

## Research Article

# Machine Learning and Reverse Methods for a Deeper Understanding of Public Roadway Improvement Action Impacts during Execution

Sohrab Mamdoohi and Elise Miller-Hooks 

*Department of Civil, Environmental and Infrastructure Engineering, George Mason University, Fairfax, VA 22030, USA*

Correspondence should be addressed to Elise Miller-Hooks; [miller@gmu.edu](mailto:miller@gmu.edu)

Received 31 August 2021; Accepted 7 September 2022; Published 27 September 2022

Academic Editor: Luis Miranda-Moreno

Copyright © 2022 Sohrab Mamdoohi and Elise Miller-Hooks. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The execution of public roadway maintenance, rehabilitation, and restoration activities disturb normal traffic flows, resulting in roadway capacity reduction, inducing travel time delays, and promoting traffic safety concerns. While they improve public roadway performance once complete, the impacts endured in executing these actions is significant. This work seeks a deeper understanding of the effects of improvement actions on traffic by juxtaposing their effects against those arising from traffic incidents that cause similar capacity reductions and related negative externalities. This is accomplished through direct and reverse comparisons with traffic incident impacts. A measure of unit delay that uses observations to determine event location extent, duration, and propagation direction was computed at both facility and corridor-wide levels to establish the degree to which improvement actions and traffic incidents are similar or dissimilar. Alternative hybrid machine-learning methods are proposed to identify and contrast those traffic characteristics that contribute greatest to correct detection of each type of downtime event. These techniques can detect traffic events and accurately distinguish between event types (whether a collision or improvement activity). The techniques were applied on seven months of data obtained from 2019 along three corridors from northern, southern, and western regions of the Commonwealth of Virginia. Those traffic characteristics that contribute greatest to correct event detection of each event type were identified and their similarities and differences were studied. General linear, multivariate regression equations were also developed for more general application.

## 1. Introduction

Public roadway maintenance, rehabilitation, and restoration activities (together improvement actions), like traffic accidents, disturb normal traffic flows, resulting in roadway capacity reduction, inducing travel time delays, promoting traffic safety concerns, and increasing net public cost. This capacity reduction may be due to blocked traffic lanes or even a roadway component (e.g. a bridge) that is temporarily taken out of commission during execution of improvement activities. This period of reduced capacity from such non-recurring events is considered, herein, as roadway downtime, and the inducing events as downtime events.

Reduction in travel time reliability, degradation in serviceability, increase in primary or secondary traffic incidents

[1], and increase in fuel consumption are other consequences of capacity reduction due to roadway downtime. In the U.S., more than 550 million gallons of fuel and 480 million hours are lost every year due to traffic congestion brought about specifically by work zones [2]. Also, an estimated 10% of congestion and 24% of unexpected freeway delays are caused by work-zones [3]. Traffic incident-related delays form another 13 to 30% of the total congestion delay over peak periods [4]. Even after roadway capacity is reinstated, the activity is complete or the event is cleared, degradation in roadway system performance can be expected for a subsequent period of time.

Improvement actions preserve or increase serviceability and ultimately improve safety, and travel time reliability for future operations. As they are undertaken, however, they negatively impact travel times and travel-time reliability for

both the facility, and, due to shifts in traffic to alternative facilities, adjacent roadways. These temporary impacts are often not explicitly accounted for in the activity planning process. Yet, in total, these actions are not infrequent and their effects not negligible; their impacts are realized during large swaths of time. Moreover, their effects are not only local but facility-level and network-wide.

This work seeks a deeper understanding of the effects of improvement actions on traffic by juxtaposing their effects against those arising from traffic incidents that cause similar capacity reductions and related negative externalities. This is accomplished through direct and reverse approaches. In the direct approach, computed vehicle unit delays are compared by event type. Computations explicitly account for location extent, duration, and negative impact propagation direction. To support the reverse engineering investigation, machine learning methods, based on concepts of Support Vector Machines (SVM) or Random Forests (RFs) with K-Nearest Neighbor (KNN), are proposed and applied to identify those traffic characteristics that contribute greatest to correct detection of both types of downtime events, i.e. roadway improvement actions and traffic incidents. Similarities and differences in these contributing factors aided in understanding their unique qualities. Additionally, findings from this investigation aided in efficiently identifying key variables for inclusion in unit delay estimation models.

Application of the developed methods and metrics on case studies allowed investigation into the: (1) most affected traffic features for event type; (2) accuracy of hybrid event detection methods and the transferability of chosen parameters; (3) influence of downtime from improvement actions on corridor performance; (4) unit delay estimates and their generalizability across corridors; and (5) relationship between event features and unit delay values. Generally, the outcomes of this work can be used to produce estimates of the public cost, whether direct through investments or indirect in the form of user costs, of highway improvement projects. These costs can then be weighed against their benefits.

A key finding of the results is that travel time reliability-related characteristics contribute more to detection of traffic incidents than to detection of improvement actions even when executed during peak hours. This indicates that improvement actions are less likely to affect traffic performance than are traffic incidents. In general, the results show that traffic incidents have more than five times the impact on traffic compared with improvement actions. This five-fold impact was found both through investigation of the contribution of traffic characteristics to correct event detection and estimated unit delays from both event types, and persists whether the analysis is conditioned on peak hours.

Also key, 60% of traffic improvement events in the case studies had nearly zero traffic impact. While not as monumental as traffic incidents, improvement actions blocking one lane and requiring approximately one hour create 9 minutes delay per vehicle on average (ranging from 0 to 16 minutes with a 90% confidence interval) when executed in the peak period. Total delays over all vehicles from improvement activities, thus, can be very substantial. However, traf-

fic incidents blocking one lane for one hour in the same peak hours create 16 minutes delay per vehicle on average. Moreover, each additional lane blocked in a collision increases the average unit delay by 67%, while in an improvement action, the average unit delay increases 17%.

The next section reviews relevant literature and further establishes the paper's contributions. This is followed by details of the methodological approaches used in this study (Section 2) and results from their application on three corridors within the Commonwealth of Virginia (Section 3). Additional findings and their implications follow in Section 4.

## 2. Literature Review

Several studies have investigated the impact of maintenance activities on the operation of real roadway facilities and the economic costs and benefits of investments in public roadway infrastructure. These works consider the effects of maintenance actions in inducing: capacity reduction (e.g., [5, 6]), secondary incidents (e.g., [7]), and traffic delays (e.g., [2, 8]). Comprehensive reviews of capacity estimation methods and work-zone impacts can also be found in ([9, 10], respectively). Findings from these studies indicate that a large number of factors contribute to the impact of maintenance activities on roadway performance. These factors when gathered across these studies include as follows: percentage of truck traffic; pavement grade; number of lanes and lane closures; lane width, work zone layout (lanes merging, lane shifting, and crossover), construction type, duration, and time (on-/off-peak, day/night); pavement/weather condition (dry, wet, icy, sunny, rainy, snowy); and traveler familiarity (commuter/noncommuter). Some of these studies provide an equation. For example, Nassiri and Aghamohammadi [5] provide a model to predict remaining roadway capacity for a work zone. These equations require input data that may be difficult to obtain, and applying their methodologies in alternative locations requires traffic volumes and/or density, in addition to speed, data as input, which may be difficult to obtain.

These studies make important, limiting assumptions on the direction of speed reduction, and the propagation of other traffic impacts, activity durations, and extent. Specifically, they include changes in traffic measures in both traffic directions equally, use preset activity durations and presume that the location of the traffic impact is limited to the work zone and, in some cases, a preset distance upstream of the work zone. Impacts of maintenance activities were investigated either within a fixed extent of time and location (e.g. [7, 8]) or within a dynamic extent defined over a fixed discharge rate. For example, Du et al. [2] defined the delay as more than 25% drop in the normal speed and then, measured the delay cost based on such criteria. The normal speed must be fed into their model as an input. The methods proposed herein use observations from the data to determine event location extent, duration, and propagation direction. They do not require prior information on normal speeds.

While methods that have been proposed for traffic incident impact analysis have similarly accounted for incident impact propagation direction and may not require preset values of incident impact duration and extent, these other methods cannot be directly applied to work zone impact analyses. They are designed for and trained on traffic events with short durations consistent with vehicular accidents.

Numerous studies provide methods for prioritizing roadway construction projects with a goal of minimizing total costs, including monetary costs of traffic delay, and safety from inaction (e.g., [11–17]). A few prior studies include the negative effects of improvement actions during their execution (i.e. downtime) in prioritization and scheduling processes [13, 18–22]. These works replicate traffic either as a user equilibrium (UE) traffic assignment (using a cell transmission model (CTM) approach or Bureau of Public Roads (BPR) function), day-to-day route choice problem, maximum flow problem, or through microsimulation of traffic.

Using an actor-critic, deep reinforcement learning solution methodology, DCMAC [23], to solve a bilevel model in which activities are scheduled in the upper level and a UE traffic assignment captures the response of traffic in the lower level, Zhou et al. [22] found that a savings in traffic delays of between 16 and 28% with up to 15% cost savings can be achieved by accounting for the construction activity downtime effects during improvement action prioritization and scheduling. In a day-to-day setting, Yang et al. [20] found that work durations of less than 60 days significantly impact the optimal schedule; while, activities with longer duration do not.

While accounting for the impacts of work zones in prioritization and scheduling, these prioritization and scheduling works used rough estimates of improvement action impacts during their execution, e.g. a percent reduction in capacity given a duration and number of lanes blocked, as inputs to their models. Event detection methods and unit delay estimates developed herein can support these approaches, providing specific impact estimates for the considered activities.

The detection of disruptions in traffic was also considered in the literature and dates back to the 1970s [24–26]. These works developed methods to detect a change in lane occupancy values beyond a threshold. They proposed methods based on standard normal deviate, decision trees, and time-series analysis, respectively. More recent detection algorithms make use of newer technologies. Data is collected from loop detectors [27], video cameras [28], probe vehicles [29], and social media [30]. Other recent works also propose the application of more involved artificial intelligence methods (e.g., [30–33]). These machine learning techniques detect anomalies or outliers within the traffic data by considering not only lane occupancy but other factors, such as speed changes. Several works extended these concepts to detect such anomalies in real-time (e.g., [34–36]).

While methods for detection of traffic incidents are plentiful, it appears that no prior study has sought to detect maintenance, rehabilitation or other construction or improvement related activities. Despite that improvement

actions are planned, and often recorded, it can be useful to have the capability to detect these events as their exact time of implementation may not be known, they may not be implemented exactly as planned due to uncertainties such as changing weather, and their occurrence may not be known to all parties. This is especially true for actions with limited duration. Moreover, it may be useful to be able to distinguish a detected potential traffic incident from an ongoing improvement activity in real-time applications. This paper fills this gap. It further proposes a reverse engineering approach for understanding differences between downtime from improvement actions and the effects of traffic incidents and introduces a concept of unit delays, computed values of which may have broad utility.

### 3. Methods

*3.1. Event Detection and Critical Factor Identification through Reverse Engineering.* The proposed reverse engineering technique uses traffic characteristics to detect the occurrence of an event. It estimates the contribution of each characteristic to detection success (correct event detection) and failure (misclassification), and with this information identifies the most critical characteristics for success under both improvement actions and traffic incidents. Investigating differences in those characteristics that most contribute to the identification of these two classes of events aids in providing a deeper understanding of how improvement actions are both similar to and different from traffic incidents in terms of their impacts on traffic performance and public welfare.

Artificial neural network (ANN), support vector machine (SVM), and random forest (RF) techniques are classical machine learning methods that have been widely applied in traffic incident detection studies (e.g. [36–38]). Xiao [39] noted that well-performing models, such as ANNs, did not perform as well when applied to a second data set. Motivated by this, Xiao proposed an ensemble learning method that integrates SVM and K-Nearest Neighbor (KNN) methods for incident detection. This work employs similar concepts to Xiao's ensemble methodology by creating a hybrid of SVM (or RF) and KNN procedures, but with a different structure and application. Here, SVM (or RF) is used first for detection and KNN second for refinement; whereas, Xiao's method seeks confirmation of correct detection by considering whether the events are selected by both methods (SVM and KNN) giving more weight to those events identified in SVM. Moreover, Xiao's method does not correct for discontinuities in space and time arising in the detection of a single event. Thus, it may identify one event as a set of small events. The proposed hybrid methodology employs four detection classes to incorporate and distinguish between traffic incidents and improvement actions. Initial runs were made with a method based on the ensemble design of SVM and KNN in Xiao, but the sequential version proposed herein outperformed the ensemble design.

While methods for detection of traffic incidents are plentiful, it appears that no prior study has sought to detect maintenance, rehabilitation or other construction

or improvement related activities. Despite that improvement actions are planned, and often recorded, it can be useful to have the capability to detect these events as their exact time of implementation may not be known, they may not be implemented exactly as planned due to uncertainties such as changing weather, and their occurrence may not be known to all parties. This is especially true for actions with limited duration. Moreover, it may be useful to be able to distinguish a detected potential traffic incident from an ongoing improvement activity in real-time applications.

As Xiao [39] found in developing an ensemble learning method for traffic incidents, combining methods was found here to enhance detection accuracy. In application to real-world data (Section 5), the RFF-KNN method outperforms the SVM-KNN method in terms of detection accuracy. Both were shown to reduce misclassification errors as compared with the proposed single-phase SVM and RF methods. The following subsections present these hybrid learning methods and their components.

**3.2. Hybrid Learning Methods.** Traffic event detection can be regarded as a classification problem as suggested in [29]. The proposed hybrid learning methods are aimed at classifying each data instance (a row or record from a dataset—here, the traffic characteristics associated with one roadway segment and one point in time) as falling into a nonevent or event class (improvement action, traffic incident, both). Each hybrid method is composed of two learning phases. Phase I starts by training a classification model on a portion, typically half, of the data. The trained SVM/RF classifier detects the type of downtime event (class) (at every location (roadway segment) and time increment (1-minute interval)). Phase II employs the KNN algorithm to further refine the results, reducing the number of misspecified instances. Gaps across time or location for a single event are identified and their classifications modified for increased consistency by maximizing homogeneity within and heterogeneity between downtime event classes.

Let  $\Theta(t_m, l)$  represent an element of a matrix of traffic characteristics (e.g., speed and volume) at time interval  $t_m$  and location  $l$ ,  $DT_{Cat\_P(I)}(t_m, l)$  and  $DT_{Cat\_P(II)}(t_m, l)$  be corresponding elements of matrices of predicted downtime event types resulting from the phase I and II classifiers for the specified  $t_m$  and  $l$ . The hybrid learning method, thus, applies a SVM or RF classifier to  $\Theta(t_m, l)$  in phase I and refines the results using the KNN algorithm in Phase II. The final (refined) prediction,  $DT_{Cat\_P(II)}(t_m, l)$ , uses values of  $DT_{Cat\_P(I)}(t_{m-k}, l)$ ,  $\dots$ ,  $DT_{Cat\_P(I)}(t_{m-1}, l)$ ,  $DT_{Cat\_P(I)}(t_{m+1}, l)$ ,  $\dots$ ,  $DT_{Cat\_P(I)}(t_{m+k}, l)$ , for  $k$  a parameter of the KNN algorithm, from Phase I.

An overview of the proposed hybrid method is depicted in Figure 1. Figure 2 illustrates the steps of this method for a generic location. Applying a classifier in Phase I of this location resulted in detection of an ongoing incident for time periods 5 to 13, 16 to 21, and 28 to 29. In Phase II, KNN refines these time intervals to a contiguous interval by adding time instances 14 and 15 to the detection results and

removing time instances 28 and 29. Descriptions of the SVM, RF and KNN subprocedures follow.

**3.3. Support Vector Machine (SVM).** SVM is a supervised machine learning method that was developed for binary [40] and extended for multi-class [41] classification applications. The multiclass SVM approach breaks the classes into multiple binary (two-class) classification problems that each included the dominant class versus all alternatives. For each, SVM generates an optimal separating hyperplane that maximizes the separating margins of the two classes (e.g. traffic incidents as the dominant class against a class that includes all other event types, here, improvement action, nonevent or both) over a linearly separable sample dataset (here, speed-related changes) [42].

Given a dataset of  $n$  samples,  $(x_1, y_1), \dots, (x_i, y_i), \dots, (x_n, y_n)$  where  $x_i \in \mathbb{R}^n$ ,  $i = 1, \dots, n$  is a training vector containing all traffic characteristics for each data sample, and  $y_i \in \{1, -1\}$  is the class type (dominant or alternative) associated with  $x_i$ , SVM generates separating hyperplanes as formulated in [40]:

$$\begin{aligned} \min_{\omega, b, \xi} \quad & \frac{1}{2} \omega^T \omega + C \sum_{i=1}^n \xi_i, \\ \text{s.t.} \quad & y_i (\omega^T \phi(x_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, \\ & i = 1, \dots, n. \end{aligned} \quad (1)$$

$\omega \in \mathbb{R}^n$  is the normal vector to the hyperplane (hyperplane function:  $f(x) = \omega x + b$ ),  $b \in \mathbb{R}$  the intercept of the hyperplane,  $C > 0$  is a penalty term, and  $\phi$  a manually chosen kernel function.

As traffic data is not generally linearly separable, nonlinear mapping functions, i.e. kernel functions, transform the traffic data into a higher dimensional space through which data points can be linearly separated. A radial basis kernel is used here and is formulated as  $K(x_i, x_j) = \phi(x_i)^T \cdot \phi(x_j) = \exp(-\gamma \|x_i - x_j\|^2)$ ,  $\gamma > 0$ , where  $\gamma$  is a parameter that defines the amount of influence given to a single training data sample. The performance of SVM relies heavily on the choice of the kernel function and its parameters ( $C$  and  $\gamma$ ). As in [43], to improve the classifier when applied to other datasets and avoid overfitting the model to the test data, incremental grid-searches on  $C$  (ranging from 1 to 10 with step size of 1) and  $\gamma$  (ranging from 0.1 to 1 with step size of 0.1) were completed. Values of  $C$  and  $\gamma$  were chosen to provide the highest average  $V$ -fold cross, with  $V = 10$ , validation accuracy. See [44] for background on cross validation method. Once optimal hyperplanes are generated for each two-class problem, the predicted class of a data sample is determined according to the two-class problem with the highest prediction score.

**3.4. Random Forests (RFs).** RFs, originally proposed by Breiman [45], are constructed on a collection of tree-structured



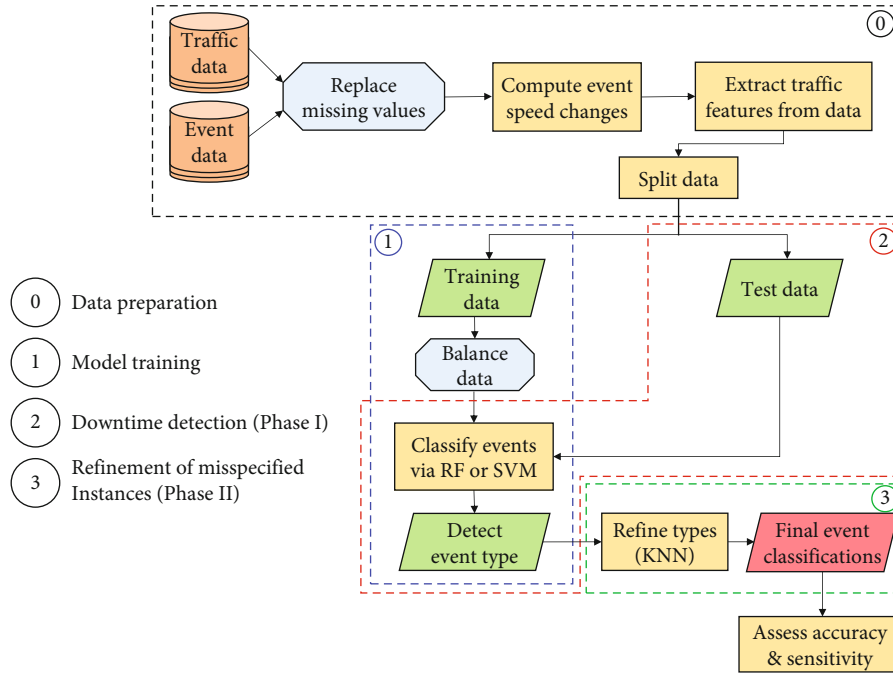


FIGURE 1: Hybrid learning method overview.

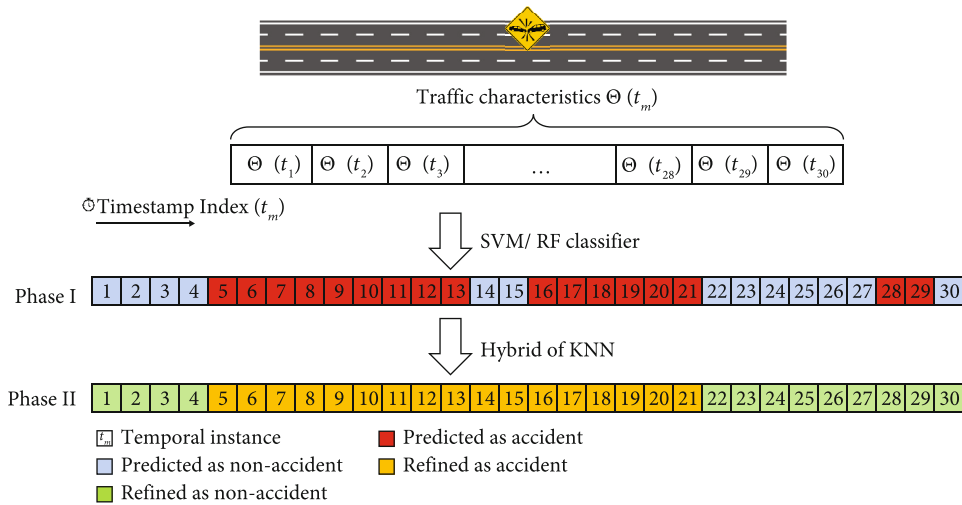


FIGURE 2: Illustrative example of the proposed hybrid method.

classifiers similar to decision trees. Each tree casts a single vote (i.e. identifies one of the possible classes (event types)) toward the most popular class via mode or mean for each data element of a given data set. A RF is formed from an ensemble of randomly generated tree classifiers. Inserting randomness into constructing the tree classifiers makes for a more robust method that is less influenced by outliers and unbalanced datasets and is protected some from overfitting. RFs have been used in traffic incident management and safety studies in recent years (e.g., [46]). Through variable elimination, it has a secondary benefit of providing a measure of variable importance [47, 48].

Here, RF is proposed for detecting event types (traffic incident occurrence, improvement action, both, neither)

given a data set containing traffic characteristics, such as speed changes. Tree classifiers are grown to produce a forest. Each tree classifier is grown by applying conditional statements to a randomly chosen, fixed number ( $m_{try}$ , a parameter) of the event's data items (e.g.  $m_{try}$  traffic characteristics from a row (a sample) in the dataset) creating a decision tree. RF constructs  $n_{tree}$  such decision trees, each predicting a sample's event type. It votes accordingly. Here, the result of each decision tree is one of the four event types (traffic incident, improvement action, both, and neither). The event type with the largest mode over decision tree predictions is taken as the final outcome.

The final chosen type for a given sample is obtained from the total collection of votes, one from each tree in the

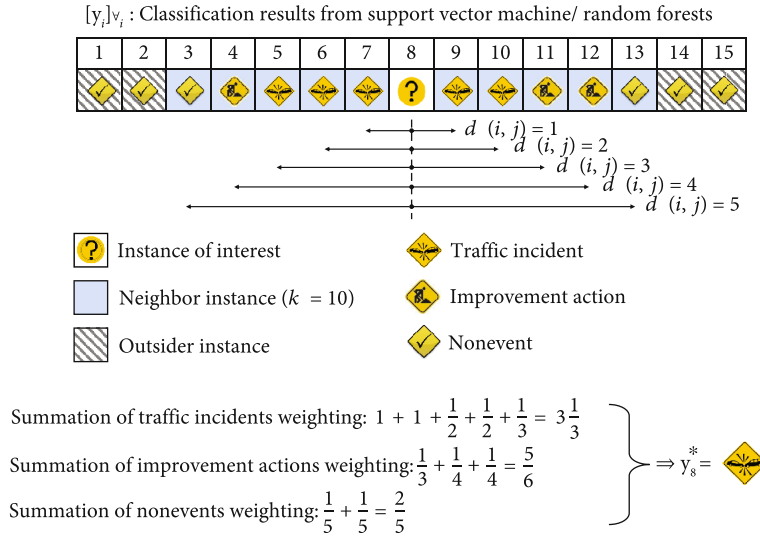


FIGURE 3: Refining classification of a data record through K-Nearest Neighbor.

RF for the test dataset. The accuracy of classification is affected by the chosen (tuned) parameters,  $m_{try}$  and  $n_{tree}$ .

3.5. *K-Nearest Neighbor (KNN)*. KNN is a classification method that labels a data instance according to the classes of the nearest  $k$  instances. The data instance is labeled based on a majority vote from these  $k$  neighbors when weighted by their respective distances. That is, those neighbors closest to the data instance will have greater contribution.

Here, KNN is used to refine classification determinations made from either SVM or RF classifier creating two hybrid techniques: SVM-KNN and RF-KNN. Consider a dataset of event classifications organized by segments over time created from running SVM or RF. For the subset of data records associated with roadway segment  $l$ , classification  $y_i^*$  is obtained from its  $k$  nearest (in time) records:  $y_i^* = [w_j] \times [y_j]_{\forall j \in \{i \pm 1, \dots, i \pm k/2, j \neq i\}}$ . In the voting process, each neighbor  $j \in \{i \pm 1, \dots, i \pm k/2, j \neq i\}$  is given a weight  $w_j$  as a function of its time difference from time  $t_i$ :  $w_j = 1/|t_i - t_j|$  for  $t_i$  and  $t_j$  the time of data instances  $i$  and  $j$ , respectively. This is illustrated on an example in Figure 3 for  $k=10$ .

### 4. Case Study

Three case studies were conducted involving corridors in Virginia, each with a freeway segment and parallel, alternative arterial. The corridors are between 12 and 30 miles in length and each varies in terms of number of lanes (capacity), speed limit, average daily traffic (ADT), percentage of truck traffic, and presence of horizontal and vertical curves. This section describes the case study locations with relevant input data, analysis outcomes, and ultimate findings.

4.1. *Locations and Data Sources*. Case studies were selected from northern, southern, and western regions of the Commonwealth of Virginia. Their locations are depicted in Figure 4 and are specified as follows:

Case study I: A 12-mile stretch of the I-66 Westbound corridor between Leesburg Pike (Exit 66) and US-29 (Exit 52) and adjacent roadway sections of US-50 Westbound (between Graham Rd and US-29/VA-237/Old Lee Hwy) and US-29 (between US-50/VA-237/Old Lee Hwy and I-66 (Centerville)) in Fairfax County.

Case study II: A 30-mile stretch of the I-81 Southbound corridor between I-64/Exit 221 and I-64/Exit 191 and adjacent roadway sections of US-11 Southbound between VA-262 and I-64 in Augusta and Rockbridge Counties.

Case study III: A 25-mile stretch of the I-64 Westbound corridor between Croaker Rd/Exit 231B and US-60/Exit 200 and adjacent roadway sections of US-60 Westbound between Croaker Rd and I-295 in James City and Henrico Counties.

4.1.1. *Traffic Data*. Traffic data, including information related to speeds and travel times, were obtained from the I-95 Vehicle Probe Project (VPP) II contract under the INRIX suite. The data, which provide space-mean speeds, were collected through the INRIX traffic message channel (TMC) monitoring platform. INRIX reports traffic data by road segment, each referred to as a TMC segment. Selected stretches of I-66, I-81, and I-64 include a total of 16, 16, and 14 TMC segments, respectively. Lengths of these TMC segments range from less than 1 mile to 6 miles. INRIX data are widely utilized in transportation studies (e.g. [51, 52]).

The case studies relied on seven-month historical TMC segment data from April 1 to October 31 in 2019. Traffic data were retrieved and aggregated at 1-minute increments (as in [53] for each TMC segment. This choice of a 1-minute resolution for event detection aligns with the reporting frequency of data in the event dataset, and is of the highest level of granularity achievable with the datasets that are used. Traffic data of each TMC segment includes a reference speed (speed limit), current speed, historical average speed, and time required to traverse the segment. For the studied TMC segments, the number of 1-minute traffic data records collected in this 214-day period

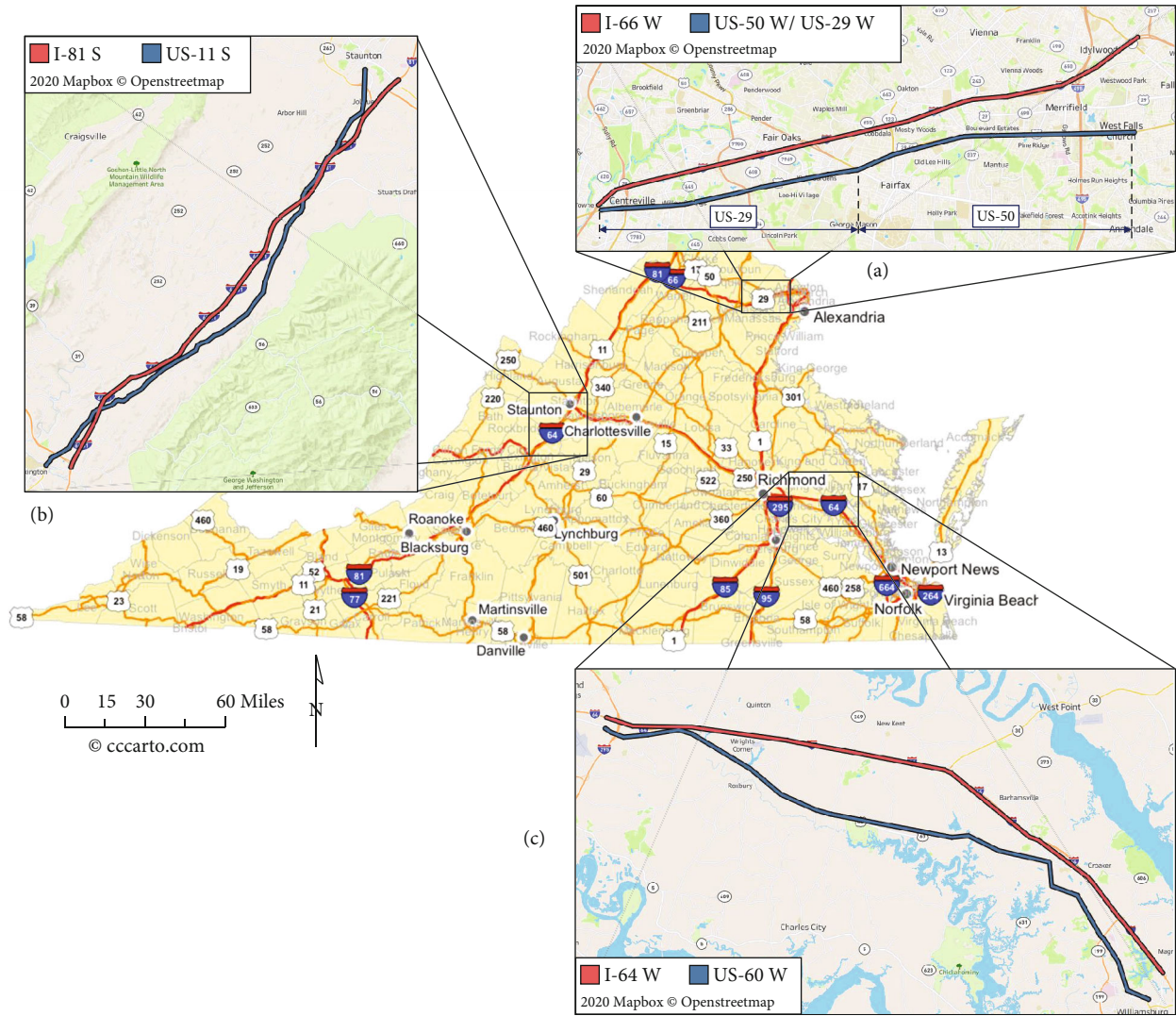


FIGURE 4: Case Study Locations across the Commonwealth of Virginia (Developed Using Tools in CHUBB Custom Cartography, [49]): (a) I-66 and US-50 Westbound, (b) I-81 and US-11 Southbound, and (c) I-64 and US-60 Westbound [50].

is 4,930,560, 4,930,560, and 4,314,240 in total for I-66, I-81, and I-64, respectively. Additional data associated with TMC segments 3 miles upstream of the study location start points and 1 TMC segment downstream of the end points were also obtained. These additional data were included to catch event impacts that may persist beyond the study locations. The format of the collected traffic data by TMC segment is shown in Table 1. Table 2 provides a snapshot of additional information from the INRIX dataset that is used in identifying locations associated with events and their impacts.

To minimize random variations and impacts from random errors in the traffic data, the traffic data given in 1-minute increments were smoothed using a weighted and centered moving-average that weights the data according to INRIX-supplied observation confidence scores. Less than 0.5% of the data were missing. Since the missing data were limited to a small number of observations and rarely occurred over consecutive periods of any extended length, the missing data were replaced by the mean values for prior and subsequent time increments. Such instances of missing

data were assumed to result from temporary failures arising within the TMC system. An approach that imputes missing values from their nearest neighbor values while considering time, location, hour of the day, weekend/weekday, and annual average daily traffic was tested, but resulted in replacement values that were inconsistent with prior and subsequent data values. To obtain speed profiles under recurrent conditions and further calculate unit delays, Recurrent Speed Profiles (RSPs) were computed through methods presented in [54].

**4.1.2. Event Data.** An event dataset was created that includes traffic incidents of vehicular collisions only, with subcategories of single-vehicle accidents, multiple-vehicle accidents, and tractor trailer accidents, and improvement actions with work zone details. Improvement actions include as follows: bridge inspection work, guardrail repairs, ITS equipment repairs, paving operations, pothole patching operations, rehabilitation project activities, resurfacing operations, survey work, median/jersey wall repair/installation,

TABLE 1: Traffic data format.

TMC code	Timestamps	Speed (mph)	Historical average speed (mph)	Reference speed (mph)	Travel time (sec)	Confidence score
110P04174	2019-04-01 00:00:00	57	60	60	30.76	30
110P04174	2019-04-01 00:01:00	53	60	60	33.08	30
⋮	⋮	⋮	⋮	⋮	⋮	⋮
110P04175	2019-04-01 09:00:00	65	61	57	52.86	30
⋮	⋮	⋮	⋮	⋮	⋮	⋮

TABLE 2: TMC segments information format.

TMC code	Freeway information		Start coordinates		End coordinates		Length (mile)
	Name	Intersection	Latitude	Longitude	Latitude	Longitude	
110P04174	I-66 W	VA-7/exit 66	38.90079	-77.19357	38.89674	-77.20097	0.487071
110+ 04175	I-66 W	I-495/exit 64	38.89674	-77.20097	38.88978	-77.21302	0.806837
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

TABLE 3: Frequency of events by type for April 1 to October 31 in 2019.

Event type	Case study		
	I-66 W	I-81 S	I-64 W
Collisions	516	72	220
Improvements	729	104	151
Total	1245	176	371

new roadway construction, and road widening project activities. The events' data used in the case studies were retrieved from a nation-wide traffic event data archiving system Regional Integrated Transportation Information System (RITIS). The Virginia Department of Transportation (VDOT) manages information about work zones, incidents and other types of traffic disruptions arising in Virginia with the aim of providing travelers with traffic information and facilitating roadway services. Table 3 synthesizes this database of studied collisions and improvement actions for the study locations during the study period. The events' data include basic information related to timing to the minute and location in terms of geographic coordinates, duration, maximum number of lanes closed, and brief text with description.

#### 4.2. Data Preparation

**4.2.1. Relating Traffic Event Occurrence to Traffic Performance.** The collected traffic data from INRIX was sorted to create a time-ordered and TMC-segment sequenced set of records. In an event occurrence, upstream traffic will often travel at slower speeds while downstream speeds may be higher in comparison. Other traffic characteristics may be similarly affected up and downstream of the event. To investigate such possible

changes over the course of an event's impact, 17 additional traffic-related variables were developed. 16 relate to traffic characteristics and the remaining is associated with time. These variables build on the speed profiles of TMC segments and relate either to traffic or the event impact's temporal characteristics. A reduction in the speed ratio (SpR) can be expected at the time and location of an event. At the very moment of the event, the average speed ratio in both upstream and downstream segments is expected to be greater than that in the location of the event. At later time increments, as delays propagate, this will not be the case. Observations of the dynamics aid in event detection.

Several of the listed traffic-related variables are defined based on a metric for travel time reliability, Extra Buffer Time Index (EBTI), previously introduced by Tavassoli Hojati et al. [55]. In general, extra buffer time is defined as the extra delay caused by an event. It indicates the extra travel time needed to arrive at a destination on time with 95% certainty in a traffic event. Details of EBTI calculation can be found in [55].

Temporal characteristic, including peak hour and weekend indicators, acts as a dummy variable. Table 4 consolidates the potential explanatory variables, including spatiotemporal traffic characteristics and time-related variables, for event detection. The response variable to these explanatory variables is the event class for a given location-time pair  $(l, t)$ .

**4.2.2. Balancing the Data.** It is reasonable to expect that for any location, there will be a disproportionate number of location-time pairs in which no event (neither traffic incident nor improvement activity) has occurred. This leads to an imbalance in data supporting to the four detection outcomes (no event, traffic incident, improvement activity, or both traffic incident and improvement activity). Consider



TABLE 4: Potential explanatory variables (at 1-minute increment for a given TMC segment) and their definitions.

Variable	Description/value
Traffic characteristics	
<i>Speed-related</i>	
SpR <sub>(t, l)</sub>	Ratio of speed at time $t$ in segment $l$ to recurrent speed at the same time and location.
SdR <sub>(t-15: t+15, l)</sub>	Standard deviation of SpR from 15 mins before to 15 mins after time $t$ in segment $l$ .
UpSdR <sub>(t-15: t+15, l+1, l)</sub>	Ratio of SdR in upstream segment (US) at the same increment of time.
DownSdR <sub>(t-15: t+15, l-1, l)</sub>	Ratio of SdR in downstream segment (DS) at the same increment of time.
AvgR <sub>(t: t+15, l)</sub>	Average of consecutive SpRs from time $t$ to 15 minutes after in segment $l$ .
UpAvgR <sub>(t: t+15, l+1)</sub>	AvgR of US within same increment of time.
DownAvgR <sub>(t: t+15, l-1)</sub>	AvgR of DS within same increment of time.
U/CAvgR <sub>(t: t+15, l+1, l)</sub>	Ratio of AvgR in US to chosen segment at the same increment of time.
D/CAvgR <sub>(t: t+15, l-1, l)</sub>	Ratio of AvgR in DS to chosen segment at the same increment of time.
D/UAvgR <sub>(t: t+15, l-1, l+1)</sub>	Ratio of AvgR in DS to US at the same increment of time.
<i>Travel time-related</i>	
EBTI <sub>(t: t+15, l)</sub>	EBTI of consecutive timestamps from time $t$ to 15 minutes after in segment $l$ .
UpEBTI <sub>(t: t+15, l+1)</sub>	EBTI of US within same increment of time.
DownEBTI <sub>(t: t+15, l-1)</sub>	EBTI of DS within same increment of time.
U/CEBTI <sub>(t: t+15, l+1)</sub>	Ratio of EBTI in US to chosen segment at the same increment of time.
D/CEBTI <sub>(t: t+15, l-1)</sub>	Ratio of EBTI in DS to chosen segment at the same increment of time.
D/UEBTI <sub>(t: t+15, l-1, l+1)</sub>	Ratio of EBTI in DS to US at the same increment of time.
Temporal characteristic	
Peak	1 = weekday peak period (5:30 – 9:30, 15:00 – 19:00); 0 = otherwise.

TABLE 5: Number of disruptions by type for case study locations.

Case	Event type	Imbalanced training	Imbalanced test	Imbalanced Total	Balanced training
I-66 W	Collision	12,391 (0.45%)	18,348 (0.85%)	30,739 (0.62%)	139,934 (14.01%)
	Improvements	188,979 (6.81%)	77,751 (3.60%)	266,730 (5.41%)	188,979 (18.92%)
	Both	584 (0.02%)	685 (0.03%)	1,269 (0.03%)	70,561 (7.06%)
	None	2,571,486 (92.72%)	2,060,336 (95.51%)	4,631,822 (93.94%)	599,386 (60.01%)
	Overall	2,773,440	2,157,120	4,930,560	998,861
I-81 S	Collision	1,846 (0.07%)	2,279 (0.09%)	4,125 (0.08%)	141,802 (16.78%)
	Improvements	36,859 (1.50%)	33,707 (1.37%)	70,566 (1.43%)	143,867 (17.02%)
	Both	76 (0.003%)	29 (0.001%)	105 (0.002%)	56,967 (6.74%)
	None	2,426,499 (98.43%)	2,429,265 (98.54%)	4,855,764 (98.48%)	502,596 (59.46%)
	Overall	2,465,280	2,465,280	4,930,560	845,232
I-64 W	Collision	25,171 (1.17%)	3,854 (0.18%)	29,025 (0.67%)	138,654 (16.15%)
	Improvements	358,359 (16.61%)	355,936 (16.50%)	714,295 (16.56%)	142,859 (16.64%)
	Both	9177 (0.43%)	2,573 (0.12%)	11,750 (0.27%)	66,654 (7.76%)
	None	1,764,413 (81.79%)	1,794,757 (83.20%)	3,559,170 (82.50%)	510,546 (59.45%)
	Overall	2,157,120	2,157,120	4,314,240	858,713

the case study on the I-66 W corridor. With its 4,930,560 data records, only 231,711 data records, 4.7% of the data set, relate to an event. This data imbalance can lead to issues of bias, misclassification, or low accuracy/sensitivity in detection outcomes.

To avoid issues of low sensitivity or accuracy, methods for balancing the data are often applied. In works in the lit-

erature, oversampling and undersampling techniques were used to balance the data by duplicating instances of the minority class and removing random instances of the majority class, respectively. A downside to these techniques is that they are known to cause issues of over- or underfitting to the data. To address this concern, [56] introduced Synthetic Minority Oversampling Technique (SMOTE). SMOTE

TABLE 6: Confusion matrix of an event type detection method.

	Event type	Predicted				Actual number of instances
		1	2	...	L	
Actual	1	$p_{11}^1$	$p_{12}$	...	$p_{1L}$	$N_1$
		$r_{11}^2$	$r_{12}$		$r_{1L}$	
	2	$p_{21}$	$p_{22}$	...	$p_{2L}$	$N_2$
		$r_{21}$	$r_{22}$		$r_{2L}$	
	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
L	$p_{L1}$	$p_{L2}$	...	$p_{LL}$	$N_L$	
	$r_{L1}$	$r_{L2}$		$r_{LL}$		

generates synthetic data records of the minority class for the K-Nearest Neighbors of a minority sample.

In this work, the dataset is split into training (including 2,773,440 records for 9 continuous segments of I-66 W, 2,465,280 records for 8 continuous segments of I-81 S, and 2,465,280 records for 7 continuous segments of I-64 W) and test (including 2,157,120 records for 7 continuous segments of I-66 W, 2,465,280 records for 8 continuous segments of I-81 S, and 2,465,280 records for 7 continuous segments of I-64 W) datasets. Parameters of the machine learning techniques are set on the training data and then applied within the machine learning method in the analysis completed on the test data. Table 5 provides a count of the event records and final number after applying SMOTE by event type for each of the three case study locations. Event records are between 0.01 and 7% of the records in the original training datasets. Records from the three event classes are considered as minority classes, while nonevent records form the majority class.

A mixed technique of adding synthetic minority records and removing majority records was applied to achieve the best compromise between data size and algorithmic performance. First, SMOTE with  $k = 5$  was applied to the training datasets to create synthetic minority records for inclusion. The application of SMOTE was completed separately for each minority class. Second, a random set of the majority class (nonevent) records was removed producing a roughly 1:4 ratio of event to nonevent records. This ratio is considered a golden standard and, thus, follows standard methods from case-control studies [57]. The parameters of the event detection models were calibrated on the balanced training dataset.

## 5. Results and Discussion

The importance of traffic features to the accurate detection of event occurrence by type is analyzed in this section. Event-based unit delays at facility- and corridor-levels were estimated for each of the case study locations. Commonalities and differences between improvement actions and traffic incidents can be discerned from event detection model results. Factors affecting these unit delays were identified through multivariate statistical analysis. Before proceeding to apply the detection methods, their performance on the case study locations is analyzed.

TABLE 7: Overall accuracy rates.

Case study	Detection method			
	RF	RF + KNN	SVM	SVM + KNN
I-66 W	95.38%	96.52%	80.95%	81.21%
I-81 S	98.42%	99.09%	87.68%	88.04%
I-64 W	90.17%	91.02%	86.74%	86.96%

To evaluate the performance of the SVM, RF, SVM-KNN, and RF-KNN methods in classifying each  $\Theta(t_m, l)$ , each data element for each time interval  $t_m$  and location  $l$  as being associated with one of the four event (or nonevent) types, a confusion matrix is created from the results of each method that includes classification errors by event type, and an overall prediction accuracy is computed from the matrix. The confusion matrix concept was first introduced in [58] and is illustrated for this application in Table 6. In the confusion matrix,  $p_{ij}$  gives the number of instances of event type  $i$  that are predicted to be of event type  $j$  and  $r_{ij}$  gives the ratio of instances of event type  $i$  that were predicted to be of event type  $j$  to the total number of instances of type  $i$ ,  $N_i$ :  $r_{ij} = p_{ij}/N_i$ . Thus,  $r_{ii}$  is known as the detection rate for type  $i$ . The overall predictive capability of a method for a given application then is determined by  $Acc_{total} = \sum_{i=1}^L p_{ii} / \sum_{i=1}^L N_i$ .

Either of the developed hybrid techniques can be designed around various input traffic characteristics. Those characteristics that contribute in the method's specification with the greatest accuracy provide additional insights into which characteristics from the data contribute most to the event type. Different characteristics may contribute more to one or another event type.

In applying either proposed hybrid method, some improvement actions were not detected. The occurrence of false positive and negative findings, in part, can be explained by the differences in the effects of improvement actions on traffic performance as a function of lanes blocked, duration of implementation, and other factors. Understanding which features make an activity less detectable can provide further understanding.

**5.1. Event Detection Method Performance.** The proposed detection methods were executed on balanced training datasets. The default value of  $m_{try}$  needed for the RF-based procedures suggested in the literature can be obtained from the square root of the number of explanatory variables in the training data set [59]. Larger values of  $m_{try}$ , though, have been found to lead to better performance (e.g., [46, 60]), but larger values result in greater computation time. Runs with  $m_{try}$  between 3 and 8 were completed as 8 variables were chosen for inclusion. It was found that an  $m_{try}$  of 6 performed best. For the SVM-based methods, parameters  $C$  (set to 10 for all locations) and  $\gamma$  (set to 0.6 for locations I and III and 1 for location II) were set to achieve the highest average 10-fold cross validation accuracy for all locations. Additional parameter settings include  $k$  (for KNN) set to

TABLE 8: Confusion matrix.

Case study	Downtime class	Method	Detected								
			Both	Col. <sup>1</sup>	$P_{ij}$ Imp. <sup>2</sup>	None	Both	Col.	$r_{ij}$ Imp.	None	
Case Study I (I-66 W)	Both <i>N</i> = 685	<i>RF</i>	685	0	0	0	<b>100%</b>	0.00%	0.00%	0.00%	
		<i>RF + KNN</i>	678	0	7	0	<b>98.98%</b>	0.00%	1.02%	0.00%	
		<i>SVM</i>	501	0	97	87	<b>73.14%</b>	0.00%	14.16%	12.70%	
		<i>SVM + KNN</i>	502	0	98	85	<b>73.28%</b>	0.00%	14.31%	12.41%	
	Collision <i>N</i> = 18348	<i>RF</i>	7	14944	275	3122	0.04%	<b>81.45%</b>	1.50%	17.02%	
		<i>RF + KNN</i>	0	16237	88	2023	0.00%	<b>88.49%</b>	0.48%	11.03%	
		<i>SVM</i>	386	9073	3068	5821	2.10%	<b>49.45%</b>	16.72%	31.73%	
		<i>SVM + KNN</i>	363	9140	3002	5843	1.98%	<b>49.81%</b>	16.36%	31.85%	
	Imp. <i>N</i> = 77751	<i>RF</i>	8	39	53282	24422	0.01%	0.05%	<b>68.53%</b>	31.41%	
		<i>RF + KNN</i>	23	5	56882	20841	0.03%	0.01%	<b>73.16%</b>	26.80%	
		<i>SVM</i>	1368	1669	28986	45728	1.76%	2.15%	<b>37.28%</b>	58.81%	
		<i>SVM + KNN</i>	1291	1626	29043	45791	1.66%	2.09%	<b>37.35%</b>	58.89%	
	None <i>N</i> = 2060336	<i>RF</i>	177	7129	64565	1988465	0.01%	0.35%	3.13%	<b>96.51%</b>	
		<i>RF + KNN</i>	82	5522	46520	2008212	0.00%	0.27%	2.26%	<b>97.47%</b>	
		<i>SVM</i>	11379	78985	262285	1707687	0.55%	3.83%	12.73%	<b>82.88%</b>	
		<i>SVM + KNN</i>	10700	78822	257779	1713034	0.52%	3.83%	12.51%	<b>83.14%</b>	
	Case Study II (I-81 S)	Both <i>N</i> = 29*	<i>RF</i>	0	0	0	29	<b>0.00%</b>	0.00%	0.00%	100.00%
			<i>RF + KNN</i>	0	0	0	29	<b>0.00%</b>	0.00%	0.00%	100.00%
<i>SVM</i>			0	0	0	29	<b>0.00%</b>	0.00%	0.00%	100.00%	
<i>SVM + KNN</i>			0	0	0	29	<b>0.00%</b>	0.00%	0.00%	100.00%	
Collision <i>N</i> = 2279		<i>RF</i>	0	1898	2	379	0.00%	<b>83.28%</b>	0.09%	16.63%	
		<i>RF + KNN</i>	0	1962	0	317	0.00%	<b>86.09%</b>	0.00%	13.91%	
		<i>SVM</i>	0	1390	111	778	0.00%	<b>60.99%</b>	4.87%	34.14%	
		<i>SVM + KNN</i>	0	1389	109	781	0.00%	<b>60.95%</b>	4.78%	34.27%	
Imp. <i>N</i> = 33707		<i>RF</i>	0	7	19324	14376	0.00%	0.02%	<b>57.33%</b>	42.65%	
		<i>RF + KNN</i>	0	0	19901	13806	0.00%	0.00%	<b>59.04%</b>	40.96%	
		<i>SVM</i>	259	69	16729	16650	0.77%	0.20%	<b>49.63%</b>	49.40%	
		<i>SVM + KNN</i>	244	61	16710	16692	0.72%	0.18%	<b>49.57%</b>	49.52%	
None <i>N</i> = 2429265		<i>RF</i>	139	1025	22987	2405114	0.01%	0.04%	0.95%	<b>99.01%</b>	
		<i>RF + KNN</i>	19	603	7654	2420989	0.00%	0.02%	0.32%	<b>99.66%</b>	
		<i>SVM</i>	40598	6096	239237	2143334	1.67%	0.25%	9.85%	<b>88.23%</b>	
		<i>SVM + KNN</i>	37883	5846	233229	2152307	1.56%	0.24%	9.60%	<b>88.60%</b>	
Case Study III (I-66 W)		Both <i>N</i> = 2573	<i>RF</i>	2295	7	64	207	<b>89.20%</b>	0.27%	2.49%	8.05%
			<i>RF + KNN</i>	2381	0	53	139	<b>92.54%</b>	0.00%	2.06%	5.40%
	<i>SVM</i>		1998	90	47	438	<b>77.65%</b>	3.50%	1.83%	17.02%	
	<i>SVM + KNN</i>		2035	80	45	413	<b>79.09%</b>	3.11%	1.75%	16.05%	
	Collision <i>N</i> = 3854	<i>RF</i>	3	2985	52	814	0.08%	<b>77.45%</b>	1.35%	21.12%	
		<i>RF + KNN</i>	0	3117	29	708	0.00%	<b>80.88%</b>	0.75%	18.37%	
		<i>SVM</i>	20	1830	40	1964	0.52%	<b>47.48%</b>	1.04%	50.96%	
		<i>SVM + KNN</i>	5	1856	40	1953	0.13%	<b>48.16%</b>	1.04%	50.67%	
	Imp. <i>N</i> = 355936	<i>RF</i>	131	124	188290	167391	0.04%	0.03%	<b>52.90%</b>	47.03%	
		<i>RF + KNN</i>	160	272	202150	153354	0.04%	0.08%	<b>56.79%</b>	43.08%	
		<i>SVM</i>	2201	10409	136838	206488	0.62%	2.92%	<b>38.44%</b>	58.01%	

TABLE 8: Continued.

Case study	Downtime class	Method	Detected							
			Both	Col. <sup>1</sup>	$P_{ij}$ Imp. <sup>2</sup>	None	Both	Col.	$r_{ij}$ Imp.	None
None $N = 1794757$		SVM + KNN	2005	9764	137300	206867	0.56%	2.74%	<b>38.57%</b>	58.12%
		RF	392	1546	55268	1737551	0.02%	0.09%	3.08%	<b>96.81%</b>
		RF + KNN	237	762	38043	1755715	0.01%	0.04%	2.12%	<b>97.82%</b>
		SVM	3438	42634	18274	1730411	0.19%	2.38%	1.02%	<b>96.41%</b>
		SVM + KNN	2959	39600	17469	1734729	0.16%	2.21%	0.97%	<b>96.66%</b>

<sup>1</sup>Collisions, <sup>2</sup>Improvement actions. \*Note that in Case Study II, there is only sparse data for the both category with 29 time-space instances.

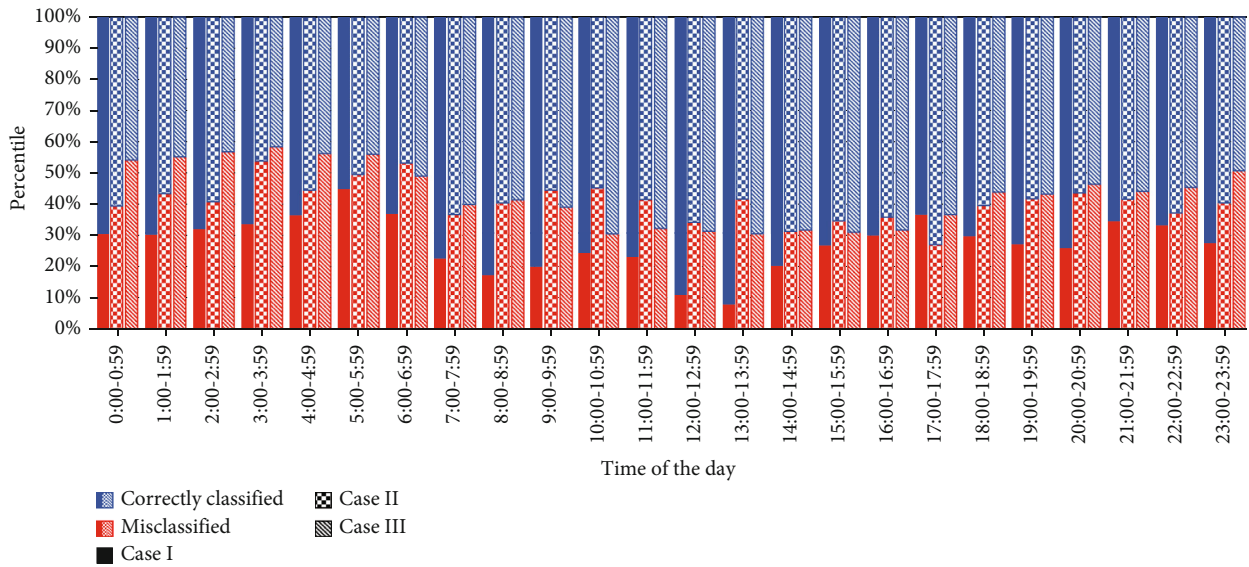


FIGURE 5: Correctly classified vs. misclassified improvement actions when conditioned on time of day for each case study location.

10 in all runs and RF's  $n_{tree}$  at 200 for location I and II and 150 for location II.

Table 7 reports detection accuracies. The accuracy of all tested methods was very high with the best performing method ranging from approximately 91 to 99%. Results of the event detection methods are reported in a confusion matrix (Table 8) giving prediction counts and detection rates for each event type and case study location with parameter settings as identified earlier. Incorporating KNN to create the hybrid approaches increased detection rates by as high as 7 percentage points (81.5 to 88.5%), and, thus, decreased misclassifications similarly. The KNN addition also improved the method's accuracy for all case studies and nearly all event types when added to both SVM and RF methods. Generally, the RF-based approaches outperformed the SVM-based methods with and without incorporating KNN.

Note that the parameters applied in Case Studies I and III are identical and differ only minimally from those used in Case Study II. If these parameters are used on Case Study location II, the results diminish in accuracy only minimally.

5.1.1. *Misclassified Improvement Event Instances.* Between 26 (Case Study I) and 41% (Case Study III) of improvement

actions were misclassified as nonevents. These results were further analyzed by conditioning on time of day (Figure 5), maximum number of lanes closed (Figure 6) and improvement type (Figure 6). From Figure 5, it can be noted that improvement actions were most likely to be misclassified when they were executed in off-peak hours. In Case Study I, the misclassification rate when aggregated over the day is 27%, however, when conditioned on the time of day, this value ranges between 8% and 45%, 45% occurring in off-peak hours. The variability over the day in misclassification rates is not as severe in the other two, more rural case study locations.

The procedures are generally able to distinguish between improvement actions and traffic incidents, implying that impact traffic differently. However, the ability to distinguish between this event types may be because improvement actions tend to be executed in off-peak hours and traffic incidents are more likely to occur in peak hours. Thus, the increased detection rates of improvement actions in off-peak hours may be from taking advantage of knowledge of the time of day in which the event occurs. That is, if the event arises in the off-peak hours, it is best to guess that it is an improvement event. Conditioning on time of day, density plots and means of AvgR (Figure 7) and EBTI (Figure 8)



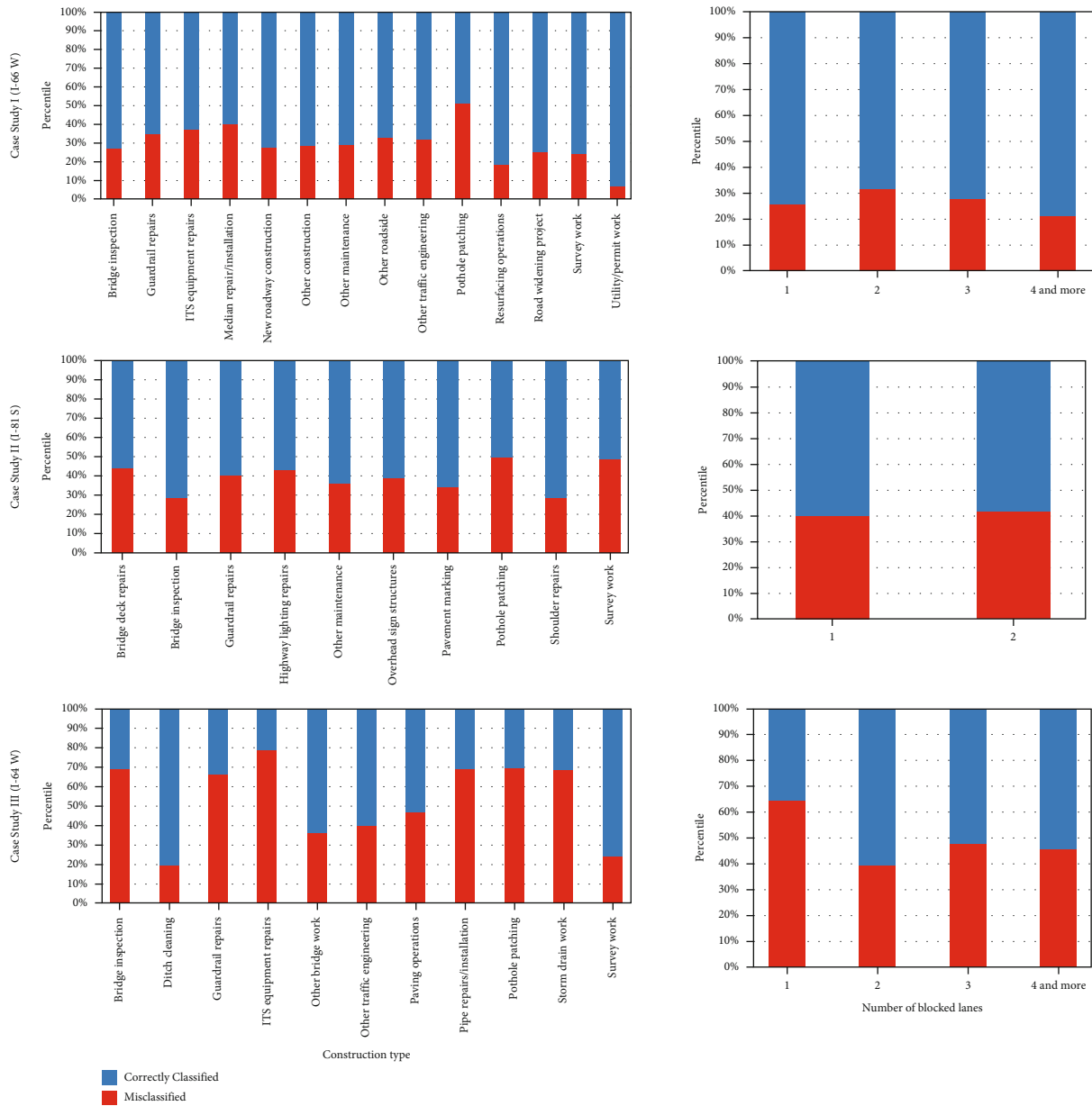


FIGURE 6: Improvement detection rates conditioned on types and number of blocked lanes.

traffic performance metrics were further studied. Note that density gives the probability scaled by bin width. This is proceeded by investigation into the travel time delays from each event type. Statistics of AvgR and EBTI conditioned on type and time of occurrence are summarized in Table 9.

Sharp peak values around a ratio of 1 for AvgR and very low EBTI values with some difference in AvgR, but no significant difference in EBTI between peak and off-peak time periods, were observed. This indicates that improvement actions have little impact on travel time reliability at any time of day, but have higher impact on speeds in peak hours. Additionally, when misclassified, regardless of time of day, the improvement activity was presumed to be a nonevent.

The results in Figure 6 from studying improvement detection rates conditioned on type and number of lanes blocked shows that the greater the reduction in roadway

capacity caused by the event, the more likely the event is to be correctly classified. Pothole patching, which typically has low impact on capacity, was further studied by comparing results in Figure 5 with the timing of events as illustrated in Figure 9. These events were misclassified as nonevents in between 50% and 70% of all occurrences for the case studies. The results indicate that time-of-day is not a relevant factor. This finding was useful in designing the unit delay estimation models.

5.1.2. Feature Importance in Event Detection by Event Type.

Features that contributed most to correct event occurrence and type detection for events involving either traffic incidents or improvement actions were identified to understand how improvement actions differ from traffic incidents in terms of their impact on traffic. For this purpose, feature

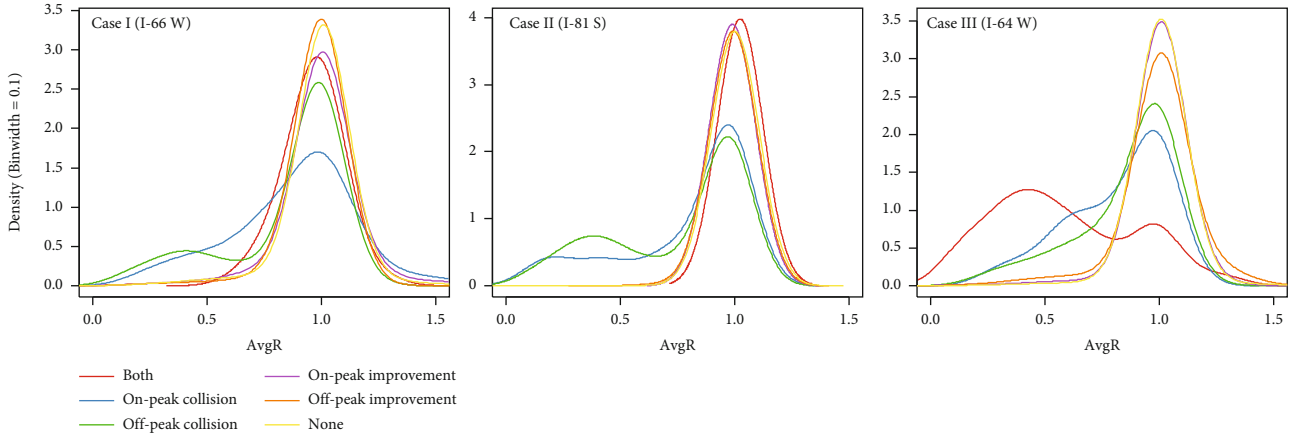


FIGURE 7: AvgR density plots.

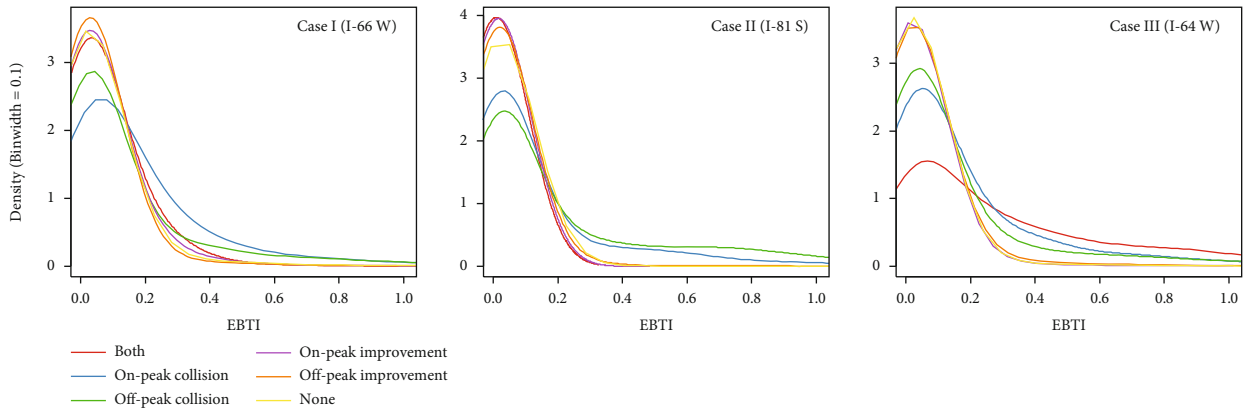


FIGURE 8: EBTI density plots.

TABLE 9: Statics of AvgR and EBTI.

Type	Case studies					
	Case Study I (I-66 W)		Case Study II (I-81 S)		Case Study III (I-64 W)	
	Mean of AvgR <sup>1</sup>	Mean of EBTI <sup>2</sup>	Mean of AvgR	Mean of EBTI	Mean of AvgR	Mean of EBTI
Both	0.9631	0.0648	1.0232	0.0102	0.5856	0.7084
On-peak collision	0.8877	0.2385	0.7975	0.3147	0.8186	0.1984
Off-peak collision	0.8557	0.2606	0.7552	0.2900	0.8596	0.1699
On-peak improvements	0.9362	0.0719	0.9498	0.0199	0.9451	0.0667
Off-peak improvements	1.0004	0.0696	1.0014	0.0202	0.9951	0.0612
None	1.0012	0.0688	1.0003	0.0249	1.0019	0.0515

<sup>1</sup>Average of consecutive speed ratios, <sup>2</sup>Extra Buffer Time Index.

importance scores were computed through a permutation-based approach employed in training the detection portion of the RF-KNN hybrid method. In this approach, a baseline model is created on a given dataset with various traffic features and its accuracy and detection rates are recorded. Then, values from one feature are randomly shuffled, and the modified dataset is passed to the model update accuracy and detection rates. The feature importance scores are computed from the difference between detection rates of the baseline and permuted models. This difference is known as the Mean

Decrease Accuracy (MDA). The higher the MDA value of a traffic feature, the more important is that feature. Table 10 reports the MDA values computed for each of the case study locations and highlights the top 7 features for each event type and case study location.

Eight features were of greatest importance over all locations for improvement and collision event types are as follows: UpSdR, U/CAvgR, D/CAvgR, AvgR, SdR, D/CEBTI, DownAvgR, and UpAvgR. These features were supplied as candidates for inclusion, and various combinations of six

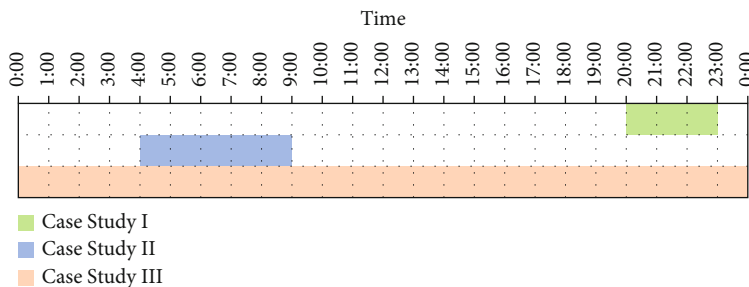


FIGURE 9: Timing of pothole patching actions.

TABLE 10: Feature importance scores from RF method.

Features	MDA in detection of type (percent):											
	Both	Case Study I			Case Study II			Case Study III				
		Coll. <sup>1</sup>	Imp. <sup>2</sup>	None	Both	Coll.	Imp.	None	Both	Coll.	Imp.	None
AvgR	29.01	32.76	8.47	<b>9.86</b>	28.16	<b>46.38</b>	16.38	<b>5.87</b>	<b>55.16</b>	<b>55.09</b>	<b>8.42</b>	<b>7.85</b>
D/CAvgR	<b>54.22</b>	<b>49.61</b>	<b>12.11</b>	<b>9.64</b>	32.53	<b>45.32</b>	<b>19.48</b>	<b>3.58</b>	<b>54.48</b>	<b>53.46</b>	<b>8.89</b>	<b>6.74</b>
D/CEBTI	30.27	17.48	4.10	2.01	<b>47.88</b>	25.77	<b>26.53</b>	<b>3.82</b>	28.14	26.02	2.29	3.10
D/UAvgR	<b>44.79</b>	<b>43.27</b>	<b>8.99</b>	<b>8.75</b>	21.97	<b>36.84</b>	18.85	<b>4.75</b>	<b>35.85</b>	<b>35.67</b>	<b>4.95</b>	<b>5.25</b>
D/UEBTI	18.12	16.28	3.59	1.96	25.32	17.93	9.59	1.83	24.25	25.90	1.57	2.20
DownAvgR	<b>51.03</b>	38.71	7.61	<b>8.32</b>	35.85	23.15	17.87	<b>5.23</b>	<b>38.07</b>	<b>37.73</b>	<b>5.17</b>	<b>6.51</b>
DownEBTI	20.99	21.31	5.28	3.14	<b>36.01</b>	14.72	14.03	2.39	31.23	34.84	2.47	<b>3.56</b>
DownSdR	<b>41.74</b>	<b>40.95</b>	<b>11.74</b>	3.98	<b>58.94</b>	26.74	<b>25.19</b>	1.76	<b>37.76</b>	34.57	3.13	2.88
EBTI	31.58	20.17	4.62	3.63	23.34	19.20	13.39	3.20	26.20	25.37	1.79	3.53
Peak	21.99	27.67	6.33	1.77	14.44	13.92	9.51	0.22	12.26	12.94	1.49	0.35
SdR	<b>45.07</b>	<b>42.08</b>	<b>11.47</b>	<b>7.01</b>	<b>41.66</b>	<b>39.66</b>	<b>31.98</b>	2.25	33.89	30.60	<b>4.09</b>	2.16
SpR	39.72	33.17	7.93	5.72	36.39	<b>46.28</b>	14.44	2.12	<b>45.88</b>	<b>42.01</b>	<b>6.34</b>	3.16
U/CAvgR	<b>45.47</b>	<b>42.82</b>	<b>12.17</b>	<b>10.00</b>	<b>44.40</b>	<b>39.89</b>	<b>26.59</b>	<b>4.55</b>	35.24	<b>36.60</b>	<b>4.80</b>	<b>4.53</b>
U/CEBTI	15.51	15.64	4.69	2.46	25.33	19.22	15.99	3.11	27.76	21.86	1.74	3.29
UpAvgR	36.45	<b>40.73</b>	<b>9.63</b>	<b>9.46</b>	<b>51.32</b>	<b>39.19</b>	<b>20.53</b>	<b>6.08</b>	33.09	<b>35.81</b>	3.58	<b>5.37</b>
UpEBTI	32.30	24.78	5.47	4.22	35.85	20.41	13.27	2.41	26.03	24.53	1.65	2.94
UpSdR	<b>43.77</b>	<b>43.54</b>	<b>13.95</b>	5.37	<b>54.78</b>	29.45	<b>23.59</b>	2.58	<b>36.94</b>	32.44	3.69	2.27

<sup>1</sup>Collision, <sup>2</sup>Improvement actions.

( $m_{try} = 6$ ) of the eight were included in the developed improvement action/collision detection methods.

The results indicate that EBTI-related metrics were not among the top important features for almost any event type and location. Additionally, these metrics were among the least important features for improvement activities. This confirms the earlier claim that improvement actions have little impact on travel time reliability.

Also, of note is that the feature importance score is generally greater (almost five times on average) for all metrics associated with collisions than for those associated with improvement actions. As expected, it was observed that traffic conditions on upstream segments were more important than for downstream segments. On the other hand, the ratio of changes of traffic conditions in downstream segments to changes of traffic conditions in the segment containing the event (D/CAvgR and D/CEBTI) played a more important role than a similar ratio for upstream segments (U/CAvgR and U/CEBTI). This may indicate that the rise in speed downstream of the event location is more significant than

the drop in speed upstream of the event location. This difference was greater for collisions than for improvement activities.

5.2. Unit Delay Estimates. Using traffic data collected by probe vehicles, traffic conditions are assessed to determine if an event has arisen and has impact. An approach proposed in [54] that uses  $k$ -means clustering was employed to classify a location- (roadway segment) time pair as having or not having a significant change in speed as per the speed ratio. Following rules of contiguity, the spatial and temporal extent of change in speed ratio as a consequence of the event is delineated and is referred to as the event’s impact area. Vehicle unit delays, a measure of extra travel time incurred due to a reduction in speed, are computed based on speed differences over the event’s impact area. Specifically, the extra travel time incurred per vehicle along each roadway segment that falls in the delineated impact area is calculated as the difference between the inverse of both average observed and average recurrent speeds multiplied by the length of that

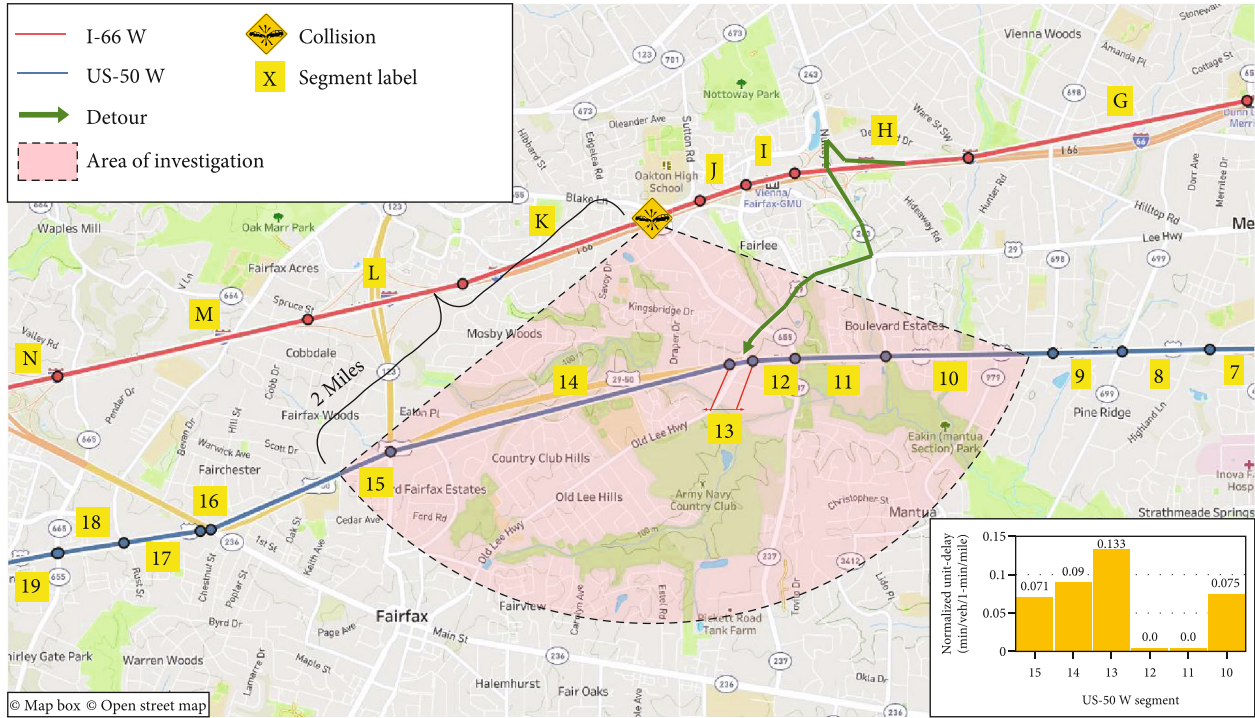


FIGURE 10: The area of investigation of a 2-mile radius, at which the potentially impacted segments of US-50 W are selected.

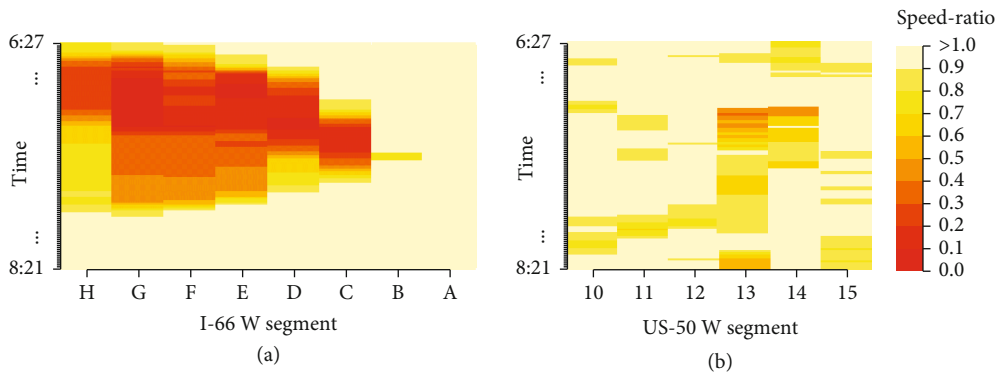


FIGURE 11: Speed ratio matrices for (a) I-66 W and (b) US-50 W for a collision occurred in segment H of I-66 W at 9/4/2019 6:27.

segment. Summing the extra travel time per vehicle incurred over all segments falling within the impact area and dividing by the event’s impact duration gives the vehicle unit delay for the event.

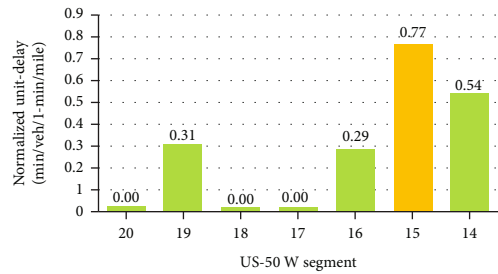
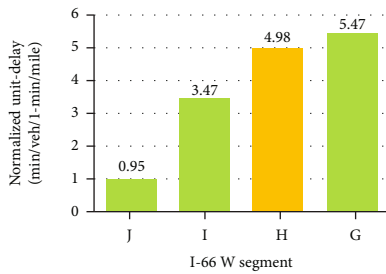
Unit delays were computed for improvement action and collision event types at both facility and corridor levels, as well as by case study location and in the aggregate. Unit delays were also computed for the minor, adjacent facilities to the main study roadways where the events took place. To study the network-wide impacts of an event, unit delay calculations are completed over a more inclusive impact area that incorporates not only the impacted segment(s) of the facility on which the incident or activity arises but portions of a parallel facility and relevant connectors (including on-

and off-ramps) to this facility to which traffic diverts. Thus, the vehicle unit delays are computed across a broader geographical area.

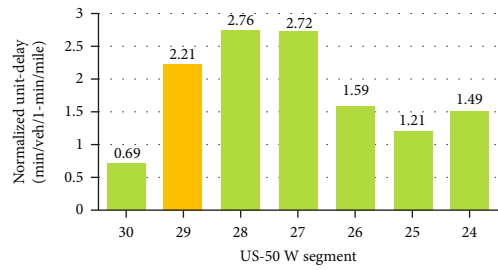
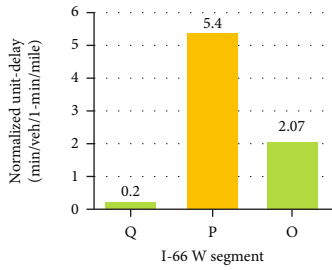
Similar to the speed-ratio matrix developed in [54], a second speed-ratio matrix was developed over observed traffic conditions at time intervals after the event on all potentially impacted segments of the minor facility. In the minor facility, the extent of the event’s impact over space and time was restricted to only those segments of the facility that fall within a 2- or 3-mile radius of the event’s location. The same time intervals as used in studying the event impact on the major facility were applied in studying the impact on the minor facility. Traffic conditions at the on- and off-ramps in affected segments of both facilities were further analyzed.



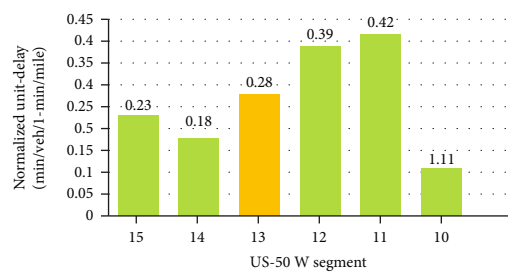
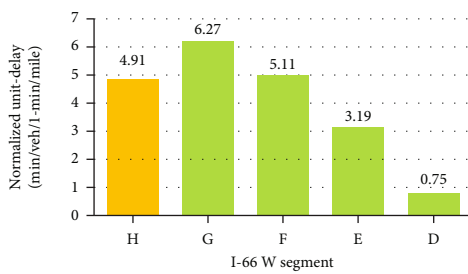
Event: paving operations -time: 2019-09-24 19:08:00 -location: I-66 W at mile marker 49.10



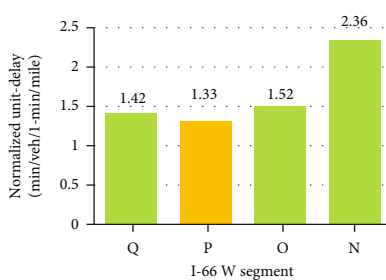
Event: tractor trailer accident -time: 2019-08-24 11:05:00 -location: I-66 w at mile marker 53.00



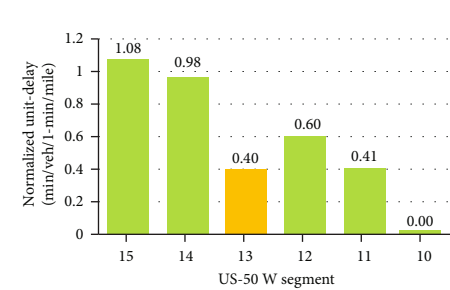
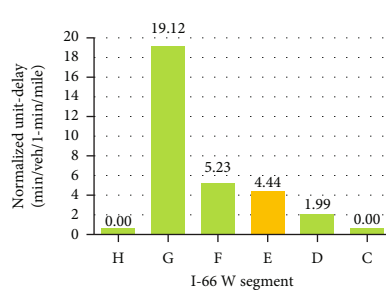
Event: multi-vehicle accident -time: 2019-10-12 17:41:00 -location: I-66 w at mile marker 60.90



Event: multi-vehicle accident -time: 2019-10-25 16:36:00 -location: I-66 w at mile marker 53.20



Event: road widening project activities -time: 2019-09-25 09:00:00 -location: I-66 w at mile marker 61.60



■ Intersection segment  
■ Other segment

FIGURE 12: Normalized unit-delays of impacted segments in Case Study I.

TABLE 11: Facility and Corridor Level Unit Delays by Event Type in the Aggregate and Across Case Studies.

	Event type	Frequency	Mean of unit-delays (mins/Veh/1-min event duration) at:		
			Main facility	Minor facility (2-mile)	Minor facility (3-mile)
Case I	Collision	516	1.677	0.948	1.209
	Improvements	729	0.439	0.349	0.524
Case II	Collision	72	2.808	0.291	0.322
	Improvements	104	0.092	0.089	0.115
Case III	Collision	220	3.099	0.179	0.221
	Improvements	151	0.417	0.146	0.161
Agg*	Collision	808	2.165	0.680	0.861
	Improvements	984	0.399	0.293	0.428

\*Aggregated (via weighted average).

Since the lengths of the segments that contain the ramps are unequal, unit-delays normalized to segment length are reported.

Figure 10 shows the process of identifying the potentially affected segments of US-50 W (minor facility of Case Study I). Segments 10 to 15 of US-50 W fall within the extent boundary (i.e. the 2-mile radius). Segment H of I-66 W and 13 of US-50 W contain interchanges that fall within the bounded area. It is hypothesized that traffic may detour, leaving the major facility (I-66 W) at segment H and entering the minor facility (US-50 W) at segment 13. Normalized unit-delays of each minor facility (US-50 W) segment are also plotted in this figure. Speed ratio matrices associated with a single collision event occurring on 9/4/2019 at 6:27:00 are given in Figure 11 for both major and minor facilities.

Normalized unit-delays associated with impacted segments (in both major and minor facilities) for a randomly selected set of events arising in Case Study I are reported in the plots of Figure 12. This case study was selected for this additional analysis, because the roadway segments were small enough to be able to investigate traffic changes near the intersections. From this additional analysis, it was found that higher normalized unit-delays in collision events occurred along the upstream segments of the segment containing the intersection connecting the minor facility (US-50 W) to the major facility (I-66 W). This supports the hypothesis that drivers are diverting to an alternative roadway in collision events. The same was not noted for improvement actions.

The results indicate that major and minor facility unit delays were greater for collisions than improvement actions. Table 11 and Figure 13 show the statistics and histograms of major and minor facility unit-delays, respectively. Figure 13 further indicates that approximately 60% of the improvement actions have almost no impact (near zero unit delay) at the facility level while 30% of collisions had similar limited traffic impact. By comparing aggregate mean unit delays on the facility itself against the delays in the adjacent facility, the relative difference between the roadways for collisions is less than that of improvement actions. This is reasonable as drivers may choose their routes with foreknowledge of improvement actions. This indicates the importance for studies of improvement activity impact on roadway perfor-

mance to investigate the impacts beyond the facility in which the activity is executed.

Broadening the area of investigation from a radius of 2 miles to one of 3 miles produced greater values of unit-delays with an increase of 26% for collisions and 46% for improvement actions. It may be that the wider radius catches more of the delays caused by the event on the major road, but it may also catch delays from other events occurring on the minor roadway or other roadways in the larger roadway network. Additionally, speed data may not be as accurate along secondary roadways, causing potential inaccuracies in the computations. Further investigation on the appropriate radius to use might consider dependencies on roadway geometry (e.g. number of on- and off-ramps) and event characteristics (e.g. severity of a collision).

To estimate the specific impact of the event type on traffic performance, an analysis of the relationship between event features (e.g. number of blocked lanes, time of day) and resulting delays at the facility level was completed.

Censored multivariate regression models, specifically Tobit regression [61], for estimating facility level unit-delays (minutes/vehicle per 1-minute interval of event duration) for improvement actions and collisions were developed to explore the significance of chosen independent variables, such as number of lanes blocked, event duration, time-of-day, percentage of trucks, and day-of-week. These variables were chosen to use insights gained from application of the event detection methods, results of which highlighted the most important features. Independent variables of significance included roadway and traffic characteristics (Annual average daily traffic (AADT), K-factor, and number of lanes) and event features (e.g. number of blocked lanes, event type, and time-of-day). A difference in parameters of these models and which independent variables are significant for traffic incidents versus improvement actions provides important additional insight into improvement action impacts. Results of the Tobit regression are reported in Table 12.

With these equations, vehicle unit delays can be estimated for a facility with similar characteristics. These equations can also be used for future planning for the same or other roadway segments with similar traffic characteristics.

Regression model coefficients indicate that for each additional lane blocked, the average change in mean unit-delay from a collision or improvement action is 1.497 (a 67%

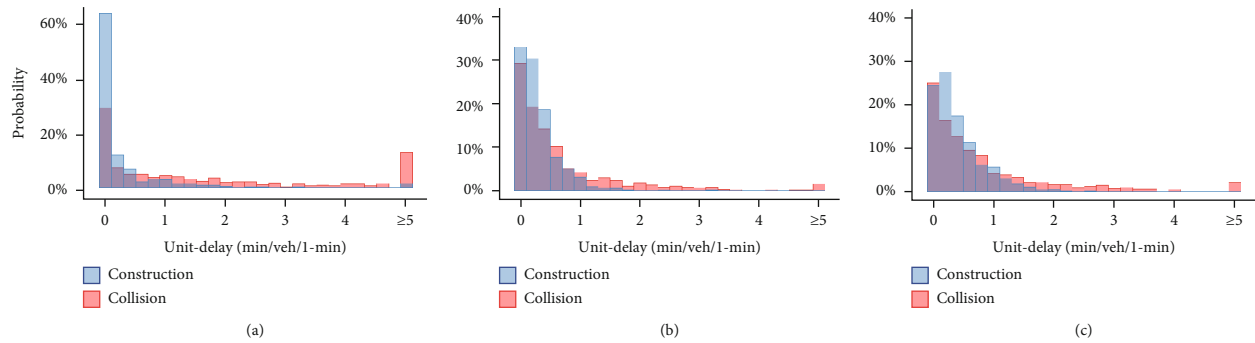


FIGURE 13: Histograms of unit-delays by event type for (a) main facility, (b) minor facility within a 2-mile radius, and (c) minor facility within a 3-mile radius bin width = 0.2.

TABLE 12: Collision and improvement actions unit delay models: coefficients from estimated tobit regression analysis (minutes/vehicle/1-minute event duration).

Variable **+	Description	Coefficient of parameter (Z value)	
		Collision	Improvements
Constant		0.778 (1.282)	-0.364 (-0.357)
NofLanes	Roadway number of lanes (2-4)	-0.661 (-4.082)	0
NofLBlocked	Number of lanes blocked (1-4)	1.497 (10.391)	0.067 (2.643)
Peak	Event occurs at least in part in peak hour 1 = yes; 0 = no	0.797 (2.741)	0.282 (2.297)
AADT/NofL	AADT divided by NofLanes	0	5.55 e-05 (2.083)
K-factor	Portion of total daily traffic volume during peak/design hour	0	-12.55 (-2.125)
MultiVehicle	Multi-vehicle collision event 1 = yes; 0 = no	1.149 (3.153)	NA

\*Statistically insignificant variables tested: distance from closest central business district, tractor trailer, type of improvements, % truck traffic. +All included variables were significant at  $\geq 95\%$ . For collisions:  $R^2 = 0.63$ . For improvements:  $R^2 = 0.48$ .

increase) or 0.067 (about a 17% increase), respectively. That is, each additional lane blocked in a collision has four times the percentage increase of an additional lane blocked in an improvement event. An improvement action event occurring in the peak hour has twice the percentage increase in average unit delay as that of a collision. Moreover, the impacts of AADT and  $K$ -factor are significant only for improvement actions. Each 10,000 vehicles per day increase in AADT (for the same number of lanes) will increase the unit delay for improvement actions by 0.55 minutes on average, while a 1% increase in the  $K$ -factor decreases unit delays of improvement actions by 0.15 minutes on average. Moreover, it was found that if multiple vehicles were involved in a collision, an increase in unit-delay by 1.149 minutes on average (a nearly 55% increase) can be expected. Given a typical road as in case study 1 (AADT = 83000,  $K$ -factor = 0.0789, NofLanes = 4), unit delay for an improvement activity executed in the peak period and blocking one lane is estimated to be 0.146 minutes per vehicle per minute of impact duration.

## 6. Conclusions

This investigation culminated in the creation of mathematical tools, insights, and outcomes. Machine learning techniques (SVM, RF, and hybrid methods SVM-KNN and RF-KNN) were introduced to detect the occurrence of both

improvement actions and traffic incidents for use both off- and on-line that rely on only widely available data. With these techniques, system operators can take action to divert traffic and/or provide appropriate services to clear an event in a timely manner when events are detected in real-time. These techniques can be extended to detect multiple, simultaneous events from one or more event type. The development of algorithms for detecting improvement activities during their execution also serves as a reverse engineering approach for better understanding the impacts of these activities on traffic. Employing these methods on the case studies revealed similarities and differences in the impacts of these activities with the impacts of traffic incidents. This deeper understanding of the effects of downtime from roadway improvement activity execution can be useful in optimal activity scheduling.

A secondary contribution to the literature on traffic event (collision and improvement action) detection is also made through the proposed hybrid learning-based methods that rely on readily available speed data. The hybrid learning methods were shown to reduce misclassification errors as compared with single-phase machine learning methods of SVM and RF. Thus, they are useful tools in detecting and distinguishing between collisions and improvement activities.

Unit-delay estimates can be used to plan and prioritize improvement actions. They can also be used in scheduling

models for prioritizing improvement actions. Findings of this work can also assist with comparison of maintenance and rehabilitation options across different locations considering their characteristics (e.g. K-factor and AADT). By plugging in the K-factor, AADT, and number of lanes of a roadway, the unit delay incurred for improvement action options can be assessed and compared by also plugging in the corresponding lanes blocked by the action and time of day the action is to be executed. The total delay of an action that, for example, closes 3 lanes for 2 days can be compared to another action that closes 1 lane for 6 days. The relative impact of these actions may also differ on other roadways with different K-factors and/or AADT values.

Application of these approaches on three case study locations led to numerous insights. In terms of the solution methodologies, it was found that the addition of KNN to create the RF-KNN method improved the sensitivity by up to 7 percentage points over the single-phase RF method. Improvement actions and collisions were detected 74 and 88 percent of the time, respectively. Improvement actions did not significantly impact travel time reliability, even if they are executed in peak periods, yet collisions impact travel time reliability by up to 32%. Pothole patching activities were found to be the least detectable type of improvement action, while other pavement-related operations, along with utility works, were found to be the most detectable types of the improvement actions. This is important, because the easier to detect, the greater the impact. The impact of events on traffic performance was estimated in terms of unit delays. Unit delay estimates indicate that more than 60% of improvement actions and more than 30% of collisions have no impact on traffic. Paving operations and multivehicle accidents were found to have the highest unit delays associated with improvement actions and collisions, respectively. Future work might examine additional variables, including weather and sight distance, in unit delay computation. The proposed RF-KNN detection model might be compared with other state-of-the-art anomaly detection models in future studies.

Network-level analysis uncovered greater tendency for roadway switching in collision events than during improvement activities. Consistent with this, facility- and network-level unit delays are closer in value for improvement actions than for collisions. It is likely that this occurs because drivers will more likely have foreknowledge of improvement activities than of collisions.

Based on feature importance scores from event detection models, as well as from the calibration of the Tobit regression model for projecting vehicle unit delays, the impact of traffic incidents on roadway performance (in term of delays) was found to be five times greater than that of improvement actions.

The developed event detection tools can aid field observation-based traffic event reporting, e.g. through observations by police or road patrol. These tools may also facilitate quick, automated event detection using streamed traffic data.

Specified unit-delay estimates, and estimates from proposed equations for computing average unit delays for spec-

ified roadways, can be used in construction activity planning and prioritizing improvement actions while accounting for facility- and corridor-level impacts arising from the event downtime.

Applying findings from this study, along with implementation of the developed methods, can lead to improved construction activity planning. With event detection and insights into the event's consequences through real-time application of the proposed tools, an agency can more readily respond to the event, providing appropriate services to clear a traffic incident and diverting traffic to alternative routes during construction activities or incidents. These actions can aid in reducing event impact and providing greater travel time reliability. They may also be used in estimating the net financial or public welfare impacts of considered transportation investments when choosing between alternatives.

In this work, the machine learning methods were trained on a portion of the data and tested on the remaining data in each of three case studies. Alternatively, the techniques could be trained on the entire dataset for one case study and then, tested on the other two case studies. Degraded performance may arise with this alternative method if the geometries and traffic characteristics of the case studies greatly differ.

## Data Availability

Some or all data, models, or code used during the study were provided by a third party. Direct requests for these materials may be made to the provider.

## Conflicts of Interest

The authors declare no conflicts of interest with respect to research, authorship, and/or publication of this article.

## Acknowledgments

This work was supported by the Virginia Transportation Research Center (VTRC) of the Virginia Department of Transportation. This support is gratefully acknowledged, but implies no endorsement of the findings. The authors are grateful to the program officers at the VTRC for their ideas and providing access to data to support this investigation.

## References

- [1] Q. Meng and J. Weng, "Evaluation of rear-end crash risk at work zone using work zone traffic data," *Accident; Analysis and Prevention*, vol. 43, no. 4, pp. 1291–1300, 2011.
- [2] B. Du, S. Chien, J. Lee, L. Spasovic, and K. Mouskos, "Artificial neural network model for estimating temporal and spatial freeway work zone delay using probe-vehicle data," *Transportation Research Record*, vol. 2573, no. 1, pp. 164–171, 2016.
- [3] Federal Highway Administration, *FHWA Work Zone Facts and Statistics - FHWA Office of Operations*, 2018, [https://ops.fhwa.dot.gov/wz/resources/facts\\_stats.htm](https://ops.fhwa.dot.gov/wz/resources/facts_stats.htm).



- [4] R. Dowling, A. Skabardonis, M. Carroll, and Z. Wang, "Methodology for measuring recurrent and nonrecurrent traffic congestion," *Transportation Research Record*, vol. 1867, no. 1, pp. 60–68, 2004.
- [5] H. Nassiri and R. Aghamohammadi, "A New Analytic Neuro-Fuzzy Model for Work Zone Capacity Estimation," in *Transportation Research Board 96th Annual Meeting*, Washington, DC, 2017.
- [6] J. Weng and Q. Meng, "Decision tree-based model for estimation of work zone capacity," *Transportation Research Record*, vol. 2257, no. 1, pp. 40–50, 2011.
- [7] A. J. Khattak, A. J. Khattak, and F. M. Council, "Effects of work zone presence on injury and non-injury crashes," *Accident; Analysis and Prevention*, vol. 34, no. 1, pp. 19–29, 2002.
- [8] Y. Chung, "Assessment of non-recurrent traffic congestion caused by freeway work zones and its statistical analysis with unobserved heterogeneity," *Transport Policy*, vol. 18, no. 4, pp. 587–594, 2011.
- [9] M. Asgharzadeh and A. Kondyli, "Comparison of highway capacity estimation methods," *Transportation Research Record*, vol. 2672, no. 15, pp. 75–84, 2018.
- [10] J. Weng and Q. Meng, "Estimating capacity and traffic delay in work zones: an overview," *Transportation Research Part C: Emerging Technologies*, vol. 35, pp. 34–45, 2013.
- [11] S. A. Bagloee and M. Asadi, "Prioritizing road extension projects with interdependent benefits under time constraint," *Transportation Research Part A: Policy and Practice*, vol. 75, pp. 196–216, 2015.
- [12] P. L. Durango-Cohen and P. Sarutipand, "Maintenance optimization for transportation systems with demand responsiveness," *Transportation Research Part C: Emerging Technologies*, vol. 17, no. 4, pp. 337–348, 2009.
- [13] A. Medury and S. Madanat, "Incorporating network considerations into pavement management systems: a case for approximate dynamic programming," *Transportation Research Part C: Emerging Technologies*, vol. 33, pp. 134–150, 2013.
- [14] D. Mizutani, Y. Nakazato, and J. Lee, "Network-level synchronized pavement repair and work zone policies: optimal solution and rule-based approximation," *Transportation Research Part C: Emerging Technologies*, vol. 120, article 102797, 2020.
- [15] W. Orabi and K. El-Rayes, "Optimizing the rehabilitation efforts of aging transportation networks," *Journal of Construction Engineering and Management*, vol. 138, no. 4, pp. 529–539, 2012.
- [16] Y. Ouyang and S. Madanat, "Optimal scheduling of rehabilitation activities for multiple pavement facilities: exact and approximate solutions," *Transportation Research Part A: Policy and Practice*, vol. 38, no. 5, pp. 347–365, 2004.
- [17] X. Tao and P. Schonfeld, "Lagrangian relaxation heuristic for selecting interdependent transportation projects under cost uncertainty," *Transportation Research Record*, vol. 1931, no. 1, pp. 74–80, 2005.
- [18] H. Y. Lee, "Optimizing schedule for improving the traffic impact of work zone on roads," *Automation in Construction*, vol. 18, no. 8, pp. 1034–1044, 2009.
- [19] M. W. Ng, D. Y. Lin, and S. T. Waller, "Optimal long-term infrastructure maintenance planning accounting for traffic dynamics," *Computer-Aided Civil and Infrastructure Engineering*, vol. 24, no. 7, pp. 459–469, 2009.
- [20] D. Yang, X. Zhao, Y. Chen, X. Zhang, and C. Chen, "Study on the day-based work zone scheduling problem in urban road networks based on the day-to-day traffic assignment model," *Transportation Research Record*, vol. 2672, no. 16, pp. 14–22, 2018.
- [21] H. Zheng, E. Nava, and Y. C. Chiu, "Measuring networkwide traffic delay in schedule optimization for work-zone planning in urban networks," *IEEE Transactions on Intelligent Transportation Systems*, vol. 15, no. 6, pp. 2595–2604, 2014.
- [22] W. Zhou, E. Miller-Hooks, K. G. Papakonstantinou, S. Stoffels, and S. McNeil, "A Reinforcement Learning Method for Multi-asset Roadway Improvement Scheduling Considering Traffic Impacts," *Journal of Infrastructure Systems*, vol. 28, no. 4, article 04022033, 2022.
- [23] K. G. Papakonstantinou and M. Shinozuka, "Planning structural inspection and maintenance policies via dynamic programming and Markov processes. Part II: POMDP implementation," *Reliability Engineering and System Safety*, vol. 130, pp. 214–224, 2014.
- [24] C. L. Dudek, C. J. Messer, and N. B. Nuckles, "Incident detection on urban freeways," *Transportation Research Record*, vol. 495, pp. 12–24, 1974.
- [25] H. J. Payne and S. C. Tignor, "Freeway Incident-Detection Algorithms Based on Decision Trees with States," *Transportation Research Record*, vol. 682, 1978.
- [26] S. R. Ahmed and A. R. Cook, "Application of time-series analysis techniques to freeway incident detection," *Transportation Research Record*, vol. 841, pp. 19–21, 1982.
- [27] A. B. Parsa, H. Taghipour, S. Derrible, and A. Mohammadian, "Real-time accident detection: coping with imbalanced data," *Accident; Analysis and Prevention*, vol. 129, pp. 202–210, 2019.
- [28] J. Ren, Y. Chen, L. Xin, J. Shi, B. Li, and Y. Liu, "Detecting and positioning of traffic incidents via video-based analysis of traffic states in a road segment," *IET Intelligent Transport Systems*, vol. 10, no. 6, pp. 428–437, 2016.
- [29] A. Kinoshita, A. Takasu, and J. Adachi, "Real-time traffic incident detection using a probabilistic topic model," *Information Systems*, vol. 54, pp. 169–188, 2015.
- [30] Z. Zhang, Q. He, J. Gao, and M. Ni, "A deep learning approach for detecting traffic accidents from social media data," *Transportation Research Part C: Emerging Technologies*, vol. 86, pp. 580–596, 2018.
- [31] C. El Hatri and J. Boumhidi, "Fuzzy deep learning based urban traffic incident detection," *Cognitive Systems Research*, vol. 50, pp. 206–213, 2018.
- [32] Y. Gu, Z. Qian, and F. Chen, "From twitter to detector: real-time traffic incident detection using social media data," *Transportation Research Part C: Emerging Technologies*, vol. 67, pp. 321–342, 2016.
- [33] R. Wang, S. Fan, and D. B. Work, "Efficient multiple model particle filtering for joint traffic state estimation and incident detection," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 521–537, 2016.
- [34] S. Chen, W. Wang, and H. van Zuylen, "Construct support vector machine ensemble to detect traffic incident," *Expert Systems with Applications*, vol. 36, no. 8, pp. 10976–10986, 2009.
- [35] M. Chong, A. Abraham, and M. Paprzycki, "Traffic Accident Analysis Using Machine Learning Paradigms," *Informatica*, vol. 29, no. 1, 2005.
- [36] F. Yuan and R. L. Cheu, "Incident detection using support vector machines," *Transportation Research Part C: Emerging Technologies*, vol. 11, no. 3-4, pp. 309–328, 2003.

- [37] X. Jin, R. L. Cheu, and D. Srinivasan, "Development and adaptation of constructive probabilistic neural network in freeway incident detection," *Transportation Research Part C: Emerging Technologies*, vol. 10, no. 2, pp. 121–147, 2002.
- [38] B. Sluban, D. Gamberger, and N. Lavrač, "Ensemble-based noise detection: noise ranking and visual performance evaluation," *Data Mining and Knowledge Discovery*, vol. 28, no. 2, pp. 265–303, 2014.
- [39] J. Xiao, "SVM and KNN ensemble learning for traffic incident detection," *Physica A: Statistical Mechanics and its Applications*, vol. 517, pp. 29–35, 2019.
- [40] C. Cortes and V. Vapnik, "Support-vector networks," *Machine Learning*, vol. 20, no. 3, pp. 273–297, 1995.
- [41] V. Vapnik, *The Nature of Statistical Learning Theory*, Springer Science & Business Media, 2000.
- [42] G. B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 42, no. 2, pp. 513–529, 2012.
- [43] I. Semanjski and S. Gautama, "Crowdsourcing mobility insights - reflection of attitude based segments on high resolution mobility behaviour data," *Transportation Research Part C: Emerging Technologies*, vol. 71, pp. 434–446, 2016.
- [44] R. Kohavi, "A Study of Cross-Validation and Bootstrap for Accuracy Estimation and Model Selection," in *Proceedings of the 14th International Joint Conference on Artificial Intelligence - Volume 2, IJCAI'95*, pp. 1137–1143, Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1995.
- [45] L. Breiman, "Random forests," *Machine Learning*, vol. 45, pp. 5–32, 2001.
- [46] A. Iranitalab and A. Khattak, "Comparison of four statistical and machine learning methods for crash severity prediction," *Accident; Analysis and Prevention*, vol. 108, pp. 27–36, 2017.
- [47] J. J. Rebollo and H. Balakrishnan, "Characterization and prediction of air traffic delays," *Transportation Research Part C: Emerging Technologies*, vol. 44, pp. 231–241, 2014.
- [48] B. Yu, S. Bao, F. Feng, and J. Sayer, "Examination and prediction of drivers' reaction when provided with V2I communication-based intersection maneuver strategies," *Transportation Research Part C: Emerging Technologies*, vol. 106, pp. 17–28, 2019.
- [49] CCCarto, "Map of Virginia cities - Virginia roads map," 2018, <https://www.cccarto.com/stateroutes/virginia/>.
- [50] OpenStreetMap, 2020, <https://www.openstreetmap.org/>.
- [51] A. Nohekhan, S. Zahedian, and K. F. Sadabadi, "Investigating the impacts of I-66 inner beltway dynamic tolling system," *Transportation Engineering*, vol. 4, article 100059, 2021.
- [52] S. Zahedian, P. Sekuła, A. Nohekhan, and Z. Vander Laan, "Estimating hourly traffic volumes using artificial neural network with additional inputs from automatic traffic recorders," *Transportation Research Record*, vol. 2674, no. 3, pp. 272–282, 2020.
- [53] C. S. Chou and E. Miller-Hooks, "Simulation-based secondary incident filtering method," *Journal of Transportation Engineering*, vol. 136, no. 8, pp. 746–754, 2010.
- [54] S. Mamdoohi and E. Miller-Hooks, *Identifying the Impact Area of a Traffic Event through K-Means Clustering*, Manuscr. Rev, 2022.
- [55] A. Tavassoli Hojati, L. Ferreira, S. Washington, P. Charles, and A. Shobeirinejad, "Modelling the impact of traffic incidents on travel time reliability," *Transportation Research Part C: Emerging Technologies*, vol. 65, pp. 49–60, 2016.
- [56] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002.
- [57] J. You, J. Wang, and J. Guo, "Real-time crash prediction on freeways using data mining and emerging techniques," *Journal of Modern Transportation*, vol. 25, no. 2, pp. 116–123, 2017.
- [58] S. V. Stehman, "Selecting and interpreting measures of thematic classification accuracy," *Remote Sensing of Environment*, vol. 62, no. 1, pp. 77–89, 1997.
- [59] V. Svetnik, A. Liaw, C. Tong, and T. Wang, "Application of Breiman's random forest to modeling structure-activity relationships of pharmaceutical molecules," in *International Workshop on Multiple Classifier Systems*, pp. 334–343, Springer, Berlin, Heidelberg, 2004.
- [60] C. Strobl, A.-L. Boulesteix, T. Kneib, T. Augustin, and A. Zeileis, "Conditional variable importance for random forests," *BMC Bioinformatics*, vol. 9, no. 1, p. 307, 2008.
- [61] B. Y. J. Tobin, "Estimation of relationships for limited dependent variables," *Econometrica: journal of the Econometric Society*, vol. 26, pp. 24–36, 1958.