# Integrated Management and Decision Support of Arterial Street Operations

# Final Report

**Florida Department of Transportation (FDOT) Project**

**Contract Number: BED29 TWO 977-11**

Prepared for

Florida Department of Transportation

Prepared by

Florida International University

August 2026

## Disclaimer

The opinions, findings, and conclusions expressed in this publication are those of the authors and not necessarily those of the State of Florida.

METRIC CONVERSION CHART

| **Approximate Conversions to SI* Units** | | | | | **Approximate Conversions from SI Units** | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Symbol | When You Know | Multiply By | To Find | Symbol | Symbol | When You Know | Multiply By | To Find | Symbol |
| **LENGTH** | | | | | **LENGTH** | | | | |
| in | inches | 25.4 | millimeters | mm | mm | millimeters | 0.039 | inches | in |
| ft | feet | 0.305 | meters | m | m | meters | 3.28 | feet | ft |
| yd | yards | 0.914 | meters | m | m | meters | 1.09 | yards | yd |
| mi | miles | 1.61 | kilometers | km | km | kilometers | 0.621 | miles | mi |
| **AREA** | | | | | **AREA** | | | | |
| $in^2$ | square inches | 645.2 | millimeters squared | $mm^2$ | $mm^2$ | millimeters squared | 0.0016 | square inches | $in^2$ |
| $ft^2$ | square feet | 0.093 | meters squared | $m^2$ | $m^2$ | meters squared | 10.764 | square feet | $ft^2$ |
| $yd^2$ | square yards | 0.836 | meters squared | $m^2$ | $m^2$ | meters squared | 1.196 | square yards | $yd^2$ |
| ac | acres | 0.405 | hectares | ha | ha | hectares | 2.47 | acres | ac |
| $mi^2$ | square miles | 2.59 | kilometers squared | $km^2$ | $km^2$ | kilometers squared | 0.386 | square miles | $mi^2$ |
| **VOLUME** | | | | | **VOLUME** | | | | |
| fl oz | fluid ounces | 29.57 | milliliters | ml | ml | milliliters | 0.034 | fluid ounces | fl oz |
| gal | gallons | 3.785 | liters | L | L | liters | 0.264 | gallons | gal |
| $ft^3$ | cubic feet | 0.028 | meters cubed | $m^3$ | $m^3$ | meters cubed | 35.315 | cubic feet | $ft^3$ |
| $yd^3$ | cubic yards | 0.765 | meters cubed | $m^3$ | $m^3$ | meters cubed | 1.308 | cubic yards | $yd^3$ |
| NOTE: Volumes greater than 1000 L shall be shown in $m^3$. | | | | | | | | | |
| **MASS** | | | | | **MASS** | | | | |
| oz | ounces | 28.35 | grams | g | g | grams | 0.035 | ounces | oz |
| lb | pounds | 0.454 | kilograms | kg | kg | kilograms | 2.205 | pounds | lb |
| T | short tons (2000 lb) | 0.907 | megagrams | Mg | Mg | megagrams | 1.102 | short tons (2000 lb) | T |
| **TEMPERATURE (exact)** | | | | | **TEMPERATURE (exact)** | | | | |
| °F | Fahrenheit | (F-32)/1.8 | Celsius | °C | °C | Celsius | 1.8C+32 | Fahrenheit | °F |
| *SI is the symbol for the International System of Measurement | | | | | | | | | |

## TECHNICAL REPORT DOCUMENTATION PAGE

<table>
<tr><td>1. Report No.</td><td>2. Government Accession No.</td><td colspan="3">3. Recipient's Catalog No.</td></tr>
<tr><td colspan="2" rowspan="2">4. Title and Subtitle<br>Integrated Management and Decision Support of Arterial Street Operations</td><td colspan="3">5. Report Date<br>June 2026</td></tr>
<tr><td colspan="3">6. Performing Organization Code</td></tr>
<tr><td colspan="2" rowspan="2">7. Author(s)<br>Mohammed Hadi, Ph.D., PE, Nattakarn Surangsrirout, and Tanchanok Sirikanchittavon</td><td colspan="3">8. Performing Organization Report No.</td></tr>
<tr><td colspan="3">10. Work Unit No.</td></tr>
<tr><td colspan="2">9. Performing Organization Name and Address<br>Florida International University<br>10555 W Flagler St<br>Miami, FL 33174</td><td colspan="3">11. Contract or Grant No.<br>BED29 TWO 977 - 11</td></tr>
<tr><td colspan="2" rowspan="2">12. Sponsoring Agency Name and Address<br>Florida Department of Transportation<br>605 Suwannee Street, MS 30<br>Tallahassee, FL 32399</td><td colspan="3">13. Type of Report and Period Covered<br>Final Report<br>June 2024 to August 2026</td></tr>
<tr><td colspan="3">14. Sponsoring Agency Code</td></tr>
<tr><td colspan="5">15. Supplementary Notes<br>The Project Managers from FDOT is Daniel Smith</td></tr>
<tr><td colspan="5">16. Abstract<br>The goal of this project is to identify and develop artificial intelligence (AI)-enabled strategies that enhance arterial traffic management while addressing the operational needs and priorities of transportation agencies. The project begins by identifying high-value applications of AI and machine learning (ML) through an extensive literature review and engagement with Transportation Systems Management and Operations (TSM&O) practitioners and signal control managers. It then evaluates the performance measures that can be derived from a variety of transportation data sources and assesses their ability to support data-driven traffic management. Building on these findings, the project develops ML models to support both offline planning and real-time operational decision-making and evaluates their performance through representative case studies to assess their effectiveness and applicability for real-world deployment.</td></tr>
<tr><td colspan="2">17. Key Words<br>Artificial Intelligence, machine learning, arterial street, signal timing, predictive analytics</td><td colspan="3">18. Distribution Statement<br>No restrictions.</td></tr>
<tr><td>19. Security Classif. (of this report) Unclassified</td><td>20. Security Classif. (of this page)<br>Unclassified</td><td>21. No. of Pages<br>145</td><td>22. Price</td></tr>
</table>

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

# EXECUTIVE SUMMARY


The increasing availability of high-quality transportation data presents a significant opportunity to transform traffic management and operations on arterial streets. By leveraging advanced data collection, analytics, and artificial intelligence (AI), transportation agencies can implement data-driven traffic management strategies that improve operations during both recurring congestion and non-recurring events, such as traffic incidents, lane closures, adverse weather, and sudden demand surges.

### Goal and Objectives

The goal of this project is to identify and develop AI-enabled arterial traffic management strategies that address operational needs and priorities of transportation agencies. The project uses available data to develop and test machine learning (ML) techniques to support planning and real-time operational decision-making.

The specific objectives are to:

- Identify opportunities for applying AI and ML to arterial traffic management through an extensive literature review and engagement with transportation systems management and operations (TSM&O) managers.
- Evaluate the performance measures that can be derived from multiple data sources and assess their ability to support active and predictive arterial traffic management strategies.
- Develop machine learning models that provide both offline planning capabilities and real-time decision support for arterial operations.
- Evaluate the developed models to validate their effectiveness and readiness for deployment in real-world traffic management applications.

### Literature Review and Stakeholder Inputs

The literature review and stakeholder meetings conducted as part of this study demonstrate that AI has the potential to significantly advance active arterial traffic management by enabling more proactive, predictive, and data-driven operational strategies. Potential machine learning applications cover a range of arterial traffic management functions including identifying variations in recurring traffic patterns, detecting and classifying non-recurring congestion events, prioritizing signal retiming efforts, and supporting a new generation of traffic-responsive and adaptive signal control strategies. Machine learning can also improve incident detection and response, predict traffic diversion during incidents and special events, manage disruptions associated with railroad crossings, detect signal timing deviations and equipment malfunctions, estimate pedestrian demand, and enhance transit, freight, emergency vehicle, railroad, and bridge priority operations. In addition, AI-enabled video analytics can improve traffic monitoring, safety analysis, and performance measurement.

Despite these opportunities, the study identified several significant challenges that must be addressed before AI can be widely deployed in arterial traffic management. The most frequently cited barrier is the availability of high-quality data. Agencies also identified limited staffing, insufficient training, software and licensing costs, and the lack of standardized data formats and interoperable platforms as major obstacles to AI implementation.

**Performance Measures Utilization in Machine Learning**

The application of machine learning to arterial traffic management requires the estimation and utilization of performance measures that support decisions at the strategic, tactical, and operational levels. This study identified mobility, reliability, and safety performance measures that can be used as inputs to, or outputs from, machine learning models. The study also addresses how these measures can be derived by integrating multiple complementary data sources

The identified performance measures support different levels of decision-making. The identified mobility measures range from traditional system-wide metrics that can be used for strategic decisions, such as vehicle miles traveled, vehicle hours of delay, and average speed; to corridor-level tactical indicators including travel time, throughput, delay, queues, and level of service. At the operational level, the measures are of higher resolutions and include detailed signal performance measures such as cycle length, green splits, phase activations, split failures, arrivals on green, queue length, pedestrian activity, and approach delay.

Reliability measures similarly span multiple decision levels. Strategic measures evaluate overall system reliability using metrics such as travel time reliability, unreliable delay, and planning time indexes, while tactical measures focus on corridor-level travel time variability, buffer index, on-time performance.

Safety measures include both traditional crash-based metrics and proactive surrogate safety measures. Strategic and tactical analyses rely on crash frequencies, crash rates, and fatality statistics to assess long-term safety performance, while operational analyses increasingly utilize high-resolution surrogate safety measures—including hard braking events, red-light running, dilemma zone entries, vehicle conflicts, and deceleration profiles—to identify emerging safety risks before crashes occur.

**Freeway Incident Diversion Impact on Arterial Streets**

This study investigated the impacts of freeway incident-related traffic diversion on parallel arterial streets through machine learning. Using a case study of southbound I-95 and the parallel Military Trail corridor in Boca Raton, Florida, the research examined diversion behavior, evaluated the ability to estimate diversion impacts, and assessed the feasibility of deploying a real-time predictive decision support system.

The analysis confirmed that high impact incidents on I-95 generate significant traffic diversion onto Military Trail and other parallel arterials, resulting in substantial increases in arterial congestion. Field observations identified a time lag between the onset of freeway congestion and the resulting deterioration of arterial traffic conditions, providing a valuable opportunity for proactive arterial traffic management. The study also found that existing time-of-day signal timing plans do not adequately respond to diversion events, limiting agencies' ability to accommodate rapidly changing traffic patterns.

To address this limitation, the study developed an XGBoost-based machine learning model capable of predicting arterial delays approximately 15 minutes before congestion develops using only existing FDOT data sources. The model demonstrated strong predictive performance and successfully identified a high proportion of significant diversion-induced delay events while maintaining a relatively low false alarm rate. Analysis of model behavior showed that freeway speed and occupancy measurements provide the strongest indicators of impending diversion.

The study also evaluated the feasibility of implementing the predictive model using live operational data from the FDOT Data Integration and Video Aggregation System (DIVAS) platform. While the machine learning model proved suitable for real-time deployment, the assessment identified several challenges related to data quality and system integration. These included inconsistencies in the provision of detector measurements, incomplete and inaccurate incident records, and limited correspondence among the SunGuide, Regional Integrated Transportation Information System (RITIS), and DIVAS platform data. These issues reduce the reliability of real-time predictions and highlight the importance of maintaining high-quality, integrated transportation data systems.

Overall, the research demonstrates that freeway diversion impacts on arterial streets can be predicted with sufficient lead time to support proactive traffic management, including the implementation of special signal timing plans before congestion develops. The findings illustrate the significant potential of AI and machine learning to improve coordinated freeway and arterial operations while emphasizing that successful deployment depends not only on predictive model performance but also on the availability of reliable, integrated, and well-maintained ITS infrastructure.

**Identification of Traffic Patterns and Associated Features**

This study developed a data-driven framework for identifying and evaluating traffic patterns across different days of the year, as well as the factors contributing to variations in operational performance among these patterns. Such analysis enables the identification of appropriate corrective actions by linking variations in the assessed performance measures to deficiencies in signal timing parameter selections for specific traffic patterns. The developed method was demonstrated using a case study segment located on Yamato Road between Military Trail and Congress Avenue in the City of Boca Raton, Florida.

The developed method uses K-means clustering in combination with principal component analysis (PCA) to successfully capture and characterize day-to-day variations in traffic conditions. In addition, by integrating insights from PCA loadings, XGBoost (a decision tree-based machine learning method), and Spearman correlation analysis, the study further identified that both traffic volume distribution and signal green time allocation are key determinants of corridor performance. In particular, the findings indicate that congestion is not driven solely by demand levels, but is also strongly affected by signal timing inefficiencies, particularly the imbalance of green time across competing movements. Further analysis suggests that a major contributor to the observed inefficiencies was the use of an adaptive signal control strategy that was subsequently disabled during the afternoon peak period, likely due to degraded performance under high-congestion conditions.

## TABLE OF CONTENTS

# 1 INTRODUCTION

This chapter provides an overview of the project, including project background, research objectives, and report organization.

## 1.1 Project Background

A major constraint to advancing traffic signal control strategies and practices has been the lack of the use of data and data analytics that allow for better management of system performance. In recent years, detailed data has started to be available from multiple sources, including advanced sensors, traffic controllers, commercial probes, and connected vehicle data (from third-party vendors). The availability of the data has allowed the derivation of Automated Traffic Signal Performance Measures (ATSPM) that can be used as inputs to active and proactive arterial traffic management. Increasingly, agencies have utilized high-resolution controller data (i.e., signal control timing and detection events), and some of them have explored advanced sensing technologies that have only become available in the market in recent years.

The availability of data creates an opportunity for major advancements in traffic management and operations on arterial streets. Data-enabled arterial management strategies can better support operations during recurrent conditions as well as non-recurrent events, such as lane blockage incidents, adverse weather conditions, or demand surges. There are several methods that can be used to support transportation agency decisions based on data. These methods range from simply reporting and visualizing summary statistics of system performance to advanced machine learning (ML) and artificial intelligence (AI), traffic flow theory and analysis, and simulation techniques. In recent years, there has been an increasing interest in using AI/ML as part of Decision Support Systems (DSS) of Transportation System Management and Operations (TSM&O).

ML can be used for monitoring, categorization, diagnosis, and prediction of system performance, in addition to providing recommendations for the setting and activation of strategies, tactics, and plans. These systems can be classified as supervised, unsupervised, and reinforcement learning (RL). Supervised learning uses training data, which involves feeding paired inputs and outputs to the model. Examples of supervised learning are statistical regressions, neural networks, Support Vector Machines (SVM), decision trees, and tree ensembles. With unsupervised learning, the input data is not associated with outputs. Examples are clustering, visualization, and association rules. RL can observe the conditions and select the best actions for a given situation.

## 1.2 Research Objectives

The goal of this project is to identify arterial traffic management strategies that can be enabled and enhanced by advanced data collection and analytics techniques based on the needs and requirements of transportation agencies. The specific objectives are:

- Identify the potential applications of AI/ML to support arterial traffic management based on an extensive literature review that investigated such applications and meetings with TSM&O and signal control managers
- Assess the performance measures that can be derived based on data from various sources and their ability to support data-enabled active and proactive arterial traffic management strategies
- Develop ML models to support off-line and real-time arterial management
- Evaluate the performance of the developed models

## 1.3 Report Organization

This report is organized as follows. Chapter 1 presents the project background, motivation, and research objectives. Chapters 2 and 3 provide a review of the relevant literature and a summary of stakeholder interviews, which informed the subsequent phases of the research. Chapters 4 and 5 describe the performance measures applicable to AI model development and the transportation data sources used to estimate these measures. Chapters 6 and 7 present the development, evaluation, and key findings of the machine learning models developed in this study for predicting route diversion during freeway incidents and identifying recurring traffic patterns.

# 2 LITERATURE REVIEW

This chapter presents a review of studies conducted to investigate the applications of ML in arterial management. This review was conducted in the beginning of this project to inform the research effort conducted in later tasks of the project including selecting the applications that became the focus of the project and conducting the actual research activities.

The reviewed studies addressed the following arterial management applications:

- Time-of-day control
- Traffic responsive plan selection (TRPS)
- Traffic adaptive control
- Signal control during non-recurrent events
- Malfunction detection and management
- Transit signal priority (TSP) and preemption applications
- AI-based video analytics

Since TSP and preemption applications and AI-based video analytics are not a focus of this research, studies on these two topics are not reviewed in this document. The remaining sections of this document summarize the results of the review of literature on the remaining topics in the above list.

## 2.1 Time-of-Day Control

Signal plan selection based on time of day is currently the dominant signal control strategy, as it is the implemented strategy for the majority of signalized intersections around the United States. The traffic signals are usually re-timed by transportation agencies at intervals (e.g., two to five years), or when receiving complaints from the public. There is a potential for using ML to support the identification of signal timing plans that better fit the traffic patterns at more frequent intervals, as needed. This section summarizes documents that discuss such uses.

***Time-of-day Patterns and Breakpoints***

Several studies have been conducted to identify signal timing patterns and the breakpoints between these patterns for use in the development of signal timing plans and the activation of plans based on time of day. The commonly used approach in identifying traffic patterns is the use K-means clustering. However, Wang et al. (2005) reported that using K-means clustering for breakpoint determination may result in “noisy” allocations of traffic patterns to clusters, and smoothing may be needed to avoid having an unacceptable number of time-of-day breakpoints. In a later study, Wang et al. (2021) reported that the traditional clustering algorithms can divide continuous time intervals into smaller segments than desirable. Thus, they proposed a cyclic weighted k-means method that extends the k-means method with centroid initialization, cluster number selection, and breakpoint adjustment. The study applied the proposed method to data from a real intersection and confirmed the feasibility and effectiveness of the method.

Chen et al. (2019) investigated the use of several classical clustering approaches, including K-means, hierarchical, and Fisher ordinal clustering, for breakpoint identification. They reported that although clustering has been used for determining time-of-day breakpoints, outliers in traffic data may lead to frequent and sometimes drastic changes in the identified patterns and, thus, the developed time-of-day plans. This can result in the recommendation of unrealistic signal timing plan breakpoints.

Rao et al. (2019) stated that many traffic state identification methods use mean values of the measured values to identify traffic states without considering traffic flow uncertainty. To account for uncertainty, they proposed an interval data-based k-means clustering method for traffic state identification at urban intersections. In this method, the uncertainty of three traffic flow variables—volume-to-capacity ratio, queue length, and delay—was represented in the form of interval data and used as input variables for clustering. The evaluation confirmed the effectiveness of the methodology.

Saha et al. (2019) investigated various clustering analysis methods to identify traffic patterns that effectively represent traffic conditions. The study applied K-means, K-prototypes, K-medoids, Hierarchical Clustering, and PCA-based K-means to cluster traffic patterns based on volume, speed/travel time, and event data (e.g., incidents and weather). The results showed that K-prototypes and K-means with Principal Components (PCs) were the most effective methods for case study. They both effectively distinguished normal traffic, incident-related congestion, and weather-impacted conditions.

Smith et al. (2001) investigated the application of data-mining tools, including statistical clustering and classification techniques, to support the development of traffic signal-timing plans. The study used hierarchical cluster analysis to identify temporal interval breakpoints. In addition, the study demonstrated the use of classification and regression trees to automatically monitor the quality of the identified time-of-day breakpoints as traffic conditions change. This use allows checking each new case of data collected from the field to determine if it is classified in the cluster that is associated with the time-of-day plan in use. A warning is given if a series of misclassified states are identified for the possible need to perform the clustering again and/or the development of new plans and interval breakpoints.

Singleton et al. (2020) used continuous pedestrian traffic signal data to estimate pedestrian volumes at signalized intersections. They then used time series clustering to identify patterns of pedestrian activity at signalized intersections using two new pedestrian activity metrics. Five non-linear regression models were also developed to predict hourly pedestrian crossing volumes.

***Prioritization of Retiming Effort***

Winfrey et al. (2023) developed a method for ranking traffic signals according to the cost-effectiveness of traffic signal retiming. The ranking methodology utilizes a ranking formula and K-means clustering to prioritize signals needing adjustments. The ranking formula is used to assign a score between 0 and 10 to intersections based on a weighting factor that accounts for congestion (travel time), reliability (planning time index), and bottleneck rank. K-means clustering was then applied to group the signals into six clusters. The weight factors were further optimized through a brute-force search approach to minimize the discrepancy between the ranking formula results and the clustering outcomes. The optimal weight combination for the formula used in the ranking was determined to be 4.62 for congestion, 4.13 for the planning time index, and 1.25 for bottleneck ranking.

Wang et al. (2022) introduced a method for ranking intersections and summarizing corridor performance using measures calculated using high-resolution controller data. The study used four performance measures in the raking: platoon ratio, split failures, arrivals on green, and red-light violations. A performance threshold for each measure was initially determined using K-means cluster analysis and then refined by expert input. The k-means clustering divided the case study data into five groups. Interestingly, it was found that cluster analysis often placed intersections with multiple red-light violations in the same cluster as those with high split failures, suggesting that motorists frequently ran red lights during split fail conditions. Sensitivity analysis showed that different weighting did not change the results significantly.

## 2.2 Traffic Responsive Control

Hadi et al. (2023) developed and evaluated a predictive TRPS strategy based on supervised ML combined with unsupervised ML (clustering) and signal timing optimization utilizing ATSPMs. The evaluated supervised learning algorithms for short-term prediction were the decision tree, Random Forest (RF), Gaussian Naïve Bayes (GNB), Multinomial Logistic Regression (MLR), Support Vector Classification (SVC), K-Nearest Neighbors (KNN), and Artificial Neural Network (ANN). The results revealed that the ANN produced the best results in terms of short-term prediction of traffic patterns compared to other evaluated algorithms. The study then assessed the performance of the predictive TRPS based on clustering and prediction by evaluating five different scenarios of signal timing plan selection. The results of the project case study showed that the predictive TRPS method can decrease travel time by 7 percent compared to existing traffic signals.

Mitrovic and Stevanovic (2019) developed a method to select an appropriate day-of-year signal timing plan for the next peak hour based on the estimated traffic profile of that peak hour. First, the method applied k-means clustering to a historical data set to identify traffic pattern profiles and develop optimal signal timing plans for the demand profiles. It then utilized near real-time traffic information to predict the profile of the next peak-hour traffic pattern and select the plan that best fits the pattern. The results of the analysis showed that adequate signal plans can be selected for 65% and 75% of the days in a year.

Wierzchowska (2016) evaluated a TRPS strategy that utilizes ML for pattern matching, where a new timing plan is implemented when the traffic volume is associated with a state that is mapped to the timing plan. K-means clustering was used to identify the patterns. KNN classification method was used as the pattern-matching mechanism. KNN is an algorithm that can classify the current traffic state to one of the stored patterns based on a similarity measure, such as distance functions. The evaluation of the developed methodology using simulation indicated a 19% reduction in travel time compared to time-of-day control.

## 2.3 Traffic Adaptive Control

Signal control system vendors have offered adaptive signal control systems that have been shown to provide significant benefits under certain traffic conditions. On the other hand, researchers have proposed the use of AI techniques such as ML/deep learning, RL, and fuzzy logic for adaptive signal control. Although developing AI-based adaptive signal control is not an objective of this study, reviewing previous research on this effort can inform the research methodology for using AI in arterial management.

Srinivasan et al. (2006) explored the use of two configurations of adaptive signal control. The first configuration was a hybrid intelligence with a multistage learning system, and the second involved Fuzzy Neural Networks (FNN) with a stochastic approximation method. The evaluation of the researched systems using a microscopic simulation program showed that both can provide significant benefits when evaluated against existing traffic signal control in terms of the mean delay stoppage time.

Araghi et al. (2015) assessed the performance of three popular AI-based methods applied in adaptive traffic signal controllers of a single intersection. The methods included Q-learning (an RL method), neural networks (NN), and fuzzy logic systems (FLS). The results from traffic simulation analysis for both the peak and non-peak period scenarios showed that the Q-learning, NN, and FLS reduced vehicle delays by 66%, 71%, and 74%, respectively, compared to the fixed-time controller.

Chiu (1992) developed an adaptive traffic control system using fuzzy logic. The system optimizes traffic flow by dynamically adjusting signal timing parameters (cycle time, phase split, and offset) using only local traffic data. A set of 40 fuzzy decision rules was implemented to independently adjust the signal timing parameters. These rules aim to improve overall traffic delays. The authors reported significant benefits when evaluating the system using traffic simulation.

Niittymäki and Pursula (2000) developed a fuzzy logic-based signal control with fuzzy rules that utilized two input variables: the number of approaching vehicles and queue length. The simulation analysis indicated 10–20% lower delays compared to the traditional vehicle-actuated control.

Guo et al. (2019) implemented an RL approach with a neural network-based Q-function approximator (Q-network) for adaptive traffic signal control. The system used vehicle positions, directions, and speeds as inputs. Simulation-based evaluation showed that, compared to traditional signal control methods, the system was able to reduce the queue lengths by 27% to 73% and decrease the average wait times by 42% to 79%.

Mannion et al. (2016) evaluated the effectiveness of RL algorithms for adaptive control. The study compared the use of three reward functions in Q-learning. The evaluation using simulation showed that all three algorithms significantly outperformed fixed-time controllers. Using queue length as a reward function produced the best results, resulting in a 61.8% increase in speed, an 88.7% reduction in speed, and a 65.7% reduction in queue length.

Hippolitus (2020) developed a traffic control system that accounts for the spatiotemporal characteristics of urban traffic flow utilizing RL with a Q-learning network. Evaluation using traffic simulation demonstrated that the RL-based traffic control system significantly reduced average vehicle delay compared to two widely used traffic signal control algorithms: the longest queue first and the fixed-time control algorithms.

Li and Miao (2024) employed an FNN algorithm for traffic signal optimization to minimize vehicle delays at intersections. The study implemented two fuzzy controllers: the Phase Sequence Optimization Fuzzy Controller and the Green Light Delay Judgment Fuzzy Controller. The Phase Sequence Optimization Controller determines the next green light phase based on waiting time, queue length, and its rate of change, while the Green Light Delay Judgment Controller decides whether to extend the current green phase using queue length data and traffic flow variations. A simulation experiment was used to evaluate the effectiveness of the AI-based fuzzy controller compared to the traditional fuzzy controller, showing that the system reduced the average vehicle delay by 20.8% to 45.67%.

The Oklahoma Department of Transportation (ODOT) launched a project that I-powered traffic signal monitoring and connected vehicle hardware (Oklahoma Department of Transportation, 2022). A key innovation was the implementation of the ATSPM system for real-time traffic monitoring and adaptive signal adjustments. AI-driven detection solutions identified traffic patterns and user types, enabling dynamic TSP adjustments to enhance safety, especially for vulnerable road users. The project also incorporated scalable AI capabilities to meet the city's needs with the flexibility to add or remove additional functionalities as needed, including SaaS-based solutions such as Econolite Centracs Mobility, NoTraffic, and Surtract. Expected benefits include up to a 40% reduction in wait times, decreased congestion, improved pedestrian and cyclist safety, and lower greenhouse gas emissions.

## 2.4 Signal Control during Non-Recurring Events

AI-based DSS can be used to provide short-term prediction of travel time and support proactive decisions to implement special plans to respond to congestion conditions due to non-recurrent events such as freeway incidents, arterial incidents, and bad weather.

Tariq et al. (2022) employed a combination of a data-driven analytical approach, predictive modeling, and traffic simulation to estimate diversion rates and optimize signal timing during freeway incidents. Traffic was first partitioned using K-means clustering analysis and the results were then used to develop predictive models. Three ML techniques—linear regression (LR), SVM, and Multilayer Perceptron (MLP)—were employed to predict diversion rates under different traffic conditions. Additionally, microscopic traffic simulation and multi-objective optimization via Non-dominated Sorting Genetic Algorithm II (NSGA-II) were used to develop and evaluate special signal timing plans. The MLP and SVM models demonstrated superior predictive accuracy for diversion rates compared to LR. The study found that the diversion rates ranged from 4% to 27% depending on factors such as incident severity, the number of lanes blocked, time of occurrence, and freeway congestion. The simulation results showed that the optimized special signal timing plans can reduce delays by up to 23% during the freeway incident diversion.

Saha et al. (2020) developed a deep learning-based model using a Long Short-Term Memory (LSTM) Neural Network to predict the performance on alternative routes during freeway incidents. The authors also proposed a methodology for developing specialized signal plans for critical intersections along the diversion routes utilizing the results from the prediction model in combination with a simulation-based signal timing optimization model. The LSTM-based prediction model successfully predicted travel times on alternative routes with an accuracy of 82%–90% during incidents and 88%–92% during normal conditions.

Yao and Qian (2020) developed a method to recommend optimal signal timing plans in real-time under incidents based on domain knowledge and learning from historical data. The recommended plans are based on the outputs from real-time traffic prediction of incident impacts. The method applied two hierarchical models—real-time traffic prediction and plan association. The two models are connected through metric learning. The results of the evaluation show that the recommendation system has a precision score of 96.75% and recall of 87.5% on the testing plan and makes recommendations an average of 22.5 minutes lead time ahead of Waze alerts.

Tariq et al. (2020) proposed automating the process of updating the signal timing plans during non-recurrent (e.g., incidents, lane blockages, surges in demand) on arterial streets. Recursive Partitioning and Regression Decision Tree (RPART) was utilized for decision tree classification and feature selection, combined with a Fuzzy Rule-Based System (FRBS) to replicate the uncertainties in traffic signal engineers' decision-making. The system learned from expert decisions and automated the process of adjusting green time allocation in response to congestion.

The model achieved an overall 77% accuracy and a mean absolute error of 5%. Microscopic traffic simulation results demonstrated that the proposed method effectively reduced delays and queue lengths under three investigated scenarios—one lane blocked out of three, two lanes blocked out of three, and a demand surge.

Lawe and Wang (2016) proposed the use of deep learning neural networks that uses a time series of traffic weather conditions to predict traffic volumes. The output (i.e., predicted traffic volume) is fed into the delay equation for use in generating the best green times to manage traffic delay. The prediction method provides better results than other investigated methods, such as the Radial Basic Function, Random Walk, SVM, and Back Propagation (BP) Neural Network.

Delaware Department of Transportation (DelDOT) implemented traffic management that monitors and predicts traffic flow performance, identifies anomalies and inefficiencies, and generates responses to existing and predicted congestion (Delaware Department of Transportation, 2024). The prediction and responses are based on time-of-day, day-of-week, seasonal, and holiday data, which are used to predict volume, speed, and occupancy for the next 5, 10, 15, 30, and 60 minutes for freeways and signalized arterial corridors. For anomaly and incident detection, separate algorithms were used for incidents on freeways and arterials. The freeway algorithm leverages travel time data from Bluetooth detectors and traffic flow data from radar detectors. The arterial algorithm is more complex due to the variable nature of arterial road networks. It utilizes per-lane volume data from system detectors to detect incidents through arterial roads and travel time data from Bluetooth detectors. The project team created incident severity scores based on detector data, travel time data, Waze alerts, recorded video clips, incident location, number of lanes blocked, and capacity reduction. The scores ranged from 0-100, with a severity score of 70 or higher indicating a possible need for Traffic Management Center (TMC) attention and possible action. The traffic flow prediction algorithm anticipates what the traffic demand will look like in up to 60 minutes in the future so the TMC can start to manage congestion before it even happens.

## 2.5 Malfunction Detection and Management

AI could be used to identify anomalies in the data received at the TMC from traffic controllers that could be due to communication or detector malfunctions. The identification of malfunction could be identified by comparing various real-time measurements with historical measurements, performing a time-series analysis, or combining the two. This section includes review of studies that investigated malfunction detection in signal control and studies other disciplines that focused on identifying anomalies.

### 2.5.1 Malfunction Detection in Signal Control

This section reviews studies that investigated the malfunctions of detection and communication components of signal control. Sunkari et al. (2010) developed modules to improve intersection operations during detector failures at isolated signalized intersections. Simulated detector failure data was used in the investigation. The objective was first to develop methods to identify detector failures, and the second was to propose a signal timing control strategy to operate the intersection in a more efficient manner during malfunctions, as follows:

- The study identified a detector failure by max presence, no activity, or erratic count. Signal controllers usually use these three parameters to identify a detector failure. The first parameter is the maximum interval that a detector is occupied (in minutes) prior to being considered a fault. The second parameter is the no activity limit diagnostic that is based on the maximum time between detector actuations (in minutes). The third parameter is an erratic count diagnostic that is based on the maximum number of actuations per minute, with typical values ranging from 40 to 70 counts per minute. The user can specify the thresholds for the three diagnostics criteria.

- Traditionally, the intersection has been operated during detector failures by placing a constant call on the failed detectors resulting in a max out on the phases mapped to these detectors every cycle. Agencies can also specify providing the minimum green for these phases. Some controllers give an option to the user to specify how long the phase should be on during a detector failure. Sunkari et al. (2010) proposed using signal timing plans identified using the average and variance of traffic demands for a rolling four-week historical operational log period combined with the green utilization for each phase that is used as an indicator of the traffic demand at the intersection.

Tarunesh and Chung (2020) developed models to estimate missing volume and occupancy values at the failed detectors using deep learning techniques such as a neural network and autoencoders. The study used a neural network to model a complete network of detector autoencoders to reduce the model size by exploiting the spatial correlation between detectors. The study defined estimation accuracy as the number of predicted values within a tolerance range around the actual values. The detector failures are identified by analyzing the variation of traffic volume over different months to find anomalous detectors based on:

- Abrupt changes in volume in consecutive 15-minute intervals over the day for multiple days
- The graph of daily volume over the year shows sudden changes.

An et al. (2017) investigated the impacts of communication loss on signal performance using event-based data, focusing on the effects on data quality and signal coordination. Two statistical tests, the one-way ANOVA test and Tukey's HSD test, were used to identify unexpected changes in traffic progression patterns caused by communication loss. The study pointed out that traffic controllers provide logs of communication failures. For example, the MaxView server logs the communication event data, including total communication attempts, failed communication attempts, average response time, and percentage of communication loss (POCL), at approximately five-minute intervals. The study used three specific criteria to characterize the missing data in phase and detector events: (a) “phase always on” looks at the completeness of phase events, (b) “unpaired on and off,” and (c) “long duration,” reflecting a time gap between a detector on–off or off–on that exceeds a certain time threshold. Communication loss can cause controller clock drifting, which can degrade signal coordination. The “cyclic occupancy in seconds” was used as a measure of traffic progression quality. The increase in control delay was also estimated using a revised break point-based model. The results indicated significant increases in the control delays between 27% and 720% when communication losses occurred.

Smaglik et al. (2022) developed a detector health monitoring procedure to identify detection performance issues beyond complete detector failure. The study mentioned that the controller is able to identify detector faults at the ends of the performance spectrum but not if the performance has degraded slightly due to increased latency or some other performance issue. The method proposed in the study uses event-based data from detectors at signalized intersections to calculate volume and density measures during the saturated flow portion of green intervals and utilize these measures to approximate the volume-density fundamental diagram.

Wang et al. (2024) explored the use of ML algorithms to predict missing and anomalous directional traffic volume data in ATSPM datasets. To assess model effectiveness, the study used two statistical methods commonly employed by traffic engineers—the mean of the past week and the mean of the previous 52 weeks—as baseline models and compared them with four ML algorithms: LSTM, MLP, decision trees, and Support Vector Regression (SVR). The results showed that LSTM outperformed all models, achieving the lowest prediction error, followed by decision trees and MLP, which ranked second and third, respectively. The study also found that adding timestamp variables improved accuracy, whereas additional features such as weather, crashes, and holidays did not significantly improve predictions.

### 2.5.2 Identifying Anomalies in Other Disciplines

Research has been conducted in other disciplines on identifying anomalies in time series data during real-time operations. Studies have used basic statistical methods, such as comparing to standard deviations in the data assuming normal distributions, comparing data points to moving averages, and/or comparing predefined thresholds. Others used ML approaches to detect anomalies.

Choi et al. (2021) categorized the anomalies in time-series into three classes: (a) point anomalies, (b) contextual anomalies, and (c) collective anomalies. By studying these anomalies, the authors pointed out the complexity of multivariate data where relationships between variables require advanced anomaly detection methods.

Darban et al. (2024) pointed out that the non-stationary nature of datasets necessitates the adaptation of deep learning models through online or incremental training approaches, enabling them to update continuously and detect anomalies in real-time. The detection of anomalies in multivariate high-dimensional time series data presents a particular challenge. In addition, the authors concluded that in the absence of labelled anomalies, unsupervised, semi-supervised or self-supervised approaches are required. 64 deep learning models were discussed and categorized in the study and time series deep anomaly detection applications across multiple domains were discussed. The authors categorized deep learning methods for time series anomaly detection into four main approaches: (a) forecasting-based models predicted the next timestamp based on historical data and detect anomalies based on the forecast, (b) reconstruction-based models operate on latent time series representations, (c) representation-based models addressed time series complexities by learning compact feature embeddings, and (d) hybrid models employed a joint objective function to combine multiple approaches for improved accuracy.

Tinawi (2019) compared LSTM, autoregression, Multi-Layer Perceptron, and Encoder-Decoder LSTM models with a dynamic error threshold for anomaly detection. The finding showed that the LSTM model achieved the best F-score of 0.76. The LSTM autoencoder models resulted in the highest precision score of 0.83. Precision measures the proportion of true positives among all predicted positives, while the F-score combines precision and recall (the proportion of true positives among all actual positives) into a single metric.

Schmidl et al. (2022) conducted a comprehensive evaluation of 71 state-of-the-art anomaly detection algorithms across 976 time-series datasets to assess their performance. The threshold-agnostic Area Under the Curve (AUC) metrics were employed to measure accuracy by focusing on anomaly scoring rather than binary labeling, ensuring a fair comparison across diverse methods. The findings demonstrated that for high-dimensional and complex datasets, deep learning models outperformed statistical approaches. While simpler methods remain effective in low-dimensional data that has well-defined patterns.

Iqbal et al. (2024) introduce deep ensemble models that integrate multiple deep learning architectures to improve anomaly detection. The study employed Convolutional Neural Networks (CNNs) to detect local patterns, LSTM networks to capture long-term temporal dependencies, and Transformer and Graph Neural Networks (GNNs) to model global relationships and interdependencies, which traditional models struggle to handle effectively. The ensemble models such as CNN-LSTM and Transformer-GNN achieve higher accuracy than single ML models.

Vajda et al. (2024) introduced models with the aim to address the requirement for fast and precise real-time anomaly detection. The core innovation involved exploiting the periodicity of data using variational mode decomposition prior to applying LSTM-based anomaly detection. This data preprocessing eliminated recurring patterns from the dataset, thus enabling the model to concentrate on identifying true anomalies.

# 3 STAKEHOLDER MEETINGS

In addition to conducting the literature review summarized in the previous Chapter, this study also conducted on-line meetings with selected stakeholders including FDOT TSM&O management teams from two FDOT Districts and two counties in South Florida (Palm Beach County and Broward County). The purpose of the meetings was to determine experience with any current applications and the needs and issues with existing arterial management and operations that can be met with the use of data analytics. The questions asked in the meeting are:

- What are the issues with your current traffic management system?
- What are the main issues that you think AI/data analytics can be used for?
- What is your experience in AI/data analytics?
- What type of data do you collect and archive? What are the main issues with the collected data quality?
- What are the limitations and constraints that can affect the successful use of AI?

Prior to asking the questions, the research team provided a summary of the project objectives and the on-going and future project activities. The findings based on the meetings are documented in this section.

## 3.1 FDOT District 2

The meeting with FDOT District 2 was conducted on 3/19/2025 and was attended by Peter Vega, FDOT District 2 TSM&O Program Manager and Adam Storm, TSM&O Engineer of FDOT District 2.

**Current Situation and AI Potentials**

The ATMS.now (Cubic system) is currently used in FDOT District 2 region as the central system. In their region, the daily and seasonal patterns of the current timing plans are sufficient to handle daily and seasonal variations since the region has more consistent (non-tourist-heavy) travel patterns.

***Corridor Ranking***: District 2 mentioned that ranking corridors for improvements is an important task that data analytics can help with. However, there is an important challenge with this ranking in that it focuses on the main street and ignores the feeder roads (or cross street), leading to inaccurate performance rankings and mitigation strategies. Improvements on the main corridor can negatively affect feeder roads. The ranking system should include effects on cross streets and upstream/downstream corridors. Corridors should not be evaluated in isolation. The ranking should consider time of day, seasonality, incident conditions, and diversion routes. For example, a corridor may rank differently during an I-95 closure than under normal conditions.

***Diversion During Incidents:*** When incidents on I-95 cause diversion to surface streets, simultaneous rail roadway crossings can worsen congestion. Even if train activity is infrequent, it is recommended to incorporate the train crossing into AI-based decision-making to develop a Plan B routing strategy in case railroad closures interfere with the diversion routes.

***Adaptive Control Systems:*** The adaptive signal control systems, according to FDOT District 2 experience, can prioritize optimizing the main corridor on the expense of cross streets. They can negatively impact side streets by increasing their delays, resulting in poorer system-wide performance. The failure to consider network-wide effects, especially upstream congestion feeding into the corridor can bias the results of the evaluation.

***Proactive and Predictive Signal Control:*** There is a need for a system that uses AI to predict congestion and adjust signal timing before issues occur, rather than reacting after the fact. This would include automatically identifying the need to activate signal plans based on traffic patterns without waiting for operator intervention.

***Automated Malfunction Detection and Prioritization:*** AI can identify detection system failures (e.g., stuck sensors, false readings) and prioritize which issues need immediate attention. The goal is to reduce the manual review of system alerts and make maintenance more efficient.

**Challenges in Using AI**

The issues mentioned in the meeting regarding challenges in using AI include:

- Regular commuters expect signals to behave in a certain way based on their day-to-day travel experience. When AI changes signal timing dynamically, it may confuse drivers and can lead to unsafe situations or even crashes.
- The system should balance system-level efficiency (e.g., throughput, delay) with user experience (e.g., reliability, consistency) in selecting the signal timing plan.
- The technologies of a large proportion of field devices (e.g., signal controllers and detectors) are outdated. High-resolution controller data is not available from old controller firmware and, therefore, may not be AI-compatible.
- There is a need for reliable communication infrastructure. If a corridor loses connectivity, AI-based management cannot work properly, and the signals will have to revert to time-of-day plans.
- There is a need to consider the software and licensing costs.
- There is a lack of standardization with different vendors using proprietary systems that don't integrate well with other systems.

**Other Experiences with Using AI and Data Analytics**

Other experiences discussed in the meetings included:

- **Bike/Pedestrian Detection Project:** This project involved deploying a system to detect pedestrian activity near University of Florida (UF) campus using AI-based video analytics to adjust signal timing based on pedestrian demand (e.g., during class changes). The system was based in existing camera technology, but the system was found to have issues including missed detections of pedestrians, merging multiple pedestrians into a single detection, and incorrectly predicting movement direction. There is a need for systems to provide sufficient accuracy for counting, trend analysis, and basic planning. Examples of implementation is the Smart St. Augustine grant for quarterly reporting on bike/pedestrian activity. The district found that using YOLO (You Only Look Once) – Open-Source AI produced good performance in truck parking detection and pedestrian tracking.
- **Vehicle Counts based on Machine Visions:** A system by Southwest Research Institute called Activision is being explored that uses computer vision to count vehicles, detect

wrong-way driving, and identify traffic queues and congestion. The system offers real-time event detection and could support enhanced safety monitoring and response.

- **Using Machine Vision in Identifying Malfunctions:** MioVision is another commercially available machine learning-based video detection system, widely used by local agencies. The system is helpful in detecting malfunctions, such as stuck loops or detection failures, even without the availability of high-resolution controller data. The system learns what normal signal operation looks like over time (e.g., typical phase durations, vehicle clear-out times, and expected on/off cycles of detectors). When real-time data deviates from the learned patterns, the system flags it as a potential malfunction (Example: If a phase stays green but vehicles are no longer present, it may indicate a stuck detection). The system can process the data to identify detection failure (no activation when expected), sticky detection (detector stays on too long, which may be due to weather and temperature), skipped phases or abnormal split patterns.

**Data Availability and Data Challenges**

Collected data includes data from the central arterial management system such as sensor data (vehicle presence, occupancy), timing logs (split failures, phase durations), alerts and alarms (e.g., communication loss, MMU/CMU conflicts). Other useful data includes SunGuide data including traffic incidents and volume counts. Video analytics data can also be collected from systems like Bosch, MioVision, YOLO, and Activision that provide pedestrian and vehicle counts, classification (e.g., cars vs. trucks), turning movements, pedestrian trajectories (where possible), alerts for wrong-way drivers, queue detection. There is a lot of promises of utilizing high-resolution controller data. However, many older controllers still in use do not support the logging of this data.

Several issues were also discussed in the meeting regarding data availability and quality including:

- Pedestrian data does not provide a good indication of pedestrian activities since only 25–30% of pedestrians press the crosswalk button. When using pedestrian detection technologies, pedestrians' intent detection is a major gap. It is hard to know if someone is about to cross or just standing nearby. In addition, shared pedestrian ramps (e.g., for northbound and westbound directions) make it hard to determine which direction the person intends to take to cross the streets.
- Count data based on video image processing has been claimed to be 95% accurate, but it hadn't yet been validated with the manual counting. The data may be good enough for trends and general analysis, but the accuracy may not be sufficient for precision-critical applications (e.g., real-time signal timing changes).
- Detection systems (loops, video, radar) often malfunction and go unrepaired for months due to staff shortages, contractor delays, and limited local agency resources.
- There is a need for a unified platform for collecting and analyzing arterial signal data across systems.
- Data gaps that the District believe should be addressed include intersection-level counts and turning movement by time of day, vehicle classification (e.g., freight/commercial vehicles), reliable pedestrian count data including pedestrians' intent (e.g., are they crossing?), and high-resolution controller data across the network.

### Prioritization of AI Features

Top priority AI applications include detecting the malfunction early, especially when a loop is stuck "on" (e.g., calls for green when no vehicle is present) or a loop is non-responsive (e.g., vehicles are present, but no call is placed).

## 3.2 Palm Beach County

The meeting with Palm Beach County was conducted on 3/11/2025 and was with Motasem Al-Turk, Director, Traffic Division, Palm Beach County / Engineering & Public Works Department

### Current Situation and AI Potentials

Currently, the agency uses the Cubic TrafficWare central system, operated through its software platform known as ATMS.now. The agency has limited experience with using AI-based data analytics. Some of the potential AI applications that can improve operations and management were discussed as documented below.

***Prediction of Incident and Diversion Impacts:*** The agency expressed strong interest in the project and expressed strong incident detection and an understanding of traffic diversion patterns. They emphasized the importance of early incident detection and proactive signal timing adjustments to minimize congestion. They discussed their involvement in the Arterial Management Program (AMP) through Active Arterial Management (AAM), which monitors key corridors such as Okeechobee Boulevard. When incidents occur, the TMC staff need to respond by adjusting signal timing not only on the affected roadway but also along the detour routes to which traffic diverts. The agency noted that diversion between arterial roads as well as diversion from freeways to arterials are important. Currently, the agency relies on staff monitoring corridor conditions, with support from CCTV camera systems, to verify incidents and assess the appropriate response. During lane blockage incidents, the agency noted that simply increasing green time in the affected direction is not always a practical solution, as it reduces time for other competing movements and doesn't effectively resolve congestion. They emphasized the need to identify where traffic is rerouting in real-time and develop strategies to manage detour routes and mitigate the increased demands

***Malfunction Alarm Management:*** Staff members find it difficult to distinguish between critical alarms and minor or false alerts, leading to challenges in accurately identifying the actual issues occurring in the system. Due to the volume of low-priority alerts, staff may be delayed in responding to more critical issues. For example, ATMS.now sends alarms for a signal that briefly enters a flash mode for just a few seconds and then recovers on its own. This event typically does not require intervention. A smarter alarm filtering mechanism is needed. The alarms needed are:

- Persistent signal flash alarms
- Impactful detector failures including those of loop detectors, video detection systems, or pedestrian push buttons that fail or behave abnormally (e.g., stuck pedestrian buttons causing continuous calls).
- Other major malfunctions, such as any failure that impacts signal coordination or causes extended traffic disruption.

***Signal Re-Timing and Signal Plan Deviation:*** The agency explained that corridors are operated almost entirely based on time-of-day. They periodically re-time the signals, ideally every two years. Although the system is designed to follow the time-of-day (TOD) plan strictly, field adjustments by technicians, such as during maintenance or after a malfunction, sometimes result in changes to signal parameters. Even minor modifications can disrupt the green time splits and coordination leading to residual queues, downstream backups, and a loss of synchronization across intersections. The agency expressed the need for a system that can automatically detect such deviations and restore the original TOD settings.

***Pedestrian Crossing Disruption:*** The signal plans typically limit the green time for side streets in the absence of pedestrian presence, which is often shorter than the time needed for the pedestrians to cross. When a pedestrian presses the button to cross, the system must extend the green phase, which disrupts the planned coordination. Once disrupted, the system takes several cycles to recover and returns to the coordinated operation. This issue becomes more severe in locations with high pedestrian volumes, such as near schools during morning drop-off hours, where frequent pedestrian activations lead to significant traffic backups. There is an interest in investigating whether AI could predict pedestrian crossing patterns, assess how often these disruptions occur, and suggest strategies to minimize their impact. For instance, AI could analyze traffic and pedestrian data to determine whether serving pedestrians every cycle might, in some cases, maintain coordination better than inconsistent interruptions.

***Digital Twins:*** The agency expressed interest in digital twin technology including the use of real-time simulation tool to predict incident impacts, evaluate signal strategies, and support proactive traffic management decisions, especially on the Okeechobee Boulevard corridor.

**Data Availability and Data Challenges**

The available data includes traffic signal data via the ATMS.now system and CCTV footages in real-time operations. Their detection system includes loop detectors and video image detection but this detection is used for real-time control only. The agency also has access to HERE data and uses Google Maps informally to monitor traffic flow, queuing, and congestion for real-time monitoring and alert confirmation.

Palm Beach County does not currently use high-resolution controller data. They have attempted to obtain turning movement counts using video image and loop detectors but found the data to be inaccurate. A recent trial using a hybrid video and radar detection system also failed to provide improved accuracy and was not more accurate than the video image detection previously tested.

**Challenges in Using AI**

The challenges mentioned in using AI include the need for additional data and sensing equipment, to support AI functionality. This will result in increased maintenance burdens associated with installing and managing this equipment.
The use of AI will also require building AI capabilities, including the need for staff training to properly operate and maintain AI-enabled systems. Finally, there are concerns about the extent and practicality of AI deployment in real-world operations.

## 3.3 Broward County

The meeting with Broward County was conducted on 3/25/2025 and was with Rasem Awwad, Director; Broward County Public Works Department/Traffic Engineering Division.

**Current Situation and AI Potentials**

Broward County currently uses Cubic ATMS.now. The vehicle detection system mainly uses video image detection. The following is a discussion of potential AI applications as discussed in the meeting.

***Malfunction Detection:*** Broward County is interested in AI-based malfunction detection to better manage signal malfunctions and traffic control issues that are not detected by the current systems. Currently, the system has standard alerts for constant calls. When the system detects continuous calls for 30 or 60 minutes, depending on the time of day, it triggers an alert. The County also performs daily screenings to check the quality and health of the detection system. There is a need for better detection and determination of the impacts of malfunction events. They want the system to be able to identify unusual traffic patterns and predict when such events might occur based on historical data and environmental conditions (e.g., time of day, weather, etc.). For example, the system should be able to recognize missed vehicle detections due to weather or shadows at a specific time resulting in skipping phases. The randomness in detection can be caused by environmental conditions (e.g., weather and shadow), camera angles, obstructions (e.g., large vehicle blocking), lighting conditions (e.g., glare during sunset), etc.

***Advanced Traffic Control during Preemptions:*** Broward County wishes to have better insights into managing traffic during railroad preemptions. The Brightline trains cause traffic disruptions, but the timing and duration of these disruptions are unpredictable. AI-based solutions could help by predicting the impact of and automatically responding to these disruptions by adjusting signal plans.

***Near-Miss Detection:*** Broward County is already testing a near-miss detection pilot with the University of Florida (UF), using video images from the GridSmart fisheye cameras at a few intersections. Portable units are used for collecting the near-miss data. Such systems could be used to potentially recommend low-cost countermeasures for pedestrian crashes including adding pedestrian crossing signs, converting the left turn to flashing yellow arrows, and so on. AI could analyze the data from these cameras to better understand near-miss events and provide insights on how to improve safety.

***Issues with Adaptive Signal Control***: There are issues with adaptive signal control that need to be considered. Such systems can cause problems with nearby non-adaptive systems. In addition, sometimes, the response of the adaptive systems can be too aggressive, especially during peak traffic times, causing issues on the side streets since they are not synchronized. It is important to fine-tune the cycle and timing limits to prevent such issues. When there is a diversion from freeways to adaptive-controlled arterials, the adaptive system can be a few cycles late in responding, potentially causing delays. This delay occurs as the system adjusts to the sudden increase in traffic from the diversion, and it may take a little time before the signals properly adapt to the new traffic conditions. In some cases, human intervention may be needed to override or assist the adaptive system.

**Data Availability and Data Challenges**

Currently ATMS.Now (the central signal software) provides data logging capabilities used for logging and monitoring traffic signal performance. This includes signal phases and timing data. However, the reporting structure of this data is not very user-friendly.

Broward County also installed an adaptive signal system (Cubic Synchro Green) that provides additional data. The system controls 20 intersections located at US1 through Downtown Fort Lauderdale, University Drive from its intersection with Broward County to I-595, and West Broward Boulevard Between SR7 and I-95. Data is collected from a MioVision video image detection system.

ATSPM data is also used including high-resolution controller data with the MioVision system embedded into the ATSPM system. The County archives both event data and performance measures through the MioVision's cloud. The data stored in the MioVision cloud is kept indefinitely. The County is also developing a dashboard that pulls information from ATSPM to continuously track performance metrics from the adaptive signal intersections. The dashboard is customizable in that operators have access to more detailed insights and other stakeholders can see only summary data.

Some of the limitations of current data collection and use include limitations of data analytics, lack of resources for data validation, insufficient storage capacity resulting in overwriting the stored data, limited data utilization due to limited staff and time making it difficult to analyze a large amount of data to extract useful insight, and inconsistencies of data from various sources with no standardization of data format.

## 3.4 FDOT District 5

A meeting was also conducted with Jeremy Dilmore on 4/02/2025 who is FDOT Emerging Technologies Manager and was until recently the Transportation Systems Management and Operation Engineer for the Florida Department of Transportation District 5.

**Current Situation and AI Potentials**

Potential AI applications discussed in the meeting as summarized below.

***Filling Gaps in the Data:*** AI can fill gaps in data caused by detection or communication failures by predicting the missing values based on historical traffic patterns and thus helping to create more complete data. FDOT District 5 worked with UF and UCF (University of Central Florida) on data filing. This project involved using AI to fill in missing traffic data caused by failures or gaps in the detection systems.

***Understanding Traffic Trends:*** AI can analyze historical traffic data and detect recurring patterns, allowing for better forecasting of future traffic trends.

***Clustering Traffic Patterns for Time-of-Day Plans:*** AI can assist with identifying clusters of traffic patterns for use in selecting time-of-day plans. Current systems focus on AM and PM peaks plans, one weekday off-peak period plan, and a weekend plan. There has been limited work done on managing traffic during the off-peak periods. AI can assist in improving traffic control during these periods, which have historically been less optimized. Clustering can be used to analyze traffic patterns over specific time intervals (e.g., 15-minute slices) across different days, weeks, or even years. The goal is to identify recurring patterns in traffic behavior that could warrant the development of more specific signal timing plans beyond the traditional peak periods. AI can recognize traffic changes during holidays (e.g., days after Thanksgiving, Christmas, etc.) and identify these as distinct patterns to be managed with dedicated signal timing plans.

***Analyzing Non-Recurrent Events:*** AI can be used to identify non-recurring traffic patterns, such as those caused by incidents, and analyzes how often they occur over time. If they happen frequently, a signal timing plan may need to be developed for these patterns. In addition, AI can recognize when special events (e.g., concerts, sports events) are impacting traffic during the off-peak period. These events can cause traffic patterns like peak periods requiring specific signal timing adjustments. The identification of patterns can also identify higher traffic volumes at logistic areas around midday when trucks are arriving or departing, which may require special plans during these periods.

***Cost Reduction in Data Collection:*** AI can reduce the time and costs associated with manual data collection, allowing for more resources to be allocated to engineering analysis.

***Detector Failures and Management:*** AI can help to identify when detection systems fail or communications break down and to recommend plans to implement during failures, instead of setting the system to maximum recall. When the detection fails, the system should try to avoid max recall by using AI-generated data to emulate detections and provide more accurate adjustments. It may be possible to provide emulated actuation—predicting the cycle-by-cycle flow based on real-time traffic conditions (providing the estimate of how many vehicles need to be granted a green light each cycle). The focus is on optimizing the actuation process to better respond to traffic needs without unnecessarily changing the signal timing plans. This would avoid underserving traffic and avoid maxing out the system at the same time. Maxing out implies that the controller allocates resources based on the 100th percentile of traffic demand. One suggested approach is to start with the 50th percentile demand and green allocation and evaluate the relative Z-score to determine whether traffic demand is significantly higher than the expected on that day. Based on this evaluation, the system could adjust the actuation, potentially going up to the 80th or 90th percentile rather than maxing out the system (i.e., using the 100th percentile) to avoid inefficiencies.

***Incident Detection:*** Incident detection on freeway systems involves detecting traffic incidents based on historical data and traffic patterns. However, this work has mainly focused on the freeway systems, not signalized arterials. AI can use ATSPM data to analyze arrival patterns and detect when these patterns change significantly, which could indicate an incident. One of the challenges is the difficulty in distinguishing between lane blockages caused by incidents and those caused by other factors like special events. AI can struggle with real-time detection of incidents, especially when the traffic pattern changes significantly due to unexpected events. It is challenging to differentiate between incidents (such as lane blockages) and events (like planned activities) that cause similar traffic disruptions.

**Data Availability and Data Challenges**
The data used in FDOT District 5 include ATSPM Data (high-resolution controller data), Flow Lab data (Used to generate turning-movement count reports by fusing probe vehicle data and the ATSPM data), probe data from HERE, and Waze data. The ATSPM data is initially stored in an SQL database and kept updated in real time. Once the data is collected, it is transferred to an Elastic database for long-term storage. The Elastic database allows for efficient searching and analysis of large volumes of traffic data. The data from the Elastic database is stored in CSV files, which makes it easier to process and analyze. These files store the traffic data in a format that can be used for further analysis or reporting.

However, the data collected is incomplete because of detector and communication failures. The frequency of failures depends on the location and the maintaining agency. In Seminole County, the detector failures are below 2%, but in other locations, the failure rate is substantially higher. Detectors are also not installed in all locations to collect data.
A main challenge is that the data is incomplete and/or inconsistent. This makes it difficult for AI to provide accurate predictions or decision support. In addition, there are difficulties in fusing data from various sources (e.g., ATSPM, probe vehicle data, Waze), as each source might have different levels of accuracy or timeliness.

## 3.5 Summary

The review of literature and meetings conducted in this study indicate that AI can play a major role in active arterial management including the following:

- Identification of recurrent traffic patterns and peak breakpoints for time-of-day signal timing plan development
- Identification of historical non-recurrent traffic patterns that have significant impacts on traffic performance
- Prioritization of retiming and other improvement efforts based on identified performance measures.
- Development of a new generation of traffic responsive strategies that identify traffic patterns in real-time and match them to historical traffic patterns identified based on clustering
- AI-enabled traffic adaptive control that addresses limitations of existing adaptive signal control systems
- Arterial incident detection and the selection of special signal timing plans to be activated during the detected and identified incidents and other non-recurrent events
- Prediction of diversion to alternative routes, the selection of special signal timing plans for proactive implementation to address diversion impacts on alternative routes, and resolving the adverse impact of railroad crossing and diversion interaction
- Detection and categorization of the impacts of detection and communication malfunctions and recommending special signal timing strategy to be implemented during malfunctions
- Estimate and predict pedestrian demands and impact on performance and recommend signal timing plan for locations with high pedestrian demands

- AI-enabled transit and freight signal priority, emergency vehicle preemption, railroad preemption, and bridge preemption
- AI-based Video analytics for volume counts, conflict analysis, incident detection, and performance measurements
- Managing traffic during railroad preemptions, which cause unpredictable traffic disruptions and automatically responding to these disruptions by adjusting signal plans
- Detecting deviations from the stored signal timing plans that happen during maintenance or after a malfunction
- Analyzing near-miss data and recommending mitigation strategies to address the identified safety problems
- Filling gaps in data caused by detection or communication failures by predicting the missing values based on historical traffic patterns and thus helping to create more complete data

The challenges with the use of AI and data analytics mentioned in the meetings are listed below.

- The data collected is incomplete and inconsistent.
- There is a need for additional sensing equipment, which will increase the capital and maintenance costs.
- ATSPM data including high-resolution controller data is not available at all locations. Many older controllers still in use do not support the logging of this data.
- Turning movement counts using automated detection technologies such as machine vision may not provide sufficient accuracy.
- There is a lack of resources for data validation.
- There insufficient storage capacity resulting in overwriting the stored data
- There are inconsistencies in the data from various sources with no standardization of data format.
- There is a lack of standardization with different vendors using proprietary systems that do not integrate well with other systems.
- Another issue is the limited staff and time and the need for staff training to properly operate and maintain AI-enabled systems.
- There are concerns about the extent and practicality of AI deployment in real-world operations.
- Pedestrian and other vulnerable road user data is lacking.
- Detection systems often malfunction and go unrepaired for months due to staff shortages, contractor delays, and limited local agency resources. This can affect AI success.
- There is a need for a unified platform for collecting and analyzing arterial signal data across systems.
- AI generated solutions should not violate driver’s expectation.

- The system should balance system-level efficiency (e.g., throughput, delay) with user experience (e.g., reliability, consistency) in selecting the signal timing plan.
- There is a need to consider the software and licensing costs.

# 4 PERFFORMANCE MEASURES

Transportation agencies make complex off-line and real-time decisions by monitoring, assessing, predicting, and responding to changing conditions. This requires the identification and assessment of key performance indicators (KPIs) for estimating system performance for use in supporting the decision processes. Agency decisions can be classified in three categories (Hadi et al., Forthcoming-a):

- Strategic decisions: These decisions involve setting direction and understanding of the agency vision, mission, goals, performance objectives; in addition to identifying strategic focus areas and priority functions.
- Tactical decisions: These decisions involve identifying, prioritizing, and implementing activities, projects, operation procedures, policies, action plans, and investment plan to improve system performance as assessed using tactical KPIs.
- Operational decisions: These decisions involve assessing and developing management and control plans, monitoring and predicting system performance, and real-time selection and activation of plans based on operation-level KPIs.

(Hadi et al., Forthcoming-a) maintained that making effective decisions based on these KPIs require the alignment of the KPIs at the strategic, tactical, and operational decision levels and to set targets for these KPIs. This will allow the agencies to have integrated decisions at the three levels. For this purpose, the authors recommended the use of a KPI tree to identify and depict the association between metrics at the strategic, tactical, and operational levels. The KPI tree can be refined and further aligned based on stakeholder inputs to define KPIs at different decision levels that are consistent with the strategic objectives of the agency.

In general performance measures can be categorized as output and outcome performance measures. Output measures include activity or processes taken (e.g., implementing lane management, increasing enforcement, providing traveler information) to achieve the desired outcomes. The output measures can also include the physical and institutional infrastructure to support the activities that contribute to achieving the desired outcomes (e.g., lane miles of roadway with incident management, number of adaptive signals, transit vehicles, qualified transportation professionals, data sharing agreements, number and locations of air quality sensors). On the other hand, the outcome measures relate to how well the agency is meeting its mission and stated goals. Examples of these output measures are mobility, reliability, safety, and environmental measures. Although this document focuses on outcome measures, both the output and outcome measures can be used as important inputs and outputs of AI/machine learning models depending on the purpose of the model.

The ITS Strategic Plan of the Florida Department of Transportation (FDOT) (FDOT, 2017) requires that FDOT districts establish Performance Enhancement Goals (PEG) for mobility performance measures (outcome measures), safety measures (outcome measures), lane clearance measures (output measures), and ITS/communication network performance measures (output measures). The plan identified the following strategic outcome measures:

- Mobility and reliability measures: These include throughput, delays, planning time index (PTI), and speeds.
- Safety measures: These measures include crash rates and crash severity.

## 4.1 Strategic Measures

The strategic level measures are system measures that can be used to allow the strategic managers to determine trends and issues of the transportation system; compare the performance to benchmarks; identify performance targets; and make strategic decisions. Tactical managers can use these measures at the facility or corridor levels, in some cases, in combination with other measures to prioritize and justify investments. Thus, these measures are aggregate measures that can be system-wide, subregion, or corridor measures for strategic decision making. They can be also calculated at the intersection, sub-facility, or facility level for use in tactical decisions.

***Mobility Measures***

With regard to mobility, candidate measures at the strategic level are the vehicle-mile-traveled (VMT), vehicle-hour delay measure (VHD), and system-wide speed calculated as follows:

$$VMT = Traffic\ Volume\ x\ Segmnet\ Length\ x\ 365 \tag{4-1}$$

$$VHD = VHT - VHT\ (FF) \tag{4-2}$$

$$Systemwide\ Speed = \frac{VMT}{VHT} \tag{4-3}$$

where $VHT$ is vehicle-hours traveled including free flow travel time and delay and $VHT(FF)$ is free flow vehicle-hours traveled computed with segment free-flow speed.

System measures have also been defined by the Moving Ahead for Progress in the 21st Century Act (MAP-21). The measures defined in MAP-21 correspond to the seven national goals (FHWA, 2012). State departments of transportation (DOTs) are required to establish performance targets for a total of 18 measures identified for the seven focus areas. The MAP-21 measure corresponding to the mobility goal is the annual hours of Peak Hour Excessive Delay Per Capita (PHED) that is calculated by summing the total excessive delay of all reporting segments in the urbanized area and dividing it by the population of the urbanized area to produce the PHED measure. A variation of the PHED that is more related to the subject of this study is to calculate the measure per vehicle at the network, corridor, or facility in a peak period. In order to calculate PHED metric for each segment *t*, a speed threshold is calculated based on "an acceptable speed" of 20 mph or 60% of the posted speed limit, and the corresponding excessive delay threshold travel time is calculated as follows.

$$Excessive\ Delay\ Threshold = \left(\frac{Travel\ Time\ Segment\ Length_s}{Threshold\ Speed_s}\right) \times 3600 \tag{4-4}$$

where *s* refers to a reporting segment. Then, the delay of a segment (i.e., Road Segment Delay, RSD) is defined as the difference between the travel time at 15-minute intervals and the above excessive delay threshold travel time. The value of RSD is between 0 and 900 seconds as the maximum delay for a 15-minute calculation interval is 900 seconds. RSD is converted into the vehicle-hour by multiplying by volume to produce the Excessive Delay measure, as shown in **Equation** 4-5.

$$Total\ Excessive\ Delay_s = \tag{4-5}$$

$$AV \times \sum_{d=1}^{TD} \sum_{h=1}^{TH} \sum_{b=1}^{TB} (Excessive\ Delay_{s,b,h,d} \times (\frac{hourly\ volume}{4})_{s,h,d}$$

where $AV$ is Average Vehicle Volume, $s$ is a travel time reporting segment, $d$ is a day of the reporting year, *TD* is total number of days in the reporting year, $h$ is single hour interval of the day where the first hour interval is 12 a.m. to 12:59 a.m., *TH* is total number of hour intervals in day "h", *b* is 15-minute bin for hour interval "h", and *TB* is total number of 15-minute bins where travel times are recorded in the travel time data set for hour interval "h"

Virginia Department of Transportation extended the MAP-21 measures to include the percent vehicle miles traveled under excessively congested conditions (PECC) (Babiceanu, 2021) calculated as follows.

$$PECC = \frac{\sum_1^C VMT}{\sum_1^T VMT}\ (\%) \qquad (4\text{-}6)$$

where $T$ is the total number of reporting segments and $C$ is the total number of segments with excessive congested conditions (i.e., operating speed is lower than a threshold such as 20 mph or 60% of the posted speed limit).

***Reliability Measures***

Although reliability measures are usually assessed at the segment, facility, or route level, they have also been assessed as system measures. The Second Strategic Highway Research Program (SHRP 2) L08 team (Zegeer et al., 2014) recommended the use of the 80th percentile highest travel time as a systemwide travel time index (TTI) according to the following calculation:

$$80\%\ TTI\ = \frac{VHT\ (80\%)\ /\ VMT\ (80\%)}{VHT\ (FF)\ /\ VMT\ (FF)} \qquad (4\text{-}7)$$

where 80%TTI = 80th percentile travel time index, VHT(80%) = 80th percentile highest vehicle hours traveled among the scenarios evaluated, VMT(80%) = vehicle miles traveled for scenario with 80th percentile highest vehicle hours traveled among scenarios evaluated, VHT(FF) = vehicle hours computed with segment free flow speeds, and VMT(FF) = vehicle miles traveled with segment free-flow speed. The SHRP 2 L08 project also defined a reliability rating, calculated as the percentage of VMT with 80th percentile TTI less than 1.33 for freeways and less than 2.5 for arterials. The definition of the TTI is presented later in this document.

MAP-21 recommended using the LOTTR to calculate the Travel Time Reliability Measure (TTRM) as a systemwide measure based on probe data. The LOTTR is calculated for each peak period as the $80^{th}$ percentile travel time divided by the $50^{th}$ percentile travel time. The travel time reliability measures are calculated using the following expression.

$$Travel\ Time\ Reliability\ Measure\ (TTRM) = 100 \times \frac{\sum_{i=1}^{R} SL_i \times AV_i \times OF_j}{\sum_{i=1}^{T} SL_i \times AV_i \times OF_j} \qquad (4\text{-}8)$$

where $R$ is the total number of reporting segments with a LOTTR less than 1.5 during all time periods, $T$ is the total number of reporting segments, $SL_i$ is the length of reporting segment $I$, $AV_i$ is the total annual traffic volume calculated as the AADT reported, and $OF_j$ is occupancy factor for vehicles in a geographic area $j$. A variation of the TTR measure can be used such as utilizing the measure recommended in the SHRP 2 L08 above instead of the LOTTR when calculating the TTRM.

The VDOT effort mentioned earlier (Babiceanu, 2021) recommended the use of the Vehicle-Hours of Delay during Unreliable Conditions (UD) as a strategic measure. This measure is calculated as follows:

$$UD = \frac{\sum_{1}^{(T-R)}(\frac{VMT}{OS} - \frac{VMT}{DTS})}{N} \tag{4-9}$$

where $T$ is the total number of reporting segments, $R$ is the total number of reliable segments (i.e., LOTTR < threshold), $VMT$ is Vehicle-Miles Traveled, $OS$ is operating speed (mph), $DTS$ is delay threshold speed (mph), and $N$ is the number of days).

***Safety Measures***

The ITS Strategic Plan of the FDOT (FDOT, 2017) specifies the use of crash rates and crash severity, as strategic measures. Thus, the crash rate in crashes per million vehicle miles (MVM), number of crash fatalities and severe injuries, and rate of fatalities and severe injuries per MVM. Similar measures are used by MAP-21, which defined the safety measures as the number of fatalities per VMT, the rate of serious injuries per VMT, and total number of fatalities and serious injuries. States are required to set targets for these measures and develop plans to achieve them.

## 4.2 Tactical Measures

These are measures used by tactical managers, but they can also be used at a fine-grained level by operation managers. Tactical managers use these measures to provide additional information to inform their investment decisions. Operation managers use these measures as a starting point to identify locations with operational issues that need further analysis using the performance measures described in the Operation Measures section.

***Mobility Measures***

Agencies use facility and segment speed as a key performance indicator to determine the level of service (LOS) of the urban street segments according to the Highway Capacity Manual (HCM) (National Academies of Sciences, 2022). The HCM procedures estimate the LOS at the urban facility and segment level based on the calculated or measured average speed and at the signalized intersection movement level based on control delay. Agencies also use a variety of other mobility measures including travel time and delay, number of stops, queue length, and throughput in their assessments of system performance.

***Reliability Measures***

There are several performance measures that have been used to quantify reliability. The SHRP 2 L05 guidance (Vandervalk et al., 2014) encourages agencies to estimate multiple reliability performance measures because different measures capture different aspects of the travel time distribution and may suggest different strategies to employ. The SHRP 2 LO3 project (Cambridge Systematics et al., 2013) evaluated six reliability metrics including the buffer index, on-time performance, 95th percentile PTI, 80th percentile TTI, skew statistics index, and misery index. The L03 project documentation reported that the 95th percentile TTI may be too extreme a value to be significantly influenced by operations strategies, and that the 80th percentile TTI is more sensitive to these improvements and thus more appropriate for assessing the reliability impacts of these strategies. The project also pointed out situations, in which the use of the buffer index may not provide a good indication of the change of reliability due to improvement alternatives.

The SHRP2 L02 project (List et al., 2014) recommends using the cumulative density function (CDF) diagrams of travel time rate in sec/mile as a primary reliability assessment tool to identify the reliability performance under different traffic regimes and influencing factors.

***Safety Measures***

The analysis of safety performance at the facility, segment, and intersection levels has been done Highway Safety Manual (HSM), which provides quantitative methods to estimate and predict crash frequencies and severities to support highway safety management and safety-based decision-making (AASHTO, 2010, AASHTO, 2014). These methods and associated performance measures are used to identify sites that have safety issues, the causes of these issues, solutions to decrease crash frequency and/or severity, and safety improvement project prioritization. In estimating and predicting safety performance, the HSM uses the average crash frequency by severity type of a network, facility, or individual site (AASHTO, 2010, AASHTO, 2014). Three-to-five-year crash data are used in the estimation of these measures. In addition to crash frequency, safety studies use segment crash rates per MVM and intersection crash rates by MVM entering the intersection. Both of these measures can also be calculated by severity level.

## 4.3 Operational Measures

Traffic operation managers, engineers, and operators can make off-line and real-time decisions based on additional measures, potentially combined with some of the measures described above. Collectively, these measures (the combinations of the measures in this section and the measures in the previous section) are referred to as ATSPM. However, the measures in this section are higher resolution than the measures discussed in the previous section.

***ATSPM Measures based on Controller Data***

The ATSPMs reviewed in this section are those computed using high-resolution controller data. These measures were defined in an integrated manner in 2014 studies conducted by Purdue University (Day et al., 2016, Day et al., 2014). These measures can be calculated using an open-source software that was originally developed by Utah Department of Transportation (UDOT) or by commercial tools produced by third-party vendors and most traffic signal control software vendors. Performance measures that can be obtained from high-resolution controllers can be categorized in four categories: Capacity Utilization, Progression Performance, Multi-Modal Measures of Effectiveness, and Maintenance Support Measures. Below is an overview of these measures.

**Capacity Utilization Performance Measures:** The capacity utilization measures allow the assessment of the green time assignment to accommodate the demands of each movement. These measures are derived based on the cycle length (C), green times (G), turn movement counts, and occupancy measurements. Examples of such derived measures for each movement include the green over cycle (g/C) ratio (phase split) distribution, phase termination per type, phase duration distribution (force-off vs. gap-out), number of phase activations and percentage of cycles with the phase activated, cycle time by time-of-day, volume/capacity (v/c) ratio (requires advanced detectors), degree of saturation, green to occupancy ratio (GOR), approach delay, queue length, yellow and red actuations, and split failure. Please note that the GOR is a measure that is supposed to reflect phase utilization as it reflects the stop bar detector occupancy during the green interval. However, it was recommended that this measure is not used for a measure of the degree of utilization but an indicator of high green utilization (e.g., when GOR is above 0.8). A split failure measure can be used to report the statistics of the termination of a phase before serving the demand in the cycle. The indication of a phase failure is a combination of a green phase exceeding a certain occupancy ratio followed by a red phase exceeding a certain occupancy ratio. A threshold occupancy of 80% for the entire duration of the green phase and 80% for the first five seconds of the red phase is usually used (Freije et al., 2014).

Progression Performance Measures: These measures can be used to assess the progression of main street movements. The computation of these measures requires advanced detectors. The Arrival on Green and Arrival on Red measures reflect the percentage arrivals on green and percentage arrivals on red. The measures can also be related to the percentage of green and percentage of red in the cycle to determine if there is a clustering of arrivals in the red or green intervals. The Platoon Ratio (Rp) and Arrival Type (AT), as defined in the HCM (National Academies of Sciences, 2022) can be used additional progression quality measures that.

**Multi-Modal Measures:** These measures include the wait time of pedestrians from the time that the push button is pushed to the time that the pedestrian phase is served. In addition, these measures include preemption and priority statistics such as the preemption provided for railroad crossings, drawbridges, or emergency vehicles; and the priority provided for transit or freight vehicles.

**Maintenance Support Measures:** These measures enable the identification of detection equipment and communications malfunctions. For example, data completeness can be assessed as the percentage of data available to quantify any data loss due to communication failure. Detection problems can also be identified based on the operation measures such as too many max-outs, force-offs, or ped-calls after midnight. A low amount of data records collected by detection channels is also an indication of problems.

***ATSPM Measures based on High Resolution Probe Data***

There are several vendors who act as mobile data aggregators and provide tools that produce high resolution traffic operation measures based on data. At least one provider developed a tool that uses data from another data aggregation vendor to report and visualize the measures. Some of the developed tools can be used for off-line analysis and others can be used off-line and in real-time. Depending on the vendor, these measures include both mobility and safety measures. Example of the reported measures include estimated volumes, travel time/delays, queue lengths, reliability measures, split failures, phase termination type, arrival on green, the percent of vehicles without a stop recorded at the intersection, hard brakes, red light running, approach speed, dilemma zone entry, fuel consumption, vehicle emissions, and vehicle trajectory plots superimposed on the intersection signal timing diagram. Split failures are estimated as the number of occurrences of a green signal failing to meet the vehicle volume demands, resulting in a vehicle stopping more than once at a traffic signal. Obviously, the accuracy of the calculations of these measures depends on the specific measure being estimated and the sample size utilized by each vendor in estimating that measure.

***Surrogate Safety Measures***

There are several issues that make crash-based HSM type of safety analysis and measures too macroscopic for operation level analysis. The HSM focuses mainly on the impacts of geometry impacts and do not consider the impacts of traffic operation scenarios; unique geometry features;

environmental variables; transportation management strategies; emerging vehicle technologies such as CVs, AVs, and cooperative vehicles; or the interactions between those factors. In addition. crash-based analysis methodologies require at least three years of crash data, which causes latency in obtaining analysis results. Other issues are associated with the HSM type of safety analysis that are beyond the scope of this study. Surrogate safety measures (SSMs) can be used to support safety analysis at the operation level.

SSMs can be categorized as distance-based proximity measures, time-based proximity measures, deceleration-based measures, energy-based measures, and other relevant categories. These measures could be estimated for different scenarios such as different congestion levels, incidents, adverse weather, and work zones. National Cooperative Highway Research Program (NCHRP) Research Report 1069 provides information on the use of SSMs based on real-world data and simulation for safety analysis (Porter et al., 2023). (Hadi et al., Forthcoming-b) reviewed several studies and provided a list of SSMs based on the review, as shown in Table 4-1.

**Table 4-1 List of Safety Surrogate Measures (SSMs) (Source: (Hadi et al., Forthcoming-b))**

| SSM Type | SSM Examples |
|---|---|
| Distance-based proximity measures | Safe stopping distance<br>Difference in spacing and stopping distance<br>Proportion of stopping distance<br>Margin to collision<br>Potential index for collision with urgent deceleration<br>Predicted minimum distance<br>Time-exposed rear-end crash risk |
| Time-based proximity measures | Headway<br>Critical gap<br>PET<br>Initially attempted PET<br>TTC<br>Relative TTC<br>TTC<br>TTCD<br>Time-exposed TTC indicator<br>Time-integrated TTC (TIT)<br>Individual braking time risk<br>Time-to-object measures<br>Time to accident |
| Deceleration-based proximity measures | Deceleration rate to avoid crash<br>Modified deceleration rate to avoid crash<br>Crash-potential index<br>Criticality index function<br>Rear-end collision risk index<br>Deceleration to safety time<br>Jerks |

| **SSM Type** | **SSM Examples** |
|---|---|
| Energy-based proximity measures | Conflict index<br>DeltaV<br>Collision energy measure |
| VRU measures | TTC |
| SSM for AVs | TTC<br>PET<br>Instantaneous safety metric<br>Inertial measurement unit<br>Lane-changing conflicts<br>Number of critical jerks<br>Maximum speed of the vehicle<br>Time-exposed TTC indicator<br>TIT<br>Time-exposed rear-end crash risk |
| Other measures | Aggregated crash index<br>Aggressive lane changing<br>Average damping ratio<br>Driving volatility<br>*J*-value<br>Key risk indicator<br>Speed differential<br>Standard deviation of lateral position |

# 5 REQUIRED DATA

This section presents a discussion of the data sources that can be used to calculate the various performance measures discussed in the previous section.

## 5.1 Detector Data

Other than stop line detectors used for signal control, detectors installed on urban arterials can be “mid-block,” advance detectors, or exit detectors that are installed either to collect volume and speed data or as part of real-time traffic control such as adaptive signal control. Detectors have been widely installed on freeways to provide demands, capacity, speed, and occupancy. For arterial applications, such detectors can be very helpful, particularly to estimate demands and capacity/throughput. The demand measurements are also used to estimate some of the strategic mobility and reliability measures described earlier. If correctly positioned, these detectors can be very helpful in detecting queue length, shockwave extents, and incidents based on speed and occupancy. Advance detectors can also be used to collect data to assess the quality of progression. The most commonly used types of traffic detectors are microwave detectors (MVDS), inductive loops, and video image detectors.

Unfortunately, due to cost constraints, the traffic detection coverage (other than stop line detection) is very low for urban street and ramp segments. Expanding detection locations to include ramps and selected critical arterial segments will result in the ability to develop more powerful models. There has been some research to identify the optimal detector placement to support the identification of traffic patterns based on limited number of detectors.

## 5.2 Automatic Vehicle Re-Identification

The estimation of travel time on the subject facility and alternative routes is critical to AI-enabled TSM&O applications. Origin-destination (O-D) demands, and the routes used between O-D pairs can be very helpful for some applications like diversion due to incidents and work zones. Several types of vehicle re-identification technologies have been used to estimate the above measures. Bluetooth readers have become the most widely used technology in the past 15 years due to their lower costs, installation flexibility, and lower privacy concerns compared to other technologies. However, there is potential for detection inaccuracy with this technology due to the large detection radius, which can result in detecting vehicles more than once, detecting vehicles on adjacent roads, and errors in the estimation of travel time for very short links. It can also be difficult to differentiate between signals from passenger cars and other modes such as trucks, buses, bicycles, and pedestrians. Some of these concerns have been addressed through data filtering and antenna settings.

## 5.3 Private Sector Data

Private sector vendors have also provided travel times, proportions of demands between O-D pairs, and the utilized routes between these pairs. Thus, private sector data have been considered as the main alternatives to automated vehicle rematching technologies like Bluetooth readers. Private sector data are generally collected from vehicles and/or mobile devices. Sufficient sample sizes of the collected data are helpful to provide better accuracy. Some of the vendors combine mobile data with data from other sources, such as traffic detectors and signal control, in their estimation of performance measures. There have been concerns in using travel time data from the private sector, particularly for real-time operations on urban arterials, due to the smaller sample size and the potential smoothing of the data. However, there has been an increase in the utilized data, and this increase is expected to continue, resulting in data quality improvements.

## 5.4 High Resolution Probe Data

As stated earlier, private sector vendors have provided travel times, the proportions of traffic demands between origin–destination pairs and the routes chosen between these pairs based on vehicle tracking (probe vehicles). In recent years, vendors have started providing more microscopic vehicle trajectory data including information such as the date, time, heading, speed, acceleration/deceleration, and location of hard-braking events. The vendors obtained the high-resolution data from original equipment manufacturers (OEM) and/or telematics providers. Some vendors developed tools that use this type of data to estimate high resolution ATSPM measures and SSMs, as discussed in the Performance Measurement section of this document. Two FHWA efforts (Pack et al., Fortcoming, Zhou et al., 2024) provided reviews of high-resolution probe data sources. The reviews listed additional features that can be obtained from some of this data including traction control engagement, antilock braking engagement, airbag deployments, stability indicators (i.e., rollover detection), wiper use, headlight use, fuel data, number of passengers, odometer readings, temperature, and road-surface conditions (e.g., pothole detection or road smoothness indicators). Obviously these measures can be used as inputs to AI modules that predict safety risks on urban arterial.

## 5.5 Crash Data

Crash datasets provide detailed information used in safety analyses and can be used as inputs to AI applications to analyze both TSM&O impacts on safety and crash impacts on management and operations, particularly in the absence of detailed incident data on arterial streets (see next section). In addition, the crash data can include details not available in the incident database such as the number and severity of injuries, number of involved vehicles and vehicle types, involved person type, cause of the crash, traffic law violation, and alcohol or drug impairment. These data are usually archived based on crash reports, commonly referred to as police accident reports, filed by State highway patrol or local police officers at the crash scene. In Florida, this data is included in the Florida Highway Safety and Motor Vehicles (FLHSMV) traffic crash database, that contains official finalized data and statistics for all reported traffic crashes. However, the finalization of this database takes a year or longer to complete, which may not be acceptable for some applications, including TSMO applications. Fortunately, another available database is available in Florida that does not have this latency. This database referred to as the Florida Signal4 Analytics, which as a web-based system that includes crash records, citations, and local traffic volumes where available.

## 5.6 Incident Data

The FDOT districts collect and archive detailed incident characteristics data that include incident type, location, timeline (detection, verification, responding agency arrival, lane clearance, and incident clearance timestamps), severity, involved number and types of vehicles, and primary vs.

secondary crash categorization. Off-line, the incident data can be obtained from traffic management center data archives and/or crowdsourced incident data providers such as Waze. In real-time, the data can be obtained from different sources including public enforcement agency notification, automated incident detection using detectors or video analytics, service patrols. incident data providers, and so on. Unfortunately, unlike the case with urban freeways managed by FDOT districts; the detection, recording, and archiving of incidents and their attributes on arterial streets is very limited. In the absence of incident data, crash data reviewed in the previous section can provide some of the information included in incident databases, although the crash databases lack a lot of important information such as the incident and lane blockage timeline and response activities timeline. There is a need for the expansion of incident management, data sharing, and information archiving on arterial streets to allow additional applications of AI to traffic management. The research team of this study attempted in previous effort to match crash records from the FLHSMV crash database and FDOT incident database was not successful. This needs to be further investigated in future study.

## 5.7 Weather Data

Adverse weather conditions such as rain and fog can reduce capacity, increase travel times, and increase the likelihood of incidents. Thus, weather data can be used as inputs to AI models to consider these impacts. Such data can be obtained from publicly available weather data or through subscriptions to detailed weather forecasts. Road Weather Information Systems (RWIS) have also been described as one of the potential intelligent transportation systems to be deployed by agencies and can add value. However, the installations of RWIS in Florida have been limited. The weather from national weather stations may not be in the vicinity of the corridor, which may reduce the accuracy of weather data. An interesting potential approach is to confirm the accuracy of weather data applied to a corridor based on the time stamps of changes in traffic patterns and changes in weather parameters such as rain precipitation.

## 5.8 Construction (Work Zone) Data

Work zone data are also needed for relevant AI applications. This data may include the anticipated and actual construction date/time, location of the start and end work zone, number of blocked lanes, number of workers at work zones, types of equipment utilized, posted speed limit, and other work zone features. Unfortunately, most existing work zone lane closure databases transportation agencies, where they exist, include the scheduled but not the actual lane closure, which is often very different from the scheduled one. Data analytics and traffic analysis will benefit from investment in lane closure databases that are dynamically and accurately updated and shared for use in off-line and real-time use. In this regard, it is useful to review and consider the use of the Work Zone Data Exchange (WZDx) Specification maintained by the United States Department of Transportation (USDOT) that supports ubiquitous access to data on work zone activity.

## 5.9 High-Resolution Controller Data

The high-resolution controller performance measures described earlier in this document are calculated based on signal timing and detection events collected at the highest time resolution of the signal controller (0.1 seconds). Simple measures like green split monitor and phase termination type, and pedestrian actuations do not need vehicle detectors. Other measures like green occupancy ratio and split failure can be calculated based on stop line detectors. Setback or exit detectors are needed to estimate traffic demands, arrival on green, and progression quality.

### 5.10 Video Analytics-Based High-Resolution Data

Video image detectors (VIDs) have been widely used for stop line and advanced detection in lieu of inductive loop detectors. Some vendors also provide the ability to measure turn movement counts based on the image processing and upload these measurements for archiving by the transportation agencies. In recent years, commercially available products have become available to track vehicles to collect trajectory data and conflict measures based on VIDs, allowing the identification of both pedestrian–vehicle and vehicle–vehicle SSMs and more detailed signal performance measures that can be used in combination with AI to classify and predict the safety and operations performance. For example, (Mishra et al., 2022) used trajectory data and braking events from video analytics combined with signal phasing data from traffic controllers to detect and analyze conflict hotspots and evaluate the efficacy of countermeasures to improve intersection safety.

### 5.11 Lidar-Based High-Resolution Data

It was reported that the accuracy and range of the data collected by VIDs can be affected by weather and lighting conditions. Lidar was used as an alternative to video analytics in collecting trajectory and conflict data. An evaluation done by (Guan, 2022) showed that lidar has a longer detection range than VIDs is not affected by poor lighting conditions. Vasudevan, Agarwala, and Several studies also developed and evaluated systems for the use of lidar data to identify vehicle–pedestrian near-crash identification.

### 5.12 Connected Vehicle Data

Data from agency's connected vehicle (CV) deployments can be used as an important input to data analytics with the increase in CV market penetration. The basic safety message (BSM), defined in SAE J2735™ standards, consists of two parts (SAE International Technical Standard, 2015). The mandatory part of the message (BSM Part 1) is to be broadcasted 10 times per second by vehicles and contains vehicle position, heading, speed, acceleration, steering wheel angle, and vehicle size. BSM Part 2 is optional and contains elements such as precipitation, air temperature, wiper status, light status, road coefficient of friction, antilock brake system activation, traction control system activation, and vehicle type. When archived, data from other message types can also provide important inputs including personal safety messages, traveler information messages, signal phase and timing messages, and MAP messages. This data can be used to support the estimation of arterial and signal performance measures, as well as safety surrogate measures.

### 5.13 Summary

Based on the discussion presented in this document, Table 5-1 presents a summary of the strategic, tactical, and operational measures that may need to be estimated and used as inputs or outputs of the developed AI/machine learning models depending on the purpose, type, and the features of the models. As stated earlier, some of the strategic measures can also be used as tactical measures but usually at higher resolution, for example using measures at the segment level in making tactical decisions compared to using area or corridor levels in making strategic measures. Tactical measures at higher resolutions can be also used to inform operational analysis and decision-making processes. Table 5-1 also includes a summary of the data sources that can be used to calculate the measures. Next to each data source, between two brackets, there is an indication of the category of the measure that could be calculated based on the source, with S indicating strategic measures, T indicating tactical measures, and O indicating operational measures.

**Table 5-1 Summary of Performance Measures and Data Sources Identified in this Study**

| Measure Category | Strategic Measures | Tactical Measures | Operational Measures | Utilized Data |
|---|---|---|---|---|
| Mobility | VMT<br>VHD<br>System-wide speed<br>PHED (per vehicle)<br>PECC | Throughput<br>Travel Time<br>Delay<br>Queue Length<br>Number of Stops<br>LOS | Cycle Length<br>Green Splits<br>Movement Counts<br>Phase termination type<br>Phase activations<br>v/c Ratio<br>GOR<br>Approach Delay<br>Queue Length<br>Split Failure<br>Arrival on Green<br>Platoon Ratio/Arrival Type<br>Pedestrian Wait Time<br>Pedestrian Activations | Sensor Data (S, T, O)<br><br>AVI (S, T, O)<br><br>Private Sector Probe (S, T, O)<br>High Res. Probe (O)<br><br>High Res. Controller (O)<br><br>VIDs and Lidar (O)<br><br>CV Data (O)<br><br>Crash, Incident, Weather, and Construction Data (O) |
| Reliability | System 80%TTI<br>% with 80th TTI > Threshold<br>TTRM based on LOTTR<br>Unreliable Delay | TTI95%<br>TTI80%<br>Buffer Index<br>On-Time Trips<br>Skew Statistics<br>Misery Index | | |
| Safety | Crash Rate in Crashes per MVM<br><br>Number of Crash Fatalities and Severe Injuries<br><br>Rate of Fatalities and Severe Injuries Per MVM | Crash Frequency by Severity Level<br><br>Crash Rates per MVM and Intersection<br>Crash Rates By MVM by Severity Level | Distance-based SSM<br><br>Time-based SSM<br><br>Deceleration-based SSM<br><br>Energy-based SSM<br><br>Delima Zone Entry<br><br>Hard Brakes<br><br>Red Light Running | Cash Data (S, T)<br><br>Incident Data (S, T, O)<br><br>Private Sector Incident Data (T, O)<br><br>High Res. Probe (O)<br><br>Lidar and VID(O)<br><br>High Res. Controller (O)<br><br>CV Data (O) |

| Measure Category | Strategic Measures | Tactical Measures | Operational Measures | Utilized Data |
| --- | --- | --- | --- | --- |
| | | | Yellow and Red Actuations | |

# 6 FREEWAY INCIDENT DIVERSION IMPACT ON ARTERIAL STREETS

When freeway incidents occur, traffic diversion can significantly impact parallel arterial streets as drivers seek alternative routes to bypass the congested freeway segments. In most cases, these impacts are not assessed and addressed in an effective manner. The FHWA's Coordinated Freeway and Arterial Operations Handbook (Urbanik et al., 2006) highlights the need to prepare arterial street operations for freeway incident diversion to mitigate expected congestion. The interdependence between freeway and arterial operations creates complex scenarios in which the performance of one directly influences the other.

Conventional arterial management typically responds to such diversion only after congestion has already formed, rather than before it develops. This section, therefore, addresses this limitation by developing a proactive machine learning model for short-term prediction of arterial conditions during freeway incident diversions. The model provides the operating agency with a real-time warning, giving operators sufficient lead time to prepare and deploy special signal timing plans before congestion develops on the arterial.

## 6.1 Study Area

The study area is located in Boca Raton, Palm Beach County, Florida, where southbound Interstate 95 (I-95) runs parallel to Military Trail (SR-809), a major north-south arterial immediately west of the freeway. Military Trail carries heavy traffic of its own and, at the same time, serves as one of the principal alternative routes during freeway incidents: motorists divert from southbound I-95 onto the southbound Military Trail corridor and rejoin the freeway at a downstream interchange such as the Palmetto Park Road interchange or the Hillsboro Boulevard interchange.

As shown in Figure 6-1, the freeway study area covers southbound I-95 between Atlantic Avenue (Exit 52) and Hillsboro Boulevard (Exit 42). The model evaluates how traffic diverting from incidents on this I-95 corridor impacts Military Trail, which is already experiencing heavy peak-hour congestion.

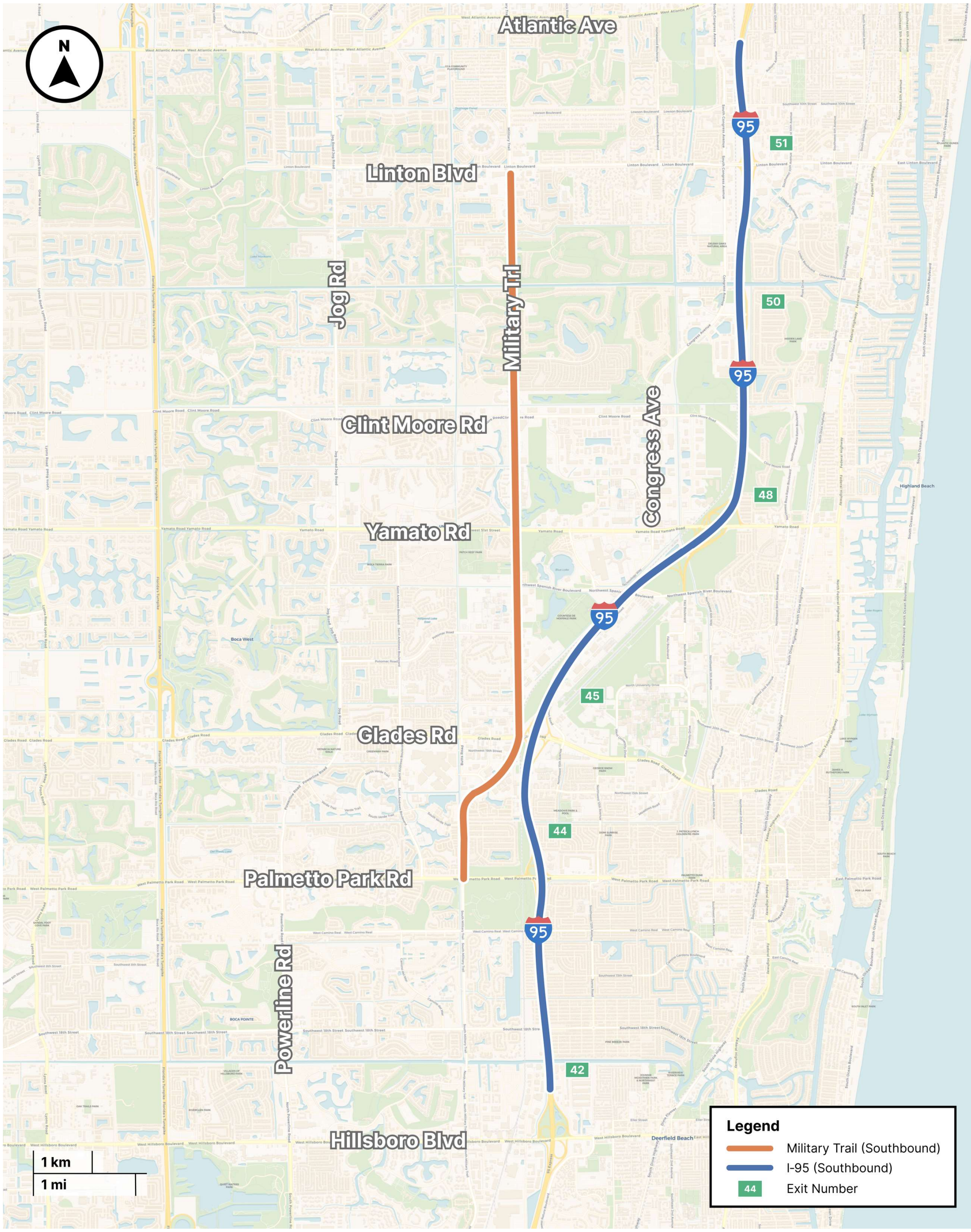



Figure 6-1 Study Area: I-95 Southbound and the Parallel Arterial Network in Boca Raton

During the January 2020 to April 2026 study period, the freeway underwent reconstruction associated with the I-95 Express Lanes program, which widened the mainline and added managed (express) lanes. The express lanes are a separate facility from the general-purpose lanes and are excluded from this study; all freeway measures and incident records used here refer to the general-purpose lanes only.

The arterial network remained structurally consistent throughout the study period, with no major changes to roadway geometry. In addition, an adaptive traffic signal control system was implemented along the Yamato Road corridor on April 2, 2024, covering twelve signalized intersections from west to east: NW 55th Diagonal, Jog Road, St Andrews Boulevard, Patch Reef Park, Military Trail, Old Course Way, Broken Sound Boulevard, T-Rex Avenue, Congress Avenue, and the three I-95 ramp terminals. While the Military Trail corridor includes the adaptively controlled intersection at Yamato Road, the remainder of the corridor operates under time-of-day (TOD) control with actuation and coordination.

## 6.2 Data Source

The predictive model developed in this section relies on three primary data categories: freeway incident records, freeway and arterial travel time estimates based on third party vendor data, and traffic sensor data. Beyond these model inputs, arterial signal operation data is used to evaluate how the existing signal control responds to incident-related diversions. This data was sourced from state-operated systems and a third-party probe vendor. This section details each data source and its role, concluding with a summary table. The specific detector selection and network segmentation are discussed in the Study Network Definition section, while detailed data screening is covered in the Data Preparation section.

- Incident Records: These records were sourced from SunGuide, the statewide traffic management system operated by FDOT and the primary platform used by traffic management centers to detect, verify, and log incidents. Each record includes the event identifier, detection time, closure time, location (latitude and longitude), lane-blockage status, event type, and weather conditions.
- Travel Time Estimates: These estimates were derived from HERE Technologies probe data and reported for directional roadway segments using Traffic Message Channel (TMC) codes for facility segmentation. For this study, the data was downloaded at a 5-minute resolution via the Regional Integrated Transportation Information System (RITIS), an archiving platform that aggregates data from different sources.
- Traffic Detector Data: These data were collected by FDOT-maintained traffic detectors on the I-95 mainline and include spot speed, volume, and occupancy measurements downloaded from RITIS.
- Signal Operation Data: These include the assigned green time splits for the signalized intersections along Military Trail. These records support the analysis of signal responsiveness to diverted traffic, as detailed in the Field Observations of Incident Diversion and Signal Response section.

The machine learning model development and training rely entirely on archived data. However, real-time operation requires live feeds from these same sources. Because the system is intended for practical operator use, data accessibility is a critical design consideration. Freeway mainline detector data and SunGuide incident records are available in real time through the Data Integration and Video Aggregation System (DIVAS), which is already part of FDOT's internal infrastructure. In addition, the HERE probe travel-time feed and access to the RITIS application programming interface (API) typically require additional subscriptions or costs. To ensure the system is practical for real-world deployment, the proposed model is specifically designed to accommodate live data inputs directly from DIVAS. Table 6-1 summarizes the data sources, their categories, real-time availability, and their specific roles in this study.

**Table 6-1 Summary of Data Sources and Their Applications**

| Data Source | Provider / Obtained Via | Data Category | Real-Time Availability (for Operators) | Role in This Study |
|---|---|---|---|---|
| SunGuide Database | FDOT | Incident records | Yes, within the traffic management center and data sharing with others | Source of incident information and features, such as onset time, location, lane blockage, event type, and related attributes |
| HERE Probe Travel Time | HERE Technologies, via RITIS | Segment travel time and speed (per TMC segments) | A real-time feed requires an additional subscription | Freeway and arterial travel time; source of the model label |
| Roadway Detectors | FDOT, via RITIS | Freeway spot speed, volume, occupancy | Through DIVAS | Freeway traffic-state inputs to the model |
| Real-Time Data Streams | FDOT (DIVAS API) | Live incident and detector feeds | Yes; credentialed API access | Real-time inputs for deployment |
| Controller Data | City of Boca Raton | Signal phase, green time, split, and adaptive control status | Yes, within the agency's signal system | Characterize the signal response to incident diversion (adaptive and TOD) |

## 6.3 Study Network Definition

The raw data described in the previous section requires further processing prior to analysis. Specifically, the corridor is equipped with multiple roadway detectors of varying quality, and the HERE travel times are reported across TMC segments. To address this, the current section addresses the selection of a single freeway detector to collect data from for each location and time period. Furthermore, this section details the aggregation of raw travel-time data into the standardized freeway and arterial segments referenced throughout the remainder of the analysis.

### 6.3.1 Freeway Detector Selection

All raw data required thorough screening to identify anomalies and ensure data quality. The detector records underwent a preliminary quality check based on raw 5-minute intervals. Readings are flagged when volume, speed, or occupancy values fall outside engineering limits. For example, data points are flagged if volumes exceed 1,000 vehicles per lane per 5 minutes (veh/ln/5-min), speeds exceed 120 miles per hour during active hours (6:00 AM–9:00 PM), occupancy exceeds 100 percent, or any values that are negative. Additionally, a zero volume during active hours is flagged, as it usually indicates a malfunctioning sensor. Given that a complete hour contains twelve 5-minute readings, an hourly quality score is calculated as the percentage of the twelve expected readings that are present and unflagged. The volume, speed, and occupancy metrics are combined into a single detector quality score.

**Figure 6-2** aggregates these quality scores by month for each detector and each individual lane from 2020 to 2025. This heatmap visualizes the timing of sensor issues with green cells indicate months with nearly complete and plausible data, while warmer colors and blank cells highlight periods of degraded or missing data. These detector measurements disrupted periods align closely with the 2021 and 2023 construction windows.


**Figure 6-2 Monthly Detector Data-Quality Heatmap by Detector and Lane (2020–2025)**

Reading the heatmap vertically for each detector shows exactly when its data became unusable. For each freeway location, the study used the detector with the most reliable record for as long as it functioned properly. Once a unit was decommissioned, it was replaced in the dataset by the next dependable detector covering the same segment.

Table 6-2 summarizes the selected freeway detectors (D1 through D6) on I-95 Southbound, detailing their milepost (MP) locations, coverage boundaries, data collection periods, and the number of general-purpose (GP) lanes.

**Table 6-2 Selected Freeway Detectors on I-95 Southbound**

| Detector | Detector MP | Coverage MP | Collection Period | No. of GP Lanes |
|---|---|---|---|---|
| D1 | 51.6 | 52 to 51 | January 2020 – April 2026 | 6 |
| D2 | 50.3 | 51 to 50 | January 2020 – June 2023 | 6 |
| | 50.1 | | August 2023 – April 2026 | 4 |
| D3 | 48.9 | 50 to 48 | January 2020 – June 2023 | 5 |
| | 48.9 | | June 2023 – April 2026 | 4 |
| D4 | 46.8 | 48 to 45 | January 2020 – June 2023 | 5 |
| | 47.0 | | September 2023 – April 2026 | 4 |
| D5 | 44.0 | 45 to 44 | January 2020 – June 2021 | 4 |
| | 44.0 | | June 2021 – April 2026 | 4 |
| D6 | 43.0 | 44 to 42 | January 2020 – May 2021 | 4 |
| | 43.0 | | June 2021 – April 2026 | 4 |

Notes: MP = milepost; GP = general-purpose. Two collection periods at one location indicate a detector replacement; the period ending in April 2026 marks the end of the predictive model's timeframe.

### 6.3.2 Network Segmentation

Roadway segments in this study are defined based on the TMC coding of the HERE probe travel-time data. On the freeway mainline, the raw HERE data divide the southbound I-95 study area into eleven TMC segments, several of which exhibit nearly identical travel-time behaviors. To eliminate redundancy, adjacent segments with travel-time profile correlations exceeding 0.80 were considered to represent identical traffic conditions and were consequently merged. This merging process involved summing up the travel times and segment lengths of the individual combined segments, followed by recomputing the travel time per mile. As illustrated in Figure 6-3**,** this grouping effectively reduced the eleven raw segments to six analytical segments (numbered in this study from FWY_1 through FWY_6) while preserving the corridor's spatial congestion patterns.

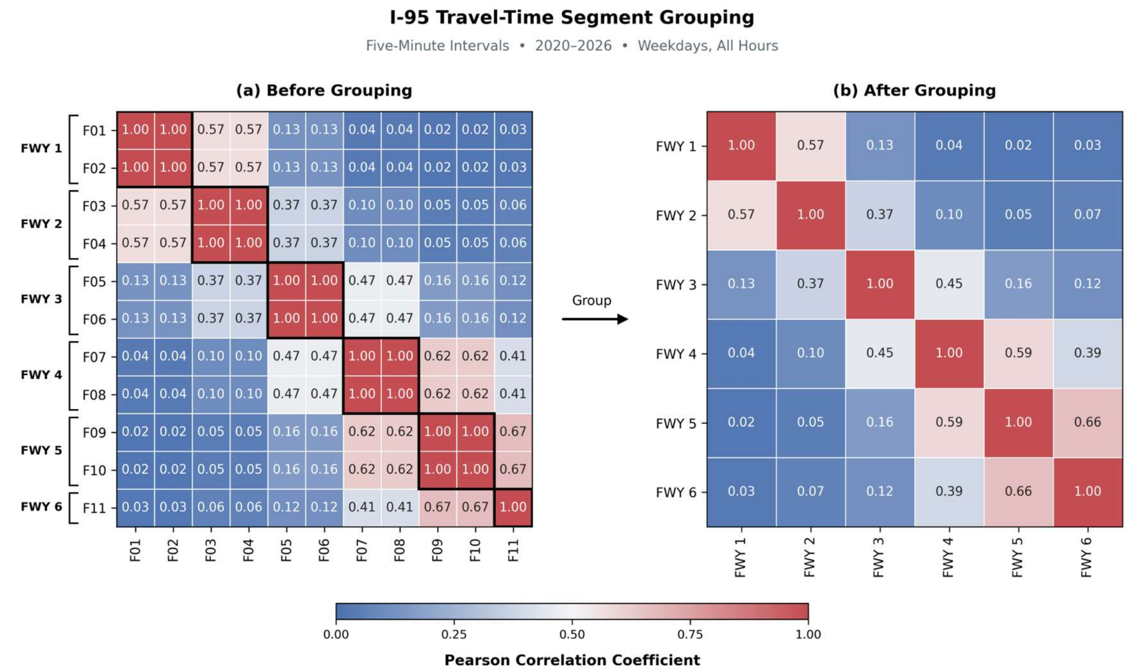



**Figure 6-3 Travel-Time Correlation Matrices Before and After Segment Grouping**

On the arterial side, the Military Trail study facility is represented by four southbound segments (numbered from MILI_1 through MILI_4). Of the Military Trail segments, the downstream segment MILI_4 is the prediction target of this study. The finalized segment configurations are detailed in Table 6-3, while the overall segment layout and detector station locations are illustrated in the study network schematic (Figure 6-4).

**Table 6-3 Definition of the Study-Area Travel-Time Segments**

| Roadway | Segment | From | To | Length (mi) |
|---|---|---|---|---|
| I-95 | FWY_1 | Atlantic Avenue (Exit 52) | Linton Boulevard (Exit 51) | 1.63 |
| | FWY_2 | Linton Boulevard (Exit 51) | Congress Avenue (Exit 50) | 1.19 |
| | FWY_3 | Congress Avenue (Exit 50) | Yamato Road (Exit 48) | 1.96 |
| | FWY_4 | Yamato Road (Exit 48) | Glades Road (Exit 45) | 2.43 |
| | FWY_5 | Glades Road (Exit 45) | Palmetto Park Road (Exit 44) | 1.12 |
| | FWY_6 | Palmetto Park Road (Exit 44) | Hillsboro Boulevard (Exit 42) | 1.56 |
| Military Trail | MILI_1 | Linton Boulevard | Clint Moore Road | 2.14 |
| | MILI_2 | Clint Moore Road | Yamato Road | 0.96 |
| | MILI_3 | Yamato Road | Spanish River Boulevard | 0.46 |
| | MILI_4 | Spanish River Boulevard | Palmetto Park Road | 2.71 |

Notes: All segments are in the southbound direction of travel. Freeway segment boundaries correspond to the I-95 interchanges, with exit numbers in parentheses.

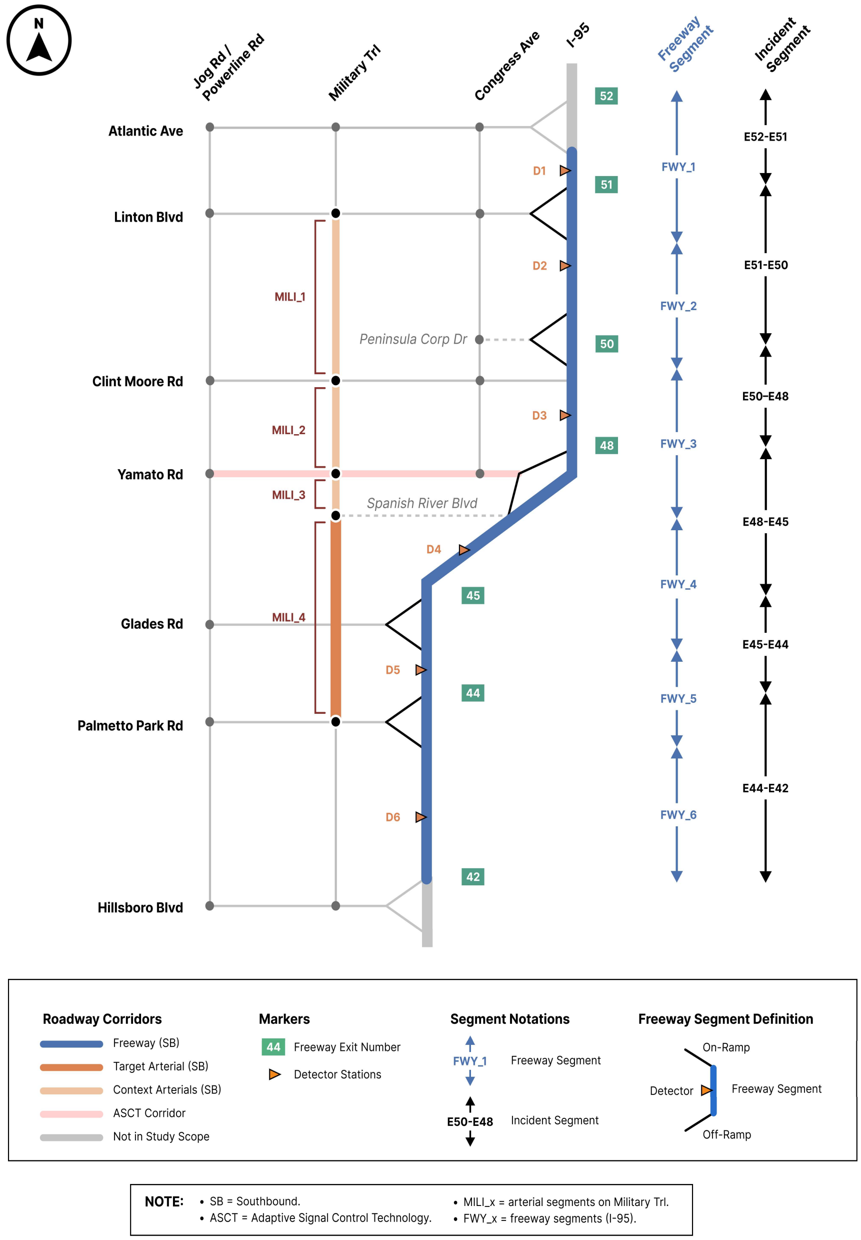



**Figure 6-4 Study Network Schematic**

## 6.4 Observations of Incident Diversion and Signal Response

This section examines the operational impact of freeway incidents by analyzing how diverted traffic affects the parallel arterial corridor. By mapping the spatial and temporal patterns of this diversion, an actionable time lag is identified between the onset of freeway congestion and subsequent arterial congestion. An evaluation of the existing signal systems further reveals their inability to mitigate the diverted demand in real time. These combined observations serve as the foundational justification for the proposed predictive model.

### 6.4.1 Incident Diversion onto the Parallel Arterials

The spatial and temporal characteristics of the diverted traffic are examined first. When a southbound I-95 incident causes freeway congestion, a portion of the demand reroutes onto parallel arterials, primarily Military Trail. This diversion pattern was observed repeatedly throughout the study period. A representative set of cases is presented here to illustrate its progression in space and time.

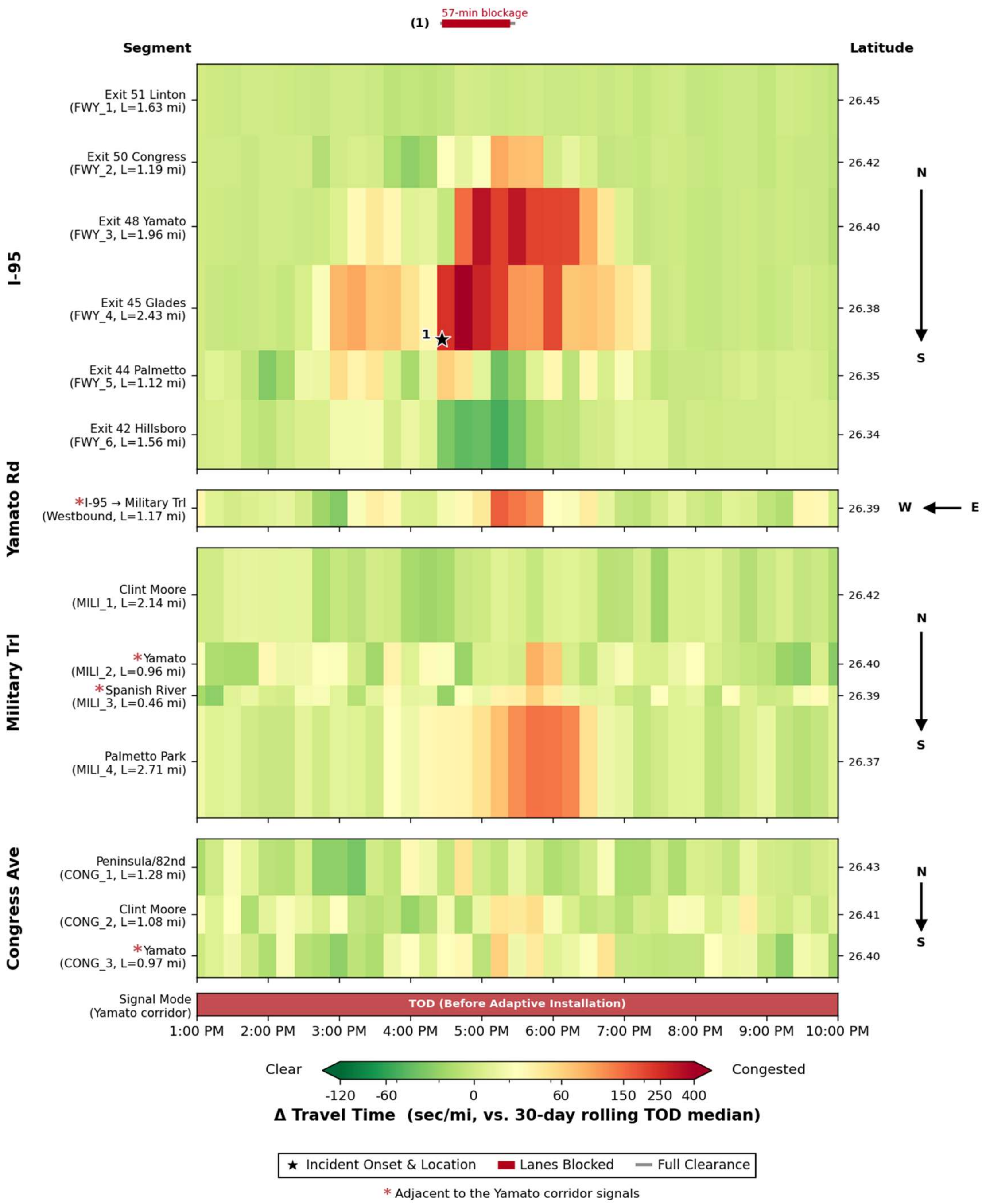
Diversion cases are categorized into two groups based on the spatial distribution of the diverted demand. **Figure 1-4** presents cases where diversion concentrates on Military Trail, while **Figure 1-5** illustrates events where the demand spreads across multiple parallel routes, including Congress Avenue. In each space-time diagram, the vertical axis represents geographic position, aligning north-to-south for freeway and arterial corridors (with north at the top) and east-to-west for the Yamato Road connecting segment. The horizontal axis denotes the time of day, while color gradients indicate travel time relative to normal conditions. Here, normal is defined as the median travel time for the same weekday and time of day over the prior 30 days. Because these visualizations isolate delay relative to this historical baseline, the identified congestion is expected to be incident-induced rather than a reflection of recurring peak-period conditions. Finally, the freeway incident is depicted as a red band on I-95 that represents congestion propagating upstream, with annotations indicating the number of blocked mainline lanes, excluding shoulders.

#### 6.4.1.1 Diversion Concentrated on Military Trail

In the first group of cases **(**Figure 6-5**(a)** to Figure 6-5**(e)),** the operational impact of the diversion is heavily concentrated on Military Trail. Following the onset of freeway congestion, elevated travel times materialize on the Military Trail segments after a distinct time gap, while Congress Avenue generally remains near its baseline conditions.

The event on January 10, 2020, illustrated in Figure 6-5**(a),** is a clear example of freeway incident diversion. A single crash at 4:26 PM blocked three of five mainline lanes near Palmetto Park Road (Exit 44), and the freeway queue grew upstream past the Glades Road and Yamato Road interchanges. The space-time diagram captures the diversion path itself: roughly an hour after the crash, the delay appeared on the short westbound Yamato Road segment that connects the freeway to Military Trail, where vehicles leaving the freeway at Exit 48 turn south onto Military Trail, and on Military Trail in the same interval. The largest and longest arterial impact fell on the segment between Spanish River Boulevard and Palmetto Park Road, which is MILI_4, the target segment of the prediction model.

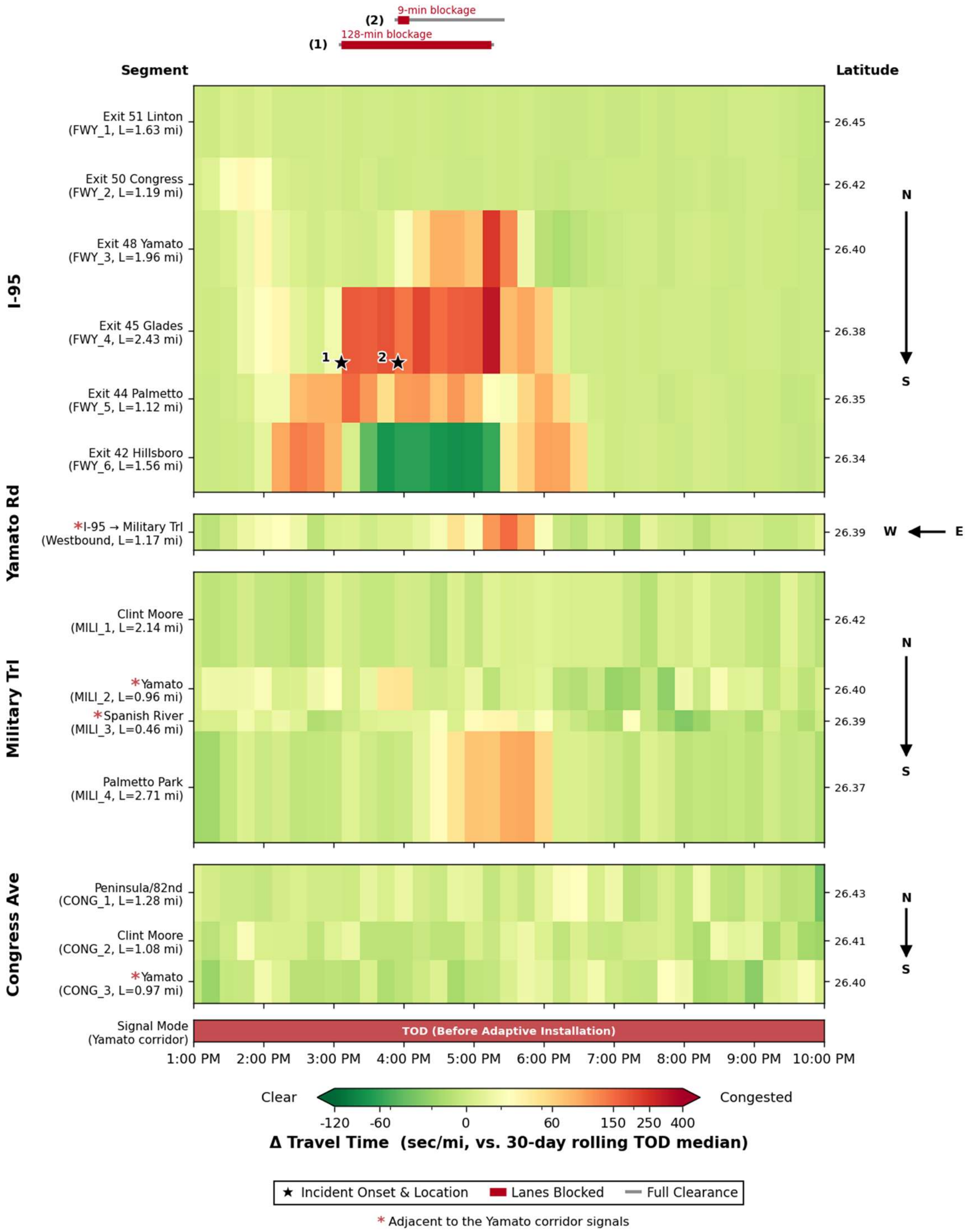
The remaining cases **(**Figure 6-5**(b)** through Figure 6-5**(e))** demonstrate that this sequential pattern persists across diverse operational conditions. Triggering events range from minor crashes to severe vehicle fires and emergency vehicle incidents, with lane blockages varying from partial to full closure of the general-purpose lanes. The lane closures that produced this congestion events lasted from roughly 40 minutes to over two hours, with full clearance frequently taking several hours longer.




**Incidents on Friday, January 10, 2020**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 4:26 PM | Before Exit 44 (FWY_4) | Crash | 3 of 5 (60%) | 57 min | 62 min |

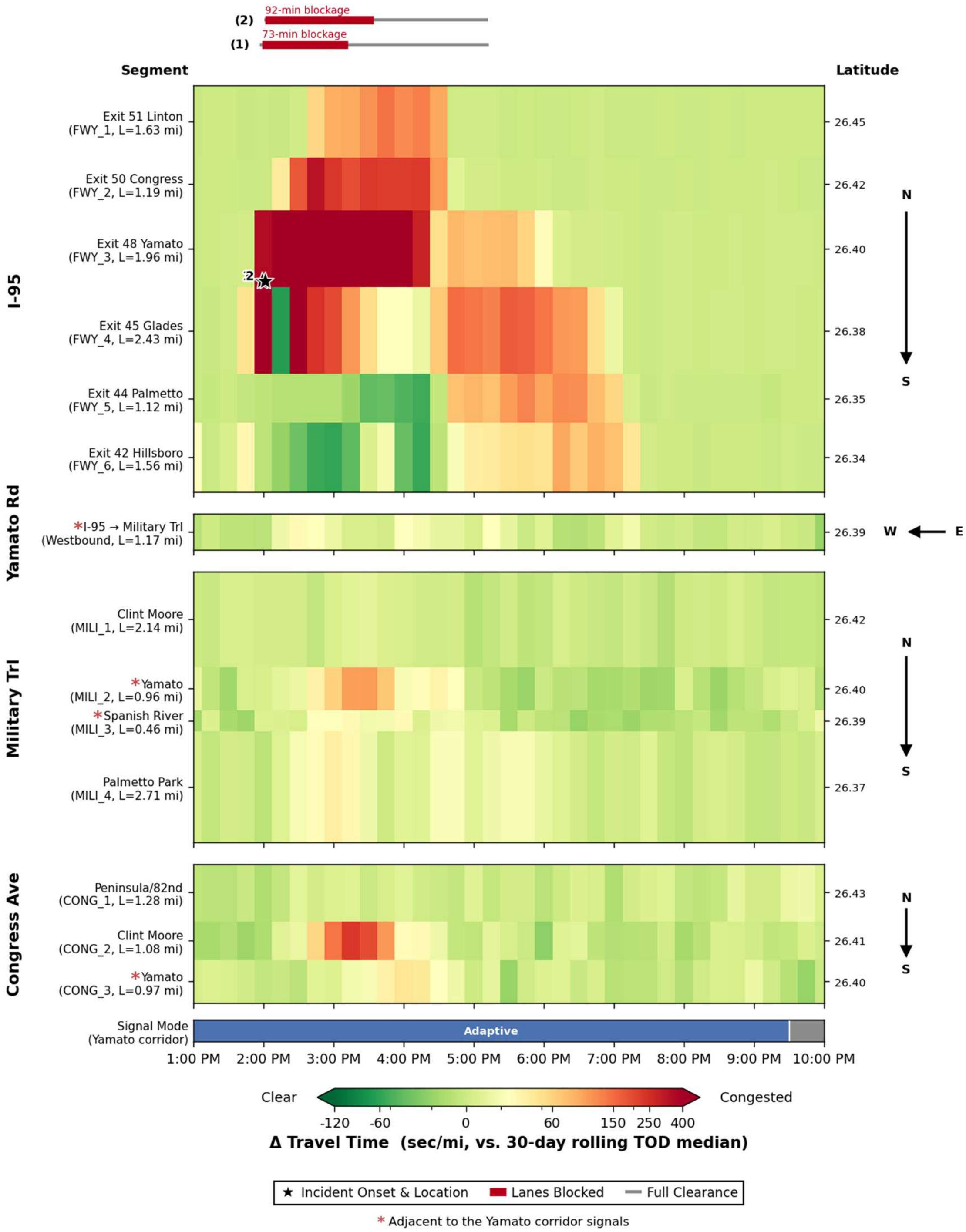
**Figure 6-5(a) Diversion Space-Time Diagram, January 10, 2020 (Friday)**




**Incidents on Tuesday, August 1, 2023**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 3:06 PM | At Exit 44 (FWY_4) | Crash | 2 of 5 (40%) | 128 min | 133 min |
| 2 | 3:54 PM | Before Exit 44 (FWY_4) | Crash | 1 of 4 (25%) | 9 min | 93 min |

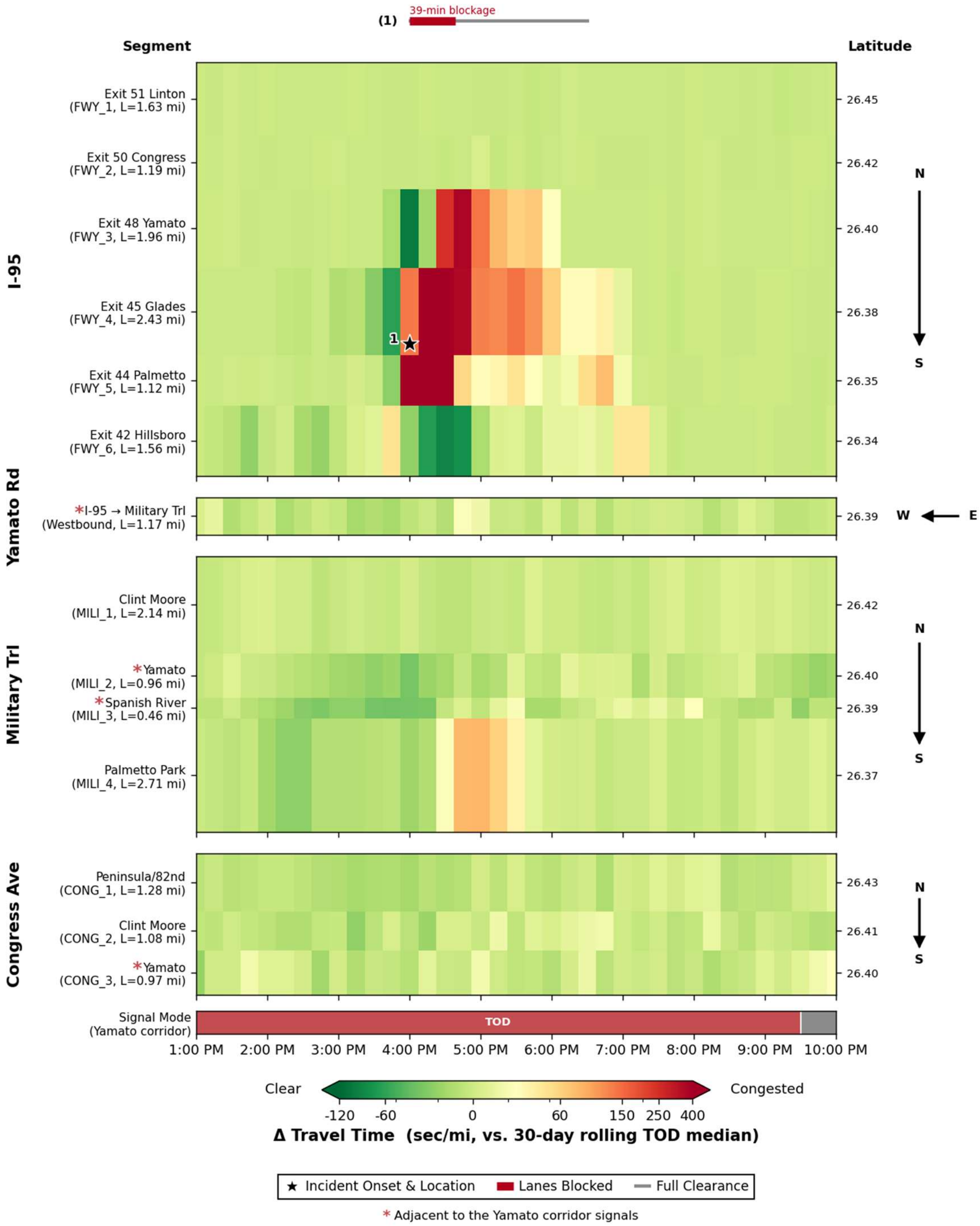
**Figure 6-5(b) Diversion Space-Time Diagram, August 1, 2023 (Tuesday)**




**Incidents on Friday, May 17, 2024**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 1:58 PM | Before Exit 45 (FWY_3) | Vehicle Fire | 4 of 4 (100%) | 73 min | 196 min |
| 2 | 2:01 PM | Beyond Exit 48 (FWY_3) | Crash | 3 of 5 (60%) | 92 min | 191 min |

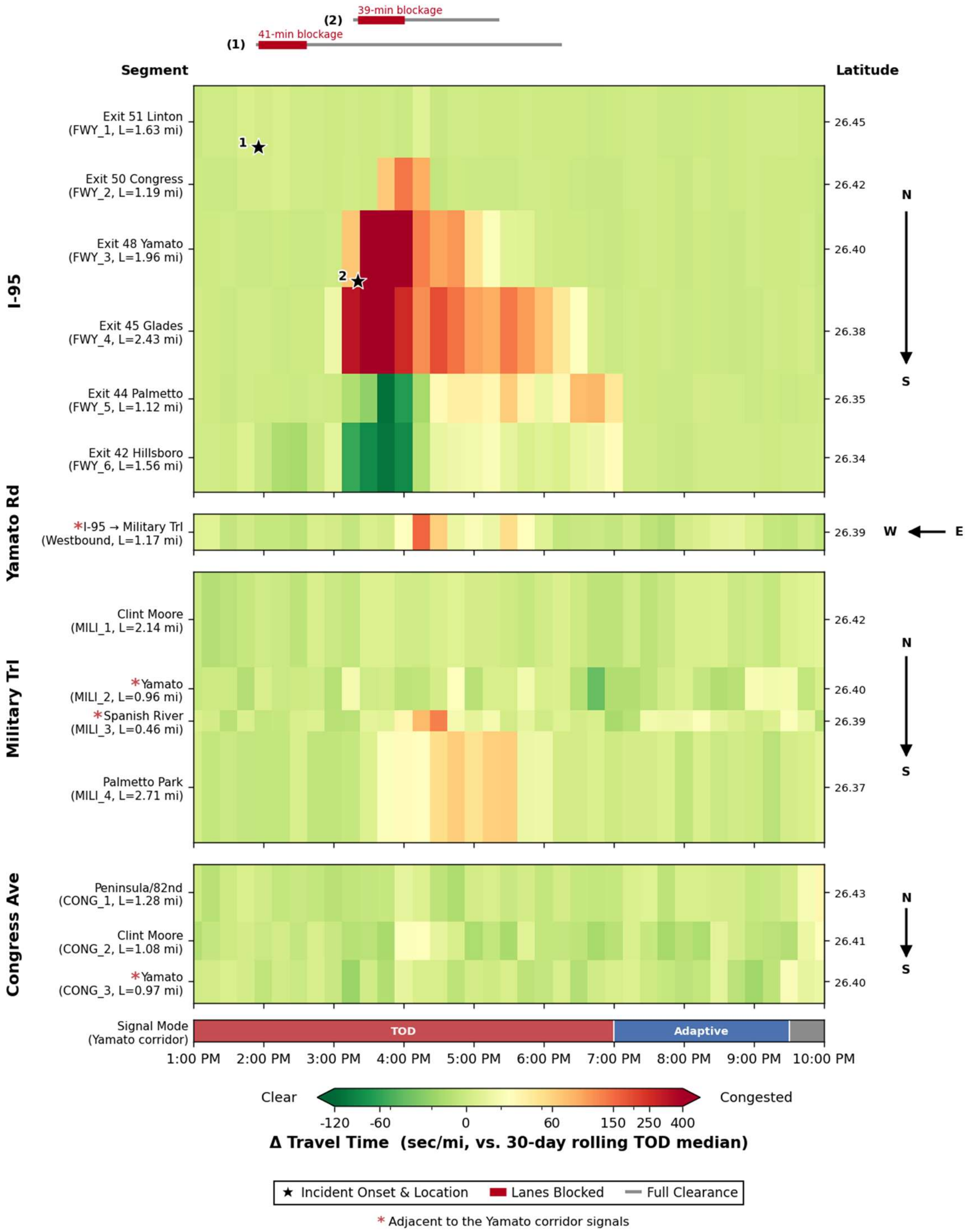
**Figure 6-5(c) Diversion Space-Time Diagram, May 17, 2024 (Friday)**




**Incidents on Friday, May 31, 2024**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 3:59 PM | At Exit 44 (FWY_4) | Emergency Vehicles | 3 of 4 (75%) | 39 min | 152 min |

**Figure 6-5(d) Diversion Space-Time Diagram, May 31, 2024 (Friday)**




**Incidents on Wednesday, April 30, 2025**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 1:55 PM | Beyond Exit 51 (FWY_1) | Crash | 2 of 5 (40%) | 41 min | 261 min |
| 2 | 3:20 PM | At Exit 45 (FWY_3) | Crash | 4 of 4 (100%) | 39 min | 124 min |

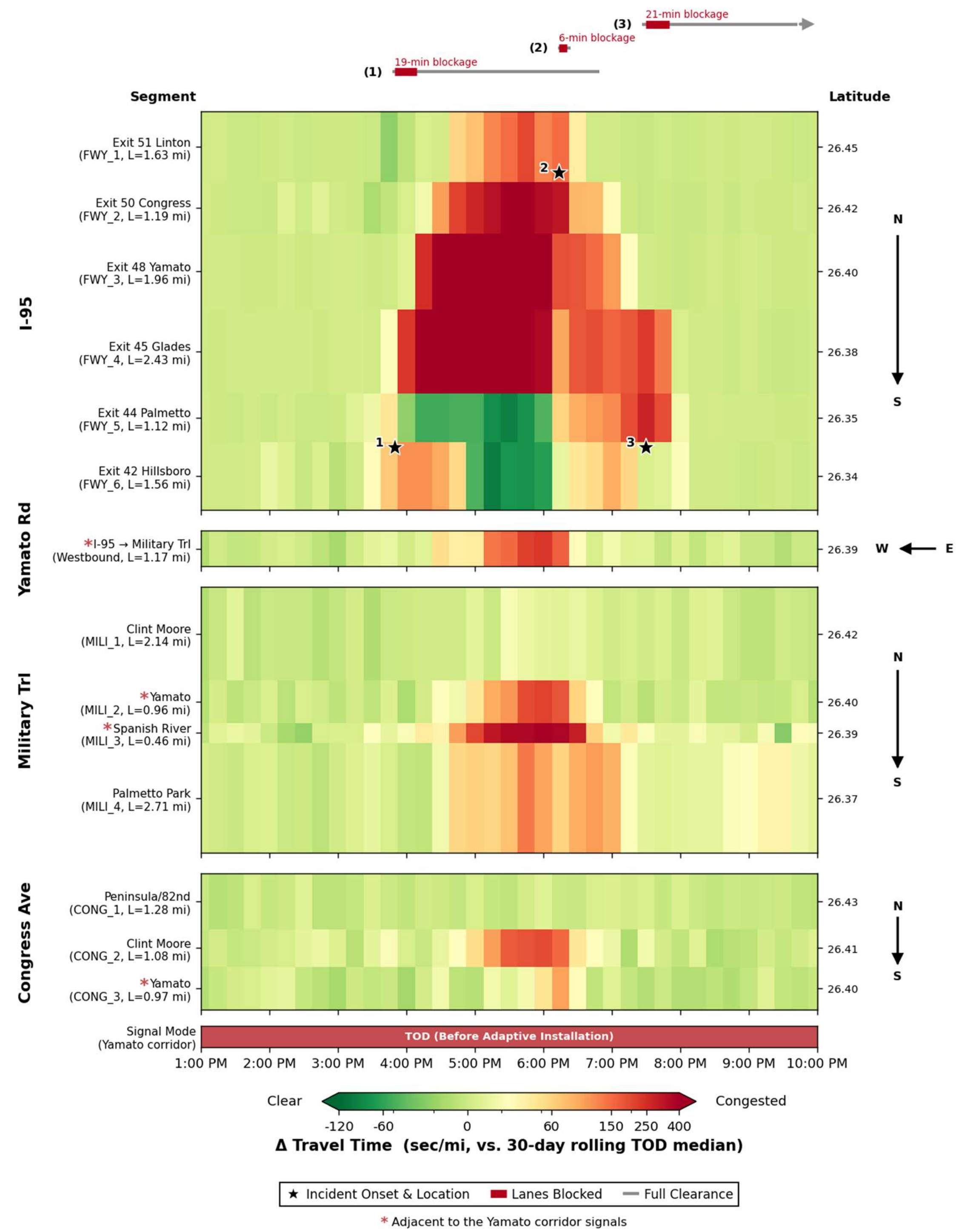
**Figure 6-5(e) Diversion Space-Time Diagram, April 30, 2025 (Wednesday)**

#### 6.4.1.2 Diversion Spread Across Multiple Arterials

In the second group of cases **(**Figure 6-6)**,** the diversion is not confined to Military Trail: the space-time diagrams show concurrent delay on other parallel routes, such as Congress Avenue, indicating a broader spread of the diverted traffic.

On February 28, 2023 **(**Figure 6-6**(a)),** three successive incidents occurring between 3:49 PM and 7:29 PM caused a sustained freeway disruption that persisted from mid-afternoon well into the evening. The capacity reduction overlapped with the peak demand period, resulting in higher congestion on I-95 and more diversion, which resulted in the delay developing simultaneously on both Military Trail and Congress Avenue during the 5:00 PM to 8:00 PM window. Furthermore, the westbound Yamato Road connector also became congested. This indicates that the diverted demand saturated all available routes within the parallel network.

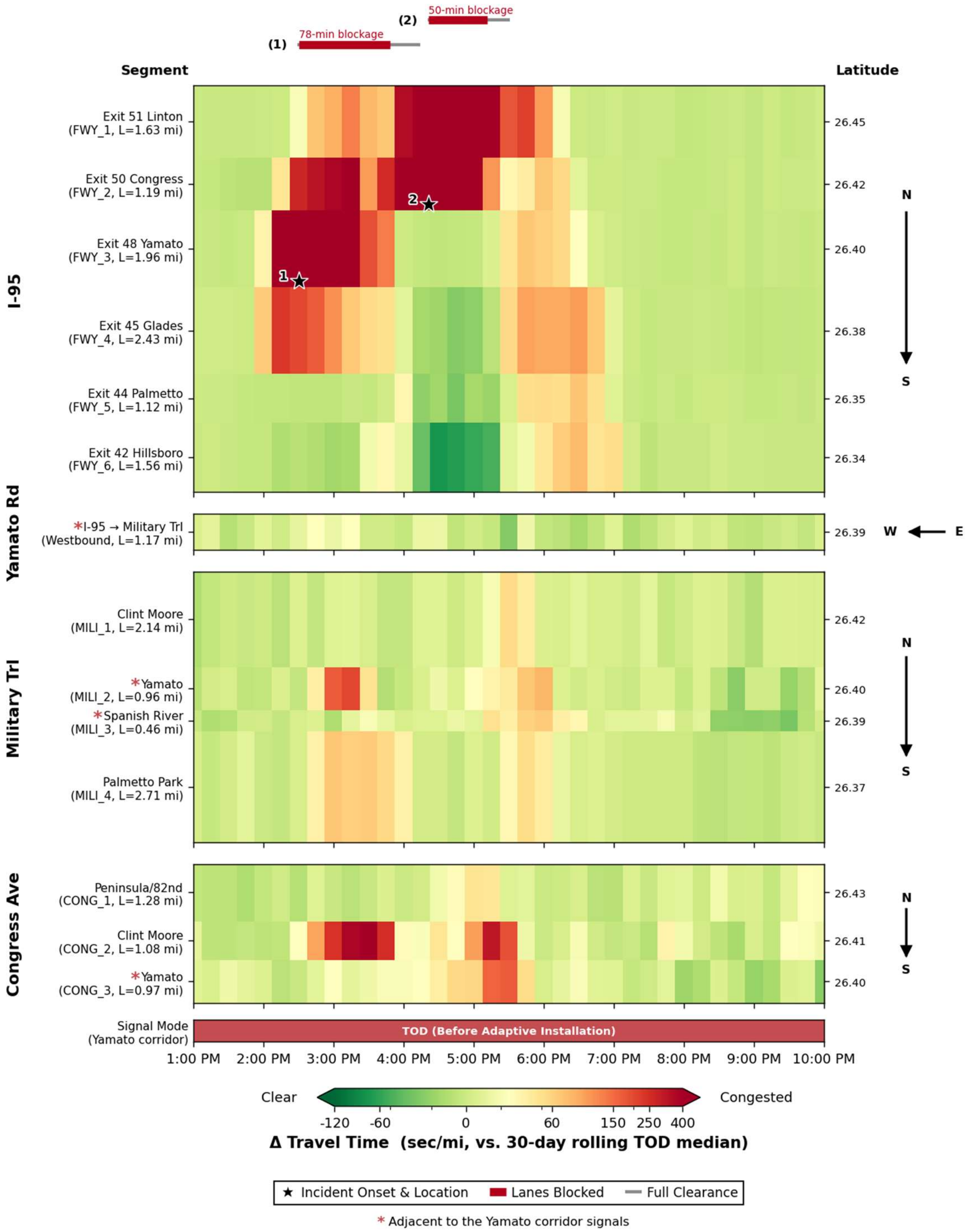
The May 3, 2023 case **(**Figure 6-6**(b))** highlights that diversion patterns are governed by network geometry as much as incident duration. Two crashes, occurring at 2:30 PM and 4:21 PM, created a freeway bottleneck at and upstream of the Yamato Road interchange. This placement effectively trapped the queue across the exits that typically feed the Military Trail diversion path. Consequently, the westbound Yamato Road connector showed no significant surge that day. Instead, the congestion developed on Congress Avenue in two distinct waves, mirroring the timing of the two crashes. This pattern confirms that motorists rerouted earlier by exiting at the Congress Avenue interchange once the Yamato Road exit became inaccessible due to the upstream queue.




**Incidents on Tuesday, February 28, 2023**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 3:49 PM | Beyond Exit 44 (FWY_6) | Disabled Vehicle | 1 of 5 (20%) | 19 min | 181 min |
| 2 | 6:13 PM | At Exit 50 (FWY_1) | Crash | 1 of 4 (25%) | 6 min | 11 min |
| 3 | 7:29 PM | Beyond Exit 44 (FWY_6) | Crash | 2 of 5 (40%) | 21 min | 159 min |

**Figure 6-6(a) Diversion Space-Time Diagram, February 28, 2023 (Tuesday)**




**Incidents on Wednesday, May 3, 2023**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 2:30 PM | Beyond Exit 48 (FWY_3) | Crash | 4 of 5 (80%) | 78 min | 105 min |
| 2 | 4:21 PM | Beyond Exit 50 (FWY_2) | Crash | 3 of 5 (60%) | 50 min | 70 min |

**Figure 6-6(b) Diversion Space-Time Diagram, May 3, 2023 (Wednesday)**

#### 6.4.1.3 Summary of the Observed Diversion Patterns

Field observations have identified two critical operational patterns that establish the empirical foundation for the proposed predictive model.

##### 6.4.1.3.1 Temporal Lag in Arterial Impact

Analysis reveals that arterial degradation does not occur instantaneously following a freeway incident. Instead, there is a consistent temporal lag, ranging from 15 to 60 minutes, between the onset of freeway congestion and the subsequent deterioration of arterial travel times. This time lag presents a critical operational window; a warning issued within this timeframe provides traffic managers with the necessary lead time to implement proactive signal timing plans before arterial congestion develops.

##### 6.4.1.3.2 Spatial Distribution of Diverted Demand

The distribution of diverted traffic varies by the incident location and existing network constraints, exhibiting three distinct patterns:

- **Primary Arterial Diversion:** Under typical conditions, where the incident occurs downstream of the Yamato Road interchange, the standard exit (Exit 48: Yamato Road) remains accessible. Traffic diverts via this interchange and travels south along Military Trail to rejoin the freeway further down. Consequently, the southern Military Trail segment, denoted as MILI_4, serves as the primary diversion path for excess freeway demand and exhibits increased travel times during these incidents.
- **Network-Wide Spillover:** If the incident is severe and the queue on the freeway extends further upstream, Congress Avenue is also used as a diversion route together with Military Trail. Under these conditions, the diverted volume exceeds the capacity of any single arterial, resulting in simultaneous degradation across multiple parallel corridors.
- **Upstream Incidents:** When incidents occur close to or upstream of the first investigated exit (Yamato exit), diversion occurs to Congress Avenue since the diverted traffic joins I-95 downstream of the Congress Avenue intersection with Yamato Road.

These observations justify the design of the predictive model because the freeway becomes congested first, and the arterial follows only after a time lag. The freeway data and the incident record can provide the information needed to anticipate the impact on arterial operations. As MILI_4 consistently serves as the primary diversion route, this segment is established as the primary target for the predictive model developed in this study.

### 6.4.2 Signal Timing Response during Incident

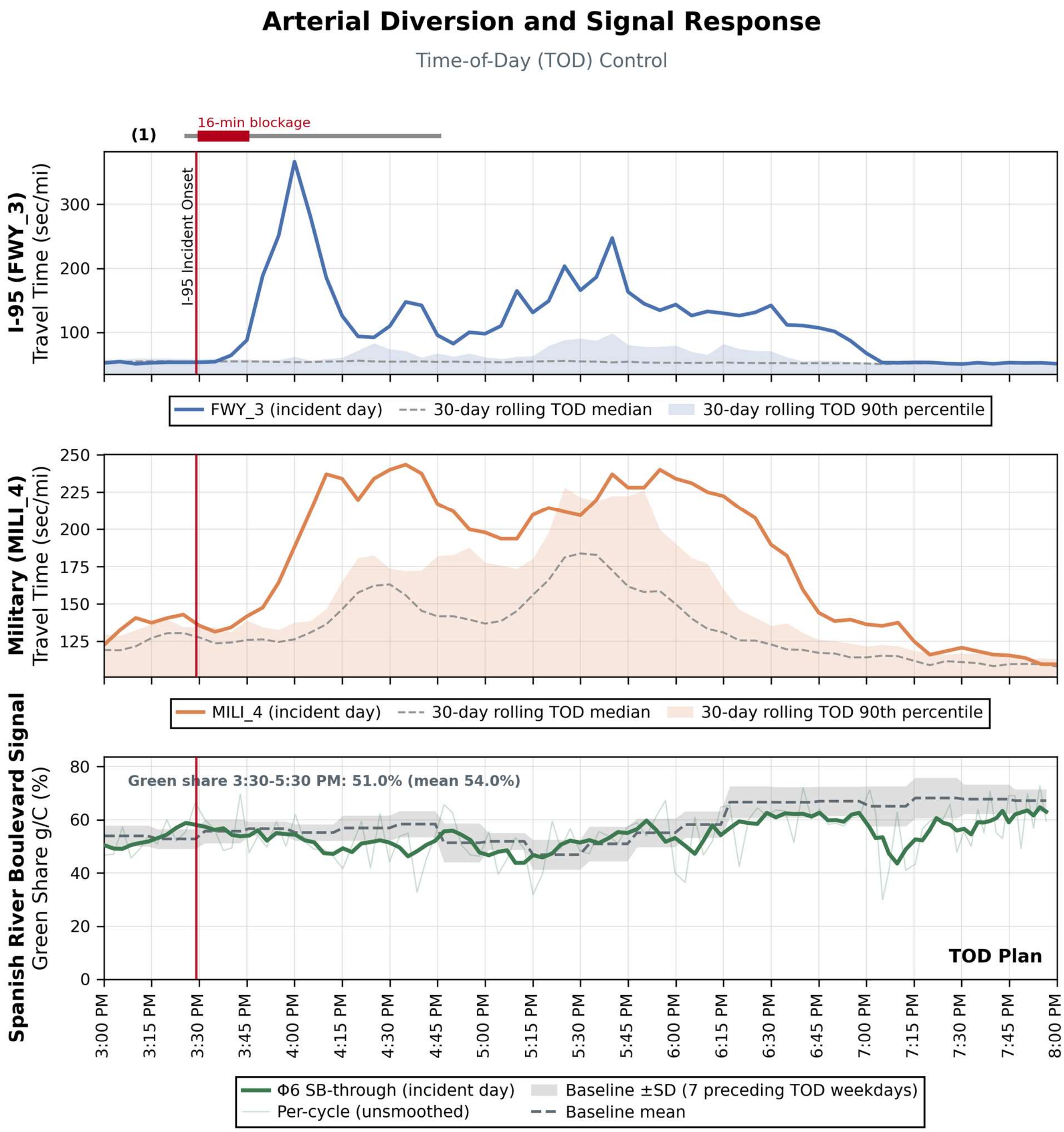
This section examines the relationship between traffic congestion and signal control related to the target segment (MILI_4). The MILI_4 segment is controlled by the signal at the Military Trail and Spanish River Boulevard intersection, which continued to operate under its standard TOD plan during the incident.

During diversion events, relying solely on a pre-programmed TOD signal plan without operator intervention often allows congestion to build up on alternate routes. A clear example of this occurred on Thursday, October 23, 2025, as presented in Figure 6-7**.** The first two panels display travel time patterns on the incident day and compare them with the median and 90th-percentile travel times from the previous 30 days. The analysis tracks the I-95 segment (FWY_3), where the incident occurred, and the adjacent Military Trail segment (MILI_4). The final panel presents the green-to-cycle ratio (g/C) for the southbound through movement (Phase 6) at the Spanish River Boulevard intersection, which regulates traffic flow for the MILI_4 segment.

At 3:29 PM, a crash blocked 3 of the 5 freeway lanes on I-95, reducing capacity by 60% and significantly increasing travel times on the freeway. Approximately 15 minutes after the incident began, travel times on Military Trail nearly doubled as diverted traffic reached the corridor. The travel time on MILI_4 surged well above the historical 90th percentile, indicating severe, non-recurring congestion.

Despite this sudden surge in travel time, the green time share for the southbound through movement (Phase 6) did not increase to accommodate the extra volume. Instead, it averaged 51.0% between 3:30–5:30 PM, which was slightly lower than the baseline historical mean of 54.0%. Consequently, delays on the arterial persisted long after the incident was fully clear.

This analysis reveals two key observations. First, once congestion builds up on the arterial, the impact persists long after the incident is cleared. Second, the TOD signal plan cannot accommodate diverted traffic to meet the surge in demand; therefore, operator intervention is necessary to respond effectively. These findings suggest that a delay warning system would be important for improving the signal system's response to abnormal demand and freeway incidents before congestion develops.




**Incidents on Thursday, October 23, 2025**

| # | Onset | Location | Type | Lanes blocked | Blocked Duration | Clearance |
|---|---|---|---|---|---|---|
| 1 | 3:29 PM | At Exit 44 (FWY_4) | Crash | 3 of 5 (60%) | 16 min | 80 min |

**Figure 6-7 Travel Time and Green Share Analysis During a Freeway Diversion Under TOD Control**

### 6.5 Descriptive Analysis of Corridor Data

This section examines the travel-time and incident records used to develop the predictive model in this study. The section characterizes travel-time distributions on the target arterial segment, temporal travel-time patterns, and incident distributions within the study area. The analysis uses HERE data to estimate travel time rate, calculated as units of travel time per mile, to ensure a consistent comparison of segments with varying lengths.

#### 6.5.1 Target Arterial Travel-Time Distribution

Figure 6-8 illustrates the target arterial (MILI_4) travel time distribution across all hours of the weekday PM peak in 2025. The figure shows a right-skewed distribution with a mean travel time of 108 seconds per mile.

In common travel-time reliability practice, the 90th and 95th percentile travel times are used to represent planning-time and near-worst-day conditions (FHWA, 2006). These values form the long tail visible on the right side of the histogram. As indicated in the graph, the 90th percentile occurs at 131 seconds per mile, and the 95th percentile reaches 146 seconds per mile.

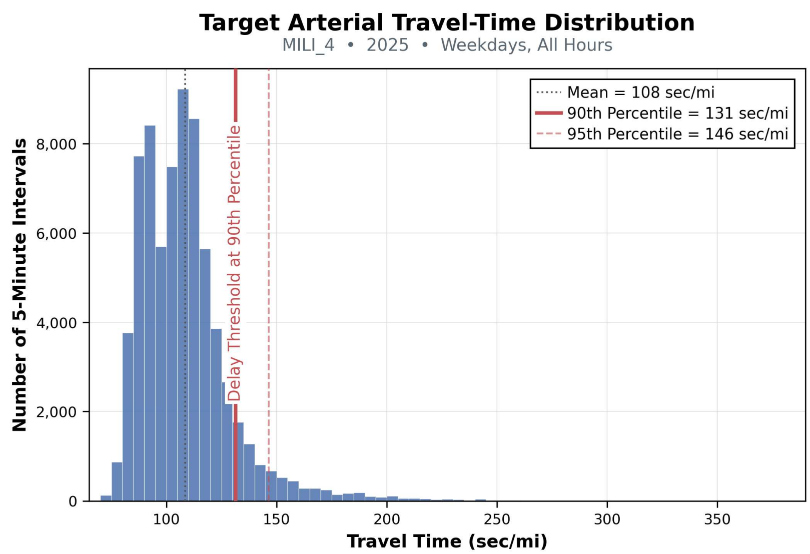



**Figure 6-8 Target Arterial Travel-Time Distribution**

#### 6.5.2 Temporal Travel-Time Patterns

##### 6.5.2.1 Travel Time Patterns Across Segments

The average hourly travel time per mile for the year 2025 was used in this study to examine changes in traffic conditions across hours of the day. On the southbound of Military Trail **(**Figure 6-9**),** Segments MILI_2, MILI_3, and MILI_4 experience the highest PM delay, especially on the MILI_2 and MILI_4 segments.

On I-95 (Figure 6-10), the southern freeway segments FWY_4 through FWY_6, spanning from Yamato Road down to Hillsboro Boulevard, experience significant congestion during the afternoon period with travel times peaking sharply between about 5:00 PM and 5:30 PM. These heavily congested freeway segments are adjacent to the Glades Road (Exit 45) and Palmetto Park Road (Exit 44) interchanges that feed Military Trail, making the downstream Military Trail segment (MILI_4) the primary target for prediction due to its greatest exposure to diverted traffic.

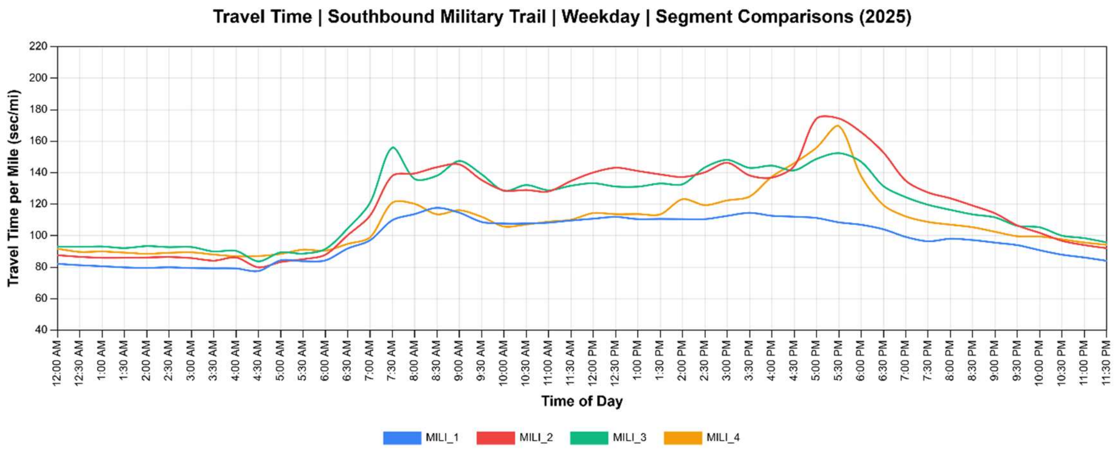



**Figure 6-9 Average Hourly Weekday Travel-Time Rate Profiles by Segment, Southbound Military Trail (2025)**

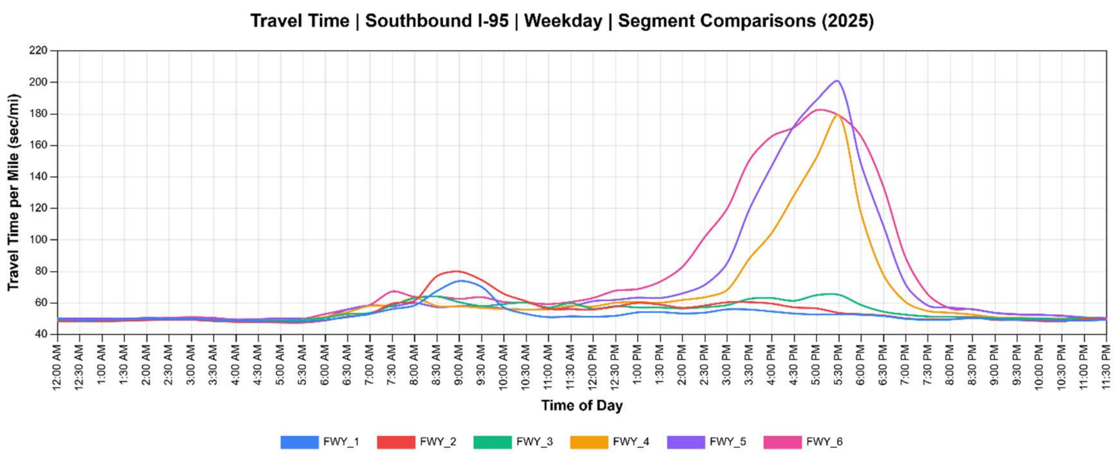



**Figure 6-10 Average Hourly Weekday Travel-Time Rate Profiles by Segment, Southbound I-95 (2025)**

### 6.5.2.2 PM Peak Period Identification

The travel time profiles in the previous section (Figure 6-9 and Figure 6-10) indicate that congestion on Military Trail begins to build up from 3:30 PM, peaks between 5:00 and 6:00 PM, and remains elevated until 6:30 PM. Based on these patterns, the core PM peak is defined as 3:30 PM to 6:30 PM. However, to ensure all incident impacts are fully captured, the prediction models were developed in this study for an extended timeframe from 3:00 PM to 9:00 PM.

#### 6.5.2.3 Year-to-Year Change

A comparison of the weekday travel time profiles from 2020–2025 shows that while the timing of the PM peak remains consistent, congestion intensity has increased over time. On Military Trail **(**Figure 6-11**),** the travel times on the MILI_4 segment rose from a lowest value in 2020–2021 to reach the highest values by 2025. Similarly, on the FWY_4 through FWY_6 I-95 segments **(**Figure 6-12**),** congestion has intensified and lasted longer through the years, with 2024 and 2025 recording the highest peak-period travel times. It is recognized that the years 2020 to 2022 were impacted by COVID. This shift in traffic patterns directly impacts how historical data is applied during model training, which is detailed in the Predictive Model Development section.

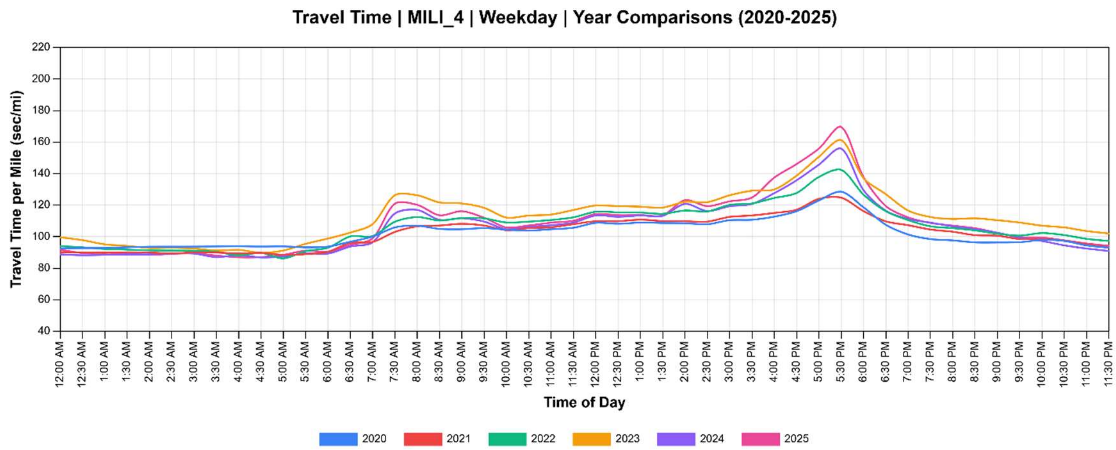



**Figure 6-11 Weekday Travel-Time Profiles by Year, Military Trail Segment MILI_4 (2020 to 2025)**

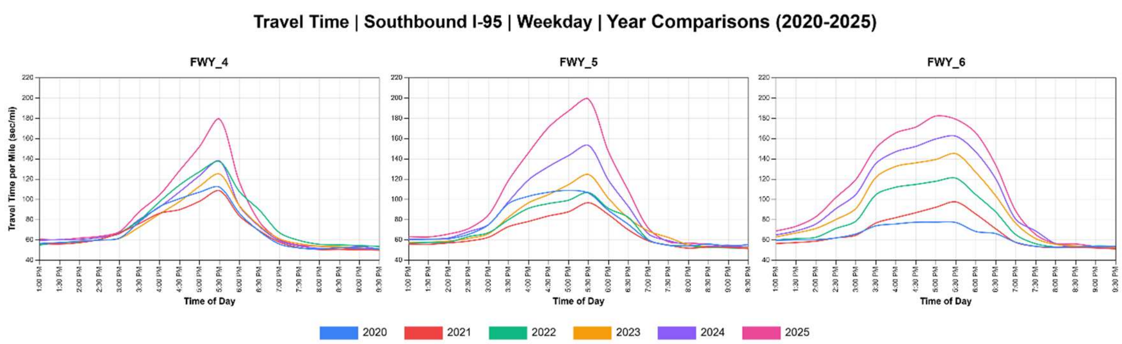



**Figure 6-12 Weekday Travel-Time Profiles by Year, Southbound I-95 Segments FWY_4 to FWY_6 (2020 to 2025)**

### 6.5.3 Incident Distribution

Figure 6-13 shows the distribution of incident occurrence along the I-95 southbound general-purpose lanes using weekday PM incident data from January 2020 to April 2026. This figure shows that incidents occur more frequently on the southern freeway segments of the studied facility. In this area, traffic is more likely to divert toward Military Trail, which is the primary and nearest parallel arterial. This concentration of freeway incidents further supports focusing the initial modeling on the MILI_4 segment, the route most exposed to freeway-to-arterial traffic shifts.

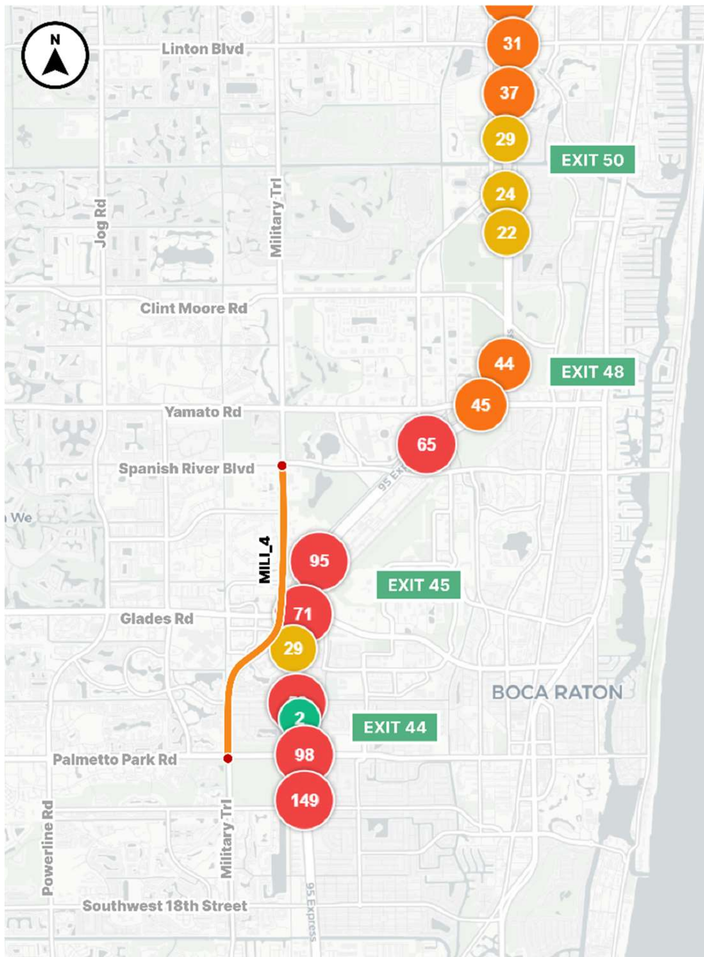



**Figure 6-13 Distribution of Freeway Incidents along the I-95 Southbound Study Corridor**

Table 6-4 breaks the total of 845 incidents down by SunGuide event type. Crashes dominate the record (732 out of 845 incidents, which is 87 percent of the total). Obstructions, mainly disabled or abandoned vehicles together with roadway debris, account for 95 incidents (11 percent), and a small remainder of other events, such as vehicle fires, police and emergency activity, and weather, make up the final 18 incidents (2 percent).

For modeling purposes, the raw SunGuide types are grouped into three categories: crash, obstruction, and miscellaneous; in line with the above findings. Scheduled road work is excluded from the model because it is a planned event that is typically active over longer, pre-announced periods rather than the sudden onset of congestion that drives diversion during other types of incidents.

**Table 6-4 Freeway Incidents by Event Type, I-95 Southbound Study Corridor**

| Event Category | SunGuide Event Type | Incidents | Percent of Total |
|---|---|---|---|
| Crash | Crash | 732 | 86.6 |
| Obstruction | – | 95 | |
| | Disabled Vehicle | 74 | 8.8 |
| | Abandoned Vehicle | 17 | 2 |
| | Debris on Roadway | 4 | 0.5 |
| Miscellaneous | – | 18 | |
| | Other | 11 | 1.3 |
| | Emergency Vehicles | 4 | 0.5 |
| | Vehicle Fire | 2 | 0.2 |
| | Police Activity | 1 | 0.1 |
| Total | | 845 | 100 |

## 6.6 Predictive Model Development

This section details the methods utilized in the development and assessment of the predictive model designed for predicting arterial traffic delay due to incident diversion. It outlines the methodological approach, including data validation and performance metrics, and documents the iterative model development process, from feature engineering to hyperparameter tuning. Finally, the model's performance is evaluated on a one-year held-out test set that remained entirely unseen during the development phases. This held-out evaluation mimics real-world deployment, providing an objective assessment of how the model performs in the field.

### 6.6.1 Model Development Framework

This section presents the model development framework before the methodology is presented in detail in the following sections.

#### 6.6.1.1 Design Objective and Scope

The design of the predictive model is set to meet three operational requirements. First, the warning must arrive early enough to allow the signal control operator to act on it. The 15-minute prediction horizon used in the model provides lead time to activate a response, such as a special signal timing plan, before congestion develops on the arterial.

Second, the model must use data that the signal operating agency already has. Its inputs are limited to the freeway detector and incident feeds available in real time through DIVAS. Thus, the implementation of the model does not require additional cost to acquire commercial data. During model development, HERE travel-time data were used to define ground-truth labels (i.e., to identify delay conditions within the historical record) to establish prediction targets. These data do not serve as model input. Once trained, the model generates real-time predictions relying solely on DIVAS detectors and incident feeds. As a result, no HERE or other commercial data subscription is required for field deployment.

Third, the tool must respect operator workload: alert fatigue is a primary concern for traffic management center personnel. When an alerting system generates frequent false alarms or alerts for minor events that are resolved without intervention, operators often begin to disregard the system. The design therefore prioritizes a low false-alarm rate over a maximum detection rate, issuing a warning only when an incident is likely to cause significant arterial delay.

The prediction task is a binary classification problem that determines whether the downstream Military Trail segment (MILI_4) will be in a delay state 15 minutes ahead. To ensure that the predictive model is responsive to current traffic conditions in real time, it operates on a 5-minute refresh cycle: as new sensor data arrives, it processes the latest readings to generate a fresh prediction for each 5-minute interval. The model is specifically designed to monitor conditions during active freeway incidents within the extended PM peak period of 3:00 PM to 9:00 PM.

#### 6.6.1.2 Delay State Definition

The target segment is considered delayed when its travel time (seconds/mile) exceeds the 90th-percentile (p90) threshold, following common travel-time reliability practice (FHWA, 2006). To reflect daily changes in traffic conditions, the model applies a delay threshold that updates dynamically. For each 5-minute time-of-day slot, the p90 of the travel time observed in that slot over the most recent 30 qualifying weekdays is computed from historical data and used as the delay-defining threshold.

The recent days were screened to exclude weekends, holidays, and days with a major incident (lasting more than 120 minutes or blocking at least two lanes for more than 30 minutes). The incident duration criterion follows the MUTCD classification of a major traffic incident as one with an expected duration of more than 2 hours (FHWA, 2023).

Once the dynamic threshold is calculated, it serves as the baseline for classifying the traffic state. The model evaluates each 5-minute interval by comparing its observed travel time against the threshold for that specific time of day. An interval receives a label of 1 ("Delay") if it predicts that the next 15-minute travel time exceeds the threshold, and 0 ("Normal") if it does not.

Figure 6-14 demonstrates how this labeling works during a real-time prediction scenario. Suppose the current time is 3:45 PM. At this moment, the model has access to traffic data recorded only up to 3:45 PM. The objective is to predict the traffic state 15 minutes ahead at 4:00 PM, which is currently unseen. The model continuously repeats this process, sliding forward every 5 minutes as new data becomes available.

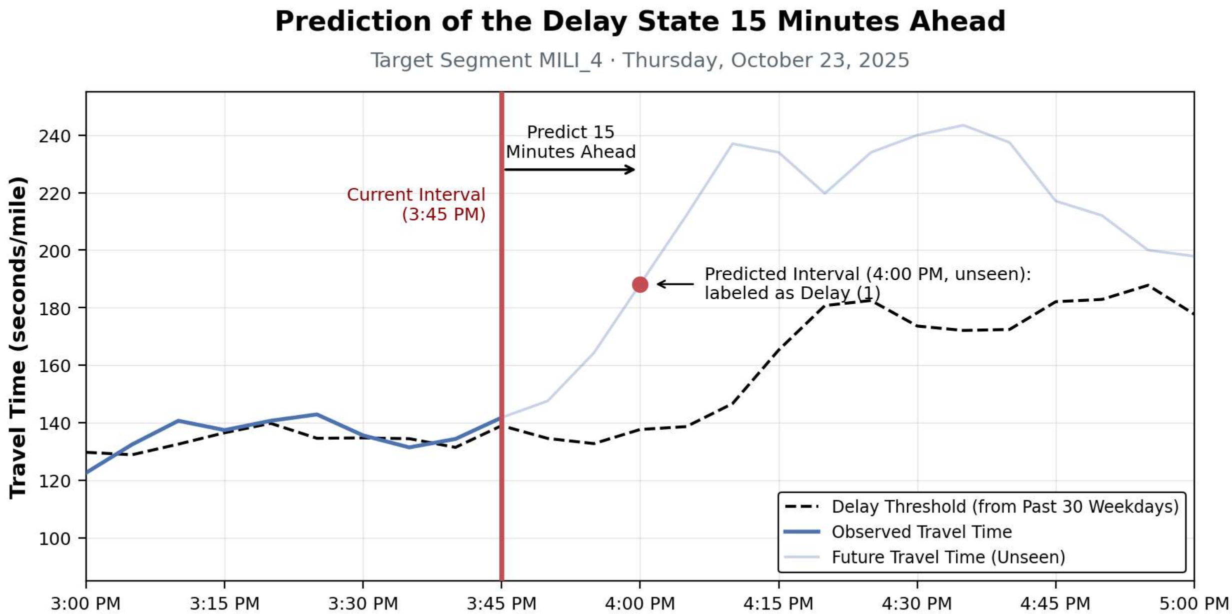



**Figure 6-14 Example of Delay State Labeling at the 15-Minute Prediction Horizon**

#### 6.6.1.3 Modeling Approach

The selected classifier is XGBoost, a machine learning model based on gradient-boosted decision trees. The model learns recurring patterns from the historical record and, for each 5-minute interval, outputs a score between 0 and 1, where a higher score indicates greater confidence that the segment will be in a delay state 15 minutes ahead. This continuous score is then converted into a binary prediction by comparing it against a fixed threshold, which defaults to 0.5. If the score meets or exceeds this threshold, the interval is flagged as a predicted delay (1); otherwise, it is recorded as no delay (0). A procedure for selecting an optimal threshold, rather than relying on the default value, is detailed in the later section.

XGBoost was selected for three reasons. First, it suits the structure of the input data: tree-based models tolerate strong correlations among input features without losing predictive performance. This is critical in this study because detector readings measured at neighboring stations and at short time lags exhibit high multicollinearity. Second, XGBoost has been applied successfully to closely related problems to the one in this study, including real-time incident detection and crash-risk prediction using freeway sensor data (Parsa et al., 2020, Li et al., 2023). Third, the model is interpretable: the contribution of each input feature to a prediction can be quantified using SHAP (SHapley Additive exPlanations) values, as shown in the feature interpretation at the end of this section.

### 6.6.1.4 Model Development Workflow

Figure 6-15 illustrates the data partitioning workflow. To develop the model and obtain an unbiased estimate of its field performance, the full dataset was divided chronologically into two parts: a full training set used to develop the predictive model, and a held-out test set reserved for final performance evaluation. The held-out test set was excluded from the model development process. This approach represents a real-world deployment scenario where the model is deployed first, and its performance is evaluated afterward.

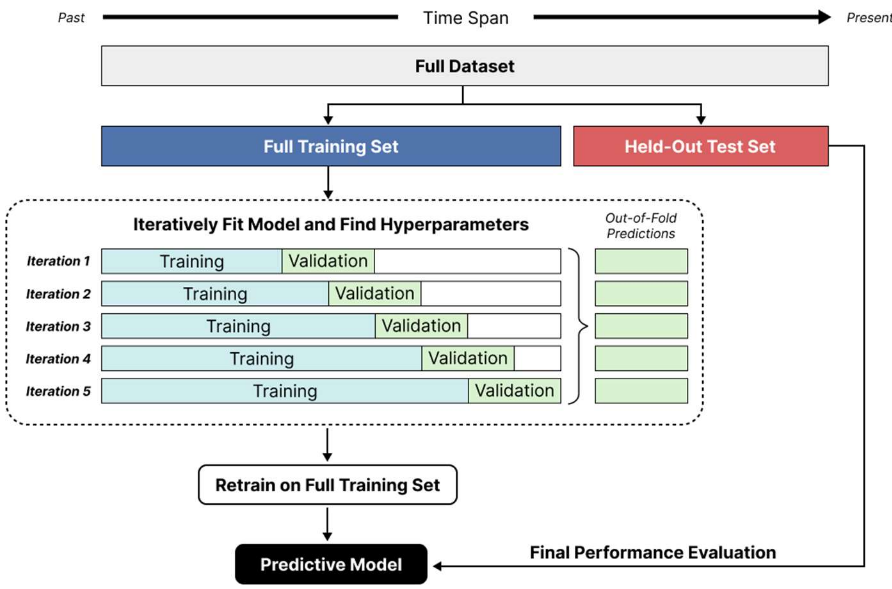



**Figure 6-15 Model Development Workflow**

The following sections detail the specific components of the model development workflow illustrated in Figure 6-15.

#### 6.6.1.4.1 Temporal Data Split

The full training set contains the historical record from January 2020 through April 2025. Two training timeframe configuration options were compared by examining the performance of models developed with the two configurations as follows.

- **Complete Historical Record:** The first model was developed for the full dataset beginning in January 2020. This option maximizes the amount of data available to train the model.
- **Post-Reconfiguration Record:** The second model only uses recent data to minimize the impacts of external factors that have changed through the years on model performance. The data used in this option covers a period starting in July 2023, unlike the first option. This reflects conditions in which the I-95 geometry and the detector layout remain consistent for the whole period.

To determine the most effective training period, both training timeframe configurations were evaluated in the experiments later in this section, and the one that yielded the highest performance was selected for use in the final model.

The held-out test set spans one full year, from May 2025 to April 2026, to ensure the evaluation captures a complete annual cycle considering seasonal traffic demand variations. Because this data was completely excluded from the model development process, it effectively represents unseen future conditions. The reported performance on the held-out test set reflects the model's behavior as if it had been deployed in May 2025 and left to run for a year.

#### 6.6.1.4.2 Cross-Validation Strategy

In the model development phase, the full training dataset is divided into sub-training and validation subsets. The trained models are fit to the sub-training data, and alternative configurations or hyperparameter tuning methods are assessed on the validation data, which remains unseen during training.

Using a single train-validation split could bias the model configuration toward the specific months present in the validation set. A cross-validation strategy reduces this risk by repeatedly splitting the data into training and validation folds (or iterations, as shown in Figure 6-15 for a 5-fold split). As a result, all development decisions, including configuration screening, feature selection, and hyperparameter selection, are validated across several time periods rather than a single instance.

#### 6.6.1.4.3 Time-Series Cross-Validation

When evaluating time-series data, the temporal order of the folds must be preserved to prevent data leakage. Data leakage occurs when future information is accidentally used during training, resulting in overly optimistic performance estimates that fail during real-world deployment. Time-series cross-validation naturally avoids this problem because, in each iteration, the model trains only on data collected before the validation period. This ensures that each fold accurately produces predictions based on past data, exactly as the system will operate in the field (Raschka and Mirjalili, 2019, Scikit-learn Developers, 2025).

Two time-series cross-validation schemes were compared, as illustrated in Figure 6-16**:**

- **Expanding Window Scheme:** This approach fixes the starting point and grows the training set with each iteration. As a result, later folds learn from progressively more historical data.
- **Rolling Window Scheme**: This approach keeps the training duration fixed and shifts both its start and end forward in time. Consequently, each fold learns only from a recent, constant-length window.

Both schemes were evaluated in the later section, and the better-performing scheme was selected for the final model.

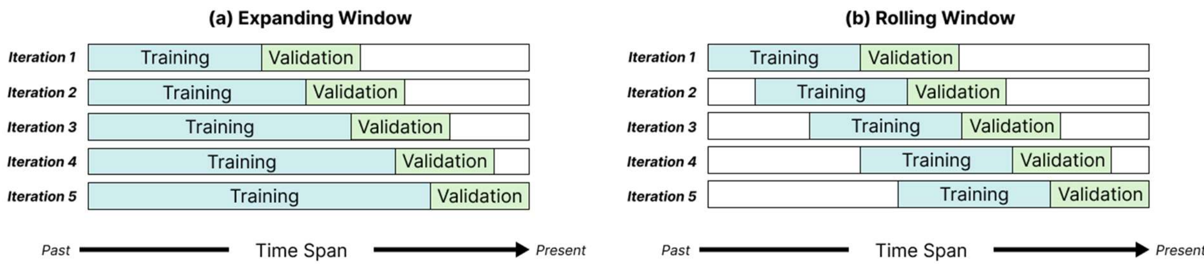



**Figure 6-16 Time Series Cross-Validation Scheme**

### 6.6.1.4.4 Imbalance Handling

Class imbalance occurs when the classes (or labels) in a dataset are unevenly distributed, with one class far larger than the others. During training, this imbalance biases a classifier toward the majority class and causes it to ignore the minority class, leading to frequent misclassification of the minority class (Altalhan et al., 2025). This condition applies directly to this study because during freeway incident diversions, intervals of delays on the alternative arterial route are far less frequent than non-delay intervals, making the delay state the minority class.

There are multiple strategies used to address the imbalance issue, and each approach has distinct advantages and disadvantages:

- **Under-Sampling:** This method removes instances from the majority class to balance the data, but it discards information that could improve the model's performance.
- **Over-Sampling (SMOTE):** This approach adds minority-class examples. A common method is the Synthetic Minority Over-sampling Technique (SMOTE), which generates synthetic minority examples from existing minority instances. However, under severe imbalance, the added synthetic data can cause overfitting (Chawla et al., 2002).
- **Class Weighting:** This strategy assigns a higher weight to the minority class and a lower weight to the majority class in the training loss. This preserves the original data distribution (Asundi et al., 2022, Bakirarar and Elhan, 2023).

Rather than choosing one of these strategies in advance, this study treats the imbalance-handling method as one of the configuration experiments. The best-performing option is selected for the final model, as reported in the later section.

### 6.6.1.4.5 Out-of-Fold Predictions

In time-series cross-validation, each iteration produces predictions on its own validation fold. In this study, these folds differ in the number of incidents they contain because the folds span different months; one month may contain very few incidents, while another may contain many. Computing a performance metric separately for each fold and then averaging the results would give each fold equal weight regardless of size, allowing a small fold with few incidents to count as much as a large one and biasing the result (Forman and Scholz, 2010).

To avoid this bias, the model performance is measured on out-of-fold (OOF) predictions, in which the model's predictions across all validation folds are pooled into a single set, and the metric is computed once over that pooled set (Forman and Scholz, 2010). Throughout the development phase, the model performances across different configurations are therefore reported and compared using OOF validation.

#### 6.6.1.4.6 Hyperparameter Tuning

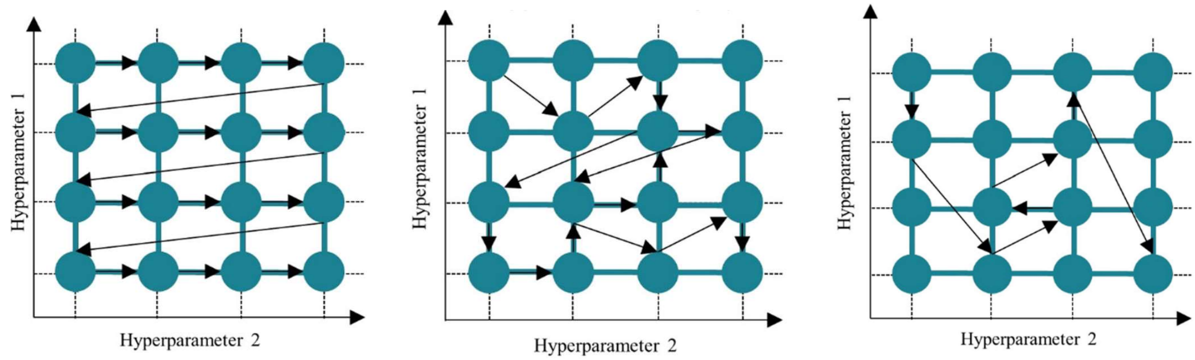
Hyperparameters are the parameter settings of the learning algorithm that are specified before training. Hyperparameter tuning is the process of searching for the combination of these settings that yields the best predictive performance, typically evaluated on a validation set. Because hyperparameters control how the model fits the data, choosing them well can substantially improve performance (Probst et al., 2019).

For XGBoost, the key hyperparameters include the learning rate, the maximum tree depth, the number of estimators (boosting rounds), and the regularization terms that penalize model complexity. Together, they determine how quickly and flexibly the model fits the training data, and therefore how well it generalizes (i.e., how well it performs on new, unseen data).

Three common search strategies for selecting these settings are illustrated in Figure 6-17**:**

- **Grid Search:** This method tests every possible combination of candidate values on a defined grid. While it guarantees finding the best setting on that grid, the number of combinations multiplies rapidly as the number of hyperparameters increases. This makes the computation time too high, especially for complex models (Bergstra and Bengio, 2012).
- **Random Search:** This method randomly tests candidate settings from a defined range for a fixed number of trials. For the same budget, it explores a wider range of values for each hyperparameter and tends to find good settings more efficiently compared to Grid Search in higher dimensions (Bergstra and Bengio, 2012).
- **Bayesian Optimization:** Rather than searching blindly, this method builds a probabilistic surrogate model of how performance varies with the hyperparameters. It then uses this model to choose the next setting most likely to improve performance, learning from previous evaluations to converge with fewer trials (Tran and Kim, 2024).

Because Grid Search is computationally expensive and often performs comparably to Random Search, it is excluded from further consideration. The choice of tuning strategy was treated as a configuration experiment with the Random Search used as the default settings (as a baseline) and compared to the use of Bayesian Optimization in model development on the validation set. The best-performing method was selected for the final model.




(a) Grid Search (b) Random Search (c) Bayesian Optimization

**Figure 6-17 Grid Search, Random Search, and Bayesian Optimization (Tran and Kim, 2024)**

#### 6.6.1.4.7 Final Model Retraining

Once the optimal configuration and hyperparameter setting are identified, the model is retrained on the training dataset using these parameters and settings. The final model is then evaluated on a separate hold-out test set to obtain an unbiased estimate of its performance of the final model.

### 6.6.2 Performance Measures

This section outlines the metrics used to evaluate the binary classification model. The evaluation framework is divided into two approaches. The first approach evaluates the model at a single classification threshold, while the second evaluates its overall discriminative power across all possible thresholds.

#### 6.6.2.1 Single-Threshold Evaluation

The XGBoost classifier outputs a continuous predicted score between 0 and 1, as illustrated in Figure 6-18**.** To make a binary decision, a classification threshold, denoted as $\tau$, must be applied to this score. By default, this threshold is typically set to 0.5. When the score exceeds $\tau$, the model classifies the interval as the positive class (delay). Conversely, if the score falls below $\tau$, it is classified as the negative class (normal condition).

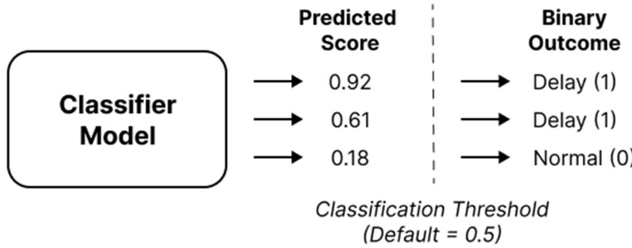



**Figure 6-18 Conversion of Continuous Predicted Scores to Binary Outcomes**

Typically, the model developer builds and delivers the predictive model, reporting baseline performance at the default threshold of 0.5 or at the threshold that maximizes the F1 score. In real-world deployment, the operator can either adopt the developer's suggested threshold or set different thresholds based on their operational requirement. For example, setting a higher threshold to require greater model confidence in predicting a delay before an alert is triggered.

Adjusting the threshold directly impacts model performance. A higher $\tau$ requires greater model confidence before an alert is triggered, which reduces false alarms but may miss actual delays. Conversely, a lower $\tau$ produces more delay alerts but at the cost of increased false alarms. Although various statistical methods exist to determine the optimal threshold, this study focuses on two specific approaches:

- **Maximize the F1 score:** This approach sets the threshold to the value that maximizes the F1 score, balancing precision and recall for the positive class. This is a common approach in machine learning practice.
- **Set a false alarm budget:** With this approach, the developer sets the threshold based on the maximum false alarm rate that the operator considers acceptable.

In either case, the threshold must be selected using the training data only and then held fixed; selecting it on the test data would leak future information into the evaluation.

Applying the chosen threshold $\tau$ categorizes the model's predictions into four fundamental outcomes, which form the confusion matrix illustrated in Figure 6-19**.** The positive class is the condition the model aims to predict. In this study, the delay condition is the positive class, and the normal condition is the negative class. Based on these definitions, each prediction falls into one of the following categories:

- **True Positive (TP):** An instance correctly identified as belonging to the positive class (a delay condition correctly predicted as a delay in this study).
- **False Positive (FP):** An instance incorrectly classified as positive (a normal condition predicted as a delay).
- **True Negative (TN):** An instance correctly identified as belonging to the negative class (a normal condition correctly predicted as normal).
- **False Negative (FN):** A positive instance that the model fails to detect (a delay condition predicted as normal).

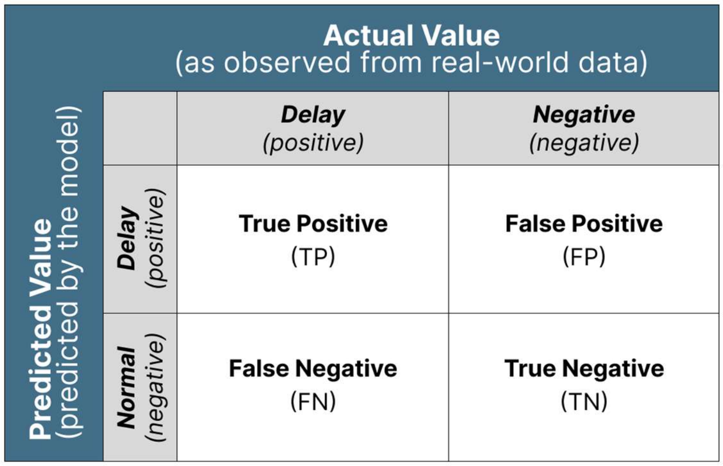



**Figure 6-19 Confusion Matrix for the Delay-Alert Task**

Once the four values of the confusion matrix are formed, they are used to calculate higher-level performance metrics. These metrics include the following:

- True positive rate (TPR), also known as recall and detection rate, which is given by **Equation Error! Reference source not found.**-1**.** It measures the proportion of actual positive instances that are correctly identified by the model. From an operational perspective, this parameter represents how effectively the model detects delay conditions that require attention.

$$TPR = \frac{TP}{TP+FN}$$

(**Error! Reference source not found.**-1)

- The false-positive rate (FPR), also known as the false-alarm rate, is given by **Equation Error! Reference source not found.**-2**.** It measures the proportion of negative instances incorrectly classified as positive. For traffic operators, this reflects how often the system produces false alarms when no delay conditions are present.

$$FPR = \frac{FP}{FP+TN}$$

(**Error! Reference source not found.**-2)

- Precision, presented in **Equation Error! Reference source not found.**-3**,** measures the proportion of instances predicted as positive that are correctly classified as such. It reflects the model's reliability in making positive predictions and its effectiveness in minimizing false positives. In operational terms, this represents alert reliability, meaning the proportion of issued warnings that correctly identify a real delay.

$$\text{Precision} = \frac{TP}{TP+FP}$$ **(Error! Reference source not found.**-3)

- The F1 score, presented in **Equation Error! Reference source not found.**-4**,** integrates PR and DR into a single number through their harmonic mean, balancing catching delay against avoiding false alarms. By considering both aspects simultaneously, the F1 score provides a more comprehensive and reliable assessment of model performance, especially when the dataset is imbalanced. In this study, the F1 score is used as the optimization objective to determine the optimal classification threshold during model training.

$$F_1\ Score = 2 \times \frac{Precision \times DR}{Precision + DR}$$ **(Error! Reference source not found.**-4)

- Prevalence measures the class imbalance in the dataset. Specifically, the positive class (delay) occurs much less frequently than the negative class (normal conditions). Prevalence is the proportion of positive instances in the dataset, calculated by dividing the number of positive instances (TP + FN) by the total number of instances (TP+FP+TN+FN).

$$Prevalence = \frac{TP + FN}{TP + FP + TN + FN}$$ **(Error! Reference source not found.**-5)

When prevalence is low, classification performance metrics influenced by the class distribution can be misleading. Accuracy, a common performance metric in machine learning, provides the clearest example of this flaw. As illustrated in **Equation Error! Reference source not found.**-6**,** accuracy measures the overall proportion of correct predictions across all classes. Because the true negatives (TN) massively outnumber all other outcomes in an imbalanced dataset, a model that simply predicts the majority (normal) class achieves high accuracy while detecting none of the delay conditions it is meant to identify (Erickson and Kitamura, 2021, Hicks et al., 2022). For this reason, this study does not rely on accuracy as the primary performance metric.

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$ **(Error! Reference source**

**not found.**-6)

#### 6.6.2.2 Threshold-Independent Evaluation

Evaluating model performance based on a single selected threshold provides a limited perspective on the model's overall discriminative capability. The overall discriminative performance is visualized by constructing continuous curves from the metric values calculated at every possible classification threshold, $\tau$. Two primary metrics are utilized in this evaluation: the area under the receiver operating characteristic curve (ROC-AUC) and the area under the precision-recall curve (PR-AUC).

- **ROC-AUC:** The ROC curve plots the TPR against the FPR, illustrating the model's overall ability to distinguish between two classes: delay and normal conditions. An area under the curve (AUC) of 1.0 indicates a perfect classifier, and a value of 0.5 represents the no-skill baseline, equivalent to random guessing. Intermediate AUC values (0.5 < AUC < 1.0) reflect varying degrees of discriminative capability, as illustrated in Figure 6-20**.**

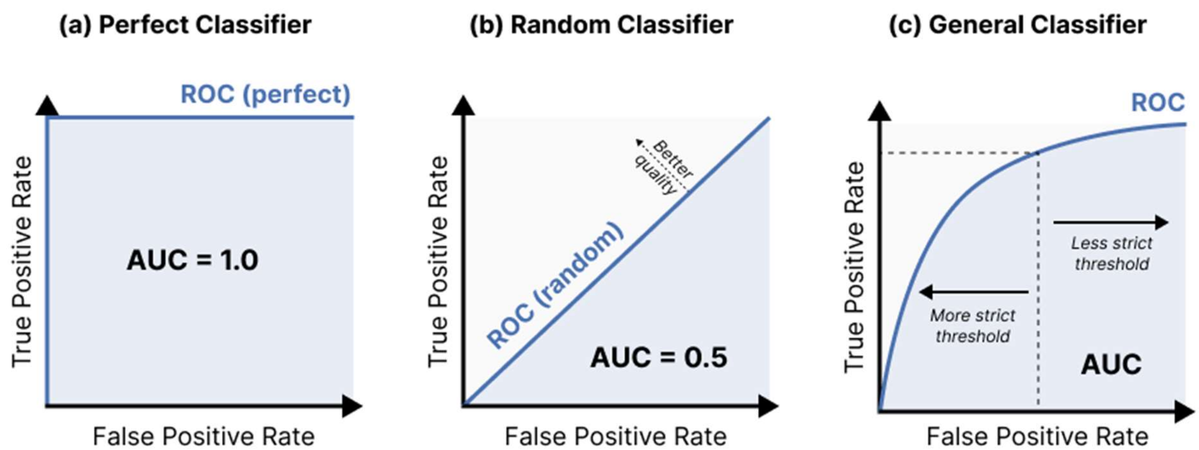



**Figure 6-20 Illustration of ROC–AUC Curves (Evidently AI, 2025)**

- **PR-AUC:** The precision-recall (PR) curve summarizes the tradeoff between precision and recall across various thresholds. If a classifier makes random predictions, its success rate in identifying positive cases will equal the overall proportion of positive cases in the dataset (i.e., the prevalence). A model with actual predictive skill must score above this no-skill prevalence baseline. Therefore, PR-AUC is interpreted relative to the prevalence, demonstrating how far the model outperforms random guessing.

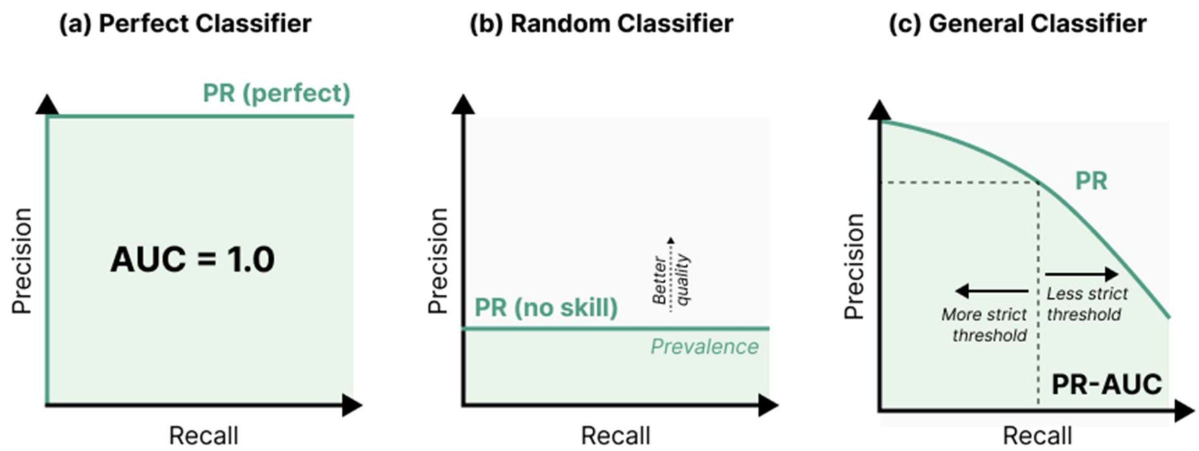



**Figure 6-21 Illustration of PR–AUC Curves**

While both curves describe the model's discriminative performance, they emphasize different aspects as follows:

- The ROC curve is invariant with class prevalence; therefore, it can be compared across different datasets and studies. However, under severe class imbalance, it can provide an overoptimistic evaluation. This occurs because the curve relies on the FPR, defined previously in **Equation Error! Reference source not found.**-2**.** The massive number of normal traffic intervals (TN) creates a large denominator, keeping the FPR persistently low even if the model generates many false alarms (FP).
- The PR curve acts as a prevalence-dependent complement to the ROC curve. The critical difference lies in the formulas for precision **(Equation Error! Reference source not found.**-3**)** and recall **(Equation Error! Reference source not found.**-1**),** as both entirely exclude true negatives (TN). Because these metrics ignore the large majority of normal traffic conditions, false alarms (FP) quickly reduce the precision score rather than blending into a large denominator. As a result, the PR curve provides a more accurate assessment of the model's ability to identify rare delay conditions. For this reason, PR-AUC is generally better suited than ROC-AUC when the data are imbalanced (Erickson and Kitamura, 2021).

In this study, both metrics are reported. ROC-AUC is included to ensure comparability with other studies, while PR-AUC serves as the primary performance measure because it accurately reflects the model's capability to detect delays under class imbalance.

### 6.6.3 Model Development

Building on the model development framework, this section outlines the complete process from data preparation to the final trained model. It details how the dataset was partitioned, how class imbalance and feature engineering were addressed, and how the model configuration and tuning strategy were chosen.

#### 6.6.3.1 Modeling Dataset and Data Partition

Table 6-5 summarizes the partitioning of the dataset used in this study. The full dataset spans from January 2020 to April 2026. The data was partitioned chronologically into a full training set and a held-out test set. As discussed earlier, two training timeframes were used in the comparative configuration investigation that follow: the complete historical database beginning in January 2020, and the recent years database (with more consistent geometry) and control beginning in July 2023, which reflects the current freeway geometry and detector layout. The held-out test set spans one full year, from May 2025 to April 2026, and was reserved exclusively for the final evaluation.

**Table 6-5 Data Partition**

| Data Set | Timeframe | Period |
|---|---|---|
| Full Training Set | Complete Data | January 2020 – April 2025 |
| | Recent Year Data | July 2023 – April 2025 |
| Held-Out Test Set | One Full Year | May 2025 – April 2026 |

#### 6.6.3.2 Class Imbalance of the Training Data

Figure 6-22 Class Imbalance of the Delay Label in the Training Set shows the imbalance in the training dataset.

It is inherently imbalanced because delay states are rare. Across all weekday intervals in the training record (2020 to April 2025), the 'Delayed' class accounts for only 12.7% of the data, so normal intervals outnumber delay intervals by nearly 7 to 1, indicating the need for methods to handle class imbalance and prevalence-aware performance measures.

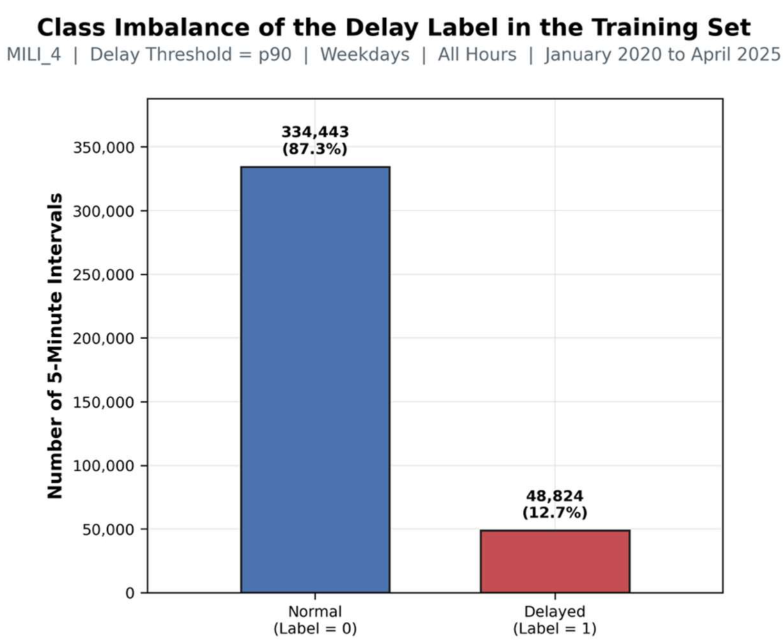



**Figure 6-22 Class Imbalance of the Delay Label in the Training Set**

#### 6.6.3.3 Feature Engineering

To prepare the data for the model development task, it is necessary to clearly distinguish between the two types of variables: features and the label, as described below:

- Features serve as the model inputs, comprising measured or derived variables that describe the corridor's conditions at a given time. The model evaluates these features to generate predictions.
- The label is the prediction target, representing the exact outcome the model is trained to forecast. For this study, the label is a binary outcome where 0 indicates normal conditions, and 1 indicates that the delay exceeds the p90 threshold.

Raw records from three input data sources had to be preprocessed before being fed into the model development pipeline. Two of these sources supply the model inputs (features): the freeway detector data provide a continuous time series of speed, occupancy, and volume, and the SunGuide incident data contribute an event log with one record per incident. The third source, the arterial travel time data, serves as the ground truth for calculating the dynamic delay threshold and assigning the binary label, as detailed previously. The feature engineering process then fused these three data sources into a single data matrix.

The model issues a prediction at a 5-minute refresh cycle. So, the unit of analysis is one 5-minute interval, with each row in the preprocessed data representing the corridor state during that interval, and every feature and label is expressed at that resolution. The detector and travel-time series were aligned to this 5-minute resolution directly since the information from these sources can be obtained at different temporal resolutions. However, the SunGuide incident database needs an extra step to be fused with the other two data sources. This is because each incident is stored as one event with a start time and a fixed set of attributes rather than as per-interval observations. Each incident record carries a detection time and a system-clearance time, and the incident is treated as active across the 5-minute intervals between them. This span is the incident duration (from detection to system clearance), not the duration that lanes are physically blocked, which is often much shorter. Each incident event record was therefore expanded across the consecutive 5-minute intervals in that active span, and its attributes (lanes blocked, type, severity, and location) were encoded into every interval. An incident active for 40 minutes, for example, was split into eight 5-minute-interval rows. The active span was capped at 120 minutes, consistent with FDOT's Open Roads Policy goal of clearing incidents within 90 minutes (FHP and FDOT, 2014) plus an allowance for the freeway queue and the resulting arterial diversion demand to dissipate after clearance. The cap also removes records that were left open in the system by mistake.

During this step, all variables were computed from the raw data feeds. As established earlier in this section, these variables fall into two distinct categories: model features (inputs) and the prediction label (target). Table 6-6 summarizes the variables derived from each source feed for model development; detailed definitions of each feature are provided in the later section.

**Table 6-6 Summary of Variable Preparation by Source Feed**

| Variable | Role | Source Feed | Processing |
|---|---|---|---|
| Detector speed, occupancy, volume (current interval) | Feature | Freeway detector time series | Aligned directly to the 5-minute data interval (reading for current interval is used as is). |
| Detector lag features (prior intervals) | Feature | Freeway detector time series | Extracted from the detector reading recorded 5, 15, and 30 minutes earlier. |
| Detector variants | Feature | Freeway detector time series | Calculated as the difference in volume, speed, and occupancy from a typical-day baseline (median) and rolling statistics (mean and standard deviation) of recent readings. |
| Active incident count, lanes blocked, total segment lanes, severity, and shoulder blockage flag | Feature | SunGuide event log | Expanded over its active intervals; each attribute encoded onto every interval in the span. |
| Incident-type group (crash / obstruction / miscellaneous) | Feature | SunGuide event log | Grouped from many raw event types into three categories, as described in the earlier sections. |
| Minutes elapsed since lane blockage began | Feature | SunGuide event log | Subtracting the first lane-blocked time from the current interval time; defined only while a lane is blocked and missing once the lanes reopen. |
| Distance from incident to reference exits | Feature | SunGuide event log | Distance from the incident location to each exit along the freeway. |
| Hour of day, month | Feature | Interval timestamp | Set based on the timestamp in the detector database. |
| Delay threshold and baseline statistics | Label construction | Arterial travel-time series (HERE) | Calculated as the rolling p90 for each 5-minute time slot over the 30 most recent qualifying weekdays (excluding non-typical days as described previously). |
| Delay Prediction label (0 / 1) | Label | Arterial travel-time series (HERE) | Comparing the travel time per mile 15 minutes ahead against the rolling delay threshold; 1 = delay, 0 = normal. |

#### 6.6.3.4 Model Configuration Experiment Setup

Developing a predictive model requires selecting specific pre-training configurations. This study sequentially evaluates these choices to determine the optimal configuration, using a validation process based on the PR-AUC. At each step, the selected setting was fixed before evaluating subsequent options. The selected setting was not necessarily the highest-scoring candidate. When performance was comparable, the option that better aligned with the deployment setting and resulted in a simpler model was chosen.

**Table 6-7** summarizes the model configuration experimental design, with the baseline setting marked for each experiment. The candidate evaluation procedure and corresponding results are presented in the following section.

**Table 6-7 Predictive Model Configuration Experiment Design**

| Configuration Variable | Description | Candidate Configurations |
| --- | --- | --- |
| **Base Model Configuration** | The initial configuration used as the starting point for sequential forward selection. | The baseline setting of every experiment below: the full record **(FullHistory)**, all weekday hours **(WkdyAllHrs)**, all intervals regardless of incident activity **(ActiveInactiveInc)**, an expanding validation window **(Expanding)**, all corridor incidents **(MiliFullExtend)**, raw detector readings only **(RawFt)**, no temporal lags **(NoLag)**, and class weighting for imbalance **(ClassWeight)**. |
| **E1 Training Data Scope** | | |
| E1.1 Historical Data Scope | Tests whether incorporating a broader historical dataset (from 2020 onward) improves predictive accuracy, or if training on the most recent data (collected after the 2023 freeway lane reconfiguration) provides better results. | FullHistory: January 2020–April 2025.<br>PostLaneConfig: July 2023–April 2025. |
| E1.2 Validation Window | Determines the training window used within each cross-validation fold, to test whether recent data provides better forecasts than the entire historical record. | **Expanding:** every fold trains on all data available before it.<br>**Rolling12m:** each fold trains on the most recent 12 months only. |
| **E2 Time-of-Day Scope** | | |
| E2.1 Time-of-Day Window | Compares training across all weekday hours against training restricted to the afternoon peak window. This evaluates whether incorporating broader all-day data improves performance, given that the evaluation set is limited to the afternoon peak to match the deployment scope. | **WkdyAllHrs:** All weekday hours.<br>**WkdyPm:** 3:00 PM–9:00 PM (deployment scope). |
| **E3 Incident-Related Scope** | | |

| Configuration Variable | Description | Candidate Configurations |
|---|---|---|
| E3.1 Incident Interval Scope | Tests whether a broader training dataset that includes intervals without active incidents outperforms a smaller training set that matches the deployment scope by containing only incident-active intervals.<br><br>An incident-active interval is the period from detection to clearance, subject to a maximum cap of 120 minutes for incidents with longer duration. | **ActiveInactiveInc:** all intervals, whether or not a freeway incident is active.<br><br>**IncActiveOnly:** only incident-active intervals. |
| E3.2 Incident Location Scope | Upstream and downstream directions are based on the southbound freeway flow and are defined relative to the Yamato Road interchange (Exit 48), located near the middle of the study corridor. Upstream refers to the freeway segments north of Yamato Road (the higher-numbered exits) and downstream refers to the segments south of it (the lower-numbered exits).<br><br>This experiment tests how far upstream and downstream the included incidents should extend by varying the set of corridor segments whose incidents are fed to the model.<br><br>The goal is to determine the freeway incident segments that carry the most predictive signal without overextending the coverage area. | **MiliDownVicinity:** the two segments immediately downstream (south) of Yamato Road (incident segments E48-E45 and E45-E44).<br><br>**MiliDownExtend:** extend MiliDownVicinity to the next downstream segment, including three incident segments (E48-E45, E45-E44, and E44-E42).<br><br>**MiliUpShortExtend:** extend MiliDownVicinity to the next upstream (north) segment, including three incident segments (E50-E48, E48-E45, and E45-E44).<br><br>**MiliUpLongExtend:** extend MiliUpShortExtend by one more upstream segment, including four incident segments (E51-E50, E50-E48, E48-E45, and E45-E44).<br><br>**MiliFullExtend:** every incident on the study corridor, all five incident segments (E51-E50, E50-E48, E48-E45, E45-E44, and E44-E42). |
| **E4 Feature Space** | | |

| Configuration Variable | Description | Candidate Configurations |
|---|---|---|
| E4.1 Supplemental Features | Tests whether two supplemental feature groups (delta and statistics) improve the performance over using only raw detector readings.<br><br>Both groups use a baseline computed from the previous 30 weekday samples for each 5-minute slot, excluding major incidents:<br><br>**Delta features:** The difference between the current reading and the normal condition (its rolling median), divided by that median. Negative values indicate better conditions, and positive values indicate worse conditions.<br><br>**Statistics features:** The rolling mean and standard deviation of those 30 samples. | **RawFt:** Raw speed, volume, and occupancy readings only.<br><br>**RawDeltaFt:** Raw readings and delta features.<br><br>**RawStatsFt:** Raw readings and statistics features.<br><br>**RawStatsDeltaFt:** Raw readings and both supplemental groups (delta and statistics). |
| E4.2 Temporal Lags | Determines whether adding short-term historical data improves the model's performance. | **NoLag:** a snapshot of the current 5-minute interval only.<br>**Lag5m:** adds the readings from 5 minutes earlier.<br>**Lag15m:** adds the 5- and 15-minute lags.<br>**Lag30m:** adds the 5-, 15-, and 30-minute lags. |
| **E5 Class Imbalance Treatment** | | |
| E5.1 Imbalance Handling | Tests methods to correct class imbalance, as delayed intervals represent a minority of the training data. | **ClassWeight, SMOTE, UnderSampling** |

#### 6.6.3.5 Model Configuration Experiment Results

Following the experimental design in the previous section, candidates within each experiment were evaluated based on their OOF-validation PR-AUC. Because the model will be deployed within PM Peak periods during active incidents, all evaluations were restricted to this specific condition. An incident active window is defined as the time from detection to clearance, with a maximum cap of 120 minutes applied to longer incidents. The highest-scoring candidate was not automatically selected. Instead, a candidate was selected only if it improved the PR-AUC score by more than 0.05. Performance differences of 0.05 or less were treated as ties. Tied candidates were resolved by prioritizing overall simplicity and fewer input features. This approach ensures that a more complex model is adopted only when it provides a substantial predictive advantage.

Figure 6-23 illustrates the OOF validation PR-AUC for candidate settings across all experiments. The shaded region in each subplot represents the tie band, which shows performance differences of 0.05 or less from the selected configuration. Each chart includes a dotted line, representing a no-skill prevalence floor.

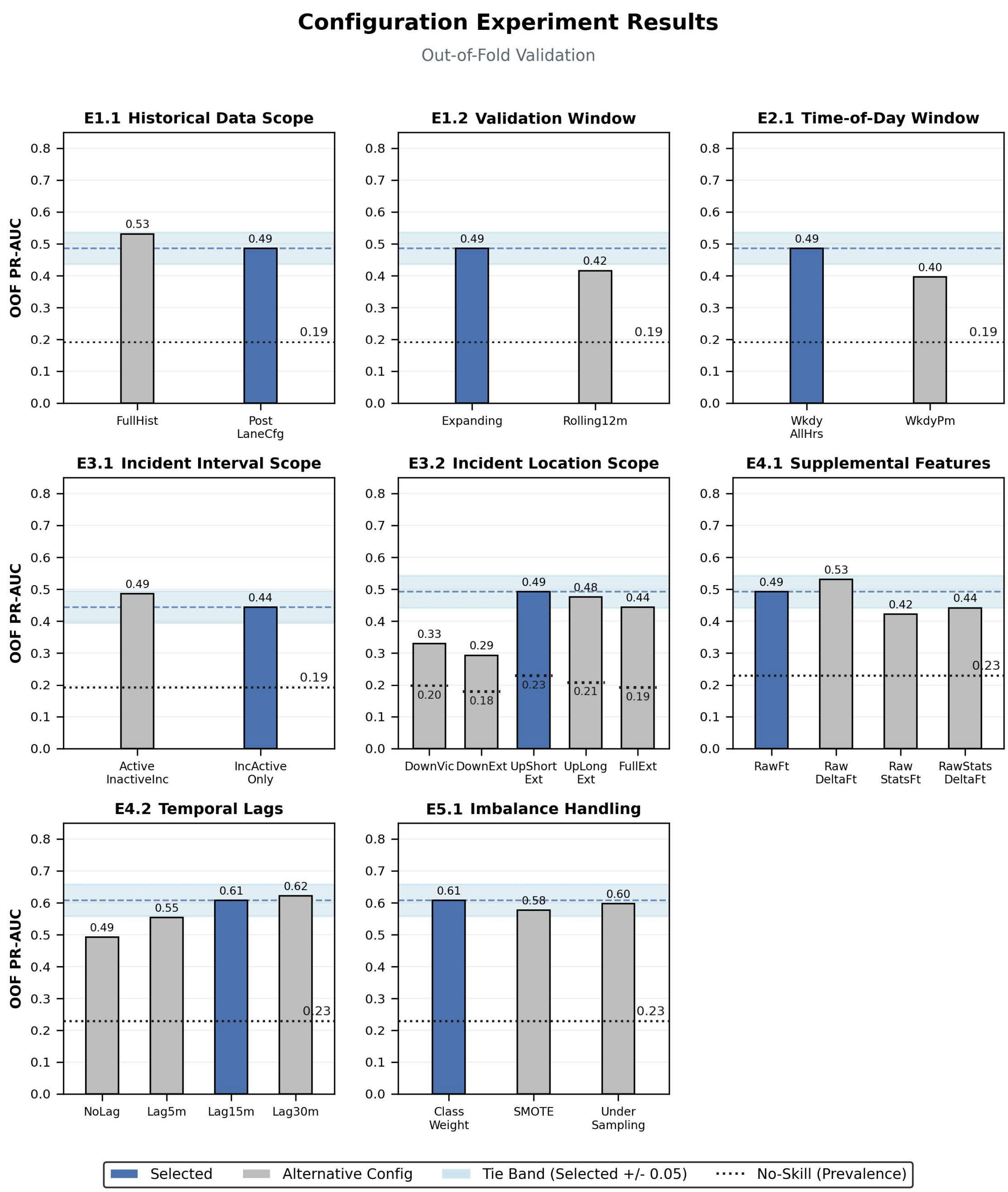



**Figure 6-23 Configuration Experiment Results**

Across all configuration categories, the models scored well above the no-skill prevalence floor, confirming that they capture a genuine predictive signal. In most categories, alternatives fell within the 0.05 PR-AUC tie band. For the tie alternatives, the final setting was chosen as noted below:

- **Historical Data Scope:** Although the full-history sample scored marginally higher, it fell within the tie band; the post-lane-configuration data was selected instead because it reflects the corridor geometry present at deployment.

- **Supplemental Features:** The raw feature set was selected because the delta and statistics features produced no meaningful gain. Raw inputs were selected for their simplicity and fewer features.
- **Incident Interval Scope:** Incident-active intervals only were preferred to align with the model's deployment scope.
- **Imbalance Handling:** Class weighting was chosen because it corrects the imbalance without altering the data distribution, avoiding both the generation of synthetic samples and the discarding of existing ones.

For certain categories, a single configuration clearly outperformed the alternatives without any ties:

- **Validation Window:** The expanding window approach substantially outperformed the 12-month rolling window. The expanding window was selected.
- **Time-of-Day Window:** Training on all weekday hours yielded a significantly higher score than restricting the data to weekday PM hours only. The all-hours configuration was selected because it provides a broader range of traffic conditions and more examples for the model to learn from.

While the remaining categories contained configurations that clearly underperformed, their top results still fell within the tie band. The final selections for these categories are as follows:

- **Incident Location Scope:** Configurations restricted to incidents downstream (south) of Yamato Road performed substantially worse, confirming that incidents on the upstream (north) segments carry the essential predictive signal for arterial delay. Adding the first upstream segment (MiliUpShortExtend) captured that signal, while extending farther upstream (MiliUpLongExtend) or across the full corridor (MiliFullExtend) produced results within the same tie band. The short upstream scope was therefore selected, as it captures the necessary information without unnecessarily extending the model's spatial coverage.
- **Temporal Lags:** Incorporating prior traffic conditions consistently improved model performance. The PR-AUC increased by 0.06 when moving from the no-lag configuration to a 5-minute lag, and by an additional 0.06 when extending the lag to 15 minutes. However, widening the window further to 30 minutes yielded only a marginal 0.01 improvement. The 15-minute lag was selected because the 30-minute configuration requires more input features without offering substantial predictive benefits.

Based on the above findings, the final configuration selected for model development is summarized as follows. The model is trained on weekday data across all hours from July 2023 to April 2025, reflecting the most recent freeway geometry and detector layout, and is restricted to the use of data from incident-active intervals only. The diversion prediction scope spans the three freeway segments between Exit 50 (Congress Ave) and Exit 44 (Palmetto Park Rd). Raw detector readings with 15-minute lags and no additional derived statistics were used as the input features. The class weighting method was used to correct class imbalance. Under this setup, the model achieves an OOF validation PR-AUC of approximately 0.61. This result is well above the no-skill prevalence baseline of 0.23, confirming the model's predictive skill beyond the base rate of a random classifier.

#### 6.6.3.6 Feature Selection

Table 6-8 lists the 105 candidate features, grouped into three categories: traffic detector variables, incident-related variables, and temporal variables.

The traffic detector variables (54 features) characterize freeway conditions along the corridor. They consist of the average speed, lane occupancy, and volume at the six freeway detectors (D1 through D6), each measured at the current interval and at lags of 5 and 15 minutes.

The incident-related variables (48 features) describe the active freeway incident within each of the three diversion segments. For each segment, they include the number of active incidents, the number of blocked lanes, the number of total lanes, whether only the shoulder is blocked, the time elapsed since the incident onset and since the lane blockage, the incident severity, and the incident type (crash, obstruction, or miscellaneous). They also include the distance from the incident location to each downstream freeway exit, which measures how close the incident is to the points where drivers can leave the freeway and divert onto the arterial.

**The temporal variables (3 features) capture time-of-day and seasonal effects through the hour of the day and the calendar month, together with a binary wet-pavement indicator that flags whether the road surface was reported as wet in the incident record, serving as a proxy for adverse weather conditions.**

Table 6-9 defines the placeholder labels (<DET>, <LAG>, <SEG>, and <EXIT>) used in the feature names.

**Table 6-8 Candidate Predictor Features and Descriptions**

| Feature Name | Description | No. of Features |
|---|---|---|
| **Traffic Detector Variables** | | **54** |
| avg_spd_<DET>_<LAG> | Average speed (mph) at freeway detector <DET>, measured over the 5-minute interval at time lag <LAG>. | 18 |
| avg_occ_<DET>_<LAG> | Average lane occupancy (percent) at freeway detector <DET>, measured over the 5-minute interval at time lag <LAG>. | 18 |
| avg_vol_<DET>_<LAG> | Average traffic volume (vehicles per 5 minutes) at freeway detector <DET>, measured over the 5-minute interval at time lag <LAG>. | 18 |
| **Incident-Related Variables (per incident segment <SEG>)** | | **48** |
| num_act_inc_<SEG> | Count of distinct incidents active in segment <SEG> during the interval. | 3 |
| num_ln_blk_<SEG> | Number of general-purpose lanes blocked by the incident in segment <SEG>. | 3 |
| tot_lns_<SEG> | Total number of general-purpose lanes in segment <SEG>. | 3 |
| blk_elapsed_<SEG> | Minutes elapsed since the lane blockage began in segment <SEG> missing when no incident is active. | 3 |
| inc_elapsed_<SEG> | Minutes elapsed since the incident onset in segment <SEG>; missing when no incident is active. | 3 |
| is_severe_<SEG> | Whether an active incident in <SEG> is flagged severe: 0 = No, 1 = Yes. | 3 |
| is_type_crash_<SEG> | Whether the incident type is a crash: 0 = No, 1 = Yes. | 3 |
| is_type_obstr_<SEG> | Whether the incident type is an obstruction (for example, debris or a disabled vehicle): 0 = No, 1 = Yes. | 3 |

| Feature Name | Description | No. of Features |
|---|---|---|
| is_type_misc_<SEG> | Whether the incident type is miscellaneous (neither a crash nor an obstruction): 0 = No, 1 = Yes. | 3 |
| is_shldr_blk_<SEG> | Whether only the shoulder is blocked in <SEG>: 0 = No, 1 = Yes. | 3 |
| mi_to_<EXIT>_<SEG> | Distance in miles measured along the I-95 mainline from the incident location in segment <SEG> to freeway exit <EXIT>. The value is computed from the incident location to each of the six exits. A positive value means the incident is still upstream of the exit and drivers can still divert there; a negative value means the incident has already passed the exit. | 18 |
| **Temporal Variables** | | **3** |
| hour_of_day | Hour of the day, based on the interval timestamp (0 to 23). | 1 |
| month | Calendar month, based on the interval timestamp (1 to 12). | 1 |
| is_wet_rd | Binary indicator of wet road-surface conditions reported in the incident record, used as a proxy for rainfall: 0 = No, 1 = Yes. | 1 |

### Table 6-9 Definitions of the Placeholder Labels Used in the Feature Names

| Placeholder | Value | Description |
|---|---|---|
| <EXIT> | E42 | Exit 42 (Hillsboro Boulevard). |
| | E44 | Exit 44 (Palmetto Park Road). |
| | E45 | Exit 45 (Glades Road). |
| | E48 | Exit 48 (Yamato Road). |
| | E50 | Exit 50 (Congress Avenue). |
| | E51 | Exit 51 (Linton Boulevard). |
| <DET> | D1 | Detector between Exit 52 (Atlantic Ave) and Exit 51 |
| | D2 | Detector between Exit 51 and Exit 50 |
| | D3 | Detector between Exit 50 and Exit 48 |
| | D4 | Detector between Exit 48 and Exit 45 |
| | D5 | Detector between Exit 45 and Exit 44 |
| | D6 | Detector between Exit 44 and Exit 42 |
| <LAG> | _t | Reading at the current interval t. |
| | _t-5m | Lagged reading 5 minutes before t. |
| | _t-15m | Lagged reading 15 minutes before t. |
| <SEG> | E50-E48 | Segment from Exit 50 to Exit 48 |
| | E48-E45 | Segment from Exit 48 to Exit 45 |
| | E45-E44 | Segment from Exit 45 to Exit 44 |

Feature correlation was evaluated using Spearman's rank correlation coefficient because the relationships among traffic variables are nonlinear. Although the XGBoost algorithm can manage multicollinearity, reducing unnecessary features simplifies the final model and improves interpretability. Figure 6-24 shows the correlations among the candidate features, with the names of the removed features highlighted in red.

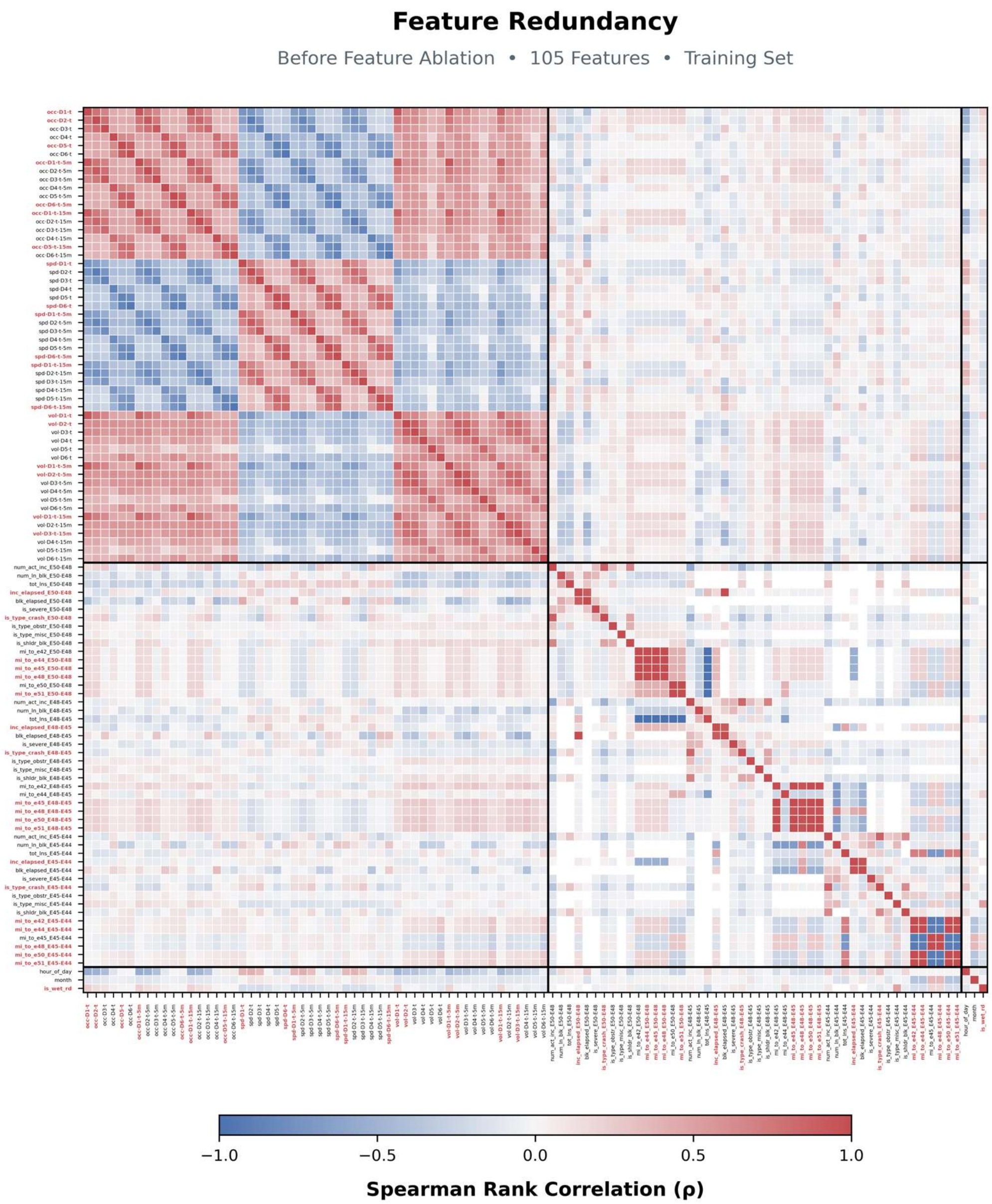



**Figure 6-24 Feature Redundancy Before Selection (105 Candidate Features)**

The redundancy was addressed in three groups: similar incident descriptors across the freeway segments and detector readings that were highly correlated with a neighboring detector or with an adjacent lag. Each group was removed only after a validation confirmed no performance degradation.

Speed, occupancy, and volume were retained despite their correlation, as they represent the fundamental variables of traffic flow and characterize complementary features of the traffic conditions.

The wet road condition (is_wet_rd) was excluded due to deployment constraints. While the DIVAS interface includes the weather field, API queries returned no data during this study, meaning the feature cannot be populated during real-time operations.

Because detector D1 is highly correlated with detector D2, it was excluded as a redundant feature with no performance degradation. Its removal was required because its live DIVAS readings drift from the RITIS training data, which could compromise real-world deployment. A detailed explanation of this discrepancy is presented in a subsequent section.

Altogether, removing the 39 features reduced the candidate set from 105 to 66 without compromising validation performance. Both the reduced set and the full candidate set yielded an OOF PR-AUC of 0.61. The remaining correlation structure is illustrated in Figure 6-25.

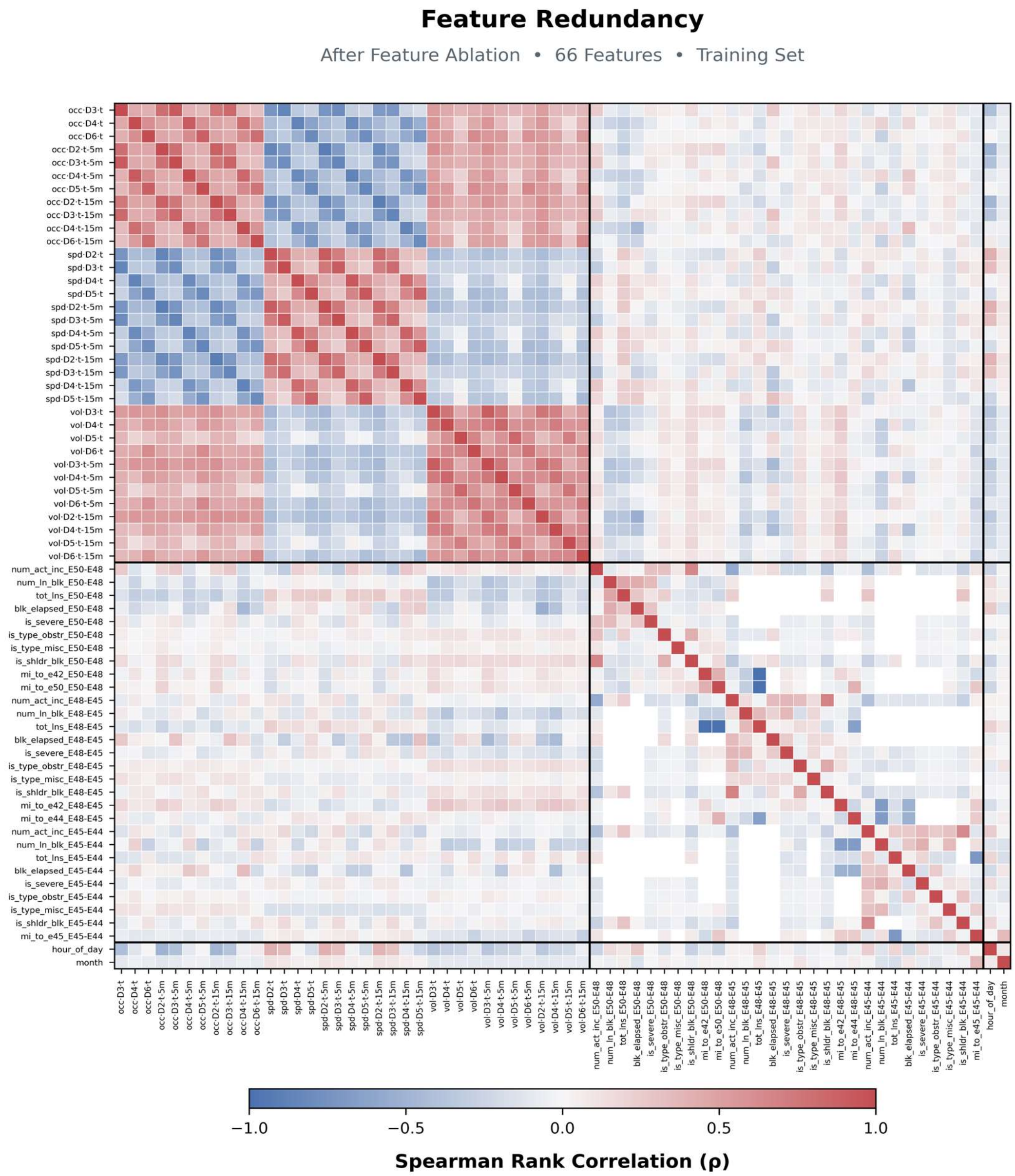



**Figure 6-25 Feature Redundancy After Selection (66 Deployed Features)**

#### 6.6.3.7 Hyperparameter Tuning

The final development phase tests whether tuning the XGBoost hyperparameters improves performance compared to the default settings. Three configurations were tested: the default settings with no tuning, a random search, and Bayesian optimization, with each of the two search methods evaluating 50 candidate configurations. These three options were compared using the 5-fold OOF PR-AUC, the same metric used for model configuration assessment and selection throughout this study. Both methods were set to select the candidate parameters that maximize this objective.

As Figure 6-26 shows, both tuning methods yielded only minor improvements over the default configuration. The default setting achieved an OOF PR-AUC of 0.61 against a no-skill prevalence baseline of 0.23. Table 6-10 lists the search space and the resulting configurations. Bayesian optimization achieved the highest score, exceeding the default by just 0.03 PR-AUC, while the random search scored 0.01 lower than the Bayesian method. Because these marginal gains fell within the 0.05 tie band, the default hyperparameters were kept for simplicity.

Finally, the previously selected model configuration with the default hyperparameters was retrained on the complete training dataset and evaluated once on the held-out test year.

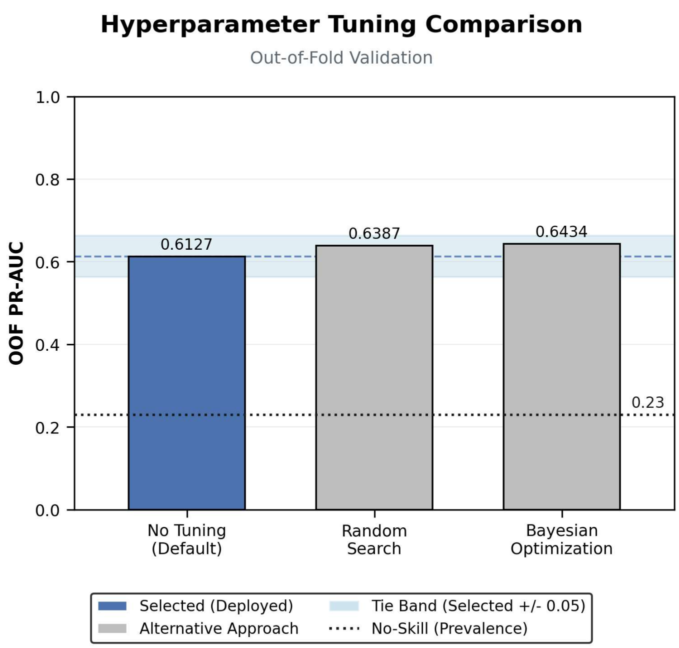



**Figure 6-26 Hyperparameter Tuning Comparison**

**Table 6-10 Hyperparameter Search Space and Selected Configurations**

| Hyperparameter | Search Space | No Tuning (Default) | Random Search | Bayesian |
|---|---|---|---|---|
| n_estimators | {200, 300, 500} | 300 | 300 | 500 |
| max_depth | {3, 4, 5, 6} | 5 | 3 | 5 |
| learning_rate | {0.01, 0.03, 0.05, 0.1} | 0.05 | 0.03 | 0.01 |
| subsample | {0.6, 0.7, 0.8, 0.9} | 0.8 | 0.8 | 0.6 |
| colsample_bytree | {0.5, 0.6, 0.7, 0.8} | 0.7 | 0.8 | 0.6 |
| min_child_weight | {5, 10, 15, 25, 50} | 15 | 5 | 50 |
| gamma | {0.1, 0.2, 0.5, 1.0, 2.0} | 0.5 | 2 | 0.5 |
| reg_alpha | {0.05, 0.1, 0.5, 1.0, 2.0} | 0.5 | 2 | 0.05 |
| reg_lambda | {1.0, 2.0, 3.0, 5.0, 10.0} | 3 | 1 | 2 |

### 6.6.4 Performance on the Held-Out Test Year

This section evaluates the finalized model on the held-out test dataset. After training on the full training set, the model was evaluated on the test year (May 2025 to April 2026), which remained untouched and unused during training. This evaluation demonstrates the operational performance the model would achieve during real-world deployment.

#### 6.6.4.1 Overall Discrimination

Figure 6-27 presents two summary measures of model performance on the held-out test year: the ROC curve in panel (a) and the PR curve in panel (b). Both curves demonstrate the model's ability to distinguish delay from non-delay intervals by plotting performance across all possible classification thresholds, $\tau$.

In panel (a), the ROC curve has an AUC of 0.77. This metric shows the model’s overall ability to distinguish between two classes: delay and normal conditions. A value of 0.50 is the no-skill baseline, equivalent to random guessing, and 1.00 is perfect separation. At 0.77, the score confirms that the model can separate these two classes well.

In panel (b), the PR curve has an AUC of 0.52. Because delays make up only about 21 percent of the intervals (prevalence), this curve provides a stricter view of performance compared to the ROC curve. It focuses on the rare delay class and sets aside the many easy non-delay intervals by ignoring correctly identified normal conditions (true negatives). For a PR curve, the no-skill baseline equals the prevalence at 0.21, so a model with no predictive value would score at this level. The model's score of 0.52 is roughly two and a half times that baseline, which shows a substantial ability to identify the rare delay intervals.

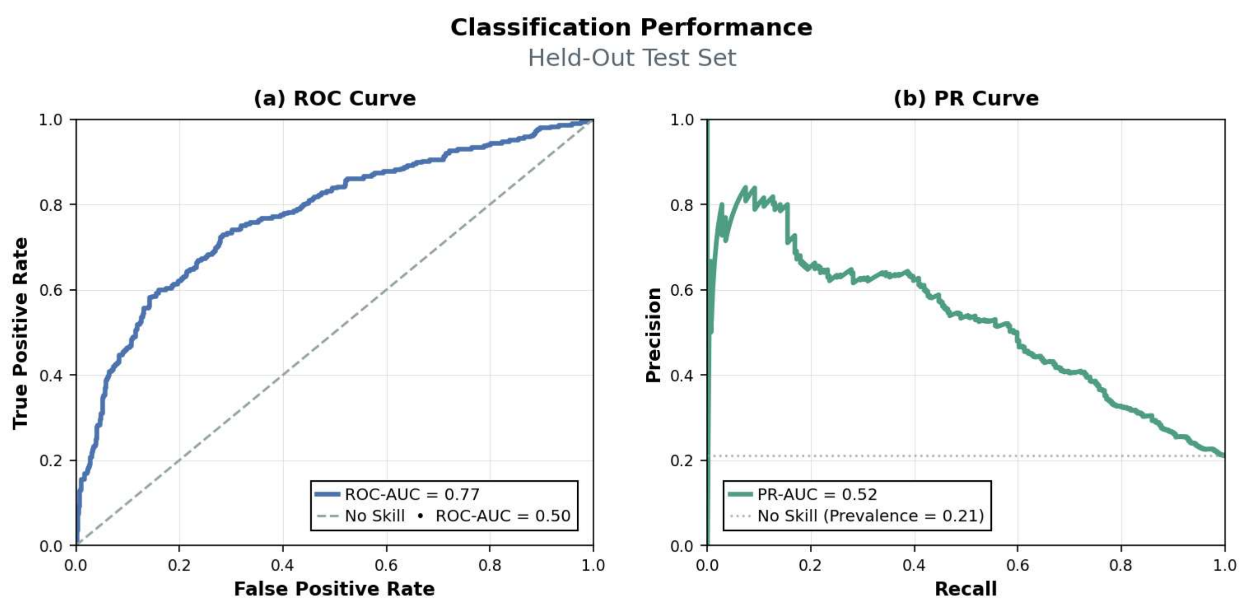
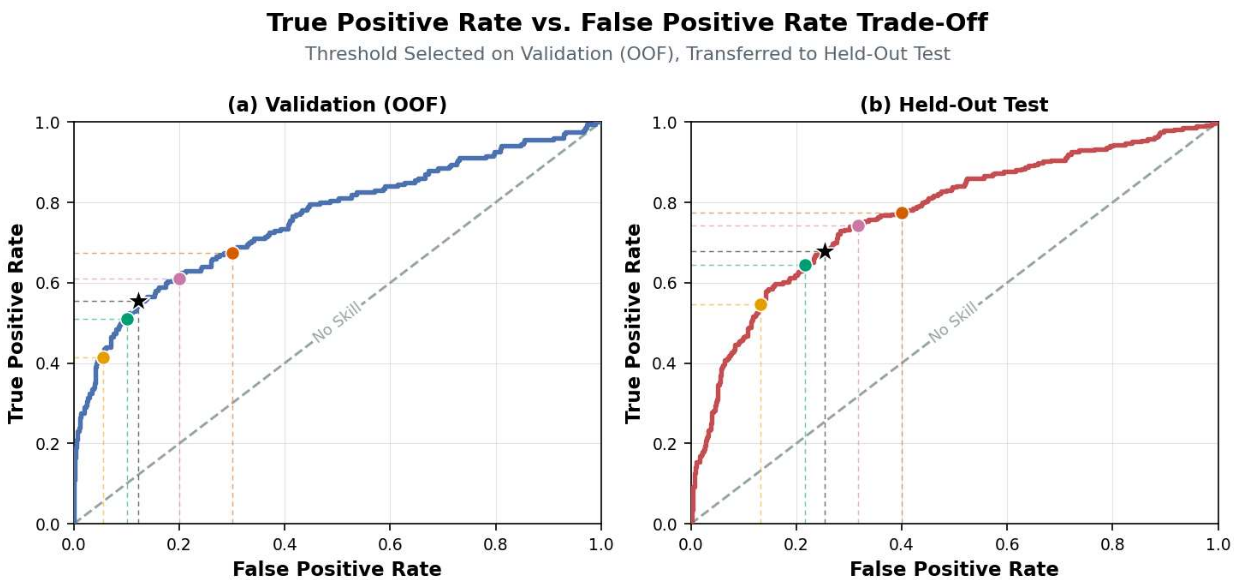



**Figure 6-27 ROC and Precision-Recall Curves on the Held-Out Test Data**

#### 6.6.4.2 Threshold Selection and Operational Trade-Offs

The ROC-AUC and PR-AUC reported above describe the model across the full range of the classification threshold, τ. In real-world deployment, a single threshold τ must be selected to turn predicted score into a decision: scores above τ are flagged as delay and trigger an alert, scores below are treated as normal.

Figure 6-28 visualizes the performance trade-offs at various candidate thresholds. Panel (a) shows the performance on the validation set, which is used to select τ. Panel (b) shows the TPR and FPR that these same fixed τ produce on the held-out test set. The performance under a fixed τ differs between the validation and test sets, reflecting how the model behaves when deployed.




| | Operating Point | τ | Validation Set TPR | Validation Set FPR | Held-Out Test Set TPR | Held-Out Test Set FPR |
|---|---|---|---|---|---|---|
| ● | FPR-budget 5% | 0.69 | 41.5% | 5.5% | 54.6% | 13.2% |
| ● | FPR-budget 10% | 0.57 | 51.0% | 10.2% | 64.4% | 21.6% |
| ● | FPR-budget 20% | 0.45 | 61.0% | 20.0% | 74.3% | 31.7% |
| ● | FPR-budget 30% | 0.37 | 67.5% | 30.0% | 77.5% | 40.0% |
| ★ | F1-max | 0.53 | 55.5% | 12.3% | 68.0% | 25.5% |

**Figure 6-28 True Positive Rate versus False Positive Rate Trade-Off across Decision Thresholds**

The choice of τ directly affects the trade-off between TPR and FPR. A higher τ makes the model more conservative. It requires higher confidence before triggering an alert, lowering both the FPR (normal intervals misclassified as delays) and the TPR (the proportion of correctly flagged delay intervals). For example, if operators want to strictly limit false alarms, they could select the yellow point in Panel (a) to cap the FPR at 5% with a threshold of τ = 0.69, which corresponds to a 41.5% TPR on the validation set. When this same threshold is applied to the test set in Panel (b), the FPR increases to 13.2%, and the TPR increases to 54.6%.

For deployment, this study used the F1-maximizing method to select the threshold, i.e., the value of $\tau$ that best balances precision and recall on the validation data. This threshold is $\tau = 0.53$, at which the validation TPR is 55.5% and the FPR is 12.3%. Applying this exact same threshold to the held-out test year allows the model to detect more delay intervals, raising the TPR to 68.0% but also increasing the FPR to 25.5%.

It is important to note that TPR and FAR are computed over individual 5-minute intervals and are only used to evaluate performance during model development. In practice, operators manage a delay as a single event. A single delay event contains multiple intervals, and one alert within that event is enough to prompt a response. After the first alert triggers a response, additional alerts every five minutes are unnecessary for that specific event. Event-level performance is discussed in a later section.

#### 6.6.4.3 Feature Interpretation

SHAP is a widely used method for machine learning model interpretation. SHAP values are derived from game theory and quantify how much each feature contributes to individual predictions

Figure 6-29 presents the contributions of the top ten features to the model output on the held-out test set. Each dot represents a single data instance, with its horizontal position indicating the SHAP value, showing how much that feature pushed the model's predicted delay score up (right side) or down (left side). The color of each dot reflects the actual feature value for that instance, with red denoting high values, blue indicating low values, and gray indicating missing values. When a feature's SHAP values cluster densely around zero, it contributes little to the model's output.

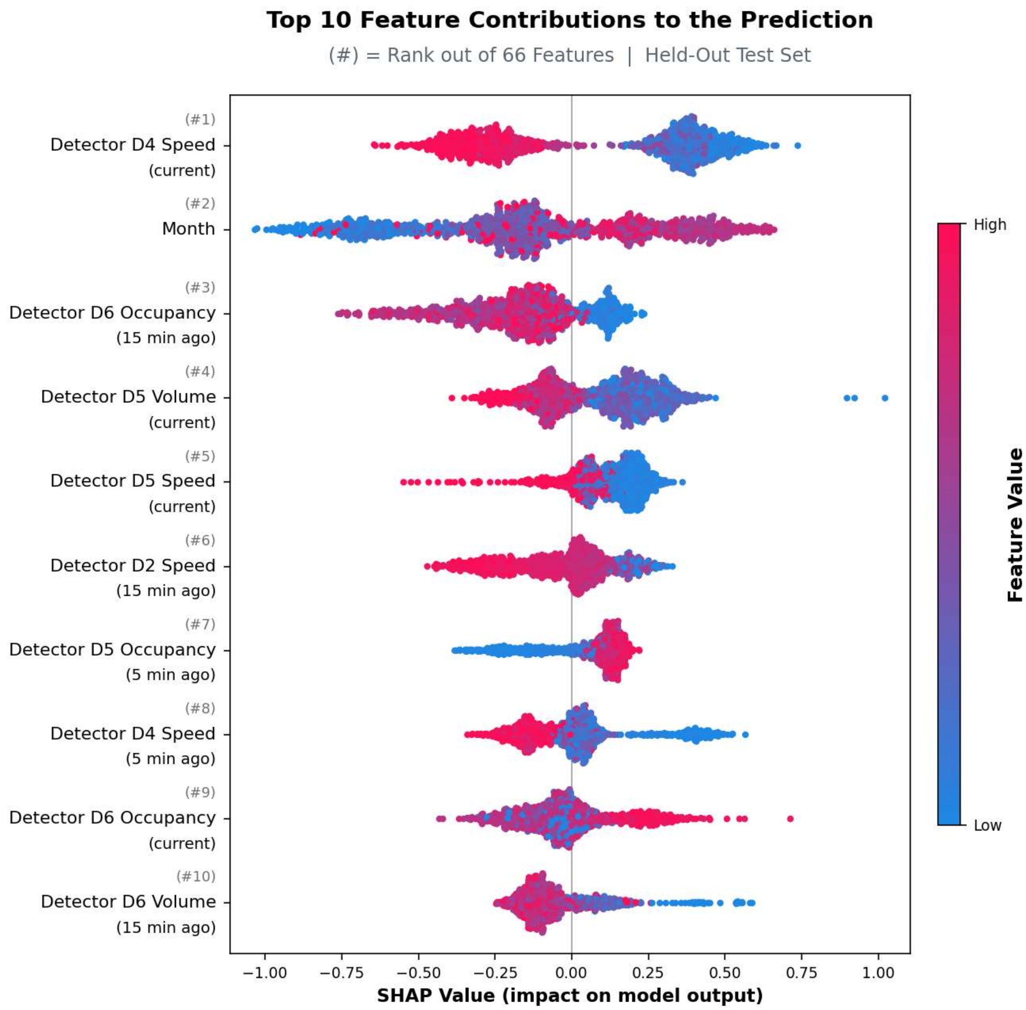



**Figure 6-29 Top 10 Feature Contributions to the Prediction**

To illustrate how individual predictors contribute to the model output, the five top features are examined below. These variables contribute to the prediction of impacts on arterial travel times:

- **Detector D4 Speed (avg_spd_D4_t):** the most influential feature. Detector D4 is located downstream of Exit 48 (Yamato Road) on the segment between Exit 48 and Exit 45. Low speeds at D4 (blue) raise the predicted delay; high speeds (red) lower it. When a bottleneck forms in this segment, the queue grows back toward Exit 48, where motorists can divert onto Military Trail. A speed drop at D4 is therefore an early sign of diversion.
- **Month of the Year (month):** The late summer months (roughly July through September) increase the predicted delay score, while the spring months (January through March) lower the score. This pattern aligns with higher freeway incident rates during the late summer, possibly due to more rainy conditions, which trigger more frequent diversions onto the arterial road.
- **Detector D6 Occupancy (avg_occ_D6_t-15m):** This feature is the lane occupancy at Detector D6, the southernmost station (between Exit 44 and Exit 42), measured 15 minutes before the prediction time. Because D6 is located at the downstream end of the study area, it only captures the traffic that has already passed through the incident segments where

blockages occur. As a result, the higher occupancy at D6 means traffic is still flowing all the way through to the south end of the corridor, so the model slightly lowers the predicted delay. Conversely, an unusually empty D6 indicates that an upstream blockage is restricting flow and that vehicles cannot reach the south-end detector.

- **Detector D5 Speed and Volume (avg_spd_D5_t, avg_vol_D5_t):** Detector D5 is adjacent to and downstream of Detector D4. It is located between Exits 44 and 45. Low speed combined with low volume on this segment is the sign of a severe blockage: when lanes are blocked, fewer vehicles get through, so a detector past the blockage records both slow speeds and a reduced count. This is what separates a real blockage from ordinary heavy traffic, which would show low speed but high volume.

Notably, no incident variables rank among the top features. This occurs because the model operates on 5-minute prediction intervals. When a single incident record is divided into these intervals, its attributes, such as the number of blocked lanes or the incident type, remain constant for the entire duration. Since these values do not change over time, they offer less predictive power to the model. In contrast, detector measurements like speed, volume, and occupancy are highly dynamic and change continuously. More importantly, these traffic state variables inherently reflect the physical impacts of an incident on the roadway. When a crash blocks a lane, the resulting loss in capacity immediately causes speeds to drop and occupancy to build up. Because the real-time detector data already captures the exact severity and consequences of blockage, the model primarily relies on these dynamic signals rather than the static incident descriptors to predict delays.

Figure 6-30 illustrates the SHAP dependence plot for freeway speed at detector D4 (avg_spd_D4_t). The feature's contribution to the model varies significantly depending on the traffic state. Lower speeds strongly drive the model to predict arterial delays, while higher speeds push the prediction toward normal conditions. A clear reversal point occurs at approximately 50 mph, marking where the SHAP value shifts from positive to negative. At speeds below this threshold, the model actively predicts delays.

By coloring the data points according to detector D4 occupancy (avg_occ_D4_t), a clear division into two traffic regimes emerges. On the top left of the plot, a dense cluster of red points represents severe congestion, characterized by low speeds of 10–20 mph and high occupancy. In this state, the model is highly confident in predicting an arterial delay. Conversely, the high-speed points on the right form a blue, low-occupancy cluster that favors normal conditions. This distinct separation confirms that the model evaluates cohesive traffic states characterized by both speed and occupancy, rather than relying on speed in isolation. Ultimately, these patterns represent the expected behavior of congestion building up and highlight the strong interdependence between freeway and arterial traffic conditions.

Figure 6-30 presents a SHAP dependence plot for avg_spd_D4_t, the freeway speed at detector D4 (downstream of Exit 48: Yamato Road). The impact of this feature depends heavily on current traffic conditions. Lower freeway speeds on the x-axis correspond to higher SHAP values on the y-axis, driving the model to predict an arterial delay. In contrast, higher speeds shift the prediction toward normal conditions. A reversal point at approximately 50 mph shows where the SHAP value changes from positive to negative, highlighting the speed at which the model alters the direction of its prediction. At freeway speeds below this point, the model starts predicting delays. The dense cluster in the top left shows low speeds of 10–20 mph, with the red coloring indicating high occupancy. When the freeway speed at D4 drops into this range, the model's confidence in predicting an arterial delay increases significantly.

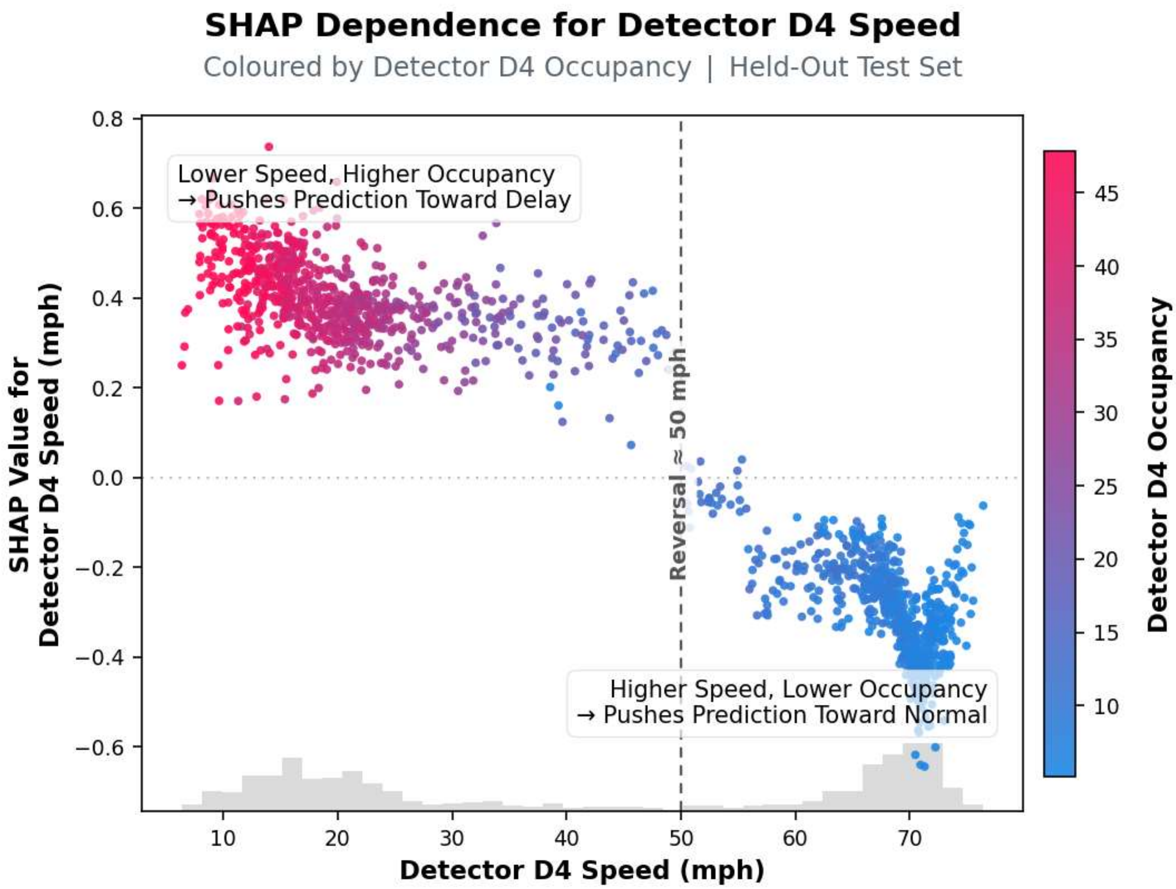



**Figure 6-30 Dependence Plot for Freeway Speed at Detector D4**

Coloring the points based on occupancy at Detector D4 (avg_occ_D4_t) reveals two distinct regimes. First, the lower-speed points form a red cluster of higher occupancy that corresponds to higher SHAP values on the y-axis, causing the model to predict an arterial delay. Second, the higher-speed points create a blue cluster of lower occupancy that shifts the SHAP values downward, leading the model to predict a normal condition. This illustrates that the model captures freeway congestion based on the combination of lower speeds and higher occupancies, rather than relying on speed alone. This freeway congestion pattern and its resulting SHAP values highlight the interdependence between the freeway and the parallel arterial.

#### 6.6.4.4 Event-Level Deployment Performance

Evaluating a model’s real-world deployment performance requires aligning the model's scoring with how it would actually be used in the field. The model operates on an interval-based framework, analyzing data and generating a prediction every five minutes. In practice, however, a traffic-management operator does not make a new decision in each five-minute interval; an ongoing incident is handled as a single event, and the operator decides once how to respond to the delay (i.e., adjusting the arterial signal timing). To align with real-world operation, the evaluation consolidated the model's interval predictions into a single decision per event and evaluated it over a single response window. This window runs from incident detection to clearance, with its duration capped at 120 minutes, since that is sufficient to cover the traffic diversion period. When multiple incidents occurred simultaneously, their overlapping windows were merged into a single event a single operational response covers all concurrent incidents along the parallel arterial.

On the model side, an event is counted as an alert if the prediction score exceeds the classification threshold ($\tau = 0.53$) at least one interval within the response window. On the ground-truth side, the observed arterial travel time determines the actual outcome: an event qualifies as a delay if the target arterial's travel time reaches the p90 threshold in at least one interval within the same window. Crossing these two outcomes places each event in one of four categories:

- **Correct Alert:** The arterial reaches the p90 line, and the model issues an alert.
- **Missed:** The arterial reaches the p90 line, but the model issues no alert.
- **False Alarm:** The model issues an alert, but the arterial stays below the p90 line.
- **Correctly Silent:** The arterial stays below the p90 line, and the model issues no alert.

The evaluation was limited to the conditions under which the model would actually operate: the core PM peak from 3:30 to 6:30 PM, including incidents on the I-95 southbound general-purpose lanes between Exits 44 and 50. Non-incident reports, such as roadwork or recurring congestion, were excluded. For each event, the model begins issuing predictions only once the incident is reported in the data feed, since the incident record is one of its inputs; this report time marks the start of the response window. Under these parameters, 64 in-scope incidents from the one-year held-out test set were consolidated into 57 independent events.

Table 6-11 summarizes the event-level evaluation across the 57 independent events in the one-year held-out test set. Of these, 28 events (49.1%) were Correct Alerts and 2 (3.5%) were Missed, the cases in which the arterial travel time reached the p90 threshold. The remaining 27 events stayed below the p90 threshold: 16 (28.1%) were False Alarms, and 11 (19.3%) were Correctly Silent.

**Table 6-11 Event-Level Operational Performance Summary**

| Observed Arterial Travel Time | Alert Outcome | Number of Events | Percentage of Events |
|---|---|---|---|
| **Reached p90 (30 Events)** | Correct Alert | 28 | 49.1% |
| | Missed | 2 | 3.5% |
| **Stayed below p90 (27 Events)** | False Alarm | 16 | 28.1% |
| | Correct Silent | 11 | 19.3% |
| **Total** | | **57** | **100.0%** |

In operational terms, an operator would see roughly 28 correct alerts and 16 false alarms over a typical year. Although these false alarms did not reach the p90 delay threshold, many still occurred while actual congestion built up. Table 6-12 examines the 16 false alarms in greater detail. For each event, it reports the highest percentile that the arterial travel time reached during the response window. As the table shows, 10 of the 16 false alarms (62.5%) coincided with congestion, with the arterial travel time peaking at or above its p70. These near-threshold alerts remain useful to the operator, since they still flag genuine congestion building on the arterial.

**Table 6-12 Distribution of False Alarms by Peak Travel Time**

| Observed Arterial Peak Travel Time | Number of False Alarm Events | Percentage of False Alarm Events |
|---|---|---|
| Between p80 and p90 | 6 | 37.5% |
| Between p70 and p80 | 4 | 25.0% |
| Between p60 and p70 | 3 | 18.8% |
| Between p50 and p60 | 1 | 6.3% |
| Below p50 | 2 | 12.5% |
| **Total** | **16** | **100.0%** |

### 6.6.5 Operational Feasibility Assessment with Live DIVAS Data

The predictive model developed, as discussed in earlier sections of this section, was trained on historical detector data from the RITIS archive. Real-time operations, however, require the model to run using real-time data stream(s) that is used as input to the model to allow the prediction. This study investigates the use of the Data Integration and Video Aggregation System (DIVAS) that was developed by the FDOT to support the integration and management of real–time information. The system has been developed as a centralized data hub for the aggregation, fusion and dissemination of near real–time transportation information and live streaming video. To assess the feasibility of utilizing DIVAS in real-world implementation of the developed model, live data was collected from the received DIVAS stream to evaluate the DIVAS feed along two operational dimensions: the latency (freshness) of the detector data it delivers and the agreement of that data with the RITIS reference against which the model was trained. Both of these dimensions affect the usability of the system for real-time application of machine learning.

#### 6.6.5.1 DIVAS Data Collection Process

The data collection covered 14 weekdays between May 26 and June 12, 2026. Throughout this timeframe, an automated Python script retrieved data from the DIVAS real-time API each day from 2:00 pm to 9:30 PM. The collection process was as follows:

- **Sensor Data:** The script retrieved data every 20 seconds. Each query returned a full snapshot of the FDOT sensor data for Palm Beach County, and the script extracted the data for six targeted detectors (D1–D6) from this snapshot that are used in this study.
- **Incident Data:** The script also queried incident data on a 5-minute cycle. Each query returned the complete list of active incidents across the district at the time of the query, rather than only newly reported events. Every incident record includes its DIVAS identifier, type, status, severity, affected lanes, location, first reported time, and last

updated time. Because each snapshot also carries the incident's last-updated timestamp and its full attributes, the incident information is captured as soon as it appears in the feed, including any revision an operator makes while the incident is active, such as a change in the number of lanes closed,

- **Logging:** The script recorded the results of every query in a timestamped log. This log captures the request time, response status, number of returned records, and round-trip latency.

Over the collection period, the logs recorded 25,648 sensor queries and 1,859 incident feed queries. Every query returned a successful response with no HTTP 404 errors, server errors, warnings, or failed requests. These logs confirm that if any data is missing, the omission originates in the DIVAS feed itself rather than from a local retrieving issue (such as a lost connection).

#### 6.6.5.2 DIVAS Detector Latency

Data latency was measured directly from the feed. For each sensor query, the collector records the request time along with the observation timestamp embedded by DIVAS in each detector record; the end-to-end latency is the difference between the two.

As shown in Figure 6-31 DIVAS Real-Time Detector Data Latency, data latency is consistent across all six detectors, with medians tightly clustered at 41–42 seconds. This observed latency is approximately twice the nominal 20-second polling rate documented for the SunGuide Transportation Sensor Subsystem (TSS) (SunGuide Software User's Manual v9.0). The 20-second nominal rate applies only to the field polling stage within the SunGuide system and does not account for the end-to-end latency of transmitting data from SunGuide to DIVAS. While the response to each API query returns instantly, DIVAS updates its data more slowly than the 20-second query interval. Consequently, consecutive queries often return the identical snapshot and carry the same DIVAS timestamp.

The maximum latency was observed at 220 seconds. No polls exceeded the 300-second failure threshold defined in this study, which corresponds to the predictive model's 5-minute refresh rate. This level of latency remains operationally acceptable.

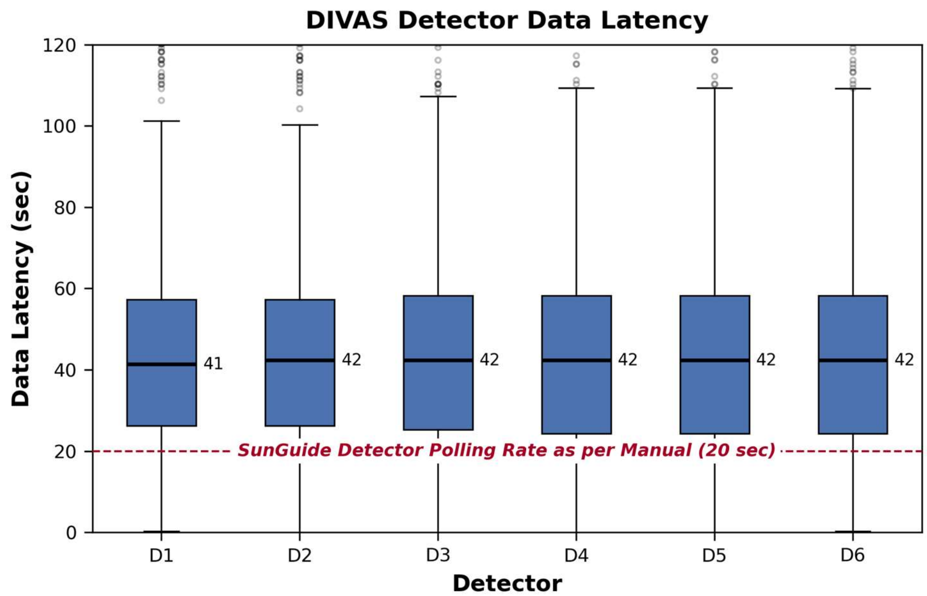



**Figure 6-31 DIVAS Real-Time Detector Data Latency**

### 6.6.5.3 DIVAS Detector Data Quality

As the model was trained exclusively on RITIS detector data, feeding it inconsistent data during real-time deployment will degrade the model's predictive performance. DIVAS and RITIS observations were therefore paired for the same six detectors over the same 14 weekday afternoon peak periods (1,260 paired 5-minute intervals per detector), and the three detector channels utilized by the model (volume, speed, and occupancy) were compared.

Figure 6-32**(a)** through Figure 6-32**(c)** present the detector-by-detector agreement; in each panel, the dashed line denotes perfect (1:1) agreement, with the Pearson correlation (r) and the median DIVAS-to-RITIS ratio annotated.

As shown in Figure 6-32**(a),** volume is the least consistent of the three detector-measured variables (volume, speed, and occupancy), with r ranging from 0.60 to 0.88. The inconsistencies vary significantly among the locations. Notably, detector D6 yields the lowest correlation (r = 0.60), and detector D1 (r = 0.62) exhibits the most severe systematic error, with the lowest median DIVAS/RITIS ratio (0.56), indicating poor overall data quality. The per-detector median ratios also vary widely, with detector D1 and D4 reporting values below the corresponding RITIS volume (with ratios of 0.56 and 0.92, respectively), whereas D2 and D3 reported higher values (with ratios of 1.14 and 1.25, respectively). Data from detectors D5 and D6 align more closely with the RITIS data (with ratios of 1.09 and 1.07, respectively). Given these discrepancies, selecting detector stations for real-time deployment requires careful evaluation of data consistency and diagnosis of the reasons for identifying inconsistencies. High divergence between RITIS and SunGuide data, such as the divergence observed in traffic volume measurements from detector D1, may necessitate excluding this data from the training, additional calibration of the sensors, or better alignment of SunGuide data (the source of DIVAS data) and RITIS archived data, which is also based on SunGuide data but with possibly additional processing and cleaning of the data.

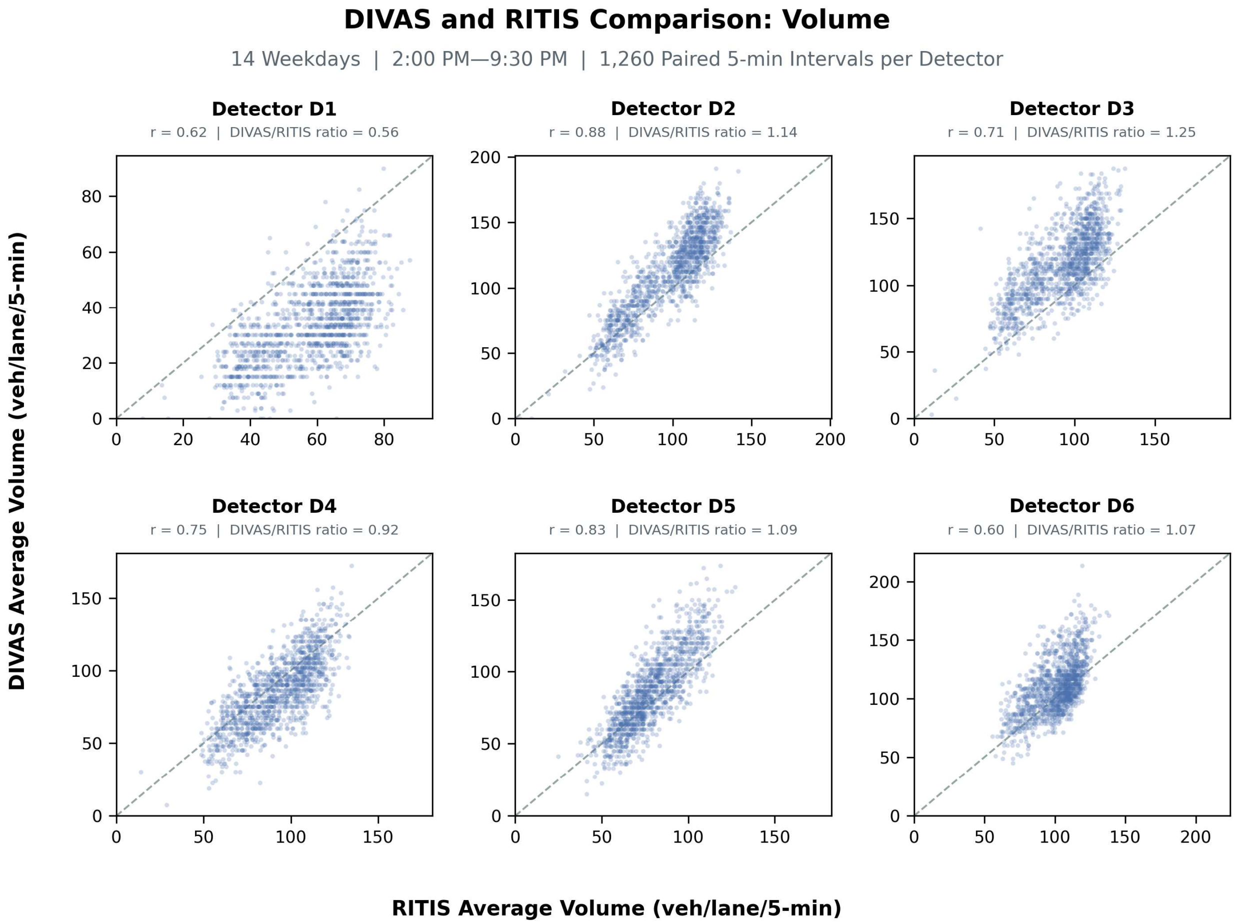



**Figure 6-32(a) Comparison of DIVAS and RITIS Volume Data**

As shown in Figure 6-32**(b),** speed is the most reliable of the three measures, showing high consistency between the two sources (with r ranging from 0.83 to 1.00 and a ratio ranging between 0.96 and 1.05). The three downstream detectors, D4 to D6, show near-perfect correlation with r values ranging from 0.99 to 1.00. While detectors D1 to D3 exhibit slightly less correlation (r ranging from 0.83 to 0.86), their median ratios are very close to 1.0. Because these values align closely across all locations, it can be concluded that the DIVAS speed measurements do not introduce data mismatches and can be used for real-time operations with high confidence.

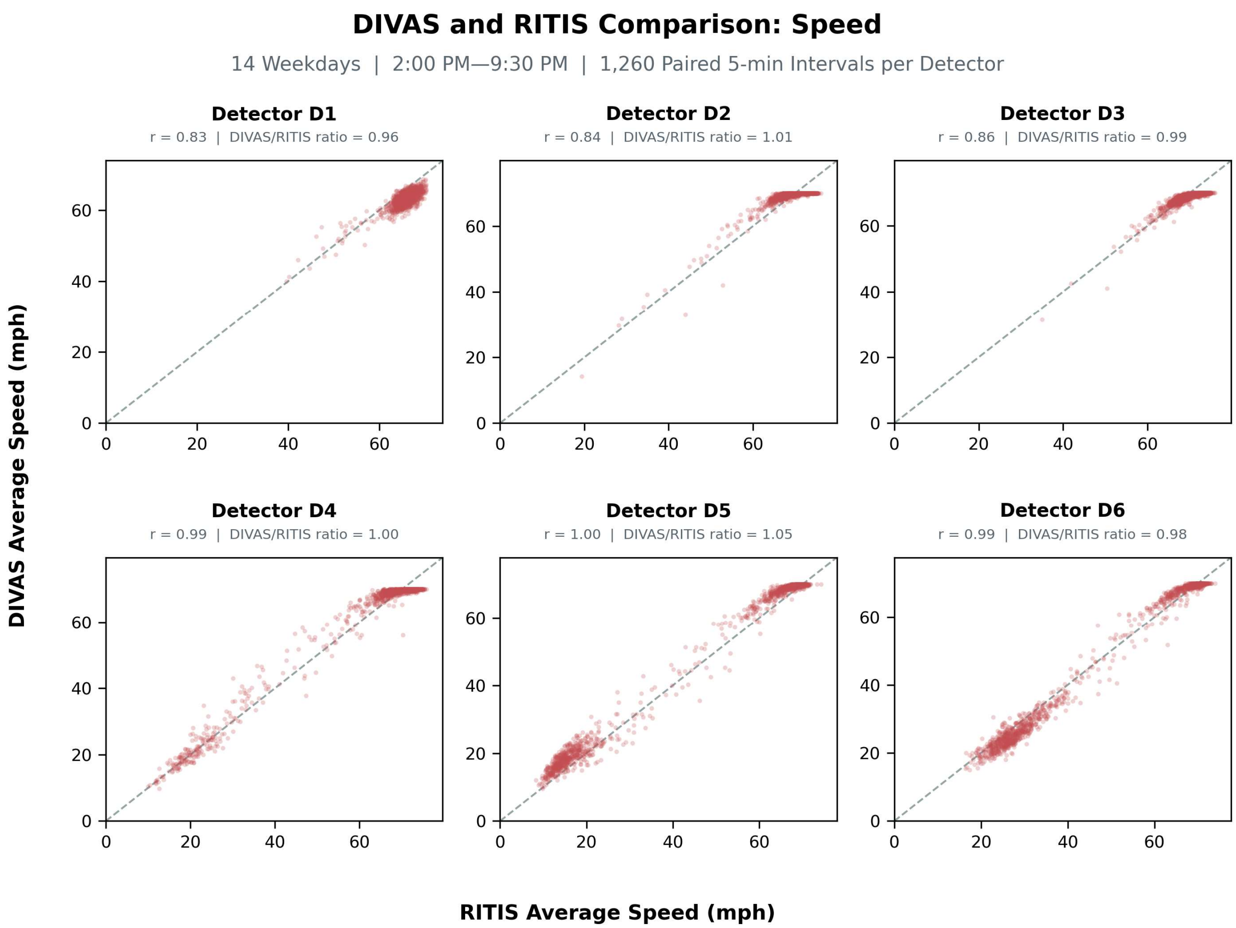



**Figure 6-32(b) Comparison of DIVAS and RITIS Speed Data**

As shown in Figure 6-32**(c),** occupancy falls between volume and speed in terms of consistency (r = 0.54–0.98 and ratio = 0.78–1.24). Four of the six detectors show strong correlation (detectors D2, D4, D5, and D6, with r ranging from 0.92 to 0.98). The exceptions are detector D3, which records values approximately 24 percent higher than RITIS (ratio = 1.24); and detector D1, which records lower values (ratio = 0.78) and exhibits the weakest correlation of any detector-channel pair (r = 0.54).

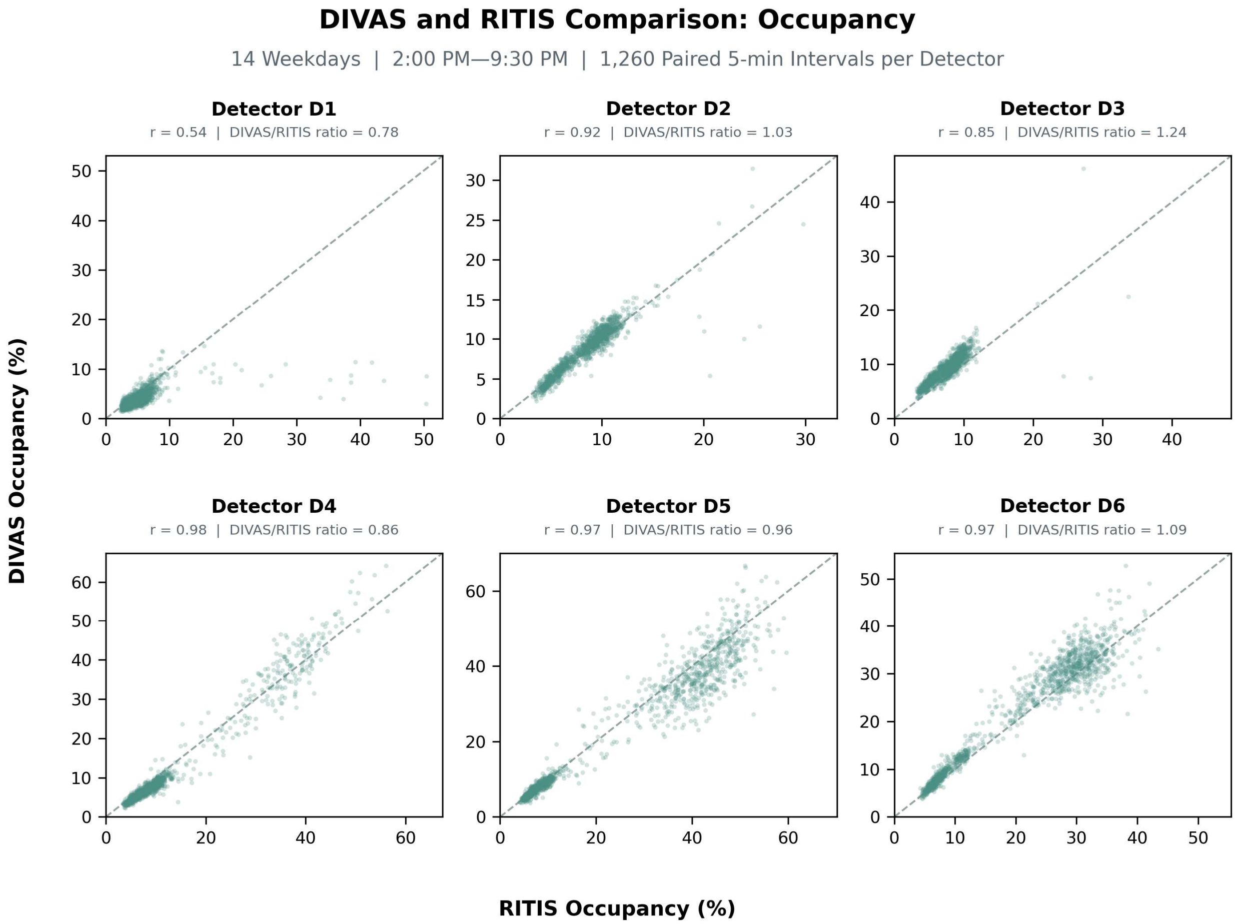



**Figure 6-32(c) Comparison of DIVAS and RITIS Occupancy Data**

Overall, data consistency between the RITIS training source and the live DIVAS feed varies significantly by the variable measured. While speed remains highly consistent and occupancy aligns well for most detectors between the two data sources, volume diverges significantly across detectors. Detector D1 shows a consistently severe deviation across all three variables and exhibits the weakest performance in both volume and occupancy. It can be concluded that while the speed and occupancy values from DIVAS are generally ready for real-time application of the developed model, the volume measurements require correction or recalibration before actual deployment. It is recommended that FDOT examine this issue.

#### 6.6.5.4 DIVAS Incident Feed

This section evaluates the quality of the DIVAS incident feed during the 14-dweekday data collection period. The incident data received using DIVAS were filtered to include only those occurring on the I-95 southbound general-purpose lanes between Exits 44 and 50. Events such as roadwork, congestion, and other non-incident information were excluded, and only incidents occurring in the PM peak (3:30 PM–6:30 PM) were included in the evaluation of incident quality.

The RITIS incident archive was collected over the same 14-weekday period to compare incident data between RITIS and DIVAS using matching incident IDs. Significant discrepancies were observed, as detailed in Table 6-13. While the RITIS archive recorded 20 incidents within the scope of the evaluation, only 4 appeared in the live DIVAS feed. However, as described later in this section, only one incident record was consistent between DIVAS and RITIS in terms of the reported incident information (location, type, lane blockage, and start time), while the remaining three incidents were inconsistent between the two data sources.

## Table 6-13 Comparison of In-Scope Incidents between RITIS Archive and DIVAS Live Feed

| Date | Incident ID | Start Time | Description (as recorded) | Type | Max Lanes Closed | Duration | DIVAS Live Feed |
|---|---|---|---|---|---|---|---|
| **May 27, 2026** | **2782362** | 4:29 PM | Exit 50: Congress Ave | Disabled Vehicle | — | 12 min | — |
| **May 28, 2026** | **2782703** | 4:55 PM | Exit 45: SR-808/Glades Rd | Crash | 1 | 3 h 10 min | — |
| | **2782706** | 5:04 PM | Exit 50: Congress Ave | Disabled Vehicle | 1 | 44 min | — |
| **May 29, 2026** | **2783051** | 5:53 PM | Exit 45: SR-808/Glades Rd | Disabled Vehicle | 1 | 2 h 31 min | — |
| **June 1, 2026** | **2783794** | 4:58 PM | Exit 45: SR-808/Glades Rd | Disabled Vehicle | 1 | 2 h 10 min | — |
| **June 3, 2026** | **2784356** | 4:33 PM | Exit 44: CR-798/Palmetto Park Rd | Crash | — | 45 min | Inconsistent with RITIS* |
| | **2784371** | 5:13 PM | Exit 50: Congress Ave | Disabled Vehicle | 1 | 2 min | — |
| **June 4, 2026** | **2784692** | 6:24 PM | Exit 44: CR-798/Palmetto Park Rd with Right shoulder blocked | Disabled Vehicle | 1 | 10 min | — |
| **June 5, 2026** | **2784928** | 3:32 PM | Exit 50: Congress Ave | Disabled Vehicle | 1 | 2 h 45 min | — |
| | **2784936** | 3:52 PM | Exit 45: SR-808/Glades Rd | Crash | 1 | 2 h 28 min | — |
| | **2784962** | 5:03 PM | Exit 44: CR-798/Palmetto Park Rd | Crash | 1 | 2 h 21 min | Consistent with RITIS |
| | **2784964** | 5:09 PM | Exit 44: CR-798/Palmetto Park Rd | Crash | — | 2 min | — |
| **June 8, 2026** | **2785677** | 2:46 PM | Exit 45: SR-808/Glades Rd | Disabled Vehicle | 1 | 44 min | — |
| | **2785715** | 4:41 PM | Exit 48: SR-794/Yamato Rd | Disabled Vehicle | 1 | 40 min | — |
| **June 9, 2026** | **2786013** | 4:28 PM | Exit 44: CR-798/Palmetto Park Rd with reported Left lane blocked | Disabled Vehicle | — | 15 min | — |
| **June 11, 2026** | **2786590** | 4:05 PM | Exit 48: SR-794/Yamato Rd | Obstruction | — | 2 min | — |
| | **2786609** | 5:16 PM | Exit 44: CR-798/Palmetto Park Rd | Disabled Vehicle | 3 | 1 h 19 min | Inconsistent with RITIS* |
| **June 12, 2026** | **2786886** | 4:07 PM | Exit 45: SR-808/Glades Rd | Crash | 1 | 41 min | — |
| | **2786887** | 4:12 PM | Exit 48: SR-794/Yamato Rd | Crash | 1 | 52 min | Inconsistent with RITIS* |
| | **2786913** | 5:43 PM | Exit 45: SR-808/Glades Rd | Disabled Vehicle | 1 | 26 min | — |

**Note:** * = Please see the inconsistency reported in **Table 6-14.**

To expand upon the findings from Table 6-13**,** Table 6-14 provides a detailed comparison of the specific incidents captured by both systems. A closer examination of these records reveals the following discrepancies:

- **Incident 2784356 (June 3, 2026):** The RITIS archive recorded a crash at Exit 44. However, the DIVAS live feed contained no crash record under this specific incident ID. Instead, the feed reported a “Congestion” entry at the exact location and at a similar time under a separate ID (2784353, logged at 4:26 PM), with no lane block recorded. During real-time operations, the predictive model developed in this study requires that the type of the incident is correctly identified as crash to work correctly and trigger an alert because it is not designed to provide alerts for “Congestion” entries.
- **Incident 2784962 (June 5, 2026):** The DIVAS feed accurately placed the event on I-95 Southbound at Exit 44 as a crash with the right lane blocked, which corresponds to the single closed lane recorded in the RITIS archive. This event was reported at 5:03 PM, matching the start time recorded in RITIS. Although these initial details were correct, the incident remained live in the DIVAS feed for only about 15 minutes before disappearing, whereas the RITIS archive recorded a total duration of 2 hours and 21 minutes.
- **Incident 2786609 (June 11, 2026):** This incident was captured at the correct exit (Exit 44, Palmetto Park Road); however, DIVAS recorded the event on "95 Express," while RITIS recorded it on "I-95" (the mainline general-purpose lanes). Both systems typically distinguish Express from general-purpose lanes, indicating a genuine conflict rather than a coarse labeling. The incident entered the feed at 5:45 PM as "Right lane blocked" and was escalated to "All lanes closed" at 6:15 PM, half an hour later, while RITIS’s record showed three closed lanes. If this event occurred in the mainline general-purpose lanes, the model would miss the trigger because incidents in the express lane facility were excluded when developing the model. The event remained live in the feed for approximately 40 minutes, whereas the RITIS archive recorded a duration of 1 hour and 19 minutes.
- **Incident 2786887 (June 12, 2026):** Although the incident ID matched the RITIS archive and it was captured promptly at 4:12 PM as a crash with the left lane blocked, DIVAS located the event at Exit 45 (SR-808/Glades Road) instead of the archive record of Exit 48 (SR-794/Yamato Road). This discrepancy spans approximately 3.5 miles. Similar to the other events, no duration was reported, and it appeared in the live feed for only a single 5-minute poll.

This comparison does not attempt to determine which system is accurate when discrepancies occur. Instead, the analysis aims to assess the practical feasibility of using information provided by DIVAS as inputs to the model in real-time deployment. Based on the assessment of this study, the DIVAS live feed exhibits significant reliability issues that must be addressed prior to operational deployment. Critical issues requiring correction include missing incidents, mismatched incident types, inconsistent locations between general-purpose and express lanes, and the excessively shorter duration of active events in the live feed compared to archive data.

**Table 6-14 Detailed Comparison of Reported Incident Attributes between the RITIS Archive and DIVAS Live Feed**

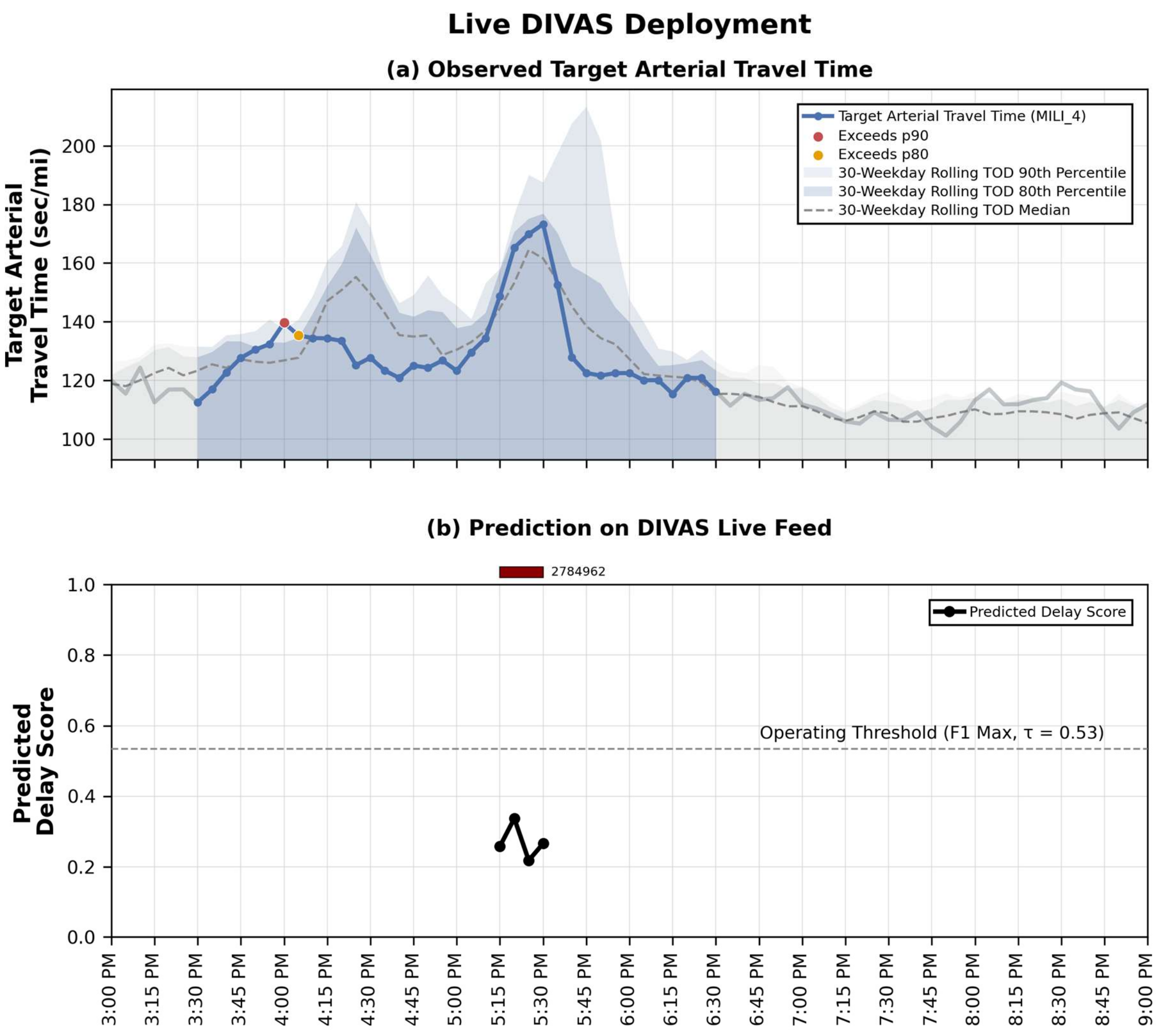
| Incident ID | Source | Description (as recorded) | Type | Lanes / Blockage | Start / Reported | Closed / Last Updated | Duration |
|---|---|---|---|---|---|---|---|
| **2784356**<br>June 3, 2026 | **RITIS** | Southbound on I-95 at Exit 44: CR-798/ Palmetto Park Rd | Crash | — | 4:33 PM | 5:19 PM | 45 min |
| | **DIVAS** | No record under this incident's own ID. The feed reported a “Congestion” entry at the same location and a nearby time, under a separate DIVAS ID (**2784353**, logged 4:26 PM). No duration was reported (live in the feed about 195 min across 40 polls). No duration reported (live in feed ~195 min) | | | | | |
| **2784962**<br>June 5, 2026 | **RITIS** | Southbound was reported at Exit 44: CR-798/ Palmetto Park Rd | Crash | 1 lane(s) closed | 5:03 PM | 7:25 PM | 2 h 21 min |
| | **DIVAS** | **Highway:** I-95<br>**Direction:** Southbound<br>**Detail:** Exit 44: CR-798/Palmetto Park Rd | Crash | Right lane blocked | 5:03 PM | 5:11 PM | No duration reported (live in feed ~15 min) |
| **2786609**<br>June 11, 2026 | **RITIS** | Southbound on I-95 beyond Exit 44: CR-798/ Palmetto Park Rd | Disabled Vehicle | 3 lane(s) closed | 5:16 PM | 6:36 PM | 1 h 19 min |
| | **DIVAS** | **Highway:** 95 Express<br>**Direction:** Southbound<br>**Detail:** Exit 44: Palmetto Park Rd | Disabled Vehicle | Right lane blocked → All lanes closed at 6:15 PM | 5:16 PM | 6:11 PM | No duration reported (live in feed ~40 min) |
| **2786887**<br>June 12, 2026 | **RITIS** | Southbound on I-95 beyond Exit 48: SR-794/ Yamato Rd | Crash | 1 lane(s) closed | 4:12 PM | 5:04 PM | 52 min |
| | **DIVAS** | **Highway:** I-95<br>**Direction:** Southbound<br>**Detail:** Exit 45: SR-808/Glades Rd | Crash | Left lane blocked | 4:12 PM | 4:16 PM | No duration reported (live in feed: a single 5-min poll) |

#### 6.6.5.5 Model Response to DIVAS Live Feed Inputs

This section demonstrates how the predictive model responds to the actual data captured from the DIVAS live feed. As previously detailed, only four incidents from the collection window appeared in the feed. Of these four, only one provided information consistent with the RITIS archive. To illustrate the model's performance, the following two cases are highlighted as case examples:

- **Case 1 (June 5, 2025):** This event was selected because the incident information correctly matched the archive, and it possessed sufficient duration to potentially trigger a real-time prediction.
- **Case 2 (June 3, 2025):** This event is worth investigating because a severe delay was observed, with travel time reaching 31% higher than the p90 threshold. Although the live feed incorrectly logged it strictly as a "Congestion" entry (causing the deployed model to miss it), this case is used to explore a hypothetical scenario. It examines how the model would have performed if DIVAS had accurately recorded the incident as a crash, matching the RITIS archive.

Figure 6-33 illustrates the actual model execution using the DIVAS live feed for June 5, 2025. On this day, the live feed showed an active incident for approximately 15 minutes. Panel (a) shows that during this specific incident active window, the observed target arterial travel time fluctuated between the historical median and the 80th percentile (p50 to p80), remaining well below the p90 threshold. Because no sustained delay occurred while the incident information was present, the model's prediction score remained below the operating threshold of 0.53, as shown in Panel (b). Therefore, the model did not issue false alarms in response to minor traffic fluctuations.



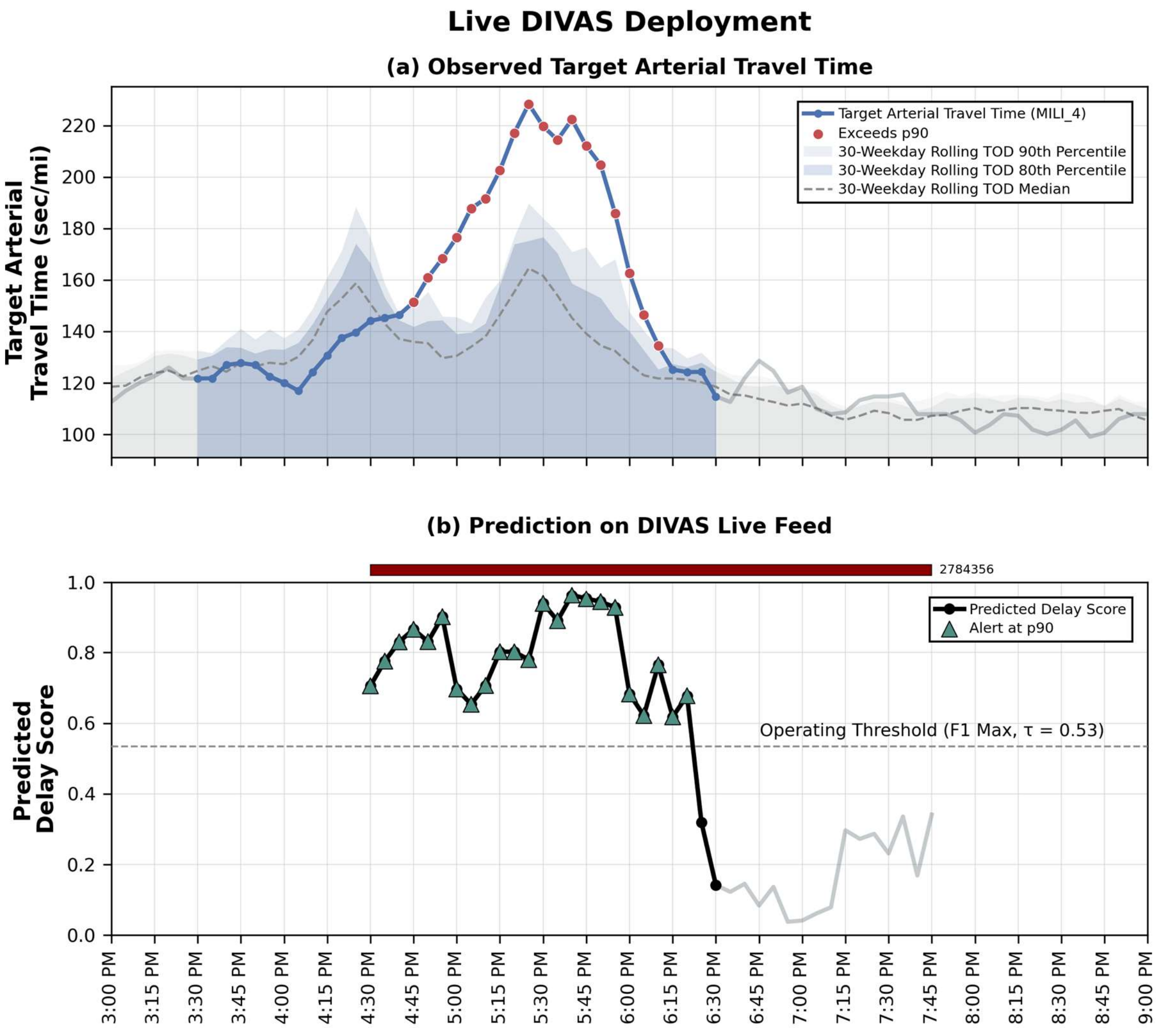

**Incident on Friday, June 5, 2026**

| ID | Source | Type | Location (as recorded) | Lanes | Start | Duration |
|---|---|---|---|---|---|---|
| **2784962** | DIVAS | Crash | I-95 SB · Exit 44: Palmetto Park Rd | Right lane | 5:03 PM | live ~15 min |

**Figure 6-33 Incident Case Study of Incident June 5, 2026 Demonstrating the Use of DIVAS Live Feed for Model Inputs**

Figure 6-34 evaluates the June 3, 2025 event to demonstrate how the model would respond if the incident type matched the RITIS archive (recorded as a "Crash"). On that afternoon, severe congestion formed on the target arterial. Panel (a) illustrates this significant delay, with observed travel times peaking at 228 seconds per mile. This measurement represents a delay approximately 31% above the p90 threshold.

Because the live DIVAS feed initially logged this event as a "Congestion" entry, the deployed model excluded it from further consideration, resulting in a missed alert. However, Panel (b) demonstrates the model's output when the live feed data was adjusted to record the incident as a "Crash" to match the RITIS archive. With the crash input, the predicted delay score immediately crosses the operating threshold of 0.53. It remains significantly elevated until the delay begins to clear, at which point the score drops back below the threshold. This outcome confirms that the model functions as intended and successfully issues an advisory when provided with the correct crash information from the live feed.




**Incident on Tuesday, June 3, 2026**

| ID | Type | Location (as recorded) | Lanes | Start | Duration |
|---|---|---|---|---|---|
| **2784356** | Congestion | I-95 SB · Exit 44: Palmetto Park Rd | — | 4:33 PM | live ~195 min |

**Figure 6-34 Incident Case Study of Incident June 3, 2026 Demonstrating the Use of DIVAS Live Feed for Model Inputs**

## 6.7 Summary

This study investigated the impacts of freeway incident-related traffic diversion on parallel arterial streets through machine learning. Using a case study of southbound I-95 and the parallel Military Trail corridor in Boca Raton, Florida, the research examined diversion behavior, evaluated the ability to estimate diversion impacts, and assessed the feasibility of deploying a real-time predictive decision support system.

The analysis confirmed that high impact incidents on I-95 generate significant traffic diversion onto Military Trail and other parallel arterials, resulting in substantial increases in arterial congestion. Field observations identified a time lag between the onset of freeway congestion and the resulting deterioration of arterial traffic conditions, providing a valuable opportunity for proactive arterial traffic management. The study also found that existing time-of-day signal timing plans do not adequately respond to diversion events, limiting agencies' ability to accommodate rapidly changing traffic patterns.

To address this limitation, the study developed an XGBoost-based machine learning model capable of predicting arterial delays approximately 15 minutes before congestion develops using only existing FDOT data sources. The model demonstrated strong predictive performance and successfully identified a high proportion of significant diversion-induced delay events while maintaining a relatively low false alarm rate. Analysis of model behavior showed that freeway speed and occupancy measurements provide the strongest indicators of impending diversion.

The study also evaluated the feasibility of implementing the predictive model using live operational data from the FDOT Data Integration and Video Aggregation System (DIVAS) platform. While the machine learning model proved suitable for real-time deployment, the assessment identified several challenges related to data quality and system integration. These included inconsistencies in the provision of detector measurements, incomplete and inaccurate incident records, and limited correspondence among the SunGuide, Regional Integrated Transportation Information System (RITIS), and DIVAS platform data. These issues reduce the reliability of real-time predictions and highlight the importance of maintaining high-quality, integrated transportation data systems.

Overall, the research demonstrates that freeway diversion impacts on arterial streets can be predicted with sufficient lead time to support proactive traffic management, including the implementation of special signal timing plans before congestion develops. The findings illustrate the significant potential of AI and machine learning to improve coordinated freeway and arterial operations while emphasizing that successful deployment depends not only on predictive model performance but also on the availability of reliable, integrated, and well-maintained ITS infrastructure..

Based on the feasibility assessment, the following actions are recommended prior to deploying the predictive model:

- **Detector Calibration:** Because the predictive model relies heavily on freeway detector measurements, the discrepancies between the DIVAS detector and the RITIS datae must be investigated to determine the root cause of the misalignment.

- **Incident Feed Review:** The DIVAS incident data feed must be reviewed to resolve the high rate of missing records and the issue of incidents disappearing from the feed prior to their actual clearance times.
- **Inter-System Connectivity:** A critical lack of correspondence was observed among the SunGuide, RITIS, and DIVAS platform data. Although newly installed replacement detectors began providing traffic measurements to the SunGuide system around June 2023, the three systems are not actively linked. Consequently, the new detector data did not appear in the RITIS archive or the DIVAS feed until the discrepancy was reported by the research team of this study during the data collection period. Improved communication and coordination protocols among the personnel managing these platforms are necessary for system integration.

# 7 IDENTIFICATION OF TRAFFIC PATTERNS AND ASSOCIATED FEATURES

This chapter presents the methods used to identify traffic patterns occurring during peak periods across different days of the year, as well as the factors contributing to variations in operational performance among these patterns. Such analysis enables the identification of appropriate corrective actions by linking variations in the assessed performance measures to deficiencies in signal timing parameter selections for specific traffic patterns. The developed methods are demonstrated using a case study segment located on Yamato Road between Military Trail and Congress Avenue in the City of Boca Raton, FL.

## 7.1 Utilized Data

The data used for the purpose of this study includes traffic data and signal controller data. The traffic data was based on the two tactical measures most used in traffic analysis to assess system performance. These are the traffic flow rate in vehicles/hour and the travel time rate in seconds per mile. The study estimated the traffic flow rate using the volume counts provided by the video image detectors at the intersections and the travel time rates utilizing travel time measurements estimated based on a third-party vendor (HERE) data. Both measures were utilized at the 15-minute resolution level. **Figure 7-1** and **Figure 7-2** show the locations of the volume measurements and the segments of travel time measurements, respectively, used in the cluster analysis. The data was collected for the period between June 1st and December 31st, 2024.


**Figure 7-1 Locations of Volume Counts Used in Cluster Analysis**

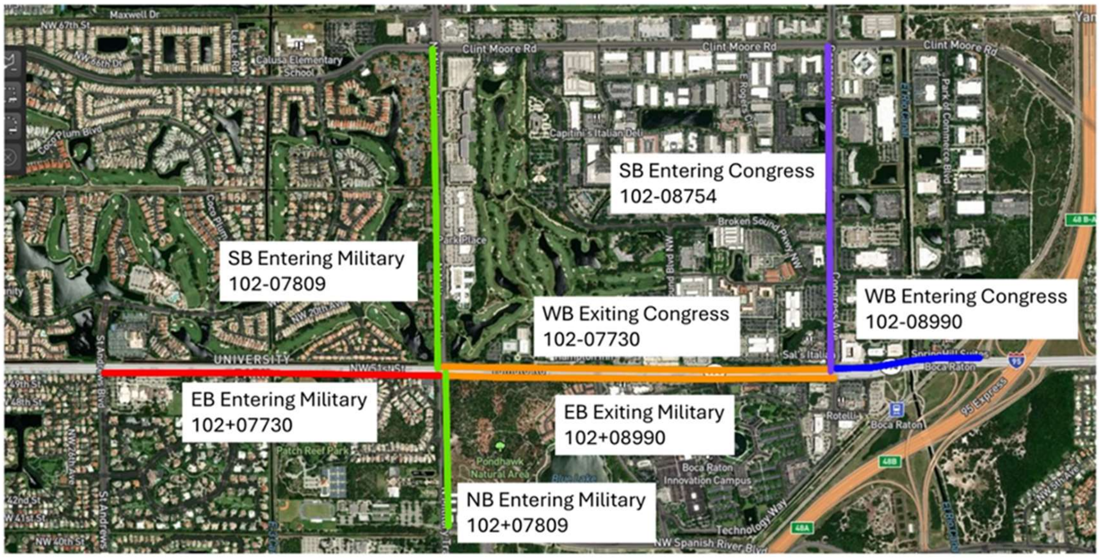



**Figure 7-2 Locations of Travel Time Measurements Used in Cluster Analysis**

For each peak period in a day, the research team assembled the travel time rate and traffic flow rate data for all 15-minute periods in each day in one record, as shown in **Table 7-1.** This grouping ensured that all 15-minute periods in the peak period of a given day are grouped in the same cluster.

**Table 7-1 Assembling 15-Minute Data for each Day Period in One Record**

| Date | EB Entering Military_T1 | EB Entering Military_T2 | EB Entering Military_T3 | ... | vol_25_ WB_T7 | vol_25_ WB_T8 | vol_25_ WB_T9 |
|---|---|---|---|---|---|---|---|
| 6/3/2024 | xxx | xxx | xxx | ... | xxx | xxx | xxx |
| 6/4/2024 | xxx | xxx | xxx | ... | xxx | xxx | xxx |
| 6/5/2024 | xxx | xxx | xxx | ... | xxx | xxx | xxx |
| 6/6/2024 | xxx | xxx | xxx | ... | xxx | xxx | xxx |
| 6/7/2024 | xxx | xxx | xxx | ... | xxx | xxx | xxx |

### 7.1.1 Clustering Analysis

The identification of the patterns was based on K-means clustering with Principal Component Analysis (PCA). A previous work of the research team (Saha et al., 2019) compared the use of the K-means, K-prototypes, K-medoids, four variations of Hierarchical Clustering, and the combination of PCA with K-means for clustering traffic patterns. Among these methods, the K-prototypes and K-means with PCA produced the best results. The K-means is an iterative algorithm that starts by identifying a set of data points as centroids. The remaining data points are then clustered by assigning them to the closest centroid using a squared Euclidian distance measure (Hartigan, 1975, Hartigan and Wong, 1979). Analysts have used PCA for dimensional reduction of the data, while retaining most of the variation in the data by converting the observations to an orthogonal system of Euclidean space (Dunteman, 1989). Researchers found that using PCA to reduce the dimensionality of the data before clustering is effective in improving the pattern recognition results, particularly when clustering high-dimension data (Ding and He, 2004, Marinai et al., 2006, Meng et al., 2015). The clustering analysis in this study was implemented using the scikit-learn library (sklearn.cluster) in Python.

The optimal number of clusters was determined using the Elbow Method. This method examines the relationship between the Within-Cluster Sum of Squares (WCSS) and the number of clusters (k).

### 7.1.2 Feature Importance Identification

Following the identification of the traffic patterns in the year, the next step was to utilize approaches to identify the characteristics of each pattern and the contributing factors to these different characteristics. This is important to better understand the patterns and the differentiating factors that make them different from the other patterns. Such understanding can be used to identify solutions to address the issues causing congestion in each pattern. The approaches utilized are discussed in the following subsections.

#### 7.1.2.1 Using PCA Loading

The first technique to identify the characteristics and contributing factors to each pattern, and the results obtained using this technique. This technique is referred to as "PCA Loading." As stated earlier, PCA was used in combination with K-means clustering to reduce dimensionality by transforming data into a lower-dimensional space (referred to as principal components). The K-means is run on these principal components not the original features to improve the K-means clustering performance. In the analysis of this study, the traffic measures (travel time rates and traffic flow rates for various segments in the facility) were transformed into two principal components: Component 1 (PC1) and Component 2 (PC2). PCA loadings reflect the eigenvectors or component weights when transforming the original variables to the principal components and indicate how original features contribute to each principal component.

#### 7.1.2.2 Using Machine Learning to Identify Feature Contributions

Although PCA loading is very useful in identifying the features differentiating different clusters, as it is clear from the discussion in the previous section, PCA can only identify features included in and contributed to clustering (travel time rates and traffic flow rates in this study). It is important to identify how the values of other features not included in clustering are different between different clusters. Such identification can allow the realization of additional underlying factors (cause-and-effect efforts) that result in the differences in travel time and throughput measurements. Thus, additional techniques are needed to identify other feature contributions. Such techniques can also further confirm the differences in travel time rates and traffic flow rates between different clusters identified using PCA Loading.

A combination of two techniques was used for identifying additional feature contributions. The first is a tree-based machine learning technique referred to as the XGBoost. The second is

Spearman Correlation, which is a nonparametric statistical technique, frequently used in combination with machine learning for feature selection. The XGBoost approach can capture potential non-linear relationships and complex interactions between variables. Spearman Correlation can measure the strength and direction of non-linear relationships between two ranked variables. The Spearman correlation was selected over Pearson correlation as it does not assume linear relationships between variable pairs. These techniques were implemented in Python using the XGBoost library (XGBClassifier), the scikit-learn library for permutation importance, and the pandas library for Spearman correlation.

#### 7.1.2.3 Confirmation Using Manual Inspection of the Data

Due to the small size of the case study, it was possible to validate the findings from the PCA Loading, Spearman correlation, and XGBoost. This inspection was based on comparing the cumulative density functions (CDFs) of traffic flow rates, travel time, and green times of different clusters.

### 7.1.3 Volume Correction

Before the turning movement counts obtained from the video image detectors were used in the analysis, their accuracy was evaluated and validated. The automatic turning counts (ATC) generated by the video detection system were compared with manual turning counts (MC) collected during three weekday afternoon peak periods to assess the accuracy and reliability of the detector data. The comparison revealed systematic discrepancies between the ATC and MC datasets, indicating that the detector-based counts required adjustment before being used in the study. To account for these differences, regression analysis was performed for each of the 24 turning movements within the study corridor. The resulting regression equations were then used to estimate adjusted traffic volumes ($V_{adj}$) from the detector-reported counts as follows:

$$V_{adj} = \alpha_1 + \beta_1\, ATC \tag{7-1}$$

The regression parameters in the above equation were estimated using the ordinary least squares (OLS) method. The results indicated that all estimated coefficients were statistically significant at the 90% confidence level (($p < 0.10$)) for each of the 24 turning movements analyzed.

### 7.1.4 Representative Day

The representative day for each cluster was selected following the procedure recommended in FHWA’s Traffic Analysis Toolbox Volume III (FHWA). The purpose of this step was to identify the observed day that best represents the typical time-dependent traffic pattern of each cluster. Obtaining the represenbtative day for each cluster, allows further examination of the differences between the identified patterns (clusters) through the comparison of the volumes, travel times, and assigned signal timing parameters of their represntative days.

To obtain the representative days, first, the average value of the two parameters used in clustering (travel time rate and traffic flow rate) for every 15 minutes across all days in the cluster is calculated as:

$$\overline{m}_{t,j} = \frac{\sum_i m_{i,j}(t)}{N_{cluster}} \quad (7\text{-}2)$$

where $\overline{m}_{t,j}$ is the the average time-variant value for the 15-minute interval $t$ across all days in a cluster, $m_{i,j}(t)$ is value of the key measure on the day $i$ in time interval $t$ at location $j$ and $N_{cluster}$ is number of days in a cluster.

The difference between the average value and the corresponding value observed in a particular day is assessed using the parameter calculated as in the following equation.

$$\dot{m}_{i,j}(t) = \frac{\sqrt{\left(\overline{m}_{t,j} - m_{i,j}(t)\right)^2}}{\overline{m}_{t,j}} \quad (7\text{-}3)$$

where $\dot{m}_{i,j}$ is a deviation parameter used to quantify the difference between the average value of each key metric and the corresponding value observed on a given day. The representative day is identified as the day that minimizes the overall deviation between its metric values and the corresponding average values across all metrics and locations.

### 7.1.5 System Performance

System performance under different traffic patterns was assessed using mobility measures including Vehicle-Hour Delay (VHD), Total Excessive Delay (ED), and the Percentage of Vehicle Miles Traveled under Excessively Congested Conditions (PECC). In addition, the volume/capacity (v/c) was used as a measure of oversaturation with the capacity calculated using the Highway Capacity Manual (HCM) procedures.

The reliability measure evaluates how often congestion occurs during the analysis period using the Green Occupancy Ratio (GOR). For each movement, the 75th percentile (P75) of GOR is used as the congestion threshold. For each movement and cluster, the number of observations exceeding this threshold is counted and divided by the total number of GOR observations to obtain the GOR Exceedance Rate, which represents how frequently the GOR exceeds the congestion threshold. This value is then divided by the number of days in each cluster to normalize the result and express the daily intensity of congestion occurrences

Performance comparisons were conducted based on (i) the average values of these measures across all days within each cluster and (ii) the values calculated for the representative day of each cluster. VHD is calculated as follows:

$$VHD = VHT - VHT\,(FF) \tag{7-4}$$

where *VHT* is the vehicle-hours traveled including free flow travel time and delay and *VHT(FF)* is free flow vehicle-hours traveled with segment free-flow speed.

ED is a MAP-21 measure and is calculated to put extra focus on extreme delays. First, an excessive delay threshold has to be calculated follows:

$$Excessive\ Delay\ Threshold = \left(\frac{Travel\ Time\ Segment\ Length_s}{Threshold\ Speed_s}\right) \times 3600 \tag{7-5}$$

where *s* refers to a reporting segment. A parameter termed Road Segment Delay (RSD) is defined as the difference between the average travel time measured over a 15-minute interval and the excessive-delay travel-time threshold described above. The threshold speed for a segment utilized in the above equation is 20 mph or 60% of the posted speed limit for that specific road segment, whichever is greater. RSD ranges from 0 to 900 seconds, since the maximum possible delay within a 15-minute analysis interval is 900 seconds. To quantify excessive delay in vehicle-hours, RSD is multiplied by the corresponding traffic volume, yielding the Excessive Delay (ED) measure, as shown in **Equation** 7-6**.**

$$Total\ Excessive\ Delay_s = \tag{7-6}$$

$$\sum_{d=1}^{TD}\sum_{h=1}^{TH}\sum_{b=1}^{TB}(Excessive\ Delay_{s,b,h,d} \times (\frac{hourl\ \ volume}{4})_{s,h,d}$$

where *s* is a travel time reporting segment, *d* is a day of the reporting year, *TD* is total number of days in the reporting year, *h* is single hour interval of the day where the first hour interval is 12 a.m. to 12:59 a.m., *TH* is total number of hour intervals in day "h", *b* is 15-minute bin for hour interval "h", and *TB* is total number of 15-minute bins where travel times are recorded in the travel time data set for hour interval "h"

The third measure used in the performance assessment is the PECC, which is a measure utilized by Virginia Department of Transportation as an extension of MAP-21 measures. PECC reflects the percent vehicle miles traveled under excessively congested conditions (PECC) (Babiceanu, 2021) calculated as follows.

$$PECC = \frac{\sum_1^C VMT}{\sum_1^T VMT}\ (\%) \tag{7-7}$$

where *T* is the total number of reporting segments and *C* is the total number of segments with excessive congested conditions (i.e., operating speed is lower than 20 mph or 60% of the posted speed limit, whichever is higher).

Another measure used in the assessment was the Green Occupancy Ratio (GOR) parameter calculated based on high resolution controller data. For each movement, the 75th percentile (P75) of the GOR is used as the congestion threshold. For each movement and cluster, the number of observations exceeding this threshold is counted and divided by the total number of GOR observations to obtain the GOR Exceedance Rate, which represents how frequently the GOR exceeds the congestion threshold. This value is then divided by the number of days in each cluster to normalize the result as a per-day rate of congestion occurrences.

### 7.1.6 Signal Control Analysis

The signal timing assigned by the adaptive signal control to each movement was also compared between the identified clusters. These assigned signal timings were also compared with the time-of-day (TOD) timing plans used before adaptive signal control activation. This comparison was used in combination with the operational performance measures described in the previous section to identify potential changes required to improve performance. The Highway Capacity Software (HCS) was then used to evaluate corridor performance under the potential signal timing adjustments. Microscopic simulation is an alternative better choice to conduct this evaluation.

## 7.2 Results and Discussion

This section summarizes the key findings of the analysis, including the identification of traffic patterns, the primary contributing factors, overall corridor performance, and the impacts of signal timing adjustments on system perfromance.

### 7.2.1 Traffic Pattern Identification

The number of clusters used based on the K-means clustering was obtained using the elbow method that identifies the optimal number of clusters at the point where the within-cluster sum of squares (WCSS) versus the number of cluster curve transitions from a steep decline to a relatively flat slope, indicating diminishing reductions in clustering error as the number of cluster increases. Based on **Figure 7-3,** five clusters were chosen as the optimal number of clusters.

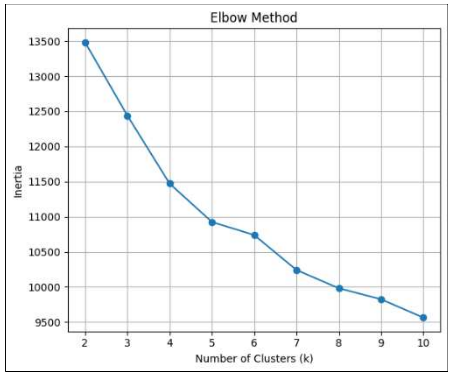



**Figure 7-3 The Relationship between The Number of Clusters and The Total Within-Cluster Sum of Squares**

Cluster 2 includes the weekend days. Thus, it is excluded from the detailed analysis of the traffic patterns later in this document. The distribution of the days by month in each of the remaining clusters is shown in **Table 7-2.** The following can be concluded based on the results shown in this table.

- Cluster 0 includes 29 days, with 10 days in August and the remaining in the Winter session, mostly in November. Thus, the days in this cluster are mixed between Summer and Winter
- Cluster 1 includes 55 days in the Summer (June and July)
- Cluster 2 includes weekend days
- Cluster 3 includes 30 days and appears to be the most frequent pattern in September, October, and December, which are school months, and are expected to be busy months.
- Cluster 4 includes 11 of the more congested days in the school months.

**Table 7-2 Cluster Analysis Results for the Case Study Corridor**

| Cluster | Month | | | | | | | | Total |
|---|---|---|---|---|---|---|---|---|---|
| | **June** | **July** | **August** | **September** | **October** | **Nov 1–20** | **Nov 21–30** | **December** | |
| 0 | 2 (6.9%) | 0 (0.0%) | **10 (34.48%)** | 0 (0.0%) | 3 (10.34%) | **9 (31.03%)** | 3 (10.34%) | 2 (6.9%) | **29** |
| 1 | **17 (30.91%)** | **22 (40.0%)** | 8 (14.55%) | 1 (1.82%) | 4 (7.27%) | 1 (1.82%) | 2 (3.64%) | 0 (0.0%) | **55** |
| 2 | 0 (0.0%) | 1 (25.0%) | 0 (0.0%) | 0 (0.0%) | 2 (50.0%) | 0 (0.0%) | 1 (25.0%) | 0 (0.0%) | **4** |
| 3 | 1 (2.56%) | 0 (0.0%) | 0 (0.0%) | **16 (41.03%)** | **13 (33.33%)** | 1 (2.56%) | 0 (0.0%) | 8 (20.51%) | **39** |
| 4 | 0 (0.0%) | 0 (0.0%) | 0 (0.0%) | 1 (9.09%) | 1 (9.09%) | **3 (27.27%)** | **1 (9.09%)** | **5 (45.45%)** | **11** |

### 7.2.2 Feature Importance Identification

As stated in the Methodology Section, the first technique to identify the characteristics and contributing factors of each pattern is "PCA Loading" with the traffic measures transformed into two principal components: Component 1 (PC1) and Component 2 (PC2). PCA loadings indicate how original features contribute to each principal component.

**Figure 7-4** indicates that in term of Component 1 (PC1), the days in Cluster 4, which is the most congested cluster, have the highest values, followed by Clusters 3 and 0 (both having about the same values), followed by Cluster 1 (the summer pattern Cluster), followed by Cluster 2 (the weekend and holidays cluster). The main differentiator between Cluster 0 and Cluster 3 are the values of PC2, which is relatively higher than other clusters for Cluster 0 and relatively lower than other clusters for Cluster 3. This information is used next in combination with PCA loading to determine the features that differentiate different clusters.

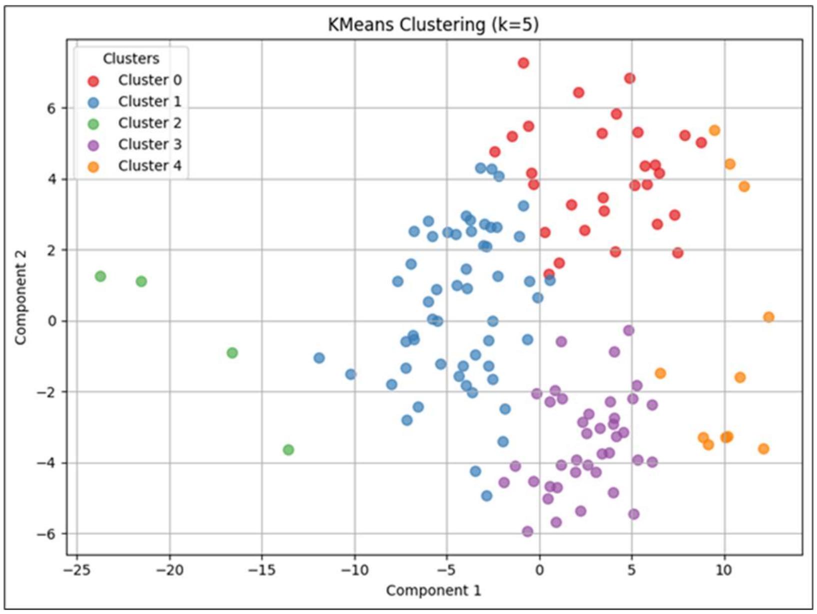



**Figure 7-4 Values of Principal Components of Different Clusters**

**Table 7-3** presents the loadings for both PCAs in relation to the features used for clustering, specifically the intersection approach volumes and link travel times at the locations shown in **Figures 7-1** and 7-2. A high absolute value in the loading matrix of a particular PC for a feature indicates a strong influence of the feature on the value of that PC. Thus, features with high loading scores for a PC used in clustering are likely contributing more to the cluster separation. Positive loading indicates that the PC score increases when the feature value increases. Negative loading indicates that the PC score increases as the feature value decreases. Combining the loading results in **Table 7-3** with the differences in the two principal component values shown in Figure 7-4 indicates the following. Cluster 2 is the weekend cluster and will not be discussed further.

- Cluster C0 is distinguished by higher volumes and travel times on the southbound (SB) approach on Congress Avenue (CA) and lower volume on the eastbound (EB) and SB through movements on Military Trail (MT). In terms of travel time, this cluster has lower travel time for SB traffic entering MT compared to the other two high demand clusters (C3 and C4) but higher westbound (WB) travel times on MT, despite that the volume of the WB is lower in C0 compared to C3 and C4 clusters. Further examination shown later indicates that this is due to higher eastbound left turn green time on Military.

- Cluster C1 is the cluster that represents most summer days and has lower travel times, particularly on MT approaches compared to other clusters. It has lower volumes, particularly on the northbound and eastbound on MT.
- Cluster C3 is the cluster that includes the highest proportion of winter days. Compared to Cluster 0, it has lower volumes and travel times on SB CA and lower travel time on WB MT but higher travel time on EB and SB MT. Compared to Cluster 4, it has lower travel times on all MT approaches and lower volumes on NB and EB MT.
- Cluster C4 includes congested days mainly in the Winter that have higher travel times on all approaches of MT and high volumes on the NB and EB on MT.

**Table 7-3 Highest Impact Traffic Measures on Clustering Results According to PCA Loading**

| **Principal** | **Features** | **Value** |
|---|---|---|
| **PC1** | EB Entering Military | 0.132 |
| | SB Entering Military | 0.124 |
| | NB Entering Military | 0.120 |
| | EB Entering Military | 0.119 |
| | vol_24_EB (Military) | 0.118 |
| | WB Exiting Congress | 0.115 |
| | vol_24_NB (Military) | 0.114 |
| **PC2** | vol_25_SB (Congress) | 0.189 |
| | vol_24_EB (Military) | -0.184 |
| | WB Exiting Congress | 0.164 |
| | SB Entering Military | -0.140 |
| | vol_24_SB (Military) | -0.133 |

**Table 7-4** shows the summary of XGBoost feature importance and spearman correlation analysis of traffic flow rate and green split (green to cycle or GOC ratio) and

**Table 7-5** shows summary of XGBoost feature importance and Spearman correlation analysis of travel time rate. The results indicate the following:

- The additional techniques confirmed that the C0 Cluster had the highest volume and travel time on southbound CA but the lowest green over cycle (GOC) ratio for this movement. The GOC for the eastbound left (EBL) is high, which may explain the reason for the lower GOC. On MT, the analysis confirmed that high GOC ratio was assigned to eastbound left (EBL) and thus EBT despite lower EB through volume resulting in lower green for WB through (WBT), causing higher travel time for the WBT movement and lower travel time for EB movement.
- The additional techniques confirmed that the C1 Cluster had the lowest travel times on all movements northbound and eastbound volumes on Military Trail, as found based on PCA Loading analysis. But it is also indicated lower southbound volumes on Military Trail and Congress Avenue and lower travel times on southbound Congress Avenue. The GOC for the eastbound and westbound lefts on Congress were also lower indicating lower demands on these movements

- Among the interesting findings regarding the C3 cluster is that it has a high GOC ratio on Southbound CA compared to Cluster 0 and 4. It also has lower GOC for the eastbound left on Military Trail despite a higher total eastbound approach volume, probably due to high eastbound left demands. Cluster 3 has high travel time on SB, EB, and NB on MT but less congestion on the WB on MT and SB CA compared to Cluster 0 and 4 due to the higher GOC assigned to these movements.
- Cluster 4 has high volumes on the WB and SB on CA combined with high GOC ratio for the eastbound left on CA. In addition, Cluster 4 has the highest travel times on all MT movements. Cluster 4 has also high SB volume with high GOC ratio on MT.

**Table 7-4 Summary of XGBoost Feature Importance and Spearman Correlation Analysis of Traffic Flow Rate and GOC Ratio**

| Intersection | Feature | Cluster 0 | | Cluster 1 | | Cluster 3 | | Cluster 4 | |
|---|---|---|---|---|---|---|---|---|---|
| | | Correlation | XGB | Correlation | XGB | Correlation | XGB | Correlation | XGB |
| **Congress Avenue** | SB Volume | 0.620 | 0.084 | -0.412 | 0.067 | -0.224 | 0.004 | 0.180 | 0.004 |
| | SB GOC | -0.642 | 0.134 | 0.016 | 0.000 | 0.570 | 0.042 | -0.008 | 0.004 |
| | EB Volume | - | - | - | - | - | - | - | - |
| | EBT GOC | 0.576 | 0.000 | 0.138 | 0.000 | -0.553 | 0.000 | -0.196 | 0.000 |
| | EBL GOC | 0.328 | 0.000 | -0.509 | 0.015 | 0.056 | 0.022 | 0.327 | 0.057 |
| | WB Volume | 0.021 | 0.000 | -0.088 | 0.000 | -0.140 | 0.060 | 0.359 | 0.057 |
| | WBT GOC | 0.348 | 0.000 | 0.121 | 0.000 | -0.310 | 0.000 | -0.226 | 0.000 |
| | WBL GOC | -0.398 | 0.000 | -0.196 | 0.001 | 0.488 | 0.004 | 0.140 | 0.000 |
| | NB Volume | - | - | - | - | - | - | - | - |
| | NB GOC | 0.317 | 0.000 | -0.165 | 0.000 | -0.127 | 0.000 | 0.029 | 0.000 |
| **Military Trail** | SB Volume | -0.154 | 0.000 | -0.500 | 0.006 | 0.550 | 0.034 | 0.218 | 0.000 |
| | SBT GOC | 0.390 | 0.000 | 0.090 | 0.000 | -0.404 | 0.000 | -0.078 | 0.000 |
| | SBL GOC | 0.436 | 0.000 | 0.073 | 0.000 | -0.427 | 0.000 | -0.077 | 0.000 |
| | EB Volume | -0.181 | 0.022 | -0.584 | 0.057 | 0.588 | 0.036 | 0.345 | 0.039 |
| | EBT GOC | 0.255 | 0.000 | 0.189 | 0.000 | -0.260 | 0.010 | -0.290 | 0.000 |
| | EBL GOC | 0.585 | 0.016 | 0.193 | 0.000 | -0.615 | 0.026 | -0.205 | 0.000 |
| | WB Volume | - | - | - | - | - | - | - | - |
| | WBT GOC | -0.470 | 0.000 | 0.028 | 0.000 | 0.377 | 0.000 | 0.031 | 0.000 |
| | WBL GOC | -0.363 | 0.022 | -0.259 | 0.000 | 0.416 | 0.000 | 0.320 | 0.048 |
| | NB Volume | 0.398 | 0.000 | -0.682 | 0.110 | 0.239 | 0.000 | 0.230 | 0.000 |
| | NBT GOC | -0.186 | 0.000 | -0.369 | 0.000 | 0.443 | 0.000 | 0.207 | 0.000 |

| Intersection | Feature | Cluster 0 | | Cluster 1 | | Cluster 3 | | Cluster 4 | |
|---|---|---|---|---|---|---|---|---|---|
| | | Correlation | XGB | Correlation | XGB | Correlation | XGB | Correlation | XGB |
| | NBL GOC | -0.253 | 0.000 | -0.297 | 0.000 | 0.422 | 0.000 | 0.213 | 0.000 |

**Table 7-5 Summary of XGBoost Feature Importance and Spearman Correlation Analysis of Travel Time Rate**

| Intersection | Feature | Cluster 0 | | Cluster 1 | | Cluster 3 | | Cluster 4 | |
|---|---|---|---|---|---|---|---|---|---|
| | | Correlation | XGB | Correlation | XGB | Correlation | XGB | Correlation | XGB |
| **Congress Avenue** | SB Entering CA | 0.561 | 0.209 | -0.517 | 0.046 | -0.114 | 0.031 | 0.273 | 0.000 |
| | WB Entering CA | 0.137 | 0.000 | -0.506 | 0.015 | 0.236 | 0.022 | 0.310 | 0.000 |
| | EB Exiting MT or EB Entering CA | 0.261 | 0.000 | -0.374 | 0.000 | 0.007 | 0.020 | 0.267 | 0.027 |
| **Military Trail** | SB Entering MT | -0.019 | 0.122 | -0.697 | 0.075 | 0.500 | 0.178 | 0.450 | 0.072 |
| | EB Entering MT | -0.067 | 0.029 | -0.623 | 0.038 | 0.497 | 0.028 | 0.394 | 0.004 |
| | WB Exiting CA or WB Entering MT | 0.486 | 0.075 | -0.434 | 0.017 | -0.152 | 0.097 | 0.300 | 0.005 |
| | NB Entering MT | 0.349 | 0.081 | -0.790 | 0.251 | 0.287 | 0.011 | 0.417 | 0.060 |

Attempts were also made to correlate the clustering membership with other variables including:

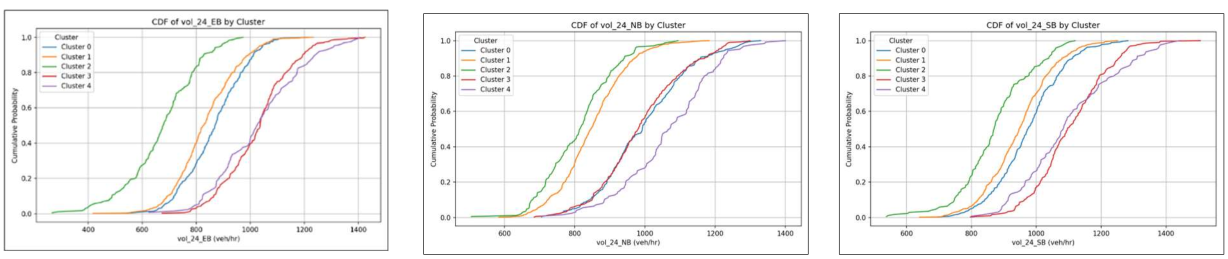
- Ped Actuation: This is defined as the number of Ped Calls exceeding twice the standard deviation of the mean of the number of the calls for a Ped Phase. The high number of Ped Actuation could be due to the Ped Actuation Stuck malfunction or high number of pedestrians.
- Malfunction Alerts: This is the number of critical alerts defined as the Unpaired On and Off events, Long Stuck defined as detector activation duration exceeding the maximum green time, and Chattering/Erratic Counts defined as the number of periods with more than 40 activations per minute.
- Green Occupancy Ratio: This is a high-resolution controller-based measure that is calculated as the percentage of green duration during which a vehicle detector is occupied by vehicles. A high GOR signifies heavier traffic leaving the stop line during green periods.

### 7.2.3 Confirmation Using Manual Inspection of the Data

Due to the small size of the case study, it was possible to confirm the findings from the PCA Loading, Spearman correlation, and XGBoost. This inspection was based on comparing the cumulative density functions (CDFs) of traffic flow rates **(Figure 7-5** and **Figure 7-6),** travel time rates **(Figure 7-7** and **Figure 7-8),** and green times **(Figure 7-9** and **Figure 7-10)** of different clusters. These examinations confirmed the findings presented in Section 7.2.2

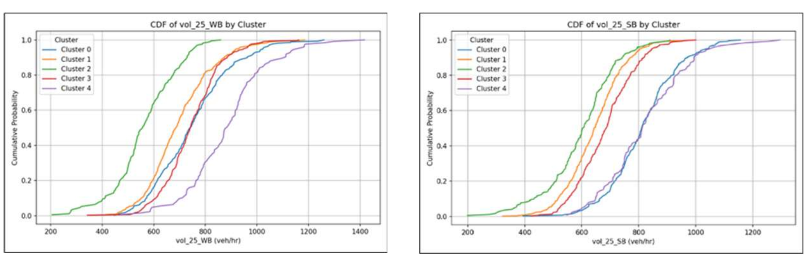
The inspection confirmed that Cluster C1 (the summer cluster) has the lowest volumes and travel times than the other cluster with an average Cycle length of 165 seconds. Cluster C0 has higher volume and travel time on Southbound CA. The examination confirmed that the volumes on MT are not exceptionally high in Cluster 0 (the SB, WB and EB, volumes are lower than Clusters 3 and 4) but the travel times on the WB and NB on MT are relatively high. The average cycle length is 170 sec for this cluster.

Cluster 3 appears to be the typical winter cluster. The volumes are heavier than Cluster 0 on all directions on MT but lower volumes and travel times on SB CA and WB MT. Cluster 4 represents the more congested days in the winter with high volumes and travel times on all approaches. Both Clusters 3 and 4 have cycle lengths ranging between 175-180 seconds.



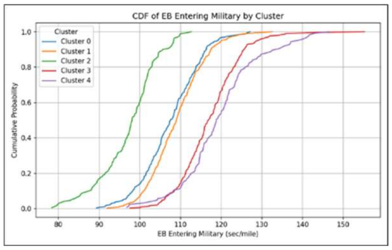

(a) EB Volume at MT (b) NB Volume at MT (c) SB Volume at MT

**Figure 7-5 CDF chart of Volume at MT**

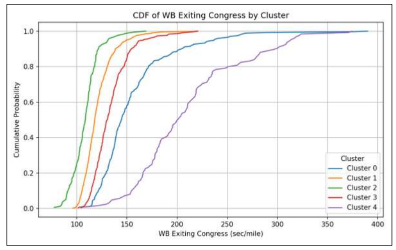



(a) WB Volume at CA (b) SB Volume at CA

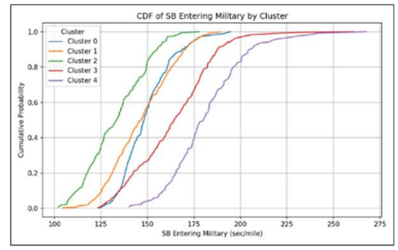
**Figure 7-6 CDF chart of Volume at CA**



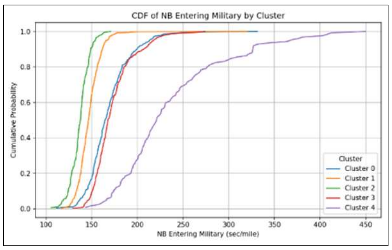

(a) EB Entering MT




(b) WB Exiting CA or WB Entering MT

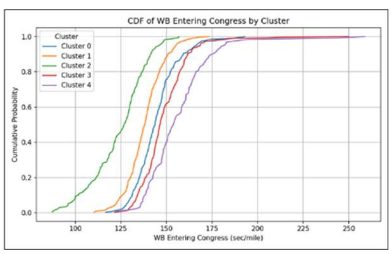



(c) SB Entering MT

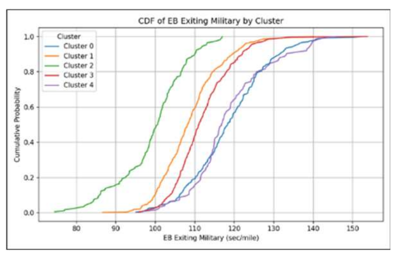



(d) NB Entering MT

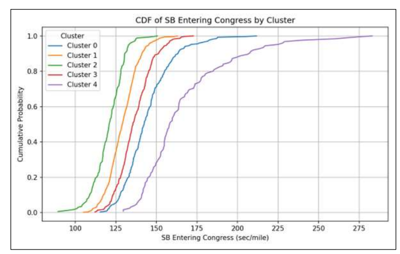
**Figure 7-7 CDF chart of Travel Time at MT**




(a) WB Entering CA

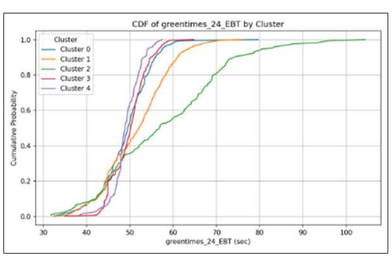



(b) EB Exiting MT or EB Entering CA

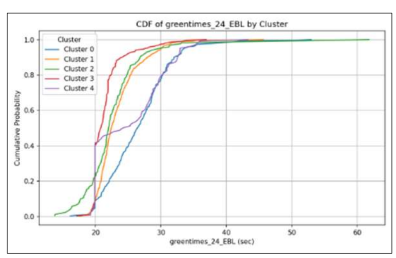



(c) SB Entering CA

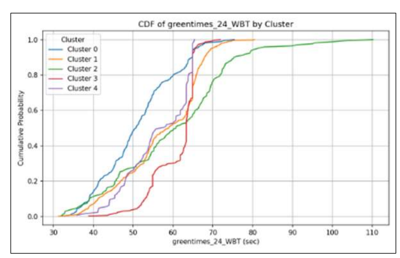
**Figure 7-8 CDF chart of Travel Time at CA**



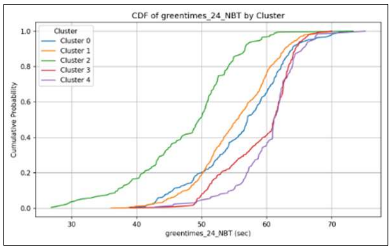

(a) EBT Green Times at MT



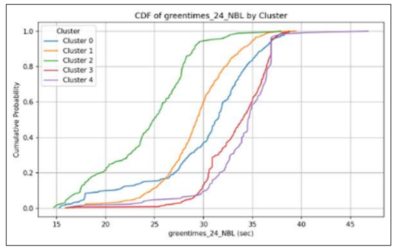

(b) EBL Green Times at MT



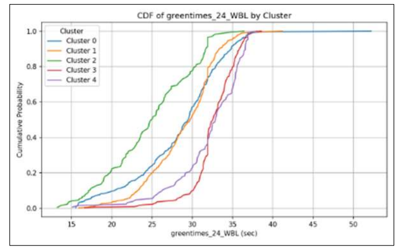

(c) WBT Green Times at MT



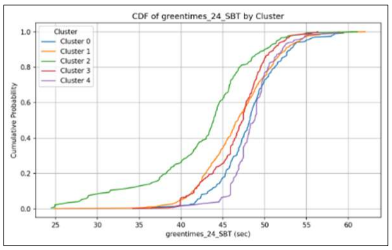

(d) NBT Green Times at MT



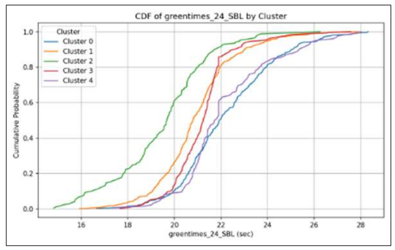

(e) NBL Green Times at MT




(f) WBL Green Times at MT

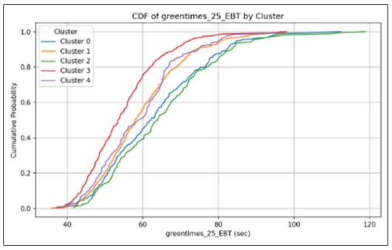



(g) SBT Green Times at MT

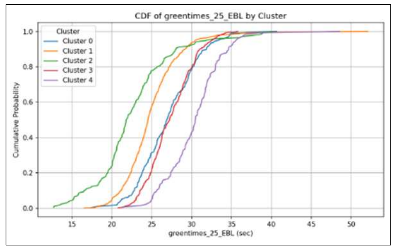



(h) SBL Green Times at MT

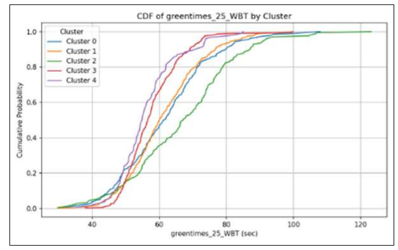
**Figure 7-9 CDF Chart of Green Times at MT**



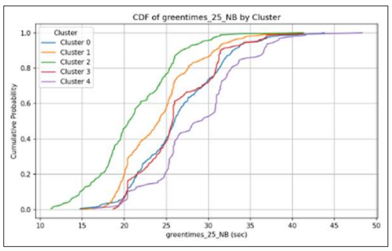

(a) EBT Green Times at CA



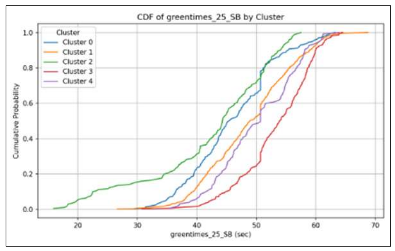

(b) EBL Green Times at CA



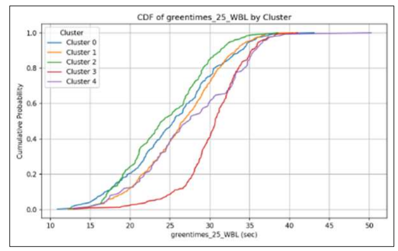

(c) WBT Green Times at CA




(d) NB Green Times at CA




(e) SB Green Times at CA




(f) WBL Green Times at CA

**Figure 7-10 CDF Chart of Green Times at CA**

### 7.2.4 Corridor Performance

The previous sections described the identification of traffic patterns and the factors contributing to those patterns based on segment-level operational measures, including travel time rate and traffic flow rate. This section extends the analysis by evaluating intersection performance using movement-level measures: vehicle-hours of delay (VHD), excess delay (ED), and performance-exceeding cycle counts (PECC). **Table 7-6** summarizes the percentage of the network-wide total VHD, ED, and PECC accrued during the study period that is attributable to each movement at the two study intersections across the four traffic patterns. Cell shading represents the relative magnitude of each movement's contribution, with red indicating the highest contributions and green indicating the lowest.

It is noteworthy that, for both VHD and ED, Cluster 0 exhibits values comparable to those of Cluster 4 for the westbound movement on MT and the southbound movement on CA. In addition, Cluster 0 records the second-highest value for the northbound movement on MT. This is particularly notable given that westbound and eastbound traffic volumes on the facility, as well as southbound volumes on MT, were lower in Cluster 0 than in Cluster 3. Despite these lower demand levels, several movements in Cluster 0 experienced poorer operational performance than Cluster 3.

These findings further support the observation that Cluster 0 performed worse operationally than Cluster 3, despite carrying lower traffic volumes on several key approaches. The results suggest that traffic demand alone does not fully explain the differences in performance among the clusters and that signal control may have contributed to the elevated delays and operational deficiencies observed in Cluster 0. The GOR analysis was initially included as another congestion measure. However, the results were found to be unreliable because of inconsistencies in the detector data.

**Table 7-6 Percentages of Total VHD, ED, and PECC Experienced during the Study Period that are Attributable to Each Movement under Each of the Four Traffic Patterns**

| Cluster | %VHD (Using Speed Limit) (%) Per Day | | | | | | | Total Per Cluster |
|---|---|---|---|---|---|---|---|---|
| | EB MT | WB MT | NB MT | SB MT | WB CA | EB CA | SB CA | |
| 0 | 1.08% | 6.55% | 5.98% | 3.00% | 3.06% | 1.33% | 3.15% | 24.15% |
| 1 | 0.89% | 3.48% | 3.82% | 2.50% | 2.97% | 1.17% | 2.34% | 17.17% |
| 2 | 0.39% | 0.96% | 2.92% | 1.70% | 1.29% | 0.56% | 1.42% | 9.24% |
| 3 | 1.28% | 3.59% | 5.56% | 3.37% | 3.49% | 1.31% | 2.51% | 21.10% |
| 4 | 1.51% | 6.38% | 7.78% | 4.15% | 3.92% | 1.46% | 3.14% | 28.34% |
| Total Per Movement | 5.15% | 20.96% | 26.06% | 14.72% | 14.73% | 5.83% | 12.56% | 100.00% |

| Cluster | ED (Using 60% Speed Limit) (%) Per Day | | | | | | | Total Per Cluster |
|---|---|---|---|---|---|---|---|---|
| | EB MT | WB MT | NB MT | SB MT | WB CA | EB CA | SB CA | |
| 0 | 0.21% | 8.34% | 8.21% | 3.41% | 3.05% | 0.50% | 3.27% | 26.99% |
| 1 | 0.08% | 2.64% | 3.88% | 2.44% | 2.63% | 0.20% | 1.78% | 13.65% |
| 2 | 0.00% | 0.03% | 2.30% | 1.06% | 0.36% | 0.02% | 0.46% | 4.23% |
| 3 | 0.40% | 2.80% | 7.37% | 4.08% | 3.61% | 0.26% | 2.12% | 20.64% |
| 4 | 0.74% | 7.88% | 12.06% | 5.70% | 4.29% | 0.58% | 3.25% | 34.50% |
| Total Per Movement | 1.43% | 21.69% | 33.82% | 16.69% | 13.94% | 1.56% | 10.88% | 100.00% |

| Cluster | PECC Per Day | | | | | | | Total Per Cluster |
|---|---|---|---|---|---|---|---|---|
| | EB MT | WB MT | NB MT | SB MT | WB CA | EB CA | SB CA | |
| 0 | 0.0% | 2.5% | 3.4% | 3.2% | 2.8% | 0.2% | 2.4% | 14.6% |
| 1 | 0.0% | 0.5% | 1.5% | 1.3% | 1.2% | 0.0% | 0.7% | 5.2% |
| 2 | 0.0% | 0.0% | 14.0% | 6.1% | 1.4% | 0.0% | 0.0% | 21.4% |
| 3 | 0.0% | 0.9% | 2.6% | 2.3% | 2.3% | 0.0% | 1.3% | 9.3% |
| 4 | 1.2% | 6.9% | 9.1% | 9.0% | 8.7% | 0.5% | 7.9% | 43.2% |
| Total Per Movement | 1.2% | 10.8% | 30.6% | 21.9% | 16.2% | 0.8% | 12.3% | 93.7% |

### 7.2.5 Signal Control Analysis

Cluster 1 represents the summer traffic pattern and exhibits relatively low levels of congestion, reflecting the reduced travel demand typically observed in South Florida during the summer months. In contrast, the performance measures for Cluster 4 indicate that it is the highest-demand traffic pattern, characterized by congested conditions across most movements, particularly on MT. Cluster 3 represents a typical school-day traffic pattern and generally exhibits lower congestion levels across all movements compared to Cluster 4. However, the northbound approach appears to experience relatively poorer performance than the other approaches within this cluster. Overall, the distribution of signal timing splits appears to provide a reasonable balance among competing movements in Clusters 3 and 4. However, the results suggest that additional fine-tuning of the signal timing plan may be warranted, particularly to allocate more green time to the northbound approach and improve its operational performance.

The poorer performance observed in **Cluster 0** relative to **Cluster 3**, despite lower traffic demand on several movements, suggests potential deficiencies in the allocation of green time during Cluster 0 conditions. Further investigation revealed that the corridor was operating under an adaptive signal control system during the days classified as Cluster 0, while the days in Clusters 3 and 4 were operating under time-of-day plans. Discussions with City staff indicated that when operating the adaptive signal control, the traffic detection system was experiencing operational issues during this period, which likely degraded the quality of the adaptive signal control decisions. According to the signal operations personnel, the adaptive system was intermittently enabled and disabled throughout portions of the six-month study period due to these detection-related challenges resulting in traffic patterns under time-of-day (Clusters 3 and 4) and traffic patterns under adaptive signal control.

The findings suggest that the underperformance observed in Cluster 0 is attributable to these system and detection issues rather than to traffic demand characteristics alone. City staff indicated that the detection problems have since been resolved and that the adaptive signal control system is now operating as intended. To further investigate this issue, the case study corridor was evaluated under Cluster 0 demand conditions using three alternative signal timing plans: (1) the average timings implemented by the adaptive signal control system during Cluster 0 periods, (2) the time-of-day (TOD) plan that was in place prior to the deployment of adaptive signal control, and (3) the average green times assigned by City operations staff during Cluster 4 conditions. The evaluation was conducted using the Highway Capacity Software (HCS) Urban Streets/Urban Arterials procedure. The analysis produced total delay estimates of 1,446, 1,037, and 852 vehicle-hours, respectively. These results indicate that both fixed-time timing plans—the original TOD plan and the timing plan based on Cluster 4 green allocations outperformed the adaptive signal control timings used during Cluster 0 periods. The findings support the conclusion that the adaptive signal control system was not operating optimally during Cluster 0 conditions, likely due to the detection and operational issues identified.

## 7.3 Conclusions

This study developed a data-driven framework for identifying and evaluating traffic patterns across different days of the year, as well as the factors contributing to variations in operational performance among these patterns. Such analysis enables the identification of appropriate corrective actions by linking variations in the assessed performance measures to deficiencies in signal timing parameter selections for specific traffic patterns. The developed method was demonstrated using a case study segment located on Yamato Road between Military Trail and Congress Avenue in the City of Boca Raton, Florida.

The developed method uses K-means clustering in combination with principal component analysis (PCA) to successfully capture and characterize day-to-day variations in traffic conditions. In addition, by integrating insights from PCA loadings, XGBoost (a decision tree-based machine learning method), and Spearman correlation analysis, the study further identified that both traffic volume distribution and signal green time allocation are key determinants of corridor performance. In particular, the findings indicate that congestion is not driven solely by demand levels, but is also strongly affected by signal timing inefficiencies, particularly the imbalance of green time across competing movements. Further analysis suggests that a major contributor to the observed inefficiencies was the use of an adaptive signal control strategy that was subsequently disabled during the afternoon peak period, likely due to degraded performance under high-congestion conditions.

## REFERENCES


AASHTO 2010. *Highway safety manual,* Washington, DC, American Association of State Highway and Transportation Officials.

AASHTO 2014. *Highway Safety Manual 1st Edition with 2014 Supplement,* Washington, DC, American Association of State Highway and Transportation Officials.

ALTALHAN, M., ALGARNI, A. & ALOUANE, M. T.-H. 2025. Imbalanced data problem in machine learning: A review. *IEEE access,* 13**,** 13686-13699.

AN, C., WU, Y.-J., XIA, J. & LU, Z. 2017. Investigating impacts of communication loss on signal performance with use of event-based data. *Transportation Research Record,* 2645**,** 38-49.

ARAGHI, S., KHOSRAVI, A. & CREIGHTON, D. 2015. A review on computational intelligence methods for controlling traffic signal timing. *Expert systems with applications,* 42**,** 1538-1550.

ASUNDI, R. V., PRAKASH, R., KUMAR, K. & INDIA, B. 2022. Class weight technique for handling class imbalance. *BMS Institute of Technology and Management*.

BABICEANU, S. 2021. BEYOND MAP-21 REQUIREMENTS-Assessing Travel Time Reliability and Congestion for Road Projects in Virginia. *National Travel Monitoring Exposition and Conference (NaTMEC 2021).*

BAKIRARAR, B. & ELHAN, A. H. 2023. Class weighting technique to deal with imbalanced class problem in machine learning: methodological research. *Türkiye Klinikleri Biyoistatistik,* 15**,** 19-29.

BERGSTRA, J. & BENGIO, Y. 2012. Random search for hyper-parameter optimization. *Journal of machine learning research,* 13.

CAMBRIDGE SYSTEMATICS, TEXAS A&M TRANSPORTATION INSTITUTE, UNIVERSITY OF WASHINGTON, DOWLING ASSOCIATES, STREET SMARTS, LEVINSON, H. & RAKHA, H. 2013. Analytical procedures for determining the impacts of reliability mitigation strategies. *SHPR 2 Report S2-L03-RR-1.* Washington, DC: Transportation Research Board of the National Academies

CHAWLA, N. V., BOWYER, K. W., HALL, L. O. & KEGELMEYER, W. P. 2002. SMOTE: synthetic minority over-sampling technique. *Journal of artificial intelligence research,* 16**,** 321-357.

CHEN, P., ZHENG, N., SUN, W. & WANG, Y. 2019. Fine-tuning time-of-day partitions for signal timing plan development: revisiting clustering approaches. *Transportmetrica A: Transport Science,* 15**,** 1195-1213.

CHIU, S. Adaptive traffic signal control using fuzzy logic. Proceedings of the Intelligent Vehicles92 Symposium, 1992. IEEE, 98-107.

CHOI, K., YI, J., PARK, C. & YOON, S. 2021. Deep learning for anomaly detection in time-series data: Review, analysis, and guidelines. *IEEE access,* 9**,** 120043-120065.

DARBAN, Z. Z., WEBB, G. I., PAN, S., AGGARWAL, C. & SALEHI, M. 2024. Deep learning for time series anomaly detection: A survey. *ACM Computing Surveys,* 57**,** 1-42.

DAY, C. M., BULLOCK, D. M., LI, H., LAVRENZ, S. M., SMITH, W. B. & STURDEVANT, J. R. 2016. Integrating traffic signal performance measures into agency business processes. *Pooled Fund Study (TPF-5(258)).* West Lafayette, Indiana: Purdue University.

DAY, C. M., BULLOCK, D. M., LI, H., REMIAS, S. M., HAINEN, A. M., FREIJE, R. S., STEVENS, A. L., STURDEVANT, J. R. & BRENNAN, T. M. 2014. Performance measures for traffic signal systems: An outcome-oriented approach. *Pooled Fund Study (TPF-5(258)).* West Lafayette, Indiana: Purdue University.

DELAWARE DEPARTMENT OF TRANSPORTATION 2024. Advanced Transportation and Congestion Management Technologies Deployment (ATCMTD). *Artificial Intelligence-enhanced Integrated Transportation Management System (AI-ITMS) Final Report 2023.* Delaware Department of Transportation.

DING, C. & HE, X. K-means clustering via principal component analysis. Proceedings of the twenty-first international conference on Machine learning, 2004. 29.

DUNTEMAN, G. H. 1989. *Principal components analysis*, Sage.

ERICKSON, B. J. & KITAMURA, F. 2021. Performance metrics for machine learning models. *Radiology: Artificial Intelligence,* 3**,** e200126.

EVIDENTLY AI. 2025. *Classification metrics guide: How to explain the ROC curve and ROC AUC score?* [Online]. Available: https://www.evidentlyai.com/classification-metrics/explain-roc-curve [Accessed].

FDOT 2017. Transportation Systems Management & Operations (TSM&O) 2017 Strategic Plan.

FHP & FDOT 2014. Open Roads Policy Agreement: Quick Clearance for Safety and Mobility. Revised January 2014 ed. Florida: State of Florida.

FHWA 2006. Travel Time Reliability: Making It There On Time, All The Time. Washington, DC: U.S. Department of Transportation, Federal Highway Administration, Office of Operations.

FHWA 2012. Moving ahead for progress in the 21st century act (MAP-21): A summary of highway provisions. Washington, DC: Federal Highway Administration

FHWA 2023. Manual on Uniform Traffic Control Devices for Streets and Highways, 11th Edition, Part 6: Temporary Traffic Control. Washington, DC: U.S. Department of Transportation, Federal Highway Administration.

FORMAN, G. & SCHOLZ, M. 2010. Apples-to-apples in cross-validation studies: pitfalls in classifier performance measurement. *Acm Sigkdd Explorations Newsletter,* 12**,** 49-57.

FREIJE, R. S., HAINEN, A. M., STEVENS, A. L., LI, H., SMITH, W. B., SUMMERS, H., DAY, C. M., STURDEVANT, J. R. & BULLOCK, D. M. 2014. Graphical performance measures for practitioners to triage split failure trouble calls. *Transportation Research Record,* No. 2439**,** 27-40.

GUAN, F. 2022. *Evaluation and Application of Roadside LiDAR-based and Vision-based Multi-model Trajectory Data.* University of Nevada, Reno.

GUO, M., WANG, P., CHAN, C.-Y. & ASKARY, S. A reinforcement learning approach for intelligent traffic signal control at urban intersections. 2019 IEEE Intelligent Transportation Systems Conference (ITSC), 2019. IEEE, 4242-4247.

HADI, M., ELEFTERIADOU, L., MATA, H. & AMINI, E. 2023. Real-Time Data-Based Decision Support System for Arterial Traffic Management.

HADI, M., HALE, D., NOYES, P., MATA, H., WANG, T., ERCISLI, S. & ABDULKADIR, F. Forthcoming-a. An Integrated Management Support System Combining Business Intelligence and Decision Support Systems. Washington, DC: FHWA Report.

HADI, M., MOGHRABI, Y. A., HUANG, Z., NUJJETTY, A. P. & WANG, Q. Forthcoming-b. State-of-the-Practice Assessment and Gap Analysis of Safety-Focused Simulation and Performance Measures. Washington, DC: FHWA Report.

HARTIGAN, J. 1975. Clustering algorithms wiley. *New York*.

HARTIGAN, J. A. & WONG, M. A. 1979. A k-means clustering algorithm. *Applied statistics,* 28**,** 100-108.

HICKS, S. A., STRÜMKE, I., THAMBAWITA, V., HAMMOU, M., RIEGLER, M. A., HALVORSEN, P. & PARASA, S. 2022. On evaluation metrics for medical applications of artificial intelligence. *Scientific reports,* 12**,** 5979.

HIPPOLITUS, J. 2020. Traffic signal control using machine learning. *INTERNATIONAL JOURNAL,* 10**,** 10849-10872.

IQBAL, A., AMIN, R., ALSUBAEI, F. S. & ALZAHRANI, A. 2024. Anomaly detection in multivariate time series data using deep ensemble models. *Plos one,* 19**,** e0303890.

LAWE, S. & WANG, R. 2016. *Optimization of Traffic Signals Using Deep Learning Neural Networks*.

LI, J., YANG, Y., HU, Y., ZHU, X., MA, N. & YUAN, X. 2023. Using Multidimensional Data to Analyze Freeway Real-Time Traffic Crash Precursors Based on XGBoost-SHAP Algorithm. *Journal of Advanced Transportation,* 2023**,** 5789573.

LI, X. & MIAO, H. 2024. Application of artificial intelligence technology in optimizing control parameters of traffic signal group systems. *International Journal of Intelligent Information Technologies (IJIIT),* 20**,** 1-19.

LIST, G. F., WILLIAMS, B. M., ROUPHAIL, N. M., HRANAC, R., BARKLEY, T., MAI, E., CICCARELLI, A., RODEGERDTS, L., PINCUS, K., NEVERS, B., KARR, A. F., ZHOU, X., WOJTOWICZ, J., SCHOFER, J. L. & KHATTAK, A. J. 2014. Establishing Monitoring Programs for Travel Time Reliability. *SHPR 2 Report S2-L02-RR-1.* Washington, DC: Transportation Research Board of the National Academies

MANNION, P., DUGGAN, J. & HOWLEY, E. 2016. An experimental review of reinforcement learning algorithms for adaptive traffic signal control. *Autonomic road transport support systems***,** 47-66.

MARINAI, S., FAINI, S., MARINO, E. & SODA, G. Efficient word retrieval by means of SOM clustering and PCA. International Workshop on Document Analysis Systems, 2006. Springer, 336-347.

MENG, S., FU, Y., LIU, T. & LI, Y. Principal component analysis for clustering temporomandibular joint data. 2015 8th International Symposium on Computational Intelligence and Design (ISCID), 2015. IEEE, 422-425.

MISHRA, A., CHEN, K., PODDAR, S., POSADAS, E., RANGARAJAN, A. & RANKA, S. 2022. Using video analytics to improve traffic intersection safety and performance. *Vehicles,* No. 4**,** 1288-1313.

MITROVIC, N. & STEVANOVIC, A. 2019. Estimating peak-hour traffic profiles for selection of appropriate day-of-year signal timing plans. *Transportation Research Record,* 2673**,** 199-213.

NATIONAL ACADEMIES OF SCIENCES, E., AND MEDICINE, 2022. *Highway Capacity Manual 7th Edition: A Guide for Multimodal Mobility Analysis,* Washington, DC, The National Academies Press.

NIITTYMÄKI, J. & PURSULA, M. 2000. Signal control using fuzzy logic. *Fuzzy sets and systems,* 116**,** 11-22.

OKLAHOMA DEPARTMENT OF TRANSPORTATION 2022. ODOT Rural AI Traffic Signal Upgrades for Increased Safety and Mobility. Oklahoma Department of Transportation,.

PACK, M., QIAN, S., IVANOV, N., HALE, D. & BAUER, J. Fortcoming. A Framework to Support Data Fusion, Analysis, and Decisionmaking. *FHWA-HOP-22-001.* Washington, DC.

PARSA, A. B., MOVAHEDI, A., TAGHIPOUR, H., DERRIBLE, S. & MOHAMMADIAN, A. K. 2020. Toward safer highways, application of XGBoost and SHAP for real-time accident detection and feature analysis. *Accident Analysis & Prevention,* 136**,** 105405.

PORTER, R. J., DUNN, M., KERSAVAGE, K., GAYAH, V., ECCLES, K., PERSAUD, B., MEDINA, J. & PERSAUD, B. 2023. Estimating Effectiveness of Safety Treatments in the Absence of Crash Data. *NHCRP Project No. 17-86.* Washington, DC: The National Academies Press.

PROBST, P., WRIGHT, M. N. & BOULESTEIX, A. L. 2019. Hyperparameters and tuning strategies for random forest. *Wiley Interdisciplinary Reviews: data mining and knowledge discovery,* 9**,** e1301.

RAO, W., XIA, J., LYU, W. & LU, Z. 2019. Interval data-based k-means clustering method for traffic state identification at urban intersections. *IET Intelligent Transport Systems,* 13**,** 1106-1115.

RASCHKA, S. & MIRJALILI, V. 2019. *Python machine learning: Machine learning and deep learning with Python, scikit-learn, and TensorFlow 2*, Packt publishing ltd.

SAE INTERNATIONAL TECHNICAL STANDARD 2015. Dedicated Short Range Communications (DSRC) Message Set Dictionary. *SAE J2735_201603.*

SAHA, R., TARIQ, M. T. & HADI, M. 2020. Deep learning approach for predictive analytics to support diversion during freeway incidents. *Transportation research record,* 2674**,** 480-492.

SAHA, R., TARIQ, M. T., HADI, M. & XIAO, Y. 2019. Pattern recognition using clustering analysis to support transportation system management, operations, and modeling. *Journal of Advanced Transportation,* 2019**,** 1628417.

SCHMIDL, S., WENIG, P. & PAPENBROCK, T. 2022. Anomaly detection in time series: a comprehensive evaluation.

SCIKIT-LEARN DEVELOPERS. 2025. *Cross-validation: evaluating estimator performance* [Online]. Available: https://scikit-learn.org/stable/modules/cross_validation.html [Accessed July 2025].

SINGLETON, P. A., RUNA, F. & HUMAGAIN, P. 2020. Utilizing archived traffic signal performance measures for pedestrian planning and analysis.

SMAGLIK, E. J., RIFFLE, K., GEHRKE, S., RUSSO, B., PROCACCIO, S., HURWITZ, D., LIU, Y., CHAI, E. & LI, X. 2022. Automated Identification of Traffic Detector Malfunctions.

SMITH, B. L., SCHERER, W. T. & HAUSER, T. A. 2001. Data-mining tools for the support of signal-timing plan development. *Transportation Research Record,* 1768**,** 141-147.

SRINIVASAN, D., CHOY, M. C. & CHEU, R. L. 2006. Neural networks for real-time traffic signal control. *IEEE Transactions on intelligent transportation systems,* 7**,** 261-272.

SUNKARI, S., SONGCHITRUKSA, P., CHARARA, H. & ZENG, X. 2010. Improved intersection operations during detector failures.

TARIQ, M. T., MASSAHI, A., SAHA, R. & HADI, M. 2020. Combining machine learning and fuzzy rule-based system in automating signal timing experts' decisions during non-recurrent congestion. *Transportation Research Record,* 2674**,** 163-176.

TARIQ, M. T., SAHA, R. & HADI, M. 2022. Data and modeling support of the management of diversion routes during freeway incidents. *Transportation research record,* 2676**,** 437-452.

TARUNESH, I. & CHUNG, E. 2020. Predicting traffic volume and occupancy at failed detectors. *Transportation Research Procedia,* 48**,** 1072-1083.

TINAWI, I. 2019. *Machine learning for time series anomaly detection.* Massachusetts Institute of Technology.

TRAN, T. D. & KIM, J. 2024. Machine learning modeling structures and framework for short-term forecasting and long-term projection of Streamflow. *Stochastic Environmental Research and Risk Assessment,* 38**,** 793-813.

URBANIK, T., HUMPHREYS, D., SMITH, B. & LEVINE, S. 2006. Coordinated Freeway and Arterial Operations Handbook. *In:* TECHNOLOGY, F. H. A. R. A. (ed.) *Coordinating, Developing, and Delivering Highway Transportation Innovations.* Federal Highway Administration: U.S. Department of Transportation.

VAJDA, D. L., DO, T. V., BÉRCZES, T. & FARKAS, K. 2024. Machine learning-based real-time anomaly detection using data pre-processing in the telemetry of server farms. *Scientific Reports,* 14**,** 23288.

VANDERVALK, A., LOUCH, H., GUERRE, J. & MARGIOTTA, R. 2014. Incorporating reliability performance measures into the transportation planning and programming processes: technical reference. *SHRP 2 Report S2-L05-RR-3.* Washington, DC: Transportation Research Board of the National Academies

WANG, B., FULDA, N., HUANG, Z., SCHULTZ, G. G., MACFARLANE, G. S., ARNESEN, J. & KHAYYAT, A. 2024. Predicting Directional Traffic Volume at Intersections with Automated Traffic Signal Performance Measures Data Using Machine Learning Algorithms. *Transportation Research Record,* 2678**,** 1736-1750.

WANG, B., SCHULTZ, G. G., MACFARLANE, G. S. & MCCUEN, S. 2022. Evaluating signal systems using automated traffic signal performance measures. *Future transportation,* 2**,** 659-674.

WANG, G., QIN, W. & WANG, Y. 2021. Cyclic weighted k-means method with application to time-of-day interval partition. *Sustainability,* 13**,** 4796.

WANG, X., COTTRELL, W. & MU, S. Using k-means clustering to identify time-of-day break points for traffic signal timing plans. Proceedings. 2005 IEEE Intelligent Transportation Systems, 2005., 2005. IEEE, 586-591.

WIERZCHOWSKA, A. 2016. *Design of Traffic-Actuated Plan Selection Road Signal Control.* Sapienza University of Rome.

WINFREY, C., MELEBY, P. & MIAO, L. 2023. Using big data and machine learning to rank traffic signals in Tennessee. *Journal of traffic and transportation engineering (English edition),* 10**,** 918-933.

YAO, W. & QIAN, S. 2020. Learning to recommend signal plans under incidents with real-time traffic prediction. *Transportation Research Record,* 2674**,** 45-59.

ZEGEER, J., BONNESON, J. A., DOWLING, R. G., RYUS, P., VANDEHEY, M., KITTELSON, W., ROUPHAIL, N. M., SCHROEDER, B. J., HAJBABAIE, A., AGHDASHI, B., CHASE, T., SAJJADI, S., MARGIOTTA, R. & ELEFTERIADOU, L. 2014. Incorporating travel time reliability into the Highway Capacity Manual. *SHRP 2 Report S2-L08-RW-1.* Washington, DC: Transportation Research Board of the National Academies

ZHOU, X., LUO, X., ABBASI, M., HUANG, Z. & TYAGI, A. 2024. Emerging Data Cleaning and Fusion for Traffic Model Calibration–Data Fusion for Microsimulation Model Calibration. *FHWA-HRT-24-142.* Washington, DC: Federal Highway Administration.