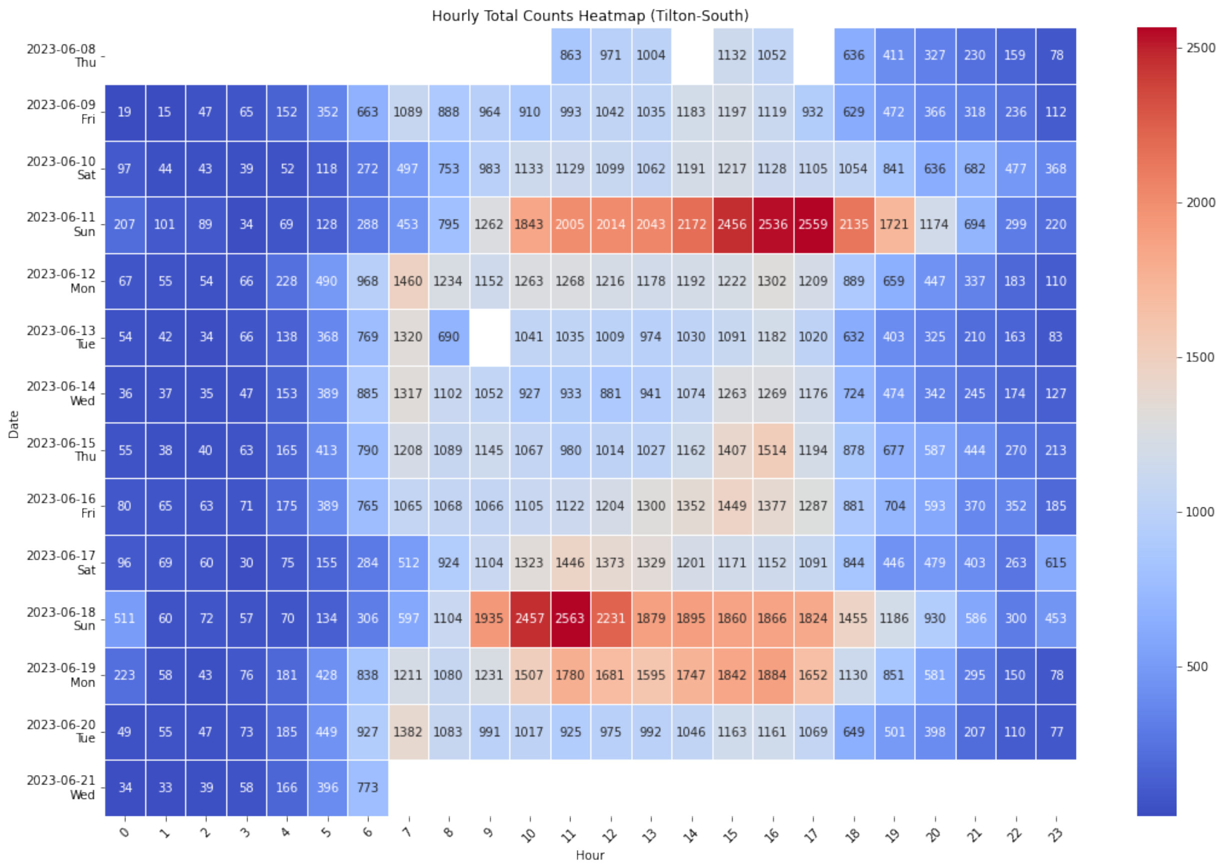


Figure 9-93. Heatmap illustrating hourly vehicle volume at Tilton South.

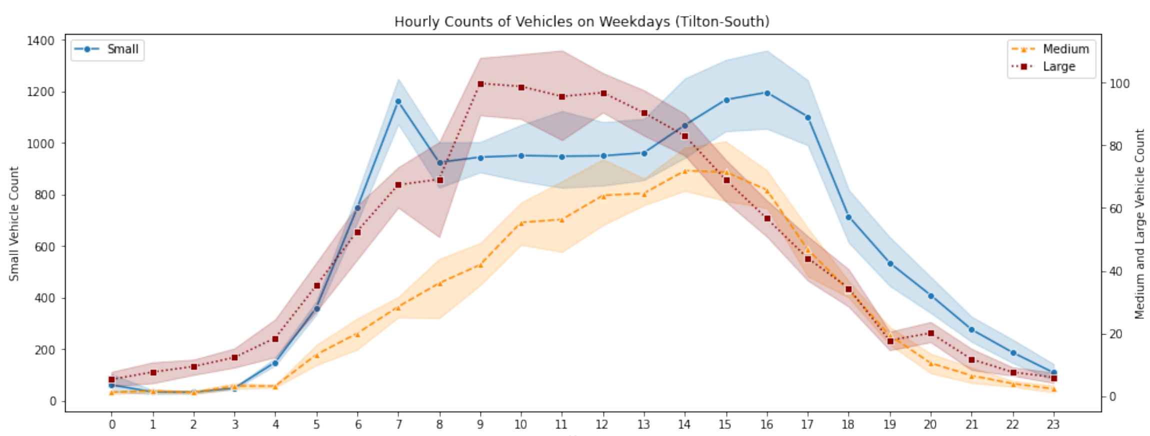


Figure 9-94. Average hourly volume of vehicles on weekdays at Tilton South.

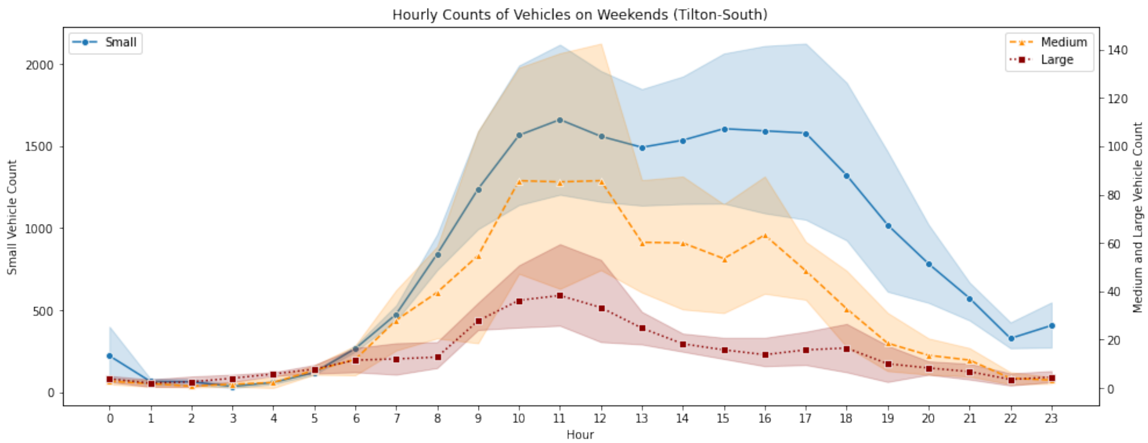


Figure 9-95. Average hourly volume of vehicles on weekends at Tilton South.

Nashua

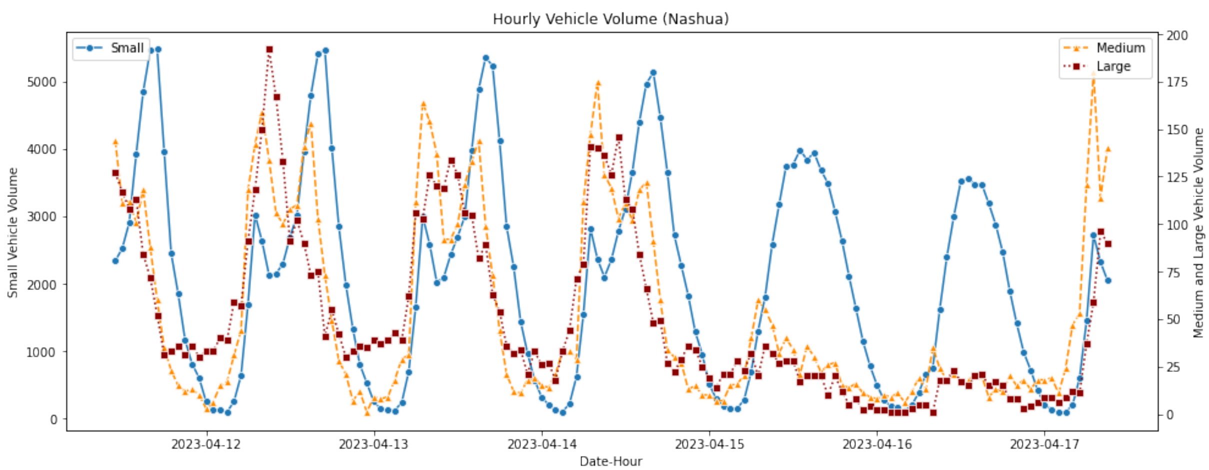


Figure 9-96. Hourly vehicle volume at Nashua for all days.

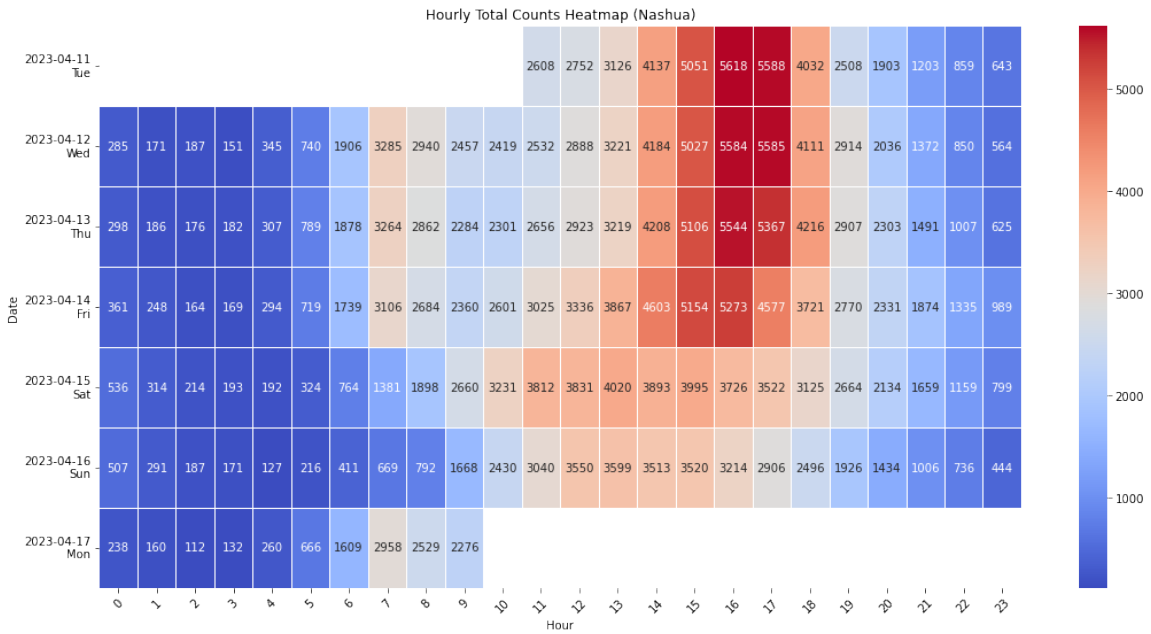


Figure 9-97. Heatmap illustrating hourly vehicle volume at Nashua.

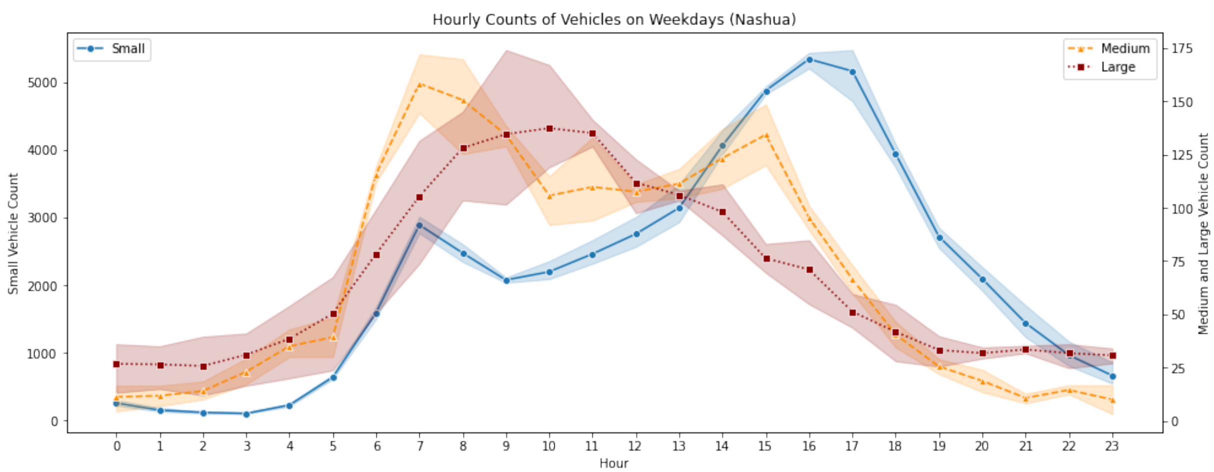


Figure 9-98. Average hourly volume of vehicles on weekdays at Nashua.

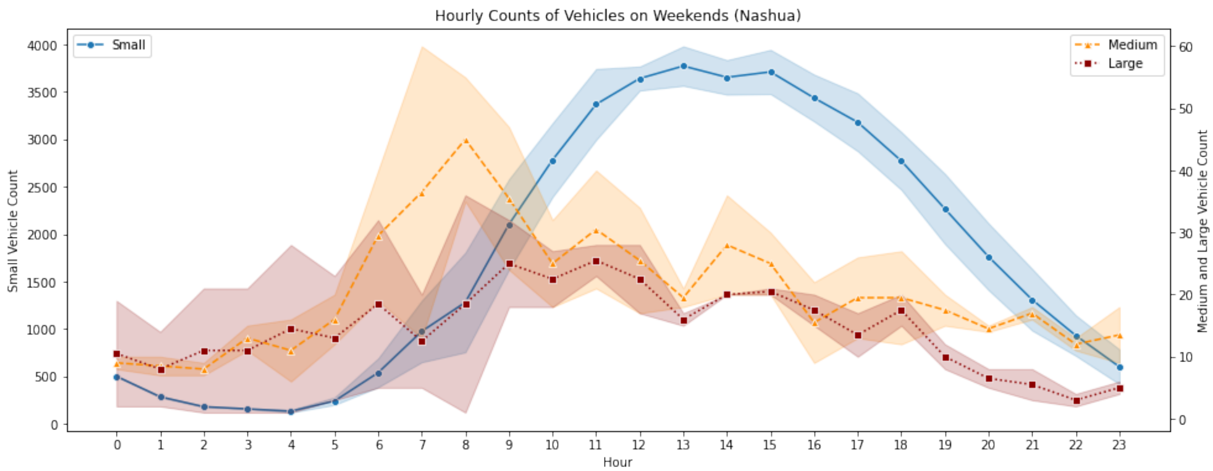


Figure 9-99. Average hourly volume of vehicles on weekends at Nashua.

Campton

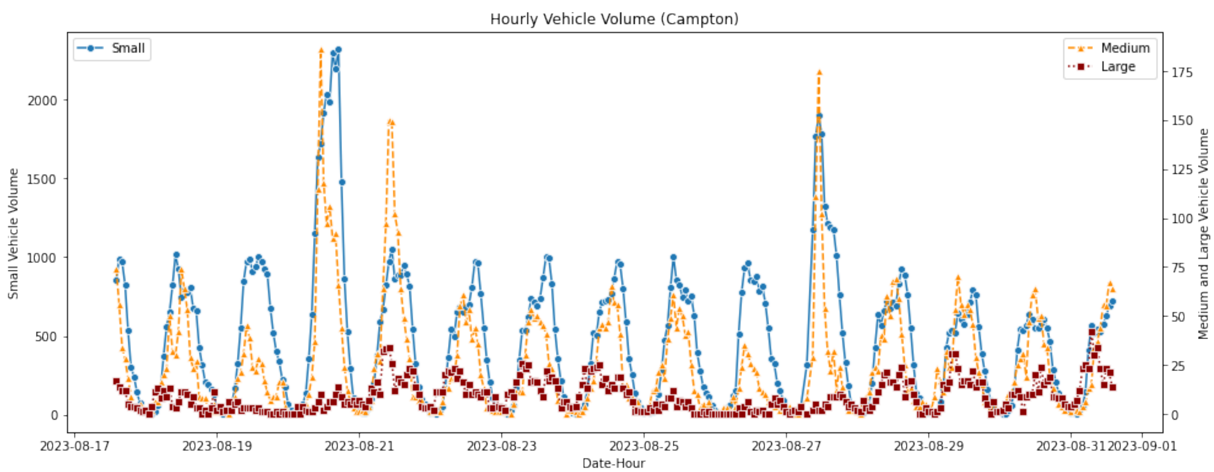


Figure 9-100. Hourly vehicle volume at Campton for all days.

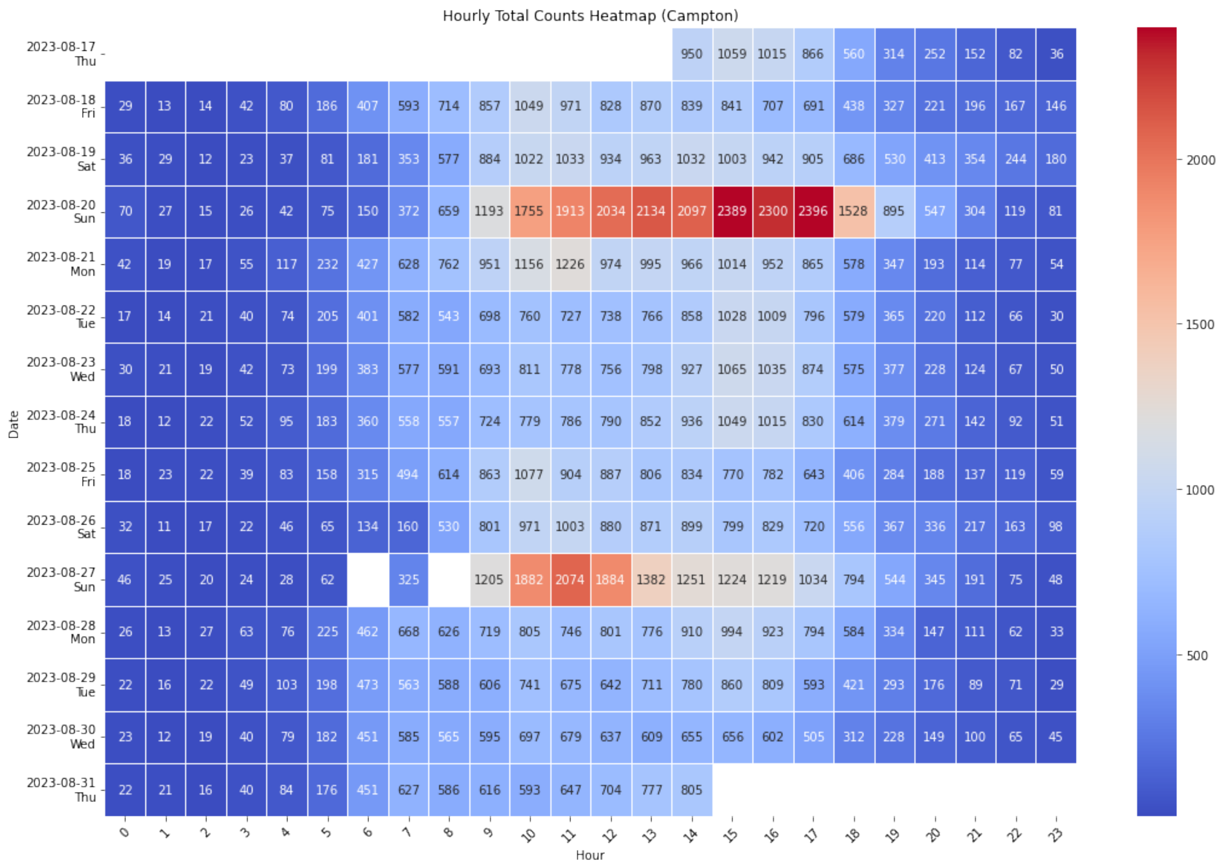


Figure 9-101. Heatmap illustrating hourly vehicle volume at Campton.

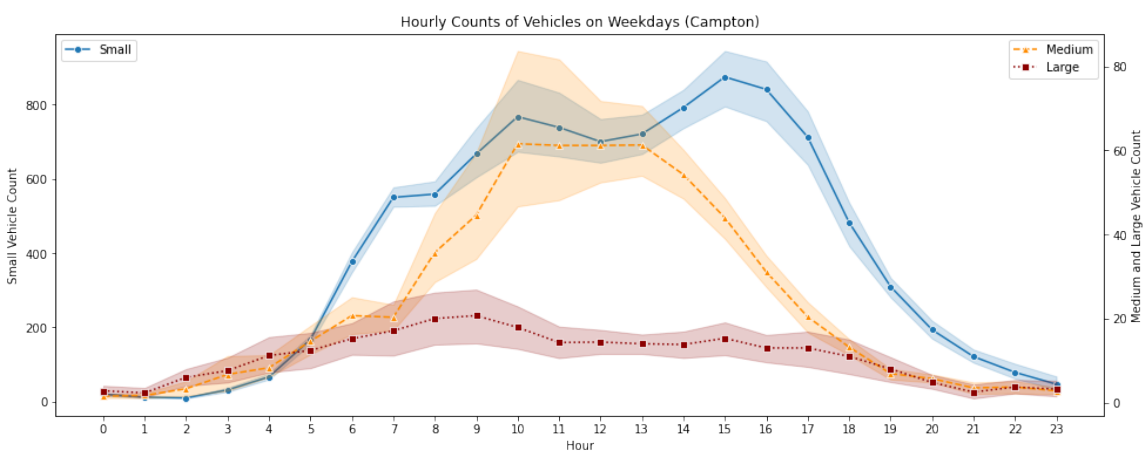


Figure 9-102. Average hourly volume of vehicles on weekdays at Campton.

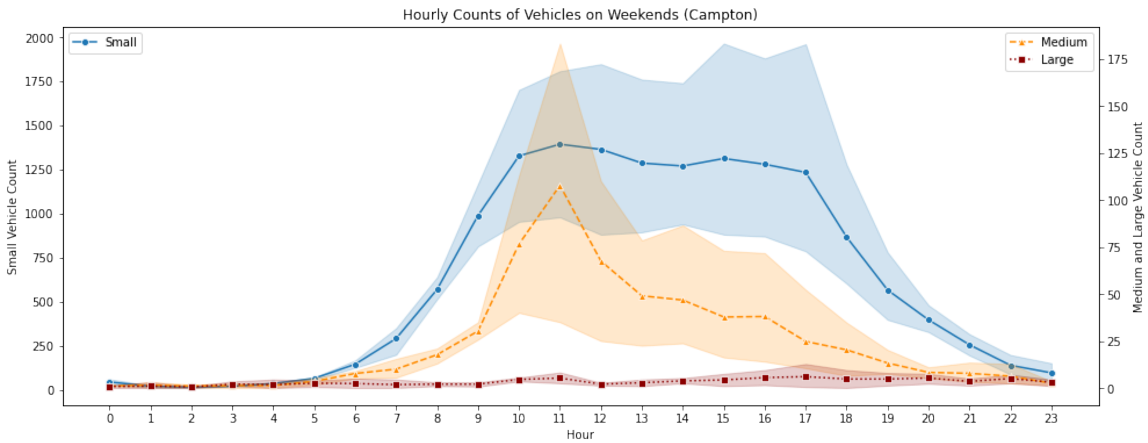


Figure 9-103. Average hourly volume of vehicles on weekends at Campton.

Analysis of merging vehicles at Campton

Two cells in the heatmap are blank because of corruption in two thermal videos. Four out of the 15 line-charts show gaps in the lines. Those gaps are due to factors like excessive camera movement due to high wind, adverse weather conditions, or other issues.

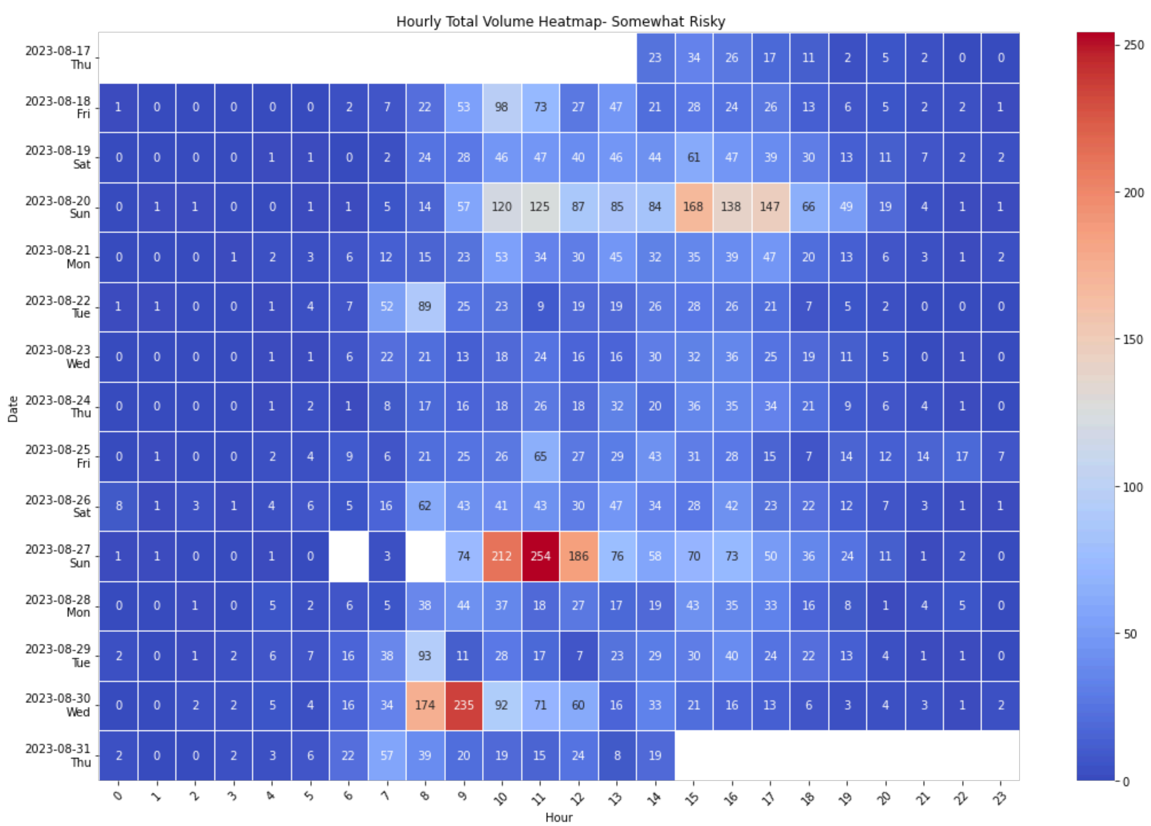


Figure 9-104. Heatmap illustrating hourly vehicle volume merging in “somewhat risky” manner (YELLOW).

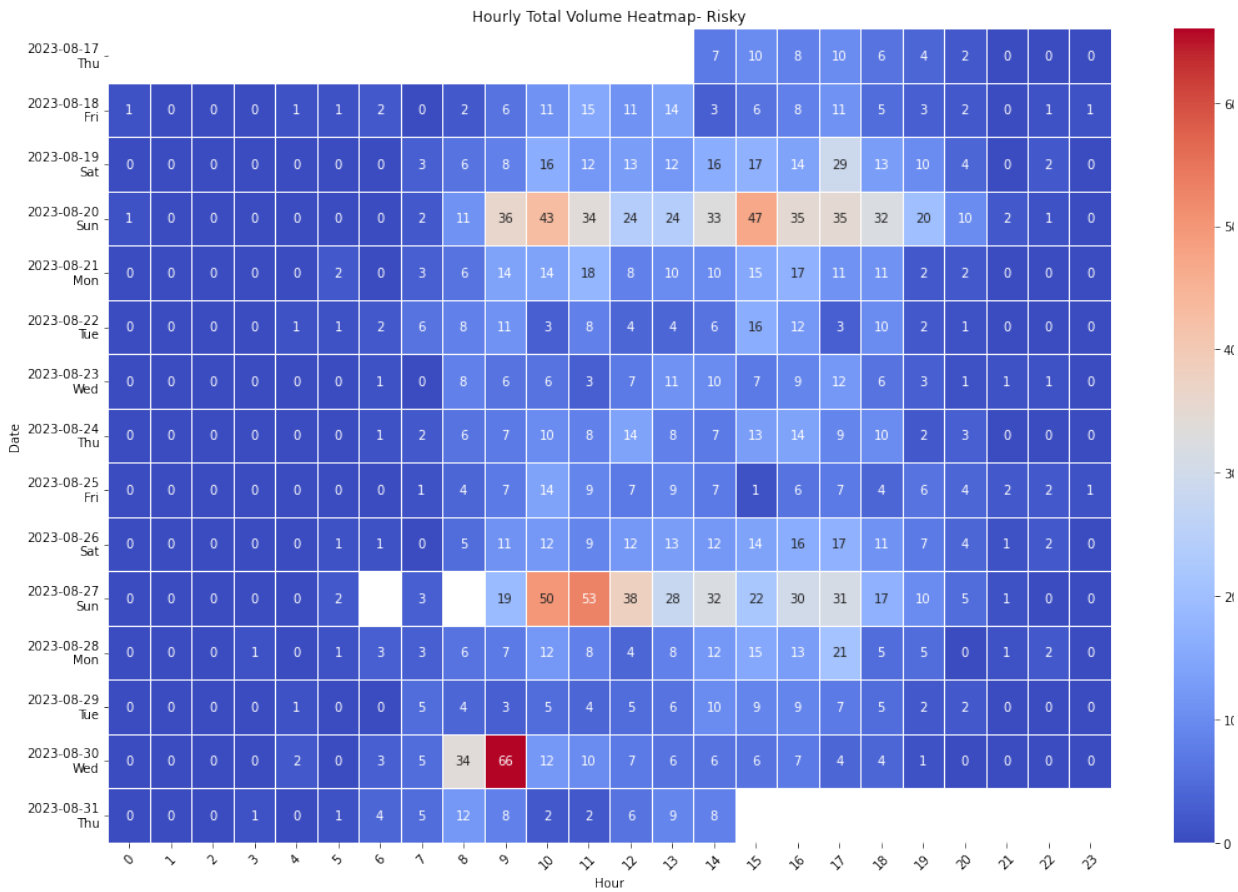


Figure 9-105. Heatmap illustrating hourly vehicle volume merging in “risky” manner (RED1).

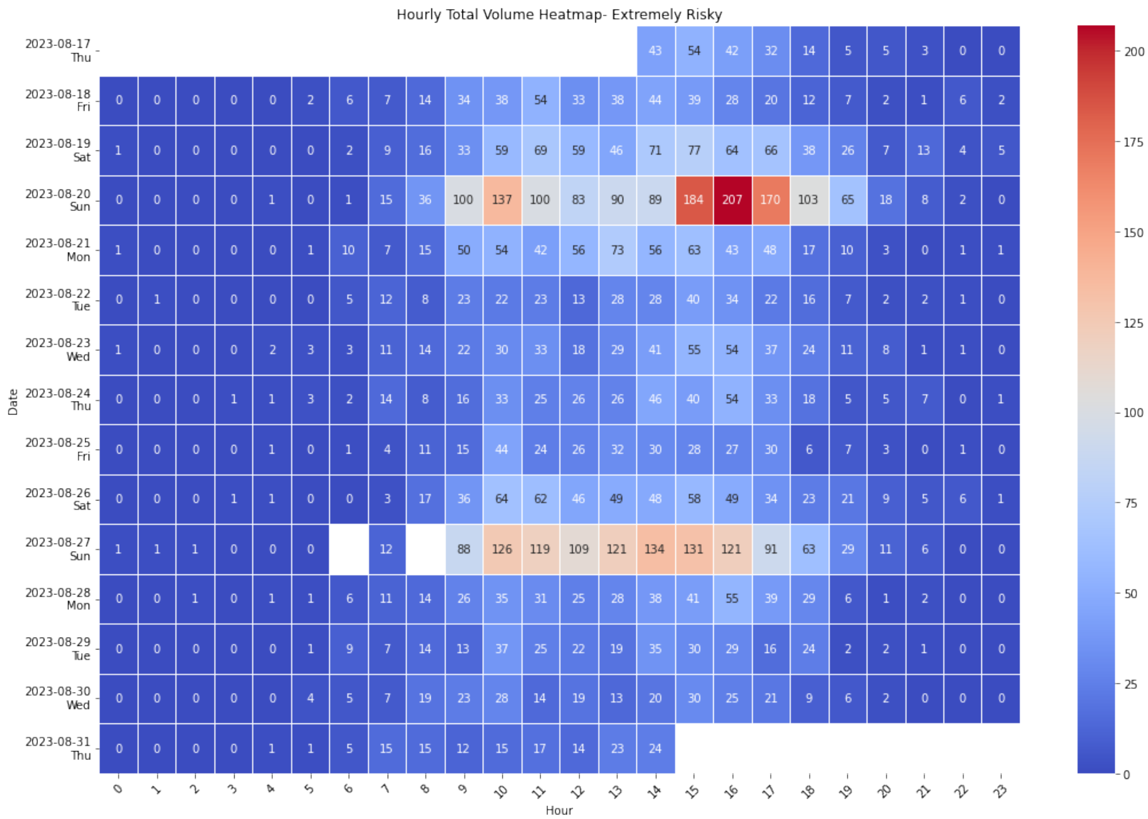


Figure 9-106. Heatmap illustrating hourly vehicle volume merging in "extremely risky" manner (RED2).

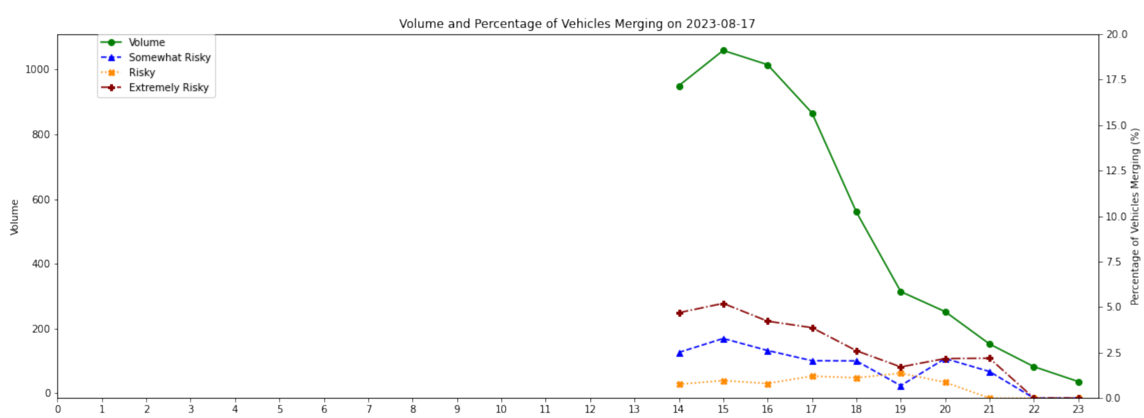


Figure 9-107. Daily traffic volume and vehicle merging pattern on 8-17-2023 in Campton NH

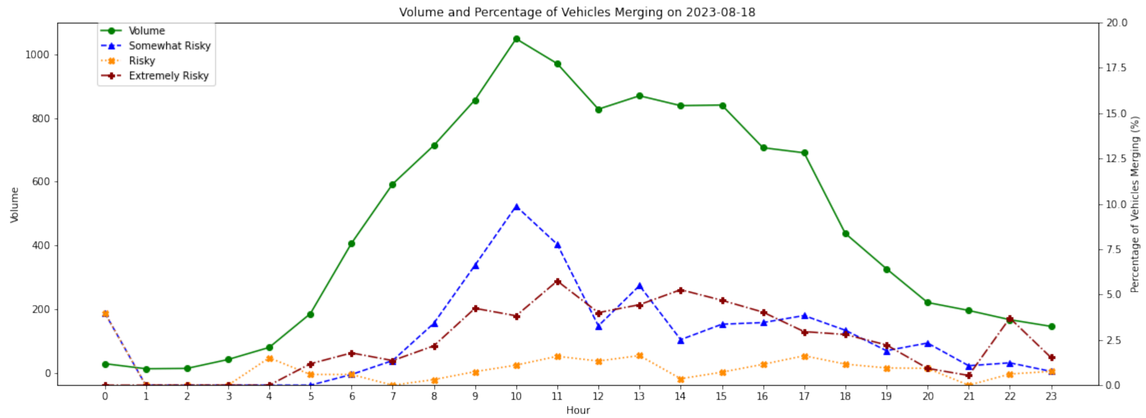


Figure 9-108. Daily traffic volume and vehicle merging pattern on 8-18-2023 in Campton NH

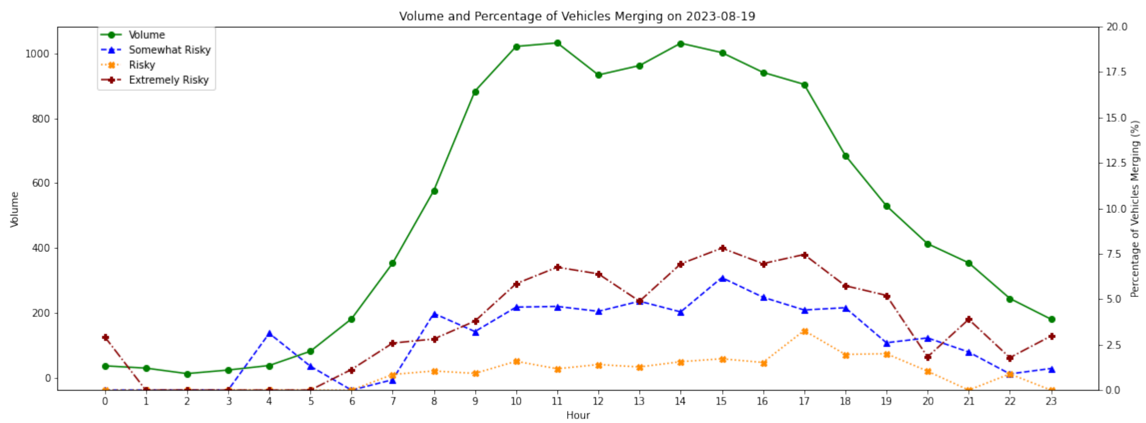


Figure 9-109. Daily traffic volume and vehicle merging pattern on 8-19-2023 in Campton NH

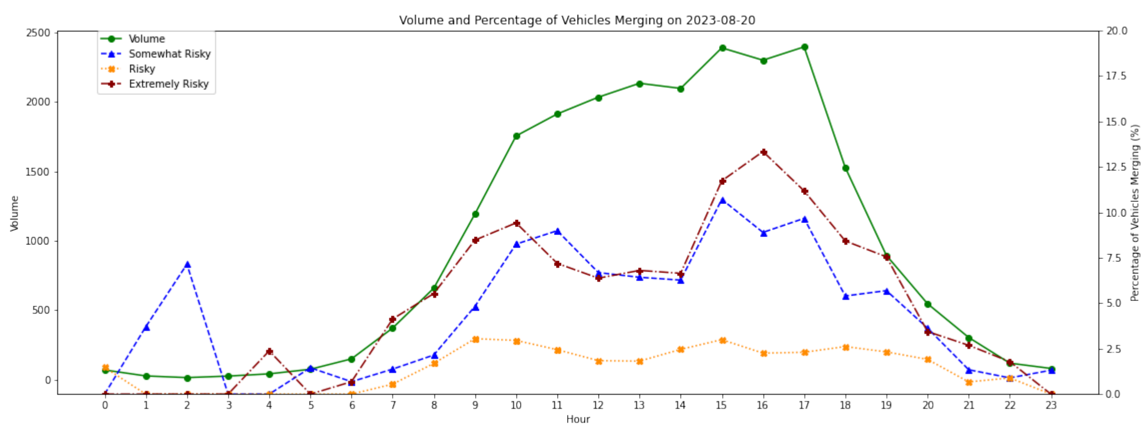


Figure 9-110. Daily traffic volume and vehicle merging pattern on 8-20-2023 in Campton NH

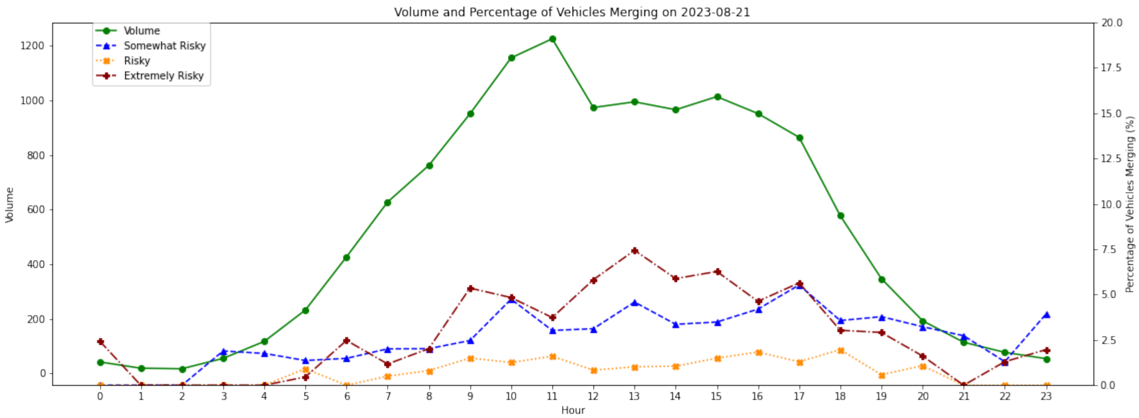


Figure 9-111. Daily traffic volume and vehicle merging pattern on 8-21-2023 in Campton NH

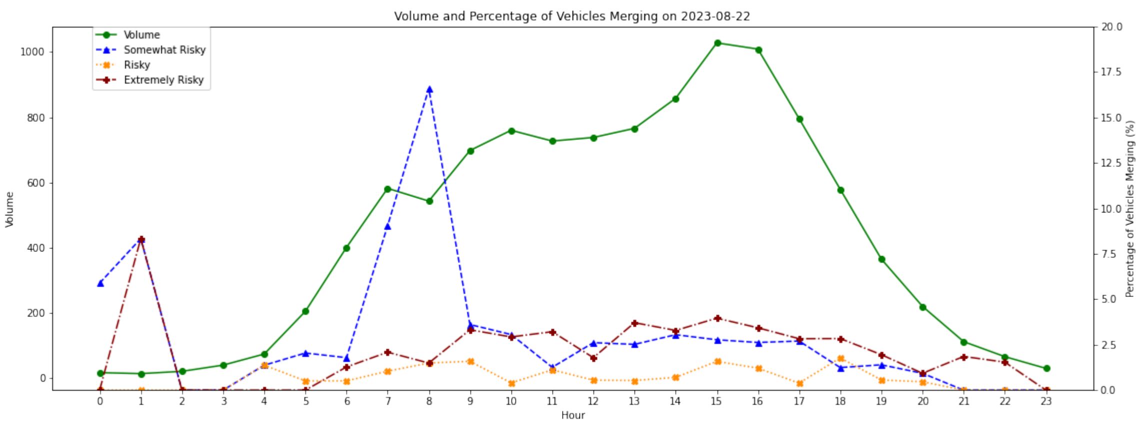


Figure 9-112. Daily traffic volume and vehicle merging pattern on 8-22-2023 in Campton NH

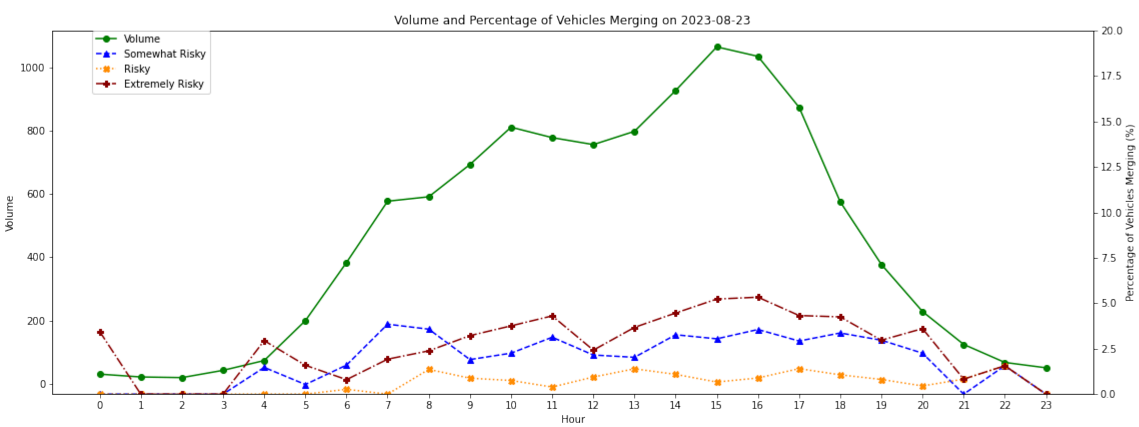


Figure 9-113. Daily traffic volume and vehicle merging pattern on 8-23-2023 in Campton NH

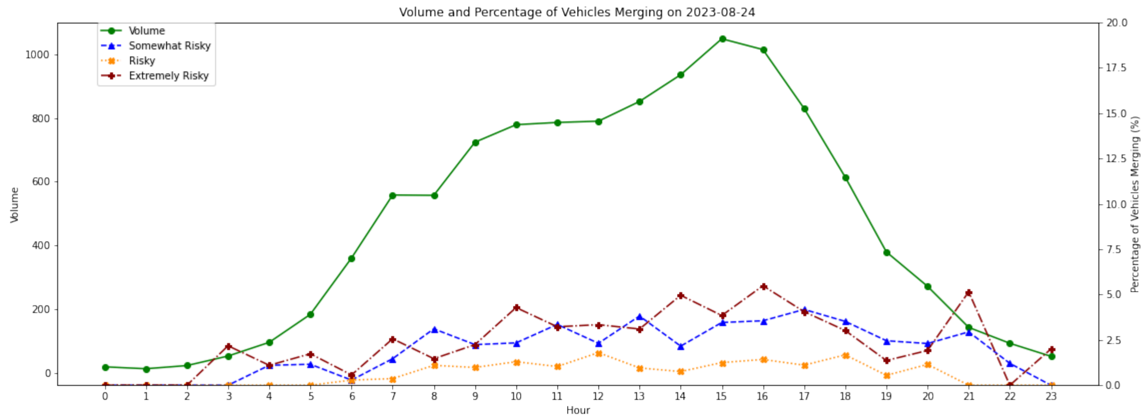


Figure 9-114. Daily traffic volume and vehicle merging pattern on 8-24-2023 in Campton NH

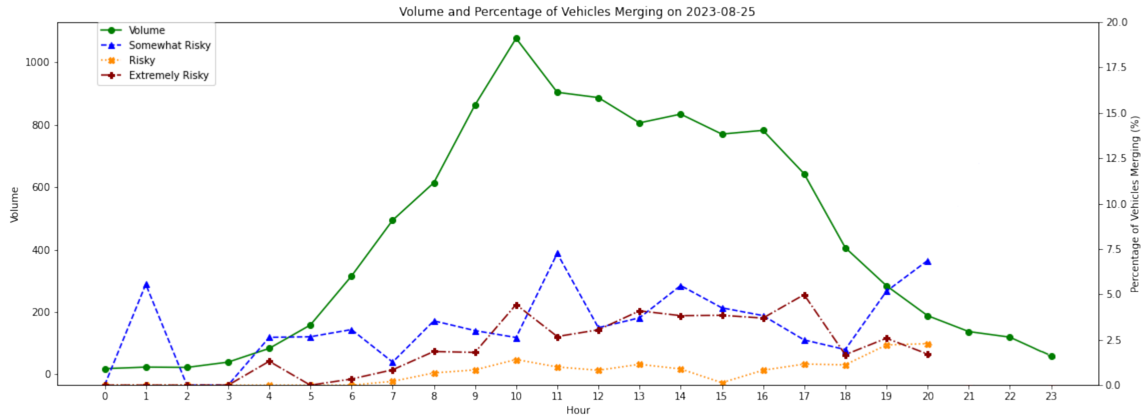


Figure 9-115. Daily traffic volume and vehicle merging pattern on 8-25-2023 in Campton NH

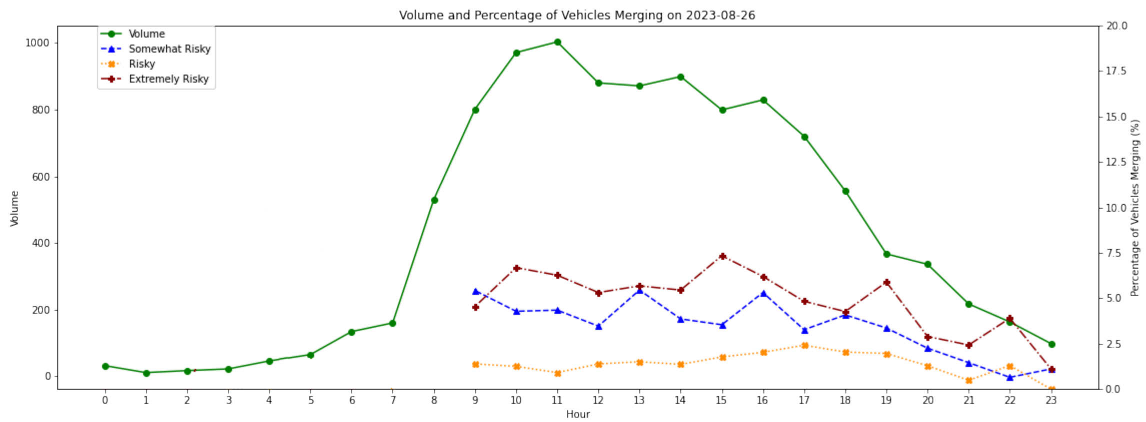


Figure 9-116. Daily traffic volume and vehicle merging pattern on 8-26-2023 in Campton NH

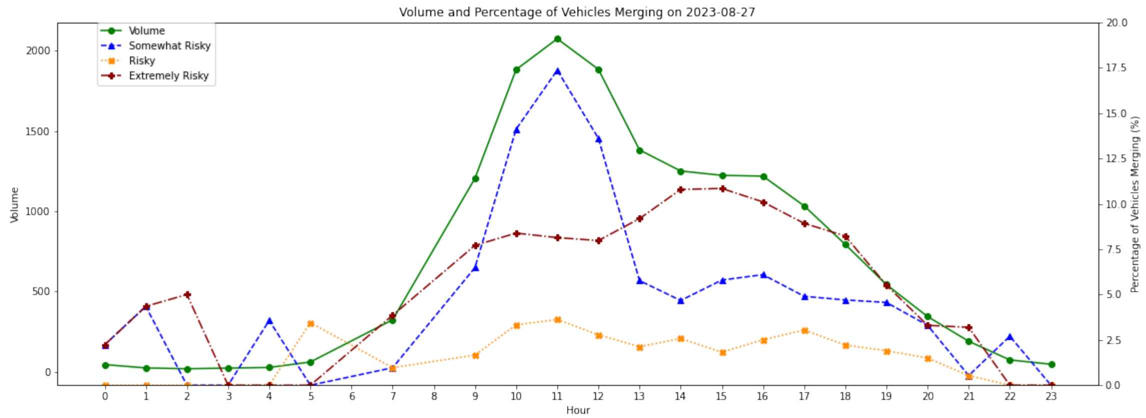


Figure 9-117. Daily traffic volume and vehicle merging pattern on 8-27-2023 in Campton NH

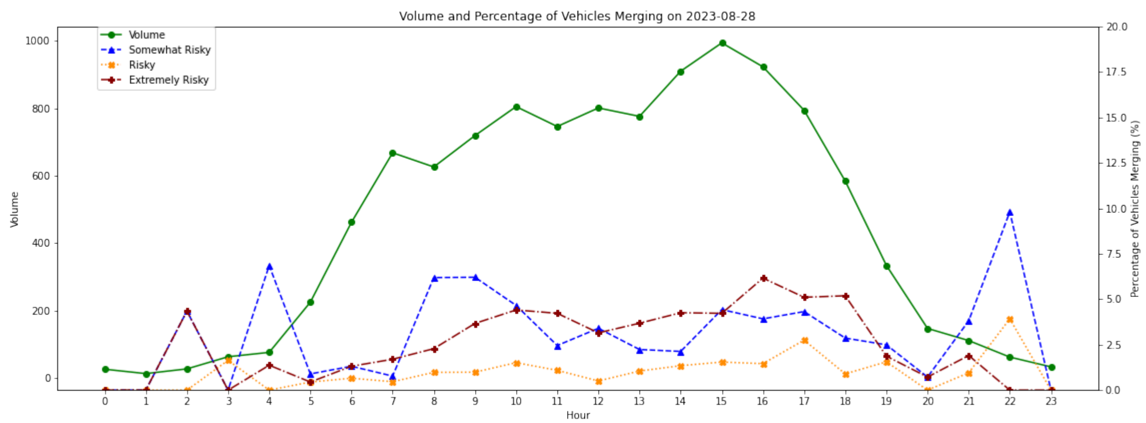


Figure 9-118. Daily traffic volume and vehicle merging pattern on 8-28-2023 in Campton NH

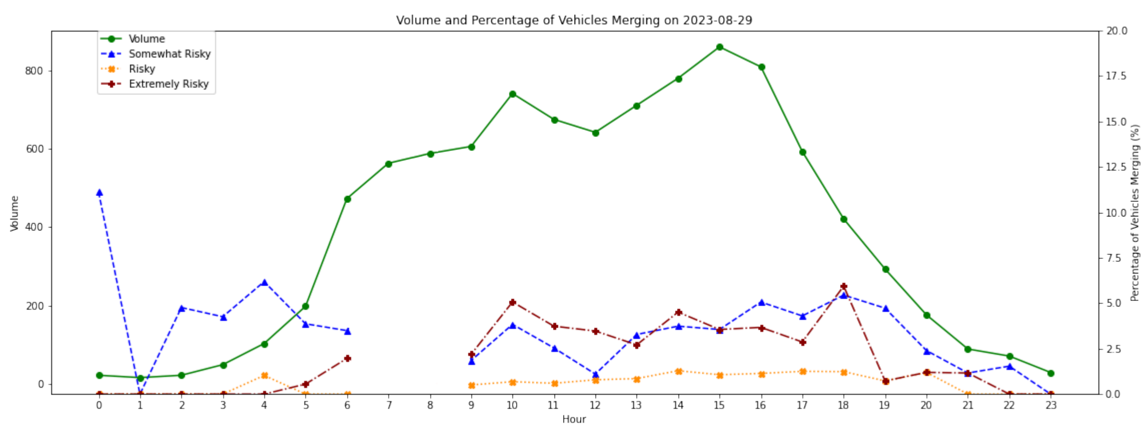


Figure 9-119. Daily traffic volume and vehicle merging pattern on 8-29-2023 in Campton NH

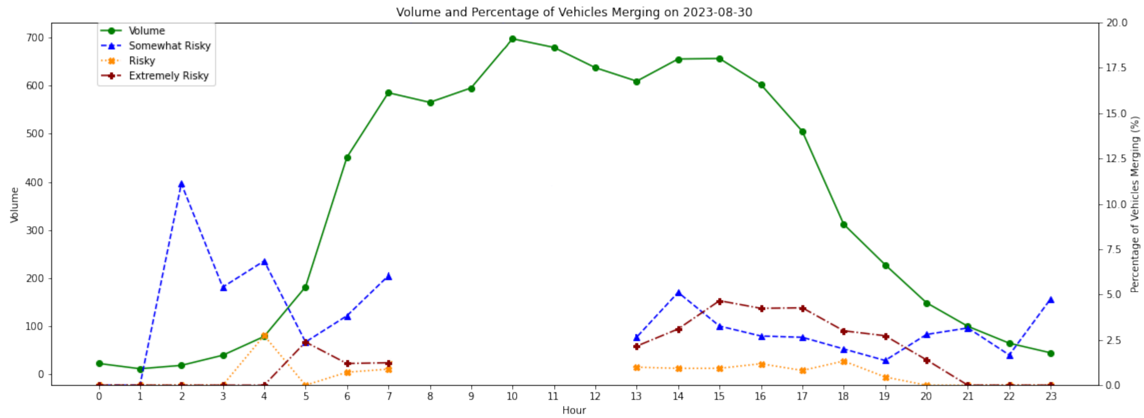


Figure 9-120. Daily traffic volume and vehicle merging pattern on 8-30-2023 in Campton NH

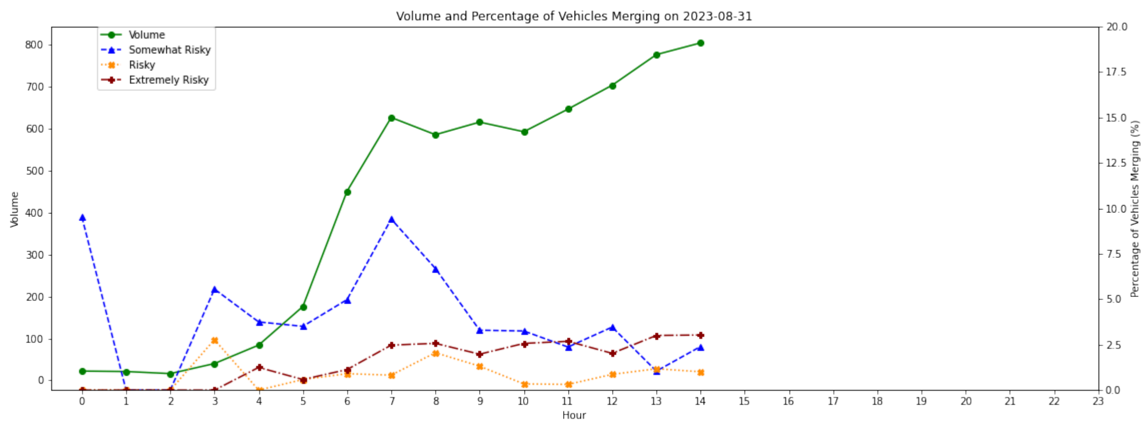


Figure 9-121. Daily traffic volume and vehicle merging pattern on 8-31-2023 in Campton NH

Table 9-1 Hourly volume of vehicles merging into single lane at the Campton location.

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/17/23	14	844	23	7	43	917	21	7	41	2	0	2	0	0	0
8/17/23	15	940	34	10	54	1038	32	9	53	2	1	1	0	0	0
8/17/23	16	918	26	8	42	994	26	8	41	0	0	1	0	0	0
8/17/23	17	768	17	10	32	827	15	10	32	1	0	0	1	0	0
8/17/23	18	507	11	6	14	538	11	5	14	0	1	0	0	0	0
8/17/23	19	281	2	4	5	292	2	4	5	0	0	0	0	0	0
8/17/23	20	218	5	2	5	230	5	2	5	0	0	0	0	0	0
8/17/23	21	132	2	0	3	137	2	0	3	0	0	0	0	0	0
8/17/23	22	75	0	0	0	75	0	0	0	0	0	0	0	0	0
8/17/23	23	33	0	0	0	33	0	0	0	0	0	0	0	0	0
8/18/23	0	23	1	1	0	25	1	1	0	0	0	0	0	0	0
8/18/23	1	13	0	0	0	13	0	0	0	0	0	0	0	0	0
8/18/23	2	13	0	0	0	13	0	0	0	0	0	0	0	0	0
8/18/23	3	37	0	0	0	37	0	0	0	0	0	0	0	0	0
8/18/23	4	65	0	1	0	66	0	1	0	0	0	0	0	0	0
8/18/23	5	166	0	1	2	169	0	1	2	0	0	0	0	0	0
8/18/23	6	327	2	2	6	337	2	2	6	0	0	0	0	0	0
8/18/23	7	501	7	0	7	515	7	0	7	0	0	0	0	0	0
8/18/23	8	607	22	2	14	645	22	2	14	0	0	0	0	0	0
8/18/23	9	709	53	6	34	802	53	6	34	0	0	0	0	0	0
8/18/23	10	846	98	11	38	993	98	11	38	0	0	0	0	0	0
8/18/23	11	798	73	15	54	940	70	15	53	0	0	1	3	0	0
8/18/23	12	754	27	11	33	825	25	11	31	2	0	2	0	0	0
8/18/23	13	757	47	14	38	856	43	12	38	2	2	0	2	0	0
8/18/23	14	768	21	3	44	836	18	3	43	3	0	1	0	0	0
8/18/23	15	760	28	6	39	833	27	5	39	1	1	0	0	0	0
8/18/23	16	635	24	8	28	695	24	7	28	0	1	0	0	0	0
8/18/23	17	620	26	11	20	677	26	10	20	0	1	0	0	0	0
8/18/23	18	399	13	5	12	429	11	4	12	2	1	0	0	0	0
8/18/23	19	299	6	3	7	315	6	3	7	0	0	0	0	0	0
8/18/23	20	206	5	2	2	215	5	2	2	0	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/18/23	21	183	2	0	1	186	2	0	1	0	0	0	0	0	0
8/18/23	22	153	2	1	6	162	2	1	5	0	0	1	0	0	0
8/18/23	23	127	1	1	2	131	1	1	2	0	0	0	0	0	0
8/19/23	0	33	0	0	1	34	0	0	1	0	0	0	0	0	0
8/19/23	1	25	0	0	0	25	0	0	0	0	0	0	0	0	0
8/19/23	2	11	0	0	0	11	0	0	0	0	0	0	0	0	0
8/19/23	3	22	0	0	0	22	0	0	0	0	0	0	0	0	0
8/19/23	4	31	1	0	0	32	1	0	0	0	0	0	0	0	0
8/19/23	5	75	1	0	0	76	1	0	0	0	0	0	0	0	0
8/19/23	6	174	0	0	2	176	0	0	2	0	0	0	0	0	0
8/19/23	7	334	2	3	9	348	2	3	9	0	0	0	0	0	0
8/19/23	8	523	24	6	16	569	23	6	16	1	0	0	0	0	0
8/19/23	9	800	28	8	33	869	27	8	31	1	0	2	0	0	0
8/19/23	10	887	46	16	59	1008	43	16	59	2	0	0	1	0	0
8/19/23	11	893	47	12	69	1021	46	12	67	1	0	2	0	0	0
8/19/23	12	811	40	13	59	923	40	13	58	0	0	1	0	0	0
8/19/23	13	837	46	12	46	941	46	12	46	0	0	0	0	0	0
8/19/23	14	892	44	16	71	1023	43	16	70	0	0	1	1	0	0
8/19/23	15	831	61	17	77	986	59	17	77	2	0	0	0	0	0
8/19/23	16	795	47	14	64	920	45	14	64	2	0	0	0	0	0
8/19/23	17	751	39	29	66	885	39	29	66	0	0	0	0	0	0
8/19/23	18	581	30	13	38	662	29	13	37	1	0	1	0	0	0
8/19/23	19	450	13	10	26	499	13	10	26	0	0	0	0	0	0
8/19/23	20	361	11	4	7	383	11	4	7	0	0	0	0	0	0
8/19/23	21	314	7	0	13	334	7	0	12	0	0	1	0	0	0
8/19/23	22	216	2	2	4	224	1	2	4	0	0	0	1	0	0
8/19/23	23	161	2	0	5	168	2	0	5	0	0	0	0	0	0
8/20/23	0	67	0	1	0	68	0	1	0	0	0	0	0	0	0
8/20/23	1	26	1	0	0	27	1	0	0	0	0	0	0	0	0
8/20/23	2	13	1	0	0	14	1	0	0	0	0	0	0	0	0
8/20/23	3	23	0	0	0	23	0	0	0	0	0	0	0	0	0
8/20/23	4	41	0	0	1	42	0	0	1	0	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/20/23	5	68	1	0	0	69	1	0	0	0	0	0	0	0	0
8/20/23	6	144	1	0	1	146	1	0	1	0	0	0	0	0	0
8/20/23	7	341	5	2	15	363	5	2	15	0	0	0	0	0	0
8/20/23	8	590	14	11	36	651	13	11	36	1	0	0	0	0	0
8/20/23	9	988	57	36	100	1181	55	35	97	2	1	3	0	0	0
8/20/23	10	1155	120	43	137	1455	113	40	130	7	3	7	0	0	0
8/20/23	11	1132	125	34	100	1391	112	30	91	10	3	8	3	1	1
8/20/23	12	1108	87	24	83	1302	79	23	81	5	1	2	3	0	0
8/20/23	13	1124	85	24	90	1323	81	23	84	4	1	4	0	0	2
8/20/23	14	1134	84	33	89	1340	80	33	87	3	0	1	1	0	1
8/20/23	15	1170	168	47	184	1569	157	45	179	10	1	3	1	1	2
8/20/23	16	1172	138	35	207	1552	127	34	201	9	1	6	2	0	0
8/20/23	17	1169	147	35	170	1521	142	35	168	1	0	1	4	0	1
8/20/23	18	1022	66	32	103	1223	62	31	103	2	1	0	2	0	0
8/20/23	19	727	49	20	65	861	48	19	63	1	1	1	0	0	1
8/20/23	20	476	19	10	18	523	19	9	18	0	1	0	0	0	0
8/20/23	21	284	4	2	8	298	4	2	8	0	0	0	0	0	0
8/20/23	22	108	1	1	2	112	1	1	2	0	0	0	0	0	0
8/20/23	23	75	1	0	0	76	1	0	0	0	0	0	0	0	0
8/21/23	0	40	0	0	1	41	0	0	1	0	0	0	0	0	0
8/21/23	1	19	0	0	0	19	0	0	0	0	0	0	0	0	0
8/21/23	2	13	0	0	0	13	0	0	0	0	0	0	0	0	0
8/21/23	3	52	1	0	0	53	1	0	0	0	0	0	0	0	0
8/21/23	4	112	2	0	0	114	2	0	0	0	0	0	0	0	0
8/21/23	5	214	3	2	1	220	3	2	1	0	0	0	0	0	0
8/21/23	6	389	6	0	10	405	6	0	8	0	0	1	0	0	1
8/21/23	7	576	12	3	7	598	11	3	7	1	0	0	0	0	0
8/21/23	8	708	15	6	15	744	15	6	15	0	0	0	0	0	0
8/21/23	9	847	23	14	50	934	21	13	49	1	0	1	1	1	0
8/21/23	10	997	53	14	54	1118	49	13	51	2	1	3	2	0	0
8/21/23	11	1032	34	18	42	1126	30	17	38	3	1	3	1	0	1
8/21/23	12	869	30	8	56	963	29	8	54	1	0	2	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/21/23	13	854	45	10	73	982	44	10	72	1	0	1	0	0	0
8/21/23	14	855	32	10	56	953	32	10	55	0	0	0	0	0	1
8/21/23	15	892	35	15	63	1005	35	15	63	0	0	0	0	0	0
8/21/23	16	831	39	17	43	930	39	17	43	0	0	0	0	0	0
8/21/23	17	745	47	11	48	851	46	11	48	1	0	0	0	0	0
8/21/23	18	513	20	11	17	561	18	11	17	2	0	0	0	0	0
8/21/23	19	319	13	2	10	344	11	2	10	2	0	0	0	0	0
8/21/23	20	175	6	2	3	186	6	2	3	0	0	0	0	0	0
8/21/23	21	107	3	0	0	110	3	0	0	0	0	0	0	0	0
8/21/23	22	75	1	0	1	77	1	0	1	0	0	0	0	0	0
8/21/23	23	48	2	0	1	51	2	0	1	0	0	0	0	0	0
8/22/23	0	16	1	0	0	17	1	0	0	0	0	0	0	0	0
8/22/23	1	10	1	0	1	12	1	0	1	0	0	0	0	0	0
8/22/23	2	19	0	0	0	19	0	0	0	0	0	0	0	0	0
8/22/23	3	40	0	0	0	40	0	0	0	0	0	0	0	0	0
8/22/23	4	71	1	1	0	73	1	1	0	0	0	0	0	0	0
8/22/23	5	191	4	1	0	196	3	1	0	0	0	0	1	0	0
8/22/23	6	378	7	2	5	392	6	2	5	1	0	0	0	0	0
8/22/23	7	505	52	6	12	575	50	6	12	2	0	0	0	0	0
8/22/23	8	432	89	8	8	537	89	8	8	0	0	0	0	0	0
8/22/23	9	634	25	11	23	693	25	11	22	0	0	1	0	0	0
8/22/23	10	704	23	3	22	752	22	3	22	1	0	0	0	0	0
8/22/23	11	676	9	8	23	716	8	8	23	1	0	0	0	0	0
8/22/23	12	692	19	4	13	728	17	4	13	1	0	0	1	0	0
8/22/23	13	704	19	4	28	755	19	4	28	0	0	0	0	0	0
8/22/23	14	793	26	6	28	853	26	6	28	0	0	0	0	0	0
8/22/23	15	927	28	16	40	1011	26	16	40	2	0	0	0	0	0
8/22/23	16	920	26	12	34	992	25	12	34	1	0	0	0	0	0
8/22/23	17	732	21	3	22	778	20	3	22	1	0	0	0	0	0
8/22/23	18	531	7	10	16	564	7	10	16	0	0	0	0	0	0
8/22/23	19	345	5	2	7	359	5	2	7	0	0	0	0	0	0
8/22/23	20	211	2	1	2	216	1	1	2	1	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/22/23	21	106	0	0	2	108	0	0	2	0	0	0	0	0	0
8/22/23	22	64	0	0	1	65	0	0	1	0	0	0	0	0	0
8/22/23	23	30	0	0	0	30	0	0	0	0	0	0	0	0	0
8/23/23	0	28	0	0	1	29	0	0	1	0	0	0	0	0	0
8/23/23	1	20	0	0	0	20	0	0	0	0	0	0	0	0	0
8/23/23	2	18	0	0	0	18	0	0	0	0	0	0	0	0	0
8/23/23	3	41	0	0	0	41	0	0	0	0	0	0	0	0	0
8/23/23	4	65	1	0	2	68	1	0	2	0	0	0	0	0	0
8/23/23	5	185	1	0	3	189	1	0	2	0	0	0	0	0	1
8/23/23	6	366	6	1	3	376	6	1	3	0	0	0	0	0	0
8/23/23	7	539	22	0	11	572	21	0	11	1	0	0	0	0	0
8/23/23	8	545	21	8	14	588	21	7	14	0	1	0	0	0	0
8/23/23	9	643	13	6	22	684	13	6	21	0	0	0	0	0	1
8/23/23	10	745	18	6	30	799	17	6	30	1	0	0	0	0	0
8/23/23	11	707	24	3	33	767	23	3	32	1	0	0	0	0	1
8/23/23	12	703	16	7	18	744	16	6	17	0	1	1	0	0	0
8/23/23	13	734	16	11	29	790	16	11	29	0	0	0	0	0	0
8/23/23	14	838	30	10	41	919	30	10	41	0	0	0	0	0	0
8/23/23	15	958	32	7	55	1052	32	7	54	0	0	1	0	0	0
8/23/23	16	914	36	9	54	1013	36	9	53	0	0	0	0	0	1
8/23/23	17	782	25	12	37	856	25	12	36	0	0	0	0	0	1
8/23/23	18	516	19	6	24	565	19	6	24	0	0	0	0	0	0
8/23/23	19	346	11	3	11	371	10	3	10	1	0	1	0	0	0
8/23/23	20	208	5	1	8	222	5	1	8	0	0	0	0	0	0
8/23/23	21	119	0	1	1	121	0	1	1	0	0	0	0	0	0
8/23/23	22	61	1	1	1	64	1	1	1	0	0	0	0	0	0
8/23/23	23	46	0	0	0	46	0	0	0	0	0	0	0	0	0
8/24/23	0	17	0	0	0	17	0	0	0	0	0	0	0	0	0
8/24/23	1	12	0	0	0	12	0	0	0	0	0	0	0	0	0
8/24/23	2	22	0	0	0	22	0	0	0	0	0	0	0	0	0
8/24/23	3	45	0	0	1	46	0	0	1	0	0	0	0	0	0
8/24/23	4	89	1	0	1	91	1	0	1	0	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/24/23	5	168	2	0	3	173	2	0	3	0	0	0	0	0	0
8/24/23	6	349	1	1	2	353	1	1	2	0	0	0	0	0	0
8/24/23	7	523	8	2	14	547	8	2	13	0	0	1	0	0	0
8/24/23	8	516	17	6	8	547	16	6	8	1	0	0	0	0	0
8/24/23	9	676	16	7	16	715	16	6	14	0	1	2	0	0	0
8/24/23	10	710	18	10	33	771	18	9	33	0	1	0	0	0	0
8/24/23	11	717	26	8	25	776	25	8	24	1	0	0	0	0	1
8/24/23	12	721	18	14	26	779	17	12	25	1	2	1	0	0	0
8/24/23	13	773	32	8	26	839	32	7	26	0	1	0	0	0	0
8/24/23	14	855	20	7	46	928	20	7	45	0	0	1	0	0	0
8/24/23	15	952	36	13	40	1041	36	12	39	0	1	0	0	0	1
8/24/23	16	882	35	14	54	985	34	14	53	1	0	1	0	0	0
8/24/23	17	738	34	9	33	814	34	9	33	0	0	0	0	0	0
8/24/23	18	548	21	10	18	597	20	10	18	1	0	0	0	0	0
8/24/23	19	351	9	2	5	367	9	2	5	0	0	0	0	0	0
8/24/23	20	247	6	3	5	261	5	3	4	0	0	0	1	0	1
8/24/23	21	125	4	0	7	136	4	0	7	0	0	0	0	0	0
8/24/23	22	82	1	0	0	83	1	0	0	0	0	0	0	0	0
8/24/23	23	49	0	0	1	50	0	0	1	0	0	0	0	0	0
8/25/23	0	17	0	0	0	17	0	0	0	0	0	0	0	0	0
8/25/23	1	17	1	0	0	18	1	0	0	0	0	0	0	0	0
8/25/23	2	17	0	0	0	17	0	0	0	0	0	0	0	0	0
8/25/23	3	37	0	0	0	37	0	0	0	0	0	0	0	0	0
8/25/23	4	73	2	0	1	76	2	0	1	0	0	0	0	0	0
8/25/23	5	146	4	0	0	150	4	0	0	0	0	0	0	0	0
8/25/23	6	284	9	0	1	294	9	0	1	0	0	0	0	0	0
8/25/23	7	461	6	1	4	472	6	1	4	0	0	0	0	0	0
8/25/23	8	557	21	4	11	593	21	4	11	0	0	0	0	0	0
8/25/23	9	785	25	7	15	832	24	6	15	1	0	0	0	1	0
8/25/23	10	907	26	14	44	991	23	13	43	3	0	1	0	1	0
8/25/23	11	798	65	9	24	896	65	9	24	0	0	0	0	0	0
8/25/23	12	793	27	7	26	853	26	6	26	1	1	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/25/23	13	712	29	9	32	782	29	8	29	0	1	2	0	0	1
8/25/23	14	704	43	7	30	784	43	7	30	0	0	0	0	0	0
8/25/23	15	668	31	1	28	728	27	1	28	4	0	0	0	0	0
8/25/23	16	670	28	6	27	731	28	5	26	0	1	1	0	0	0
8/25/23	17	550	15	7	30	602	15	7	30	0	0	0	0	0	0
8/25/23	18	340	7	4	6	357	6	4	6	1	0	0	0	0	0
8/25/23	19	244	14	6	7	271	14	6	7	0	0	0	0	0	0
8/25/23	20	156	12	4	3	175	12	4	3	0	0	0	0	0	0
8/25/23	21	110	14	2	0	126	14	2	0	0	0	0	0	0	0
8/25/23	22	85	17	2	1	105	17	2	1	0	0	0	0	0	0
8/25/23	23	48	7	1	0	56	7	1	0	0	0	0	0	0	0
8/26/23	0	23	8	0	0	31	8	0	0	0	0	0	0	0	0
8/26/23	1	7	1	0	0	8	1	0	0	0	0	0	0	0	0
8/26/23	2	11	3	0	0	14	3	0	0	0	0	0	0	0	0
8/26/23	3	17	1	0	1	19	1	0	1	0	0	0	0	0	0
8/26/23	4	34	4	0	1	39	4	0	1	0	0	0	0	0	0
8/26/23	5	57	6	1	0	64	6	1	0	0	0	0	0	0	0
8/26/23	6	122	5	1	0	128	5	1	0	0	0	0	0	0	0
8/26/23	7	122	16	0	3	141	16	0	3	0	0	0	0	0	0
8/26/23	8	435	62	5	17	519	62	5	17	0	0	0	0	0	0
8/26/23	9	703	43	11	36	793	43	11	34	0	0	1	0	0	1
8/26/23	10	841	41	12	64	958	40	12	62	1	0	2	0	0	0
8/26/23	11	876	43	9	62	990	42	9	62	1	0	0	0	0	0
8/26/23	12	778	30	12	46	866	29	12	46	1	0	0	0	0	0
8/26/23	13	754	47	13	49	863	47	13	49	0	0	0	0	0	0
8/26/23	14	787	34	12	48	881	33	12	48	1	0	0	0	0	0
8/26/23	15	691	28	14	58	791	28	13	58	0	1	0	0	0	0
8/26/23	16	684	42	16	49	791	41	16	49	1	0	0	0	0	0
8/26/23	17	630	23	17	34	704	23	17	34	0	0	0	0	0	0
8/26/23	18	483	22	11	23	539	22	11	23	0	0	0	0	0	0
8/26/23	19	317	12	7	21	357	11	7	21	1	0	0	0	0	0
8/26/23	20	291	7	4	9	311	7	4	9	0	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/26/23	21	197	3	1	5	206	3	1	5	0	0	0	0	0	0
8/26/23	22	145	1	2	6	154	1	2	6	0	0	0	0	0	0
8/26/23	23	88	1	0	1	90	1	0	1	0	0	0	0	0	0
8/27/23	0	43	1	0	1	45	1	0	1	0	0	0	0	0	0
8/27/23	1	21	1	0	1	23	1	0	1	0	0	0	0	0	0
8/27/23	2	19	0	0	1	20	0	0	1	0	0	0	0	0	0
8/27/23	3	21	0	0	0	21	0	0	0	0	0	0	0	0	0
8/27/23	4	27	1	0	0	28	1	0	0	0	0	0	0	0	0
8/27/23	5	56	0	2	0	58	0	2	0	0	0	0	0	0	0
8/27/23	7	293	3	3	12	311	3	3	12	0	0	0	0	0	0
8/27/23	9	961	74	19	88	1142	72	19	84	2	0	4	0	0	0
8/27/23	10	1115	212	50	126	1503	192	42	112	19	5	14	1	3	0
8/27/23	11	1037	254	53	119	1463	225	47	104	27	5	15	2	1	0
8/27/23	12	1034	186	38	109	1367	175	35	99	10	2	9	1	1	1
8/27/23	13	1093	76	28	121	1318	75	27	118	0	1	3	1	0	0
8/27/23	14	1018	58	32	134	1242	58	32	134	0	0	0	0	0	0
8/27/23	15	985	70	22	131	1208	69	22	129	1	0	2	0	0	0
8/27/23	16	974	73	30	121	1198	73	30	120	0	0	1	0	0	0
8/27/23	17	848	50	31	91	1020	48	29	90	1	2	1	1	0	0
8/27/23	18	650	36	17	63	766	33	17	62	3	0	1	0	0	0
8/27/23	19	463	24	10	29	526	23	10	29	1	0	0	0	0	0
8/27/23	20	306	11	5	11	333	11	5	11	0	0	0	0	0	0
8/27/23	21	180	1	1	6	188	1	0	6	0	1	0	0	0	0
8/27/23	22	72	2	0	0	74	2	0	0	0	0	0	0	0	0
8/27/23	23	47	0	0	0	47	0	0	0	0	0	0	0	0	0
8/28/23	0	25	0	0	0	25	0	0	0	0	0	0	0	0	0
8/28/23	1	13	0	0	0	13	0	0	0	0	0	0	0	0	0
8/28/23	2	21	1	0	1	23	1	0	1	0	0	0	0	0	0
8/28/23	3	60	0	1	0	61	0	1	0	0	0	0	0	0	0
8/28/23	4	67	5	0	1	73	5	0	0	0	0	0	0	0	1
8/28/23	5	220	2	1	1	224	2	0	1	0	1	0	0	0	0
8/28/23	6	442	6	3	6	457	6	3	6	0	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/28/23	7	628	5	3	11	647	5	3	11	0	0	0	0	0	0
8/28/23	8	556	38	6	14	614	38	6	13	0	0	1	0	0	0
8/28/23	9	632	44	7	26	709	43	7	25	0	0	1	1	0	0
8/28/23	10	710	37	12	35	794	36	12	35	1	0	0	0	0	0
8/28/23	11	677	18	8	31	734	18	7	31	0	1	0	0	0	0
8/28/23	12	738	27	4	25	794	26	4	24	0	0	1	1	0	0
8/28/23	13	708	17	8	28	761	17	8	28	0	0	0	0	0	0
8/28/23	14	824	19	12	38	893	18	11	36	1	1	2	0	0	0
8/28/23	15	872	43	15	41	971	42	14	40	1	0	1	0	1	0
8/28/23	16	790	35	13	55	893	34	13	55	1	0	0	0	0	0
8/28/23	17	671	33	21	39	764	32	21	38	1	0	1	0	0	0
8/28/23	18	509	16	5	29	559	14	5	28	2	0	1	0	0	0
8/28/23	19	302	8	5	6	321	8	5	5	0	0	1	0	0	0
8/28/23	20	137	1	0	1	139	1	0	1	0	0	0	0	0	0
8/28/23	21	98	4	1	2	105	4	1	2	0	0	0	0	0	0
8/28/23	22	44	5	2	0	51	5	2	0	0	0	0	0	0	0
8/28/23	23	27	0	0	0	27	0	0	0	0	0	0	0	0	0
8/29/23	0	16	2	0	0	18	2	0	0	0	0	0	0	0	0
8/29/23	1	12	0	0	0	12	0	0	0	0	0	0	0	0	0
8/29/23	2	20	1	0	0	21	1	0	0	0	0	0	0	0	0
8/29/23	3	45	2	0	0	47	2	0	0	0	0	0	0	0	0
8/29/23	4	90	6	1	0	97	6	1	0	0	0	0	0	0	0
8/29/23	5	173	7	0	1	181	7	0	1	0	0	0	0	0	0
8/29/23	6	433	16	0	9	458	16	0	9	0	0	0	0	0	0
8/29/23	7	506	38	5	7	556	38	5	7	0	0	0	0	0	0
8/29/23	8	471	93	4	14	582	90	4	14	2	0	0	1	0	0
8/29/23	9	567	11	3	13	594	10	3	13	1	0	0	0	0	0
8/29/23	10	660	28	5	37	730	26	5	37	1	0	0	1	0	0
8/29/23	11	623	17	4	25	669	17	4	23	0	0	2	0	0	0
8/29/23	12	600	7	5	22	634	7	5	22	0	0	0	0	0	0
8/29/23	13	653	23	6	19	701	21	6	19	2	0	0	0	0	0
8/29/23	14	702	29	10	35	776	28	10	34	1	0	1	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/29/23	15	775	30	9	30	844	27	9	30	2	0	0	1	0	0
8/29/23	16	712	40	9	29	790	40	9	29	0	0	0	0	0	0
8/29/23	17	511	24	7	16	558	23	7	16	1	0	0	0	0	0
8/29/23	18	353	22	5	24	404	22	5	24	0	0	0	0	0	0
8/29/23	19	258	13	2	2	275	13	2	2	0	0	0	0	0	0
8/29/23	20	159	4	2	2	167	4	2	2	0	0	0	0	0	0
8/29/23	21	84	1	0	1	86	1	0	1	0	0	0	0	0	0
8/29/23	22	64	1	0	0	65	0	0	0	0	0	0	1	0	0
8/29/23	23	27	0	0	0	27	0	0	0	0	0	0	0	0	0
8/30/23	0	19	0	0	0	19	0	0	0	0	0	0	0	0	0
8/30/23	1	12	0	0	0	12	0	0	0	0	0	0	0	0	0
8/30/23	2	16	2	0	0	18	1	0	0	0	0	0	1	0	0
8/30/23	3	35	2	0	0	37	2	0	0	0	0	0	0	0	0
8/30/23	4	66	5	2	0	73	5	2	0	0	0	0	0	0	0
8/30/23	5	160	4	0	4	168	4	0	4	0	0	0	0	0	0
8/30/23	6	394	16	3	5	418	16	3	5	0	0	0	0	0	0
8/30/23	7	518	34	5	7	564	34	5	6	0	0	1	0	0	0
8/30/23	8	326	174	34	19	553	173	34	19	1	0	0	0	0	0
8/30/23	9	250	235	66	23	574	230	66	23	4	0	0	1	0	0
8/30/23	10	529	92	12	28	661	85	12	28	7	0	0	0	0	0
8/30/23	11	559	71	10	14	654	66	9	13	5	1	1	0	0	0
8/30/23	12	540	60	7	19	626	56	7	19	4	0	0	0	0	0
8/30/23	13	568	16	6	13	603	16	6	11	0	0	1	0	0	1
8/30/23	14	586	33	6	20	645	32	6	19	1	0	0	0	0	1
8/30/23	15	587	21	6	30	644	20	6	30	0	0	0	1	0	0
8/30/23	16	542	16	7	25	590	16	7	25	0	0	0	0	0	0
8/30/23	17	454	13	4	21	492	13	4	21	0	0	0	0	0	0
8/30/23	18	281	6	4	9	300	6	4	9	0	0	0	0	0	0
8/30/23	19	210	3	1	6	220	3	1	6	0	0	0	0	0	0
8/30/23	20	137	4	0	2	143	4	0	2	0	0	0	0	0	0
8/30/23	21	92	3	0	0	95	3	0	0	0	0	0	0	0	0
8/30/23	22	59	1	0	0	60	1	0	0	0	0	0	0	0	0

Date	Hour	Safe -GREEN	Somewhat Risky - YELLOW	Risky - RED1	Extremely Risky - RED2	Total	Small- Somewhat Risky	Small- Risky	Small- Extremely Risky	Medium- Somewhat Risky	Medium- Risky	Medium- Extremely Risky	Large-Somewhat Risky	Large_ - Risky	Large Extremely Risky
8/30/23	23	40	2	0	0	42	1	0	0	0	0	0	1	0	0
8/31/23	0	19	2	0	0	21	1	0	0	0	0	0	1	0	0
8/31/23	1	20	0	0	0	20	0	0	0	0	0	0	0	0	0
8/31/23	2	16	0	0	0	16	0	0	0	0	0	0	0	0	0
8/31/23	3	33	2	1	0	36	2	1	0	0	0	0	0	0	0
8/31/23	4	76	3	0	1	80	3	0	1	0	0	0	0	0	0
8/31/23	5	163	6	1	1	171	6	1	1	0	0	0	0	0	0
8/31/23	6	411	22	4	5	442	21	4	4	1	0	1	0	0	0
8/31/23	7	529	57	5	15	606	56	5	15	1	0	0	0	0	0
8/31/23	8	517	39	12	15	583	39	12	14	0	0	0	0	0	1
8/31/23	9	567	20	8	12	607	19	8	11	1	0	0	0	0	1
8/31/23	10	548	19	2	15	584	17	2	15	2	0	0	0	0	0
8/31/23	11	599	15	2	17	633	12	2	17	3	0	0	0	0	0
8/31/23	12	648	24	6	14	692	23	6	14	0	0	0	1	0	0
8/31/23	13	725	8	9	23	765	7	8	23	1	1	0	0	0	0
8/31/23	14	741	19	8	24	792	16	8	24	2	0	0	1	0	0