Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones

PUBLICATION NO. FHWA-HRT-20-010

DECEMBER 2019

U.S. Department of Transportation

Federal Highway Administration

Research, Development, and Technology
Turner-Fairbank Highway Research Center
6300 Georgetown Pike
McLean, VA 22101-2296

FOREWORD

In 2016, the Federal Highway Administration posted a Broad Agency Announcement (BAA) to conduct research on potential safety improvements using the Naturalistic Driving Study (NDS) and Roadway Information Database (RID) collected during the research phase of the Second Strategic Highway Research Program. Phase 1 served as a "proof of concept" to determine if meaningful conclusions or countermeasures can be developed using NDS and RID databases. Phase 2 enabled researchers to conduct more indepth analyses, leading to specific highway safety improvements.

The following final report describes the methodology and results of one of six BAA projects to characterize normal and abnormal driving behavior in work zones. In this study, the researchers successfully used the NDS and RID databases to quantify the role of traffic management, work zone activities, and traffic conditions on driver behavior such as speed, merging, and so forth. The results suggest that "nudging" drivers to comply with work zone speed limits and safe following distances would be effective at reducing the number of safety-critical events. This report will be of interest to State and local departments of transportation professionals that are responsible for managing work zones, setting guidelines and policies to implement in work zones, and developing applications and communication protocols for autonomous vehicles.

Brian P. Cronin, P.E.
Director, Office of Safety and Operations
Research and Development

TECHNICAL REPORT DOCUMENTATION PAGE

1. Report No. FHWA-HRT-20-010	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones		5. Report Date December 2019 6. Performing Organization Code
7. Author(s) Carol A. Flannagan (ORCID: 0000-0001-8484-4187)*, Selpi (ORCID: 0000-0003-2800-4479)†, Pinar Boyraz Baykas (ORCID: 0000-0002-3665-1775)†, Andrew Leslie (ORCID: 0000-0001-7233-6644)*, Jordanka Kovaceva (ORCID: 0000-0002-7445-3489)†, and Robert Thomson (ORCID: 0000-0002-8847-6753)†		8. Performing Organization Report No.
9. Performing Organization Name and Address University of Michigan Transportation Research Institute* 2901 Baxter Rd Ann Arbor, MI 48109 Chalmers University of Technology† Chalmersplatsen 4, 412 96 Gothenburg, Sweden Virginia Tech Transportation Institute 3500 Transportation Research Plaza Blacksburg, VA 24061		10. Work Unit No. (TRAIS) 11. Contract or Grant No. DTFH61-16C-00006
12. Sponsoring Agency Name and Address U.S. Department of Transportation Federal Highway Administration 6300 Georgetown Pike McLean, VA 22101-2296		13. Type of Report and Period Covered Final Report; October 2014–October 2019 14. Sponsoring Agency Code HRDS-20

15. Supplementary Notes
The COR/FHWA POC is Yusuf Mohamedshah (HRDS-20, ORCID: 0000-0003-0105-5559).

16. Abstract
This research project used the Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) to improve highway safety by using statistical descriptions of normal driving behavior to identify abnormal driving behaviors in work zones. SHRP2 data used in these analyses included 50 safety-critical events (SCEs) from work zones and 444 baseline events selected on a matched case-control design.
Principal components analysis (PCA) was used to summarize kinematic data into "normal" and "abnormal" driving. Each second of driving is described by one point in three-dimensional principal component (PC) space; an ellipse containing the bulk of baseline points is considered "normal" driving. Driving segments with out-of-ellipse points have a higher probability of being an SCE. Matched case-control analysis indicates that the specific individual and traffic flow made approximately equal contributions to predicting out-of-ellipse driving.
Structural Topics Modeling (STM) was used to analyze complex categorical data obtained from annotated videos. The STM method finds "words" representing categorical data variables that occur together in many events and describes these associations as "topics." STM then associates topics with either baselines or SCEs. The STM produced 10 topics: 3 associated with SCEs, 5 associated with baselines, and 2 that were neutral. Distraction occurs in both baselines and SCEs.
Both approaches identify the role of individual drivers in producing situations where SCEs might arise. A countermeasure could use the PC calculation to indicate impending issues or specific drivers who may have higher crash risk, but not to employ significant interventions such as automatically braking a vehicle with out-of-ellipse driving patterns. STM results suggest communication to drivers or placing compliant vehicles in the traffic stream would be effective. Finally, driver distraction in work zones should be discouraged.

17. Key Words Work zones, crashes, safety, statistics, data analysis, methodology	18. Distribution Statement: No restrictions. This document is available to the public through the National Technical Information Service, Springfield, Virginia 22161. http://www.ntis.gov.		
19. Security Classif. (of report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 98	22. Price

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

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yard	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	mL
gal	gallons	3.785	liters	L
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3
		NOTE: volumes greater than 1,000 L shall be shown in m^3		
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2,000 lb)	0.907	megagrams (or "metric ton")	Mg (or "t")
		TEMPERATURE (exact degrees)		
°F	Fahrenheit	5 (F-32)/9 or (F-32)/1.8	Celsius	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	lx
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
lbf/in^2	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	yd^2
ha	hectares	2.47	acres	ac
km^2	square kilometers	0.386	square miles	mi^2
		VOLUME		
mL	milliliters	0.034	fluid ounces	fl oz
L	liters	0.264	gallons	gal
m^3	cubic meters	35.314	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg (or "t")	megagrams (or "metric ton")	1.103	short tons (2,000 lb)	T
		TEMPERATURE (exact degrees)		
°C	Celsius	1.8C+32	Fahrenheit	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	2.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	lbf/in^2

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

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

ACC	automatic cruise control
ARMA	autoregressive-moving-average
ASAP	Appropriate Speed Saves All People
AUC	area under the curve
BART	Bayesian additive regression trees
CAN	controller area network
CMF	crash modification factor
CR	context-related
CRE	crash-related event
FREX	frequency and exclusivity
GPS	global positioning system
I2V	infrastructure-to-vehicle
LDA	latent Dirichlet allocation
MDOT	Michigan Department of Transportation
MUTCD	*Manual on Uniform Traffic Control Devices*
NDD	naturalistic driving data
NDS	Naturalistic Driving Study
PC	principal component
PCA	principal components analysis
PTM	Probabilistic Topic Modeling
RADAR	radio detection and ranging
SCE	safety-critical event
SHRP2	Second Strategic Highway Research Program
STM	Structural Topic Modeling
STR	secondary-task-related
TFR	traffic flow restrictions
TMA	truck-mounted attenuator
V2I	vehicle-to-infrastructure
VAR	vector autoregression
VSL	variable speed limit
WZR	work-zone-related

EXECUTIVE SUMMARY

This research project used the Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) to improve highway safety by analyzing data to improve understanding of normal and abnormal driving behaviors in work zones. The goal of the project was to use statistical descriptions of normal driving behavior to identify abnormal behavior as the basis for countermeasures. The focus is work zone safety.

This project focused on two statistical methods that are meant to characterize normal driving behavior and, by contrast, identify abnormal driving behavior. The two methods were Principal Components Analysis (PCA) and Structural Topic Modeling (STM). In addition, to target the work toward practical applications, work zone safety was the general topic of interest. Work zone speed limits were the application focus for the PCA approach, and work zone guidance was the application focus for the STM approach.

The analyses used data from the SHRP2 NDS in which work zones were present. Specifically, the data included 50 safety-critical events (SCEs) coded as occurring in a work zone. Supplementary data consisted of additional baseline driving (no SCE) epochs selected based on a matched case-control design. Each SCE location was matched with up to 12 baseline passes through the same work zone (within 1 week of the SCE), three by the same driver and nine by other drivers. Within each driver sample, the request included at least one pass occurring at the same time of day (also considering weekday/weekend) and the remaining passes at other times of day. This sampling strategy allowed identification of driving and work zone characteristics that were associated with unsafe (SCEs) versus safe (baseline) conditions. The resulting dataset consisted of 50 SCEs and 444 baselines.

Forward video was coded by the research team to document the presence of a variety of work zone features, including workers present, specific work zone treatments (e.g., barrels, cones, chicanes), and traffic conditions (e.g., free flow, queue forming). Face video was coded by the subcontractor that archives the SHRP2 dataset for the presence of a variety of secondary tasks, the presence of driver impairments, and whether the driver's hands were on the wheel.

The PCA approach was used to efficiently summarize kinematic data (speeds and accelerations) into "normal" and "abnormal" driving. The PCA of speed and longitudinal and lateral acceleration indicated that variation occurred primarily in longitudinal measures, and that each second of data (10 observations) could be summarized with three principal components (PCs). Each second of driving could then be described as successive points in three-dimensional PC space; an ellipse drawn around the bulk of the distribution of those points can be considered to contain "normal" driving.

The PCA-based concept is illustrated in figure 1. The black dots represent all driving in SCE and baseline events, while the red line connects dots occurring during a particular SCE. An ellipsoid shape can define a boundary within this space so that excursions from that ellipse tend to be associated with atypical and SCE-related driving.

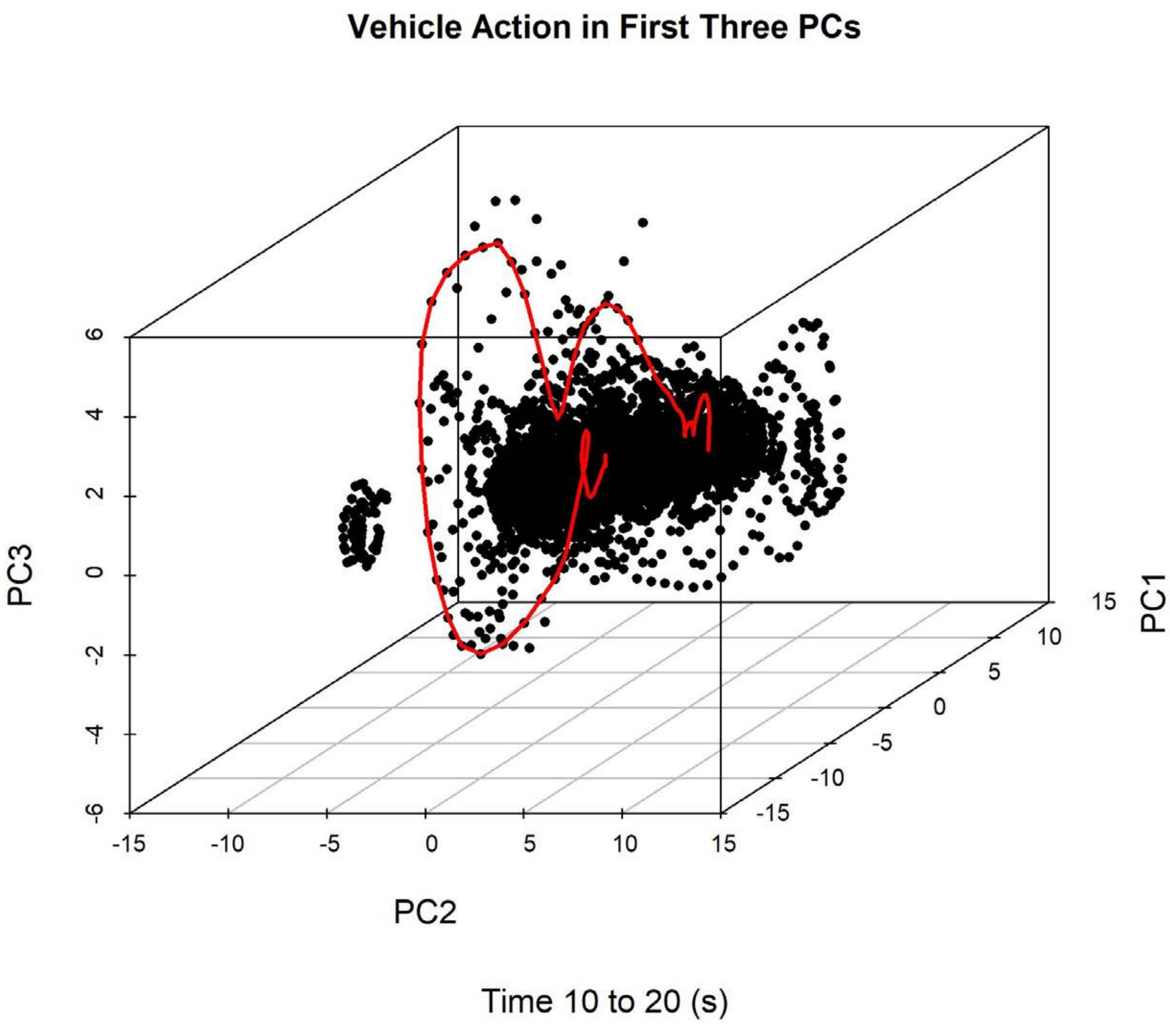


Figure 1. Graph. Example of driving data in PC space (modified from Flannagan et al. 2016).

More detailed analyses of PC space indicated that PC2 and PC3, which are more closely related to acceleration patterns, were the best predictors of SCEs. Thus, for an ellipse defined in the PC2/PC3 plane, points outside the ellipse are categorized as out-of-ellipse driving. Out-of-ellipse excursions indicate a greater probability of an event being an SCE rather than baseline. Ellipse boundaries vary for each work zone because of natural differences in variation of accelerations in different situations.

Analysis of the matched case-control data indicated that the specific individual and the time of day (i.e., traffic flow) made approximately equal contributions to predicting out-of-ellipse driving. In addition, while out-of-ellipse driving is a predictor of SCEs, its performance is not perfect, and multiple attempts to predict ellipse locations into the future failed. In other words, unpredictability is an inherent feature of SCEs, so indications of increased risk in the PC space are identifiable, but the PC space cannot pinpoint when it will occur (in the next few seconds).

The unpredictability feature of SCEs suggests that a practical countermeasure could use the PC calculation to indicate impending issues or specific drivers who might be at increased risk of crashing, but not to employ significant interventions such as automatically braking a vehicle with out-of-ellipse behavior.

The STM method is best suited for analyzing the complex categorical data from annotated video. This method, which was developed for text processing, treats each baseline or SCE driving event as a "document," and specific categories of variables as "words." The method finds words, or categorical values, which occur together in many documents (events) and describes these associations as "topics." Thus, a topic will consist of several associated words or characteristics of a driving event. Importantly, STM can then associate topics with either baselines or SCEs, thus giving some idea of the collection of features that might be either safe or unsafe, depending on the associations.

The STM produced 10 topics, 3 of which were associated with SCEs, 5 of which were associated with baselines, and 2 of which were neutral. Interestingly, all topics include some secondary tasks, indicating that distraction occurs in both baselines and SCEs and is a major feature of driving among SHRP2 participants. Specific work zone treatments associated with SCEs and distraction include workers present, chicanes, queues, and lanes ending or reducing. Baselines are more associated with free flow conditions.

Results from both analysis approaches emphasized the role of the individual in producing situations where SCEs might arise. Thus, recommendations emphasize communication with individuals and attempts to elicit compliant behavior from them. The PCA results are discussed in terms of variable speed limits (VSLs). The presence of out-of-ellipse driving might indicate the need for reduced speed, but in general, existing literature provides guidance for determining optimal speed limits. Instead, out-of-ellipse driving identifies an individual whose driving might either result in an SCE or potentially just create suboptimal throughput because that driver is braking or accelerating too suddenly for optimal flow. There are VSL models (e.g., Yu and Abdel-Aty 2014) that include driver compliance components, a role for which PCA is well suited.

The STM results are discussed primarily in the context of work zone treatment guidance. The project recommendations revolve around the need to "nudge" drivers to comply with VSLs and appropriate headways and distances to nearby vehicles. Nudging might be accomplishment by frequent general messaging through digital signs (which were associated with safe driving in the STM), audio messages, or infrastructure-to-vehicle (I2V) messages. Additional nudging could come from targeted messages to individual drivers (e.g., those exceeding the ellipse boundary), either using I2V or possibly on general signage. Finally, another approach might be to constrain traffic flow by ensuring that compliant vehicles are present in the traffic stream. Davis (2016) showed that if only one-third of the traffic stream were automatic cruise control (ACC)-equipped vehicles with compliant speed and headway settings, the whole traffic stream would move more efficiently by reducing queue buildup. In his paper, Davis (2016) assumed that this would be accomplished through I2V control of the ACC settings, but a more near-term solution might be achieved by encouraging heavy trucks and/or ACC users to drive in a compliant way. The fact that only one-third of the traffic stream needs to be compliant means that drivers who might be inclined to drive in an out-of-ellipse manner would be constrained.

Finally, the need to encourage drivers to resist distraction in work zones is also evident. Distraction countermeasures are widely implemented in the United States and other countries, yet it remains a persistent issue in driving. That said, targeted messaging (possibly including personal messages) might still help in the short-run work zone situation.

CHAPTER 1. INTRODUCTION

This research project used the Second Strategic Highway Research Program (SHRP2) Naturalistic Driving Study (NDS) to improve highway safety by analyzing data to advance understanding of normal and abnormal driving behaviors in work zones. The project goal used statistical descriptions of normal driving behavior to identify abnormal behavior (i.e., deviations from normal) as the basis for countermeasures. The focus is work zone safety.

Phase 1 of the work zone project began with exploration of two novel methodological approaches for analyzing naturalistic driving data from SHRP2. The two methods are meant to characterize normal driving behavior and, by contrast, identify abnormal driving behavior. The two methods are principal components analysis (PCA) and Probabilistic Topic Modeling (PTM). With promising results from Phase 1, Phase 2 explored applications of the two methods. Work zone speed limits were the application focus for the PCA approach, and work zone guidance was the application focus for the PTM approach. Phase 2 determined that among different PTM methods, Structural Topic Modeling (STM) was a more appropriate form of topic modeling for this application than latent Dirichlet allocation (LDA).

This report is organized as follows:

- Chapter 2 describes the data used in this project. New data were requested for Phase 2, so the description addresses the data and how they were coded and aligned.
- Chapter 3 reviews PCA-based methods, described previously in detail in the *Methods Interim Report: Reporting Task D3.1 for Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones* (Kovaceva et al. 2018).
- Chapter 4 describes applications of STM, which is a text-mining technique developed after a series of improvements to LDA.
- Chapter 5 and chapter 6 explain potential applications of the statistical methods.
- Chapter 7 summarizes and discusses the results of the entire project.

CHAPTER 2. DATA DESCRIPTION

DATA REQUESTED AND RECEIVED

Phase 1 work involved analysis of time series data for 243 safety-critical events (SCEs) and 420 baselines that occurred in work zones. For the SCEs, the data consisted of 31-s passes through work zones, with the timing aligned so that the precipitating event occurred at the 20th s. The precipitating event was, as the name implies, the event that occurred (often braking by a lead vehicle) that then led to the SCE a short time later. The baseline data were 21-s passes through work zones, but were not sampled from the same work zones or drivers as the SCEs, making it difficult to assess the effects of work zone characteristics versus driver behavior.

To examine the effects of work zone characteristics versus driver behavior, the Phase 2 data request was based on a matched case-control design, where matching was done for a subset of the original SCEs for which there were baseline passes through the same work zones available. The goal was to find several passes through the same work zones in the same timeframe to allow comparison of the behavior of the same driver versus other drivers and the behavior of the same versus different drivers at different times of day.

Based on preliminary results from the subcontractor that archives the SHRP2 dataset on the availability and frequency of baseline passes, 52 SCEs were selected as the basis for matching in the Phase 2 dataset. For each SCE, video data were requested, including 10 min prior to the event and 5 min after the event, except when that would overlap with the origin or destination buffers applied by the archiving subcontractor. After review of the videos, two of the SCEs were dropped from the dataset because there was not a work zone present or because the SCE occurred after the work zone ended. For the remaining 50 cases, the bounds for the data request were set at 150 s pre-SCE and 5 s post-SCE for any case where the work zone starting location was more than 170 s before the SCE. For any work zone where the start was closer to the SCE than that, the pass ran from 20 s before the work zone start to 5 s post-SCE.

Once the start and end timestamps for each SCE were defined, the locations of these timestamps were identified by the archiving subcontractor and used to select the baseline passes for the request. In other words, rather than matching baseline and SCE passes in terms of time, they were matched in location, regardless of how long it took drivers to traverse from start to endpoint. The remaining criteria for the baseline passes are described in the following paragraphs.

The 50 SCEs included 12 crash events and 38 near crashes. A breakdown of the event types is provided in table 1.

Table 1. Event type for SCEs requested.

Crash Type	Total	Crashes	Near Crashes
Rear-end, striking	33	7	26
Rear-end, struck	2	1	1
Road departure (left or right)	1	1	0
Sideswipe, same direction (left or right)	7	1	6
Turn across path	2	1	1
Turn into path (same direction)	3	0	3
Other	2	1	1

Selection of the baselines was based on the driver identification number and the approximate time of the work zone pass, with times categorized as follows:

- Weekday peak: Monday through Friday, 6:01 a.m. to 9:00 a.m. or 4:01 p.m. to 7:00 p.m.
- Weekday off-peak: Monday through Friday, 12:01 a.m. to 6:00 a.m., 9:01 a.m. to 4:00 p.m., or 7:01 p.m. to 12:00 a.m.
- Weekend/holiday: Saturday and Sunday all day and holidays all day where identifiable.

All baselines were required to have the same heading as the SCE pass and must have occurred within 1 week of the SCE. For each SCE, 12 baselines were requested according to the following time/driver breakdown and count (in parentheses):

- Same driver and same time (1).
- Same driver and different time bin, selected at random from those available (2).
- Different driver and weekday peak (3).
- Different driver and weekday off-peak (3).
- Different driver and weekend/holiday (3).

Where possible, all nine of the different-driver baselines were requested to come from unique drivers. In full, the data request covered 650 passes (50 events and 600 baselines) and included forward video, speed limit data, processed radio detection and ranging (RADAR) data, and face video annotations of all passes.

Due to process difficulties, the archiving subcontractor was unable to provide the full set of 600 baselines, instead providing only 444 baselines. Table 2 shows the distribution of SCEs by the number of same-driver matched baselines found, while table 3 shows the distribution of SCEs by the number of different-driver matched baselines found. The target for same-driver baselines was 3, and only 36 percent (18) of SCEs had 3 of these. In contrast, the target for different-driver baselines was 9, and 70 percent (35) of SCEs had 9 such baselines. Combined, only 30 percent (15) of the SCEs had the full set of 12 requested baselines. Three SCEs had baselines only from the SCE driver, while six had only different drivers, but all SCEs had at least one baseline.

Table 2. Distribution of SCEs by the number of same-driver baselines found.

Number of Same-Driver Baselines Provided	Count of Matched SCEs	Percent of Matched SCEs
3	18*	36
2	11	22
1	15	30
0	6	12

*Number of SCEs matching request of 3.

Table 3. Distribution of SCEs by the number of different-driver baselines found.

Number of Different-Driver Baselines Provided	Count of Matched SCEs	Percent of Matched SCEs
9	35*	70
6	3	6
1–4	9	18
0	3	6

*Number of SCEs matching request of 9.

Table 4 and table 5 show the distribution of SCEs by time bin of same-driver and different-driver baselines. One same-driver same-time-bin pass was found for 76 percent (38) of SCEs. However, the same-driver different-time-bin baselines were more difficult to find, with only 46 percent (23) of SCEs having the requested 2 baselines and another 14 percent (7) having a single baseline. In the different-driver same-time case, 6 percent (3) of SCEs had more than the requested baselines, 62 percent (31) had the requested 3 baselines, and 16 percent (8) each had either too few baselines or no baselines. In the different-driver different-time case, 4 percent (2) had more than the requested baselines, 60 percent (30) had the requested 6 baselines, 26 percent (13) had fewer than the requested number of baselines, and 10 percent (5) had no different-driver different-time baselines. Although the specific reason for the lack of baselines is not known, the combination of the 1-week window and the specifics of time and day probably reduced the number of qualifying baselines that were in the dataset to below the requested number. Delays in receiving data prevented the request of alternate baselines, but 444 cases were deemed to be sufficient for the analysis.

Table 4. Distribution of SCEs by the number of matched baselines in each time bin for same-driver baselines.

Number of Same-Driver Baselines Provided	Count of SCES in Same Time Bin as SCE (1)	Count of SCEs in Different Time Bin as SCE (2)
2	0	23*
1	38*	7
0	12	20

*Number of SCEs achieving the request.
Note: The number requested is in parentheses.

Table 5. Distribution of SCEs by the number of matched baselines in each time bin for different-driver baselines.

Number of Different-Driver Baselines Provided	Count of SCEs in Same Time Bin as SCE (3)	Count of SCEs in Different Time Bin as SCE (6)
>6	1	2
6	0	30*
4-5	2	6
3	31*	0
1–2	8	7
0	8	5

*Number of SCEs achieving the request.
Note: The number requested is in parentheses.

The time series and processed RADAR data were provided for all 494 cases to the extent that they were populated in the SHRP2 dataset. The requested supplementary information provided is outlined below:

- Forward video: available for all 50 SCEs and 443 of the baselines, with 1 baseline being lost due to problems with the forward camera. For six of the SCEs, the requested video data overlapped the origin buffer and were truncated by the archiving subcontractor to avoid exposing personal information about the drivers.
- Speed limit data: available for 48 of the SCEs and 443 of the baselines. For the 491 cases with any available speed limit data, the speed limit was known for over 75 percent of the time range in all but 8 cases.
- Face video annotations: data reduction for the face video was provided for all 494 cases.

VIDEO CODING

Forward Video Coding

The new Phase 2 forward view videos were coded by two annotators. The coding protocol was developed in Phase 1 and was implemented with some additions and refinements in Phase 2. The primary goal was to label work-zone-specific variables and other variables based on the videos that were considered useful for this project. Appendix A contains the full list of coded variables.

Face Video Coding

Face video coding was done by the subcontractor that archives the SHRP2 dataset. Due to the cost of coding, the request was limited to three key variables from the new (longer) passes. The variables were:

- Secondary tasks: if any secondary task in the SHRP2 data dictionary was observed, it was coded, along with start and end times of engagement in the secondary task (Hankey et al. 2016). The complete list of possible secondary tasks for coding is in appendix A.
- Impairments: the impairment was also coded based on the variable driver impairments used in SHRP2 dictionary (Hankey et al. 2016). Specific code values are in appendix A. The variable was coded in time series in the same way as for secondary tasks.

- Hands ever off (wheel): this variable could have values of "both hands off," "unknown," or "not applicable" (at least one hand was always on wheel). When the value was "both hands off" or "unknown," the timestamps at the start and the end were coded for each event where this was observed.

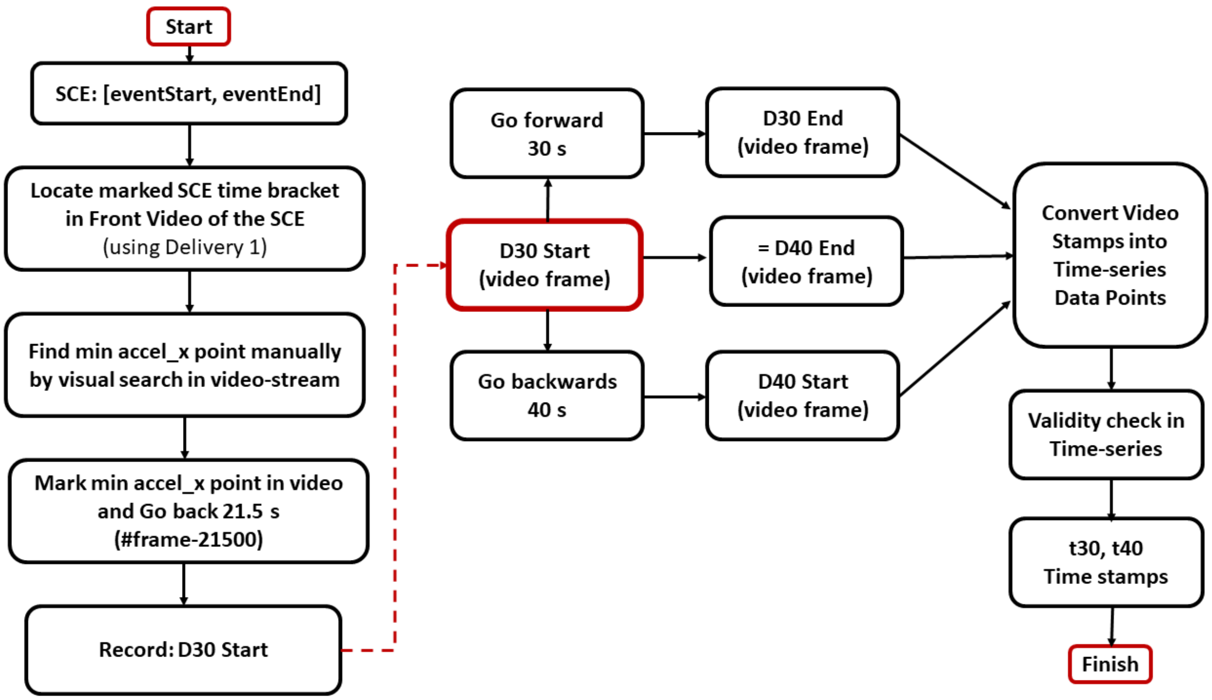
SPATIAL ALIGNMENT

Overview

Each SCE and its matched baselines occurred in the same spatial location on the same path, but each had different timestamps at each point in the path. Moreover, the provided time series did not contain global positioning system (GPS) location of the events because GPS is considered personally identifiable information. Thus, spatial alignment of the associated passes had to be done computationally using odometry, which is defined as the estimation of position from speed and acceleration. Because of partial data delivery over time and incomplete time series data for SCEs (before June 2018), the spatial alignment efforts were performed using two different odometry methodologies to make the best use of available data.

Video-Based (Visual Odometry)

First, video-based (visual odometry) was applied (June–August 2018) when the full-length time series from the original 50 SCEs was not available. It is largely a manual process of finding exact spatial coordinates by comparing the visual images of the SCE location with the images in the baseline video stream. The procedure used in video-based odometry is illustrated in figure 2. First, the SCE location image in the video stream was located using the minimum acceleration event that falls between the archiving subcontractor notated time markers: eventStart and eventEnd. The actual datapoint where the longitudinal acceleration of the vehicle was at a minimum (i.e., abrupt and intense deceleration) was always located within the time bracket of these markers. Therefore, using the already-marked time bracket [eventStart, eventEnd], the SCE front video was manually scanned to identify the minimum acceleration spatial location. This point was marked as the 215th datapoint (corresponding to 21.5 s with a data sampling rate of 10 Hz). Starting from this point and going back 21.5 s defined the start of the D30 zone. The D30 zone was defined as a 30-s clip containing the SCE, corresponding to the basic convention of 30-s clips used by the archiving subcontractor. In the video, the start of the D30 zone was also determined as a visual location. The D40 zone was coded by going back another 40 s before the start of the D30 zone and marking the video location for long-term analysis on time series. However, for this report, only the D30 zone defined the spatial boundaries of analyses. After marking the D30 (i.e., SCE-related zone) and D40 (i.e., preview zone) zones, the video times were converted into datapoints in time series to locate the corresponding time labels. Finally, to use this process to identify the corresponding locations in the baselines, video was used to visually identify the start and end of the D30 zone in each baseline. In other words, the first step included a visual comparison of the D30 start/end frames from the original SCE video to the corresponding (same-location) D30 start/end frames of the matched baselines. In the second step, a validity check was performed to make sure that the marked video frames agree with the time series data using vehicle dynamics.



t30 = time window describing the 30 s time duration after the ego-vehicle reaches the end of the D40 zone location.
t40 = time window describing the 40 s time duration before the ego-vehicle reaches the start of the D30 zone location.

Figure 2. Chart. Video-based odometry methodology/procedure (before the complete time series data were available).

Blind Odometry of Vehicle Dynamics

The second approach used blind odometry of vehicle dynamics. In the blind odometry methodology, a similar approach was employed, but instead of viewing video to find the matching points in the baseline, the endpoint of the baseline data was used as an anchor point. Figure 3 shows the structure of the baseline time series requested. Each baseline ended at the same location as the one that corresponded to 5 s after the SCE in the matched SCE (for that baseline), marked as a blue dot and labeled “End of Time-Series and Video for Baseline” in figure 3. The starting point for each baseline varied (see request description above and illustration below). Performing the blind alignment required a single known matching point, which was the endpoint of the baseline case. However, this point was not necessarily the endpoint of the SCE passes, so the anchor points for the SCEs were eventStart and eventEnd (labelled with blue triangles). Once the SCE anchor points were found in the SCEs, they were considered to be aligned with the baseline video endpoint and the rest of the information for the spatio-temporal place of D30, D40, and buffer zone before D40 was computed based on kinematics-based odometric distance.

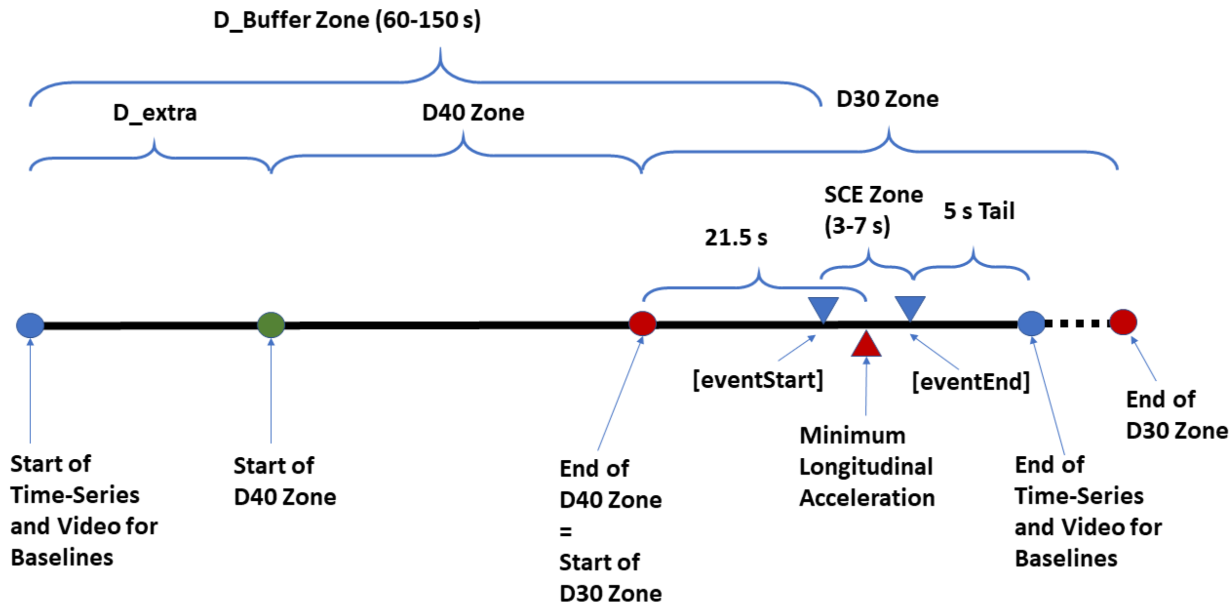


Figure 3. Diagram. Baseline time series composition.

For this report, the odometry calculations were restricted to the D30 zone (i.e., the SCE-related zone including the minimum acceleration value at 21 s) since it was possible to locate with good accuracy. The alignment algorithm in the baseline was performed using integration of the longitudinal speed based on trapezoidal integration and back-stepping with the calculated distances to find the corresponding timestamps, named as t30 and t40, in baseline data. Since the SCE D30 zone could extend beyond the endpoint of the baseline pass, the D30 zone was updated to be the D26.5 zone. When calculating the D26.5 zone, the baselines were assumed to be constructed with 5 s of data after the minimum acceleration event. This resulted in a total of 26.5 s when aligned to the 30-s SCE passes. The workflow of this algorithm is illustrated in figure 4.

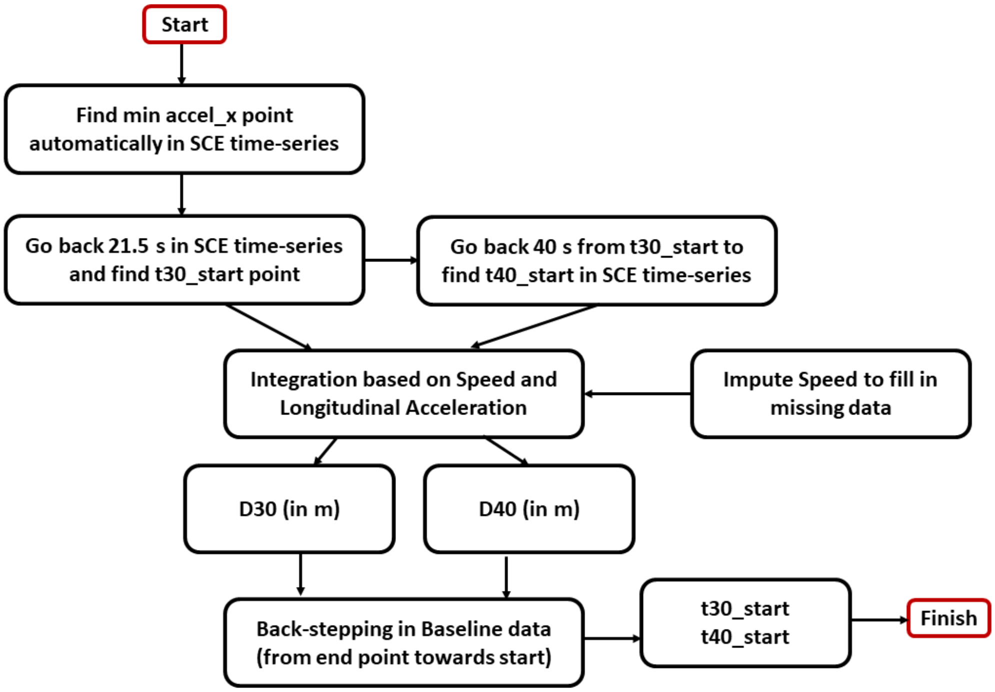


Figure 4. Chart. Blind odometry based on longitudinal vehicle dynamics using numerical integration and back-stepping.

Errors in Blind Odometry Computation and Accuracy

There are two known sources of errors in the current odometry computations. The first source of error is the result of the computation via integration based on discrete time-speed data. The vehicle dynamics data are available at 10 Hz at best; therefore, imputed speed data are used to fill the missing datapoints using a spline function between the discrete datapoints. In addition to this, the numerical integration was performed using a trapezoidal rule. Because of these limitations of the current algorithm, the corresponding datapoints (in time) in the baseline for the D26.5 distances sometimes fall between two discrete datapoints. In such cases, there is a certain roundoff error. According to analysis based on 347 baselines, this type of computation error can be at maximum 15 m, corresponding to five car lengths. Second, since the odometry computations are performed only based on the longitudinal vehicle dynamics (longitudinal acceleration, speed from GPS, or controller area network (CAN)-bus), in the scenarios where the car is performing a lane change or right/left turn, the odometry calculations can result in additional errors.

The accuracy of this procedure could be improved in the future by applying the following approaches:

- Streamlining the validation step using detection algorithms that automatically compare the locations in the video using the computer vision and image processing techniques for image matching and location recognition.

- Incorporating the lateral vehicle dynamics to improve the accuracy of odometry results that were calculated using only longitudinal dynamics data. Accuracy would only be improved by a small amount for relatively few cases.

CHAPTER 3. PCA APPROACH

OVERVIEW OF PCA METHOD

Phase 1 PCA

The PCA method used in this project was discussed at length in the *Methods Interim Report: Reporting Task D3.1 for Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones* (Kovaceva et al. 2018). PCA summarizes periods of driving (1 s at a time) in terms of three PC scores that capture most of the variance in longitudinal speed and acceleration measured at 10 Hz. Since each 1-s interval of driving can be summarized with three numbers, it can also be plotted in three-dimensional PC space. Appendix B provides further detail on PCA.

A depiction of a sequence of driving intervals was first documented in the *Phase 1 Final Report* and included here in figure 1 (Flannagan et al. 2016). In figure 1, the black dots are 1-s intervals of driving, while the dots joined by a red line are 3 s of driving during an SCE. The SCE clearly extends outside of the core area of driving behavior. Figure 5 through figure 7 provide alternate views of the same type of data, only the ellipse is shown in two dimensions from three viewpoints (projections on pairs of PC axes) and all of the crash-event driving points (blue triangles) and baseline-event driving points (brown circles) are shown. Figure 5 through figure 7 also show a sample ellipse in each view; blue points are often farthest outside each ellipse relative to baseline driving.

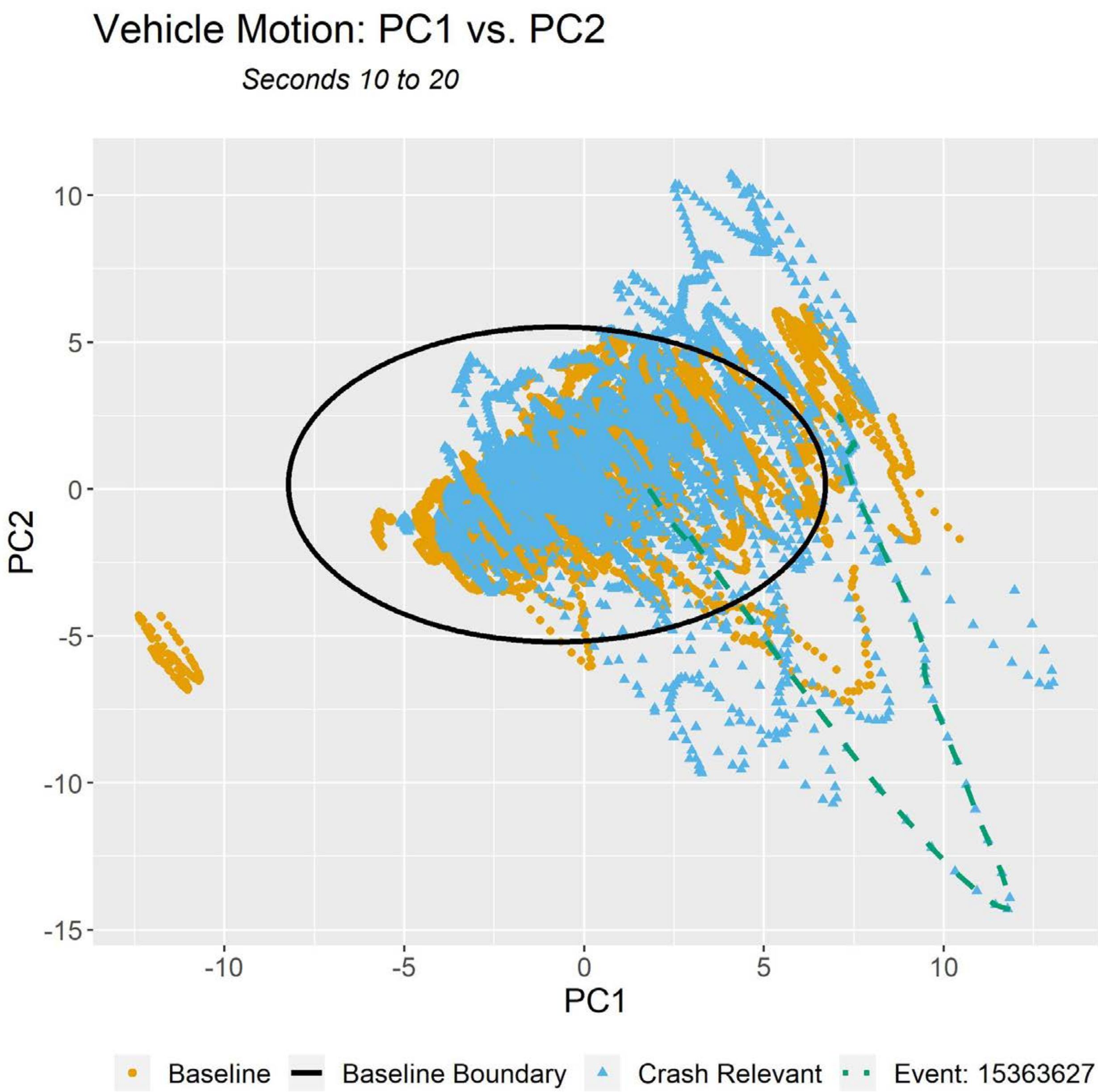


Figure 5. Graph. PC1 versus PC2 view of baseline and crash-relevant events with example "normal" driving bounding ellipse (modified from Flannagan et al. 2016).

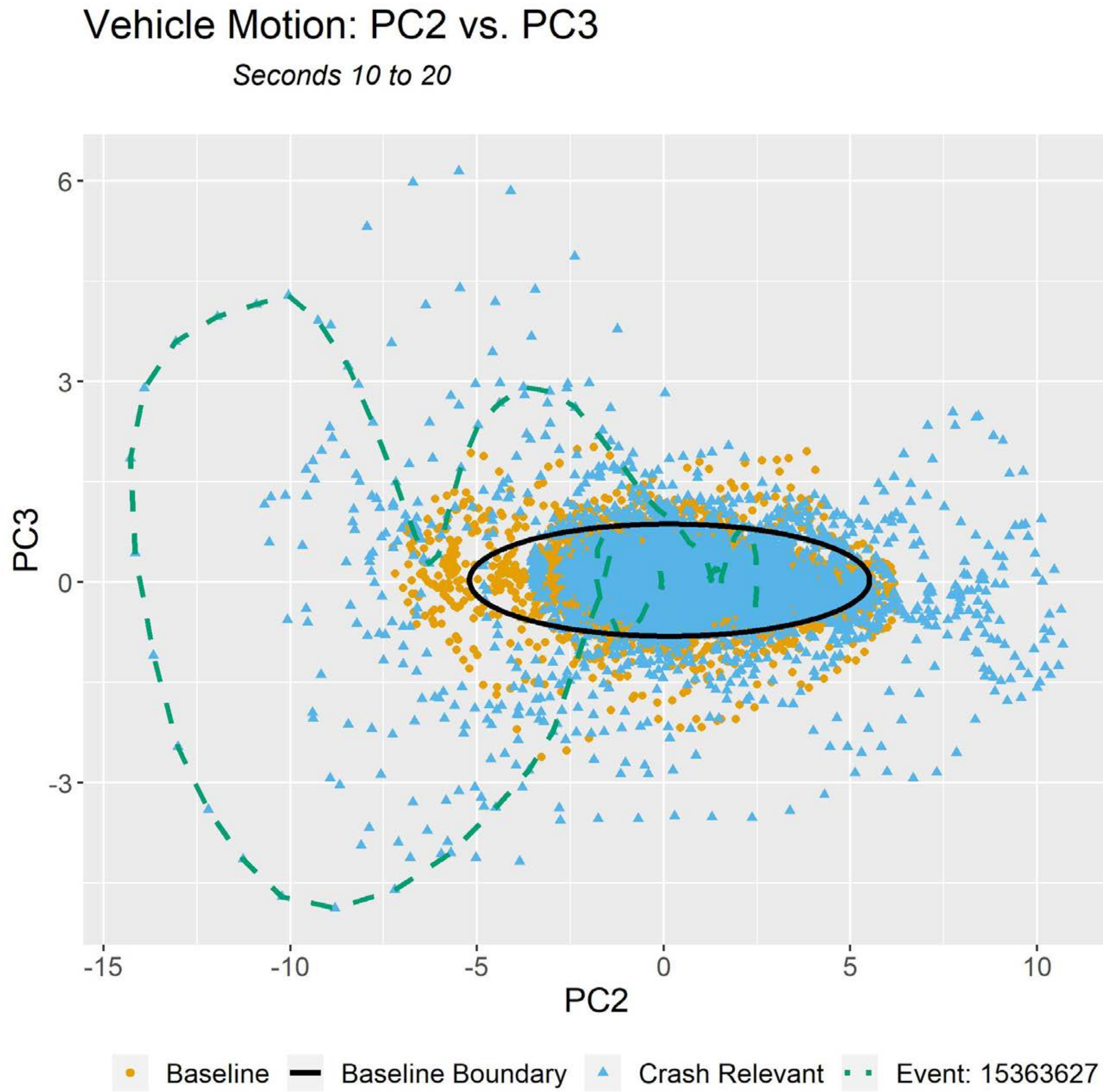


Figure 6. Graph. PC2 versus PC3 view of baseline and crash-relevant events with example "normal" driving bounding ellipse (modified from Flannagan et al. 2016).

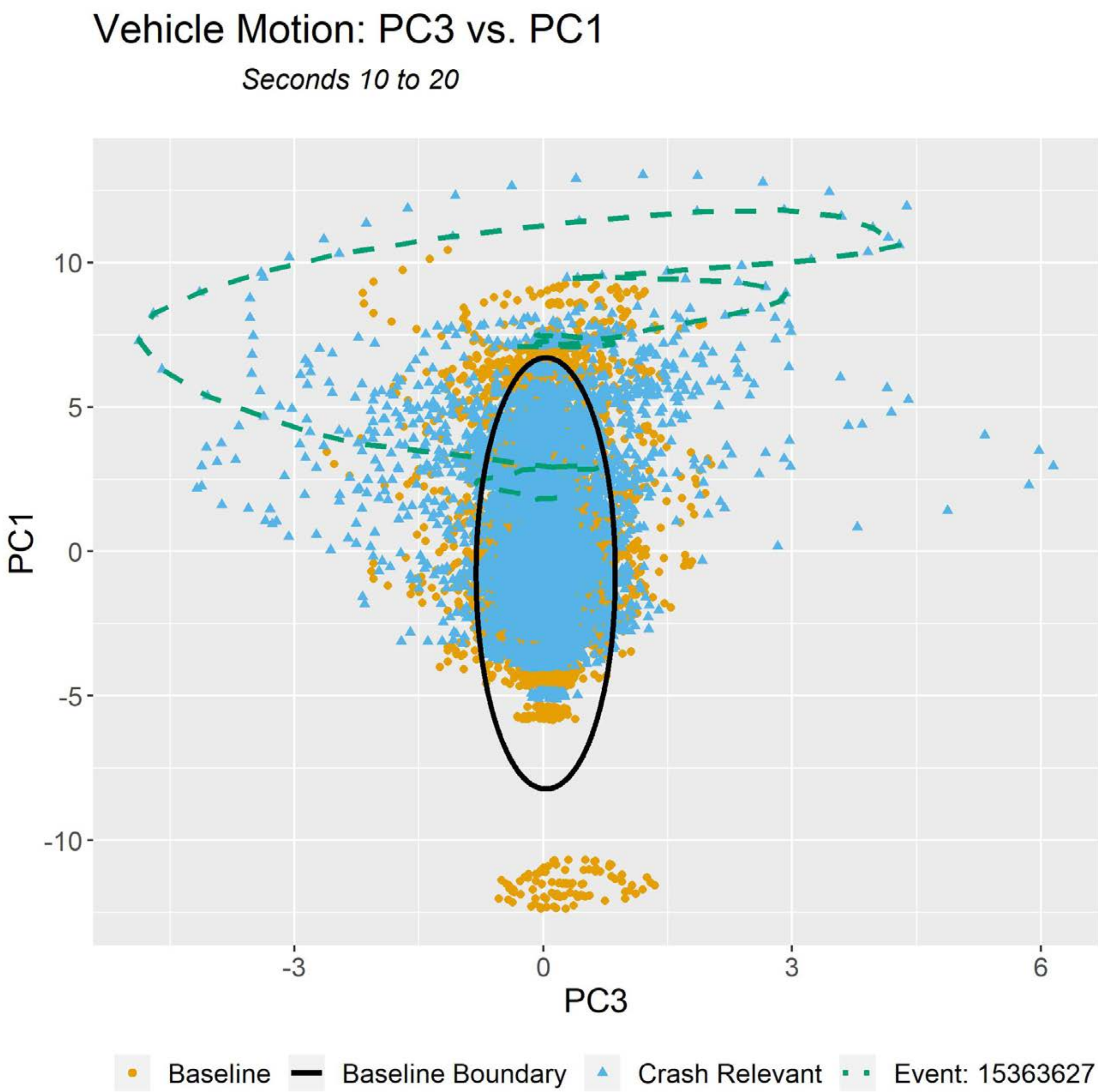


Figure 7. Graph. PC3 versus PC1 view of baseline and crash-relevant events with example "normal" driving bounding ellipse (modified from Flannagan et al. 2016).

When exploring the relationship between excursions from the normal-driving ellipse baseline driving events and SCE versus baseline event status, out-of-ellipse excursions could predict with reasonable accuracy (about 70 to 75 percent) which events would become crashes before the precipitating event occurred. Phase 2 continued to refine understanding of the performance of PCs as predictors of crashes. This included trying to generate earlier (pre-precipitating event) predictions of increased crash risk and further investigating the properties of out-of-ellipse excursions beyond the boundary of the normal-driving ellipse.

Predicting Future PCs to Identify Impending SCEs Earlier

One of the key explorations of the PCA method was to determine whether it was possible to predict the PCs themselves 1 s into the future, thus gaining more advance notice of impending issues. Working primarily with time series methods did not show promise at generating good predictions of the future. In general, out-of-ellipse excursions (beyond the ellipse boundary) were unpredictable and so had to be observed as they happened.

The results of these efforts to understand the prediction ability of PCs are shown in figure 8, which shows the area under the curve (AUC) profiles for three models developed using Bayesian additive regression trees (BART). When interpreting this plot, AUC refers to the area under the receiver operating characteristic curve; values above 0.5 indicate that results are better than random chance. The "predicted" model is an autoregressive-moving-average (ARMA), which is a model for time series data that includes an autoregressive component and a moving-average component. The vector autoregression (VAR) is an autoregressive model for vector-valued time series. For two time series models (predicted and VAR), time point 17 on the x-axis means that the observed PCs from 2–16 s and the time series-predicted PCs for the 17th s were used to generate the model predicting whether the event was baseline or SCE. This process could be described as "time-traveling" 1 s into the future and using the predicted future state to predict whether a crash would occur later (within a few s). (Recall that the precipitating event always occurs at 20 s and SCEs occur sometime shortly after that.) Similarly, for time point 18 s, 2–17 s of observed PCs and the 18th s of predicted PCs classified the event. The rest of the time points were similar. The time series "extended" represents the original data. That is, at 17 s on the x-axis, 2–17 s of the PC data were used; at 18 s, 2–18 s, and so on. Results suggest that the time series models do as well as the reference model (extended) in terms of prediction up to the 19th s; but, for the 20th s and beyond, they do not do as well as the original data. Although the time series model does not do a good job of predicting when things will go awry, it does predict the general conditions that increase the risk of an SCE. That is, the AUC is well above chance (where AUC = 0.5) and the time series models provide the same prediction level as the original data.

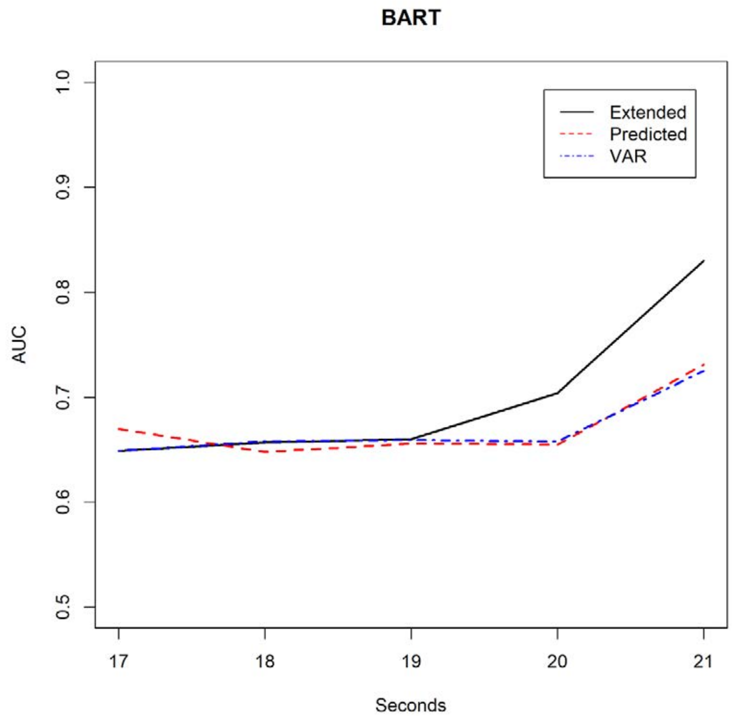


Figure 8. Graph. AUC profile of reference (original data), predicted (ARMA) method, and VAR method.

Several more approaches attempted to improve predictions at the 20th and 21st s. First, adding road characteristics to the SCE model allowed assessment of which characteristic improved SCE prediction the most. Second, the last observation of the predicted PCs were included in the SCE model to determine if it boosted prediction performance. Whether the vehicle ever went out of the PC boundary was also included as a predictor in the SCE model. Finally, additional time series methods were evaluated. The result of these explorations is that the best model was actually just using the original PC data plus traffic free flow and intersection presence to predict whether a crash would occur in the future.

Investigating the Properties of Out-of-Ellipse Excursions

After finalizing the SCE model, analysis turned to evaluating the properties of out-of-ellipse excursions beyond the PC boundary with a focus on how excursions relate to SCEs and abnormal driving behavior. The hypothesis guiding this analysis was that excursions to a certain quadrant of PC space outside the PC boundary may be related to "unsafe" driving. An example of a quadrant is the space defined by PC1 > 0, PC2 > 0, and PC3 > 0. In total, there are eight such quadrants defined by PCs 1 through 3.

A first step assessed the possibility of further reducing the PC space and hence reducing the number of quadrants under consideration. This suggested some form of dimensional reduction by removing one of the PCs from the analysis. Each PC was removed one at a time, followed by assessing the prediction performance of only the other two. Based on this analysis, removing PC1 from the three-dimensional space produced the smallest decrease in the number of events that traversed outside the PC boundary. This was closely followed by removing PC2 from the three-dimensional space. Removing PC3 from the three-dimensional space produced the largest change in number of out-of-ellipse events traversing outside the PC boundary. These results suggest that removing PC1 from the three-dimensional space to focus on an ellipse in the PC2-PC3 space would provide reasonable prediction performance.

Under the newly-defined PC boundary, examination turned to the proportion of out-of-ellipse excursions using the four quadrants defined by the PC2-PC3 space, illustrated in figure 9. Events included those that did not have free flow, defined as flow with some restrictions, stable flow, maneuverability and speed more limited, and unstable flow or forced conditions stratified by the event being baseline or nonbaseline. No free flow is a topic of interest because very few SCEs occur under the free flow scenario. Consistent with previous results, a higher proportion of nonbaselines will traverse out-of-ellipse compared to baselines. In addition, excursions tend to occur in the +- and -- scenarios (i.e., quadrants I and IV) 1 to 2 s before the precipitating event. For the +- scenario, this suggests that the vehicle was accelerating but had to exert a large decrease in acceleration, presumably to avoid a nonbaseline event. For the -- scenario, this suggests that although the vehicle was already decelerating, it was not decelerating enough to avoid a nonbaseline event and therefore had to further decrease acceleration.

Quadrant	PC3>0	PC3<0
PC2>0	I (++)	II (+-)
PC2<0	III (-+)	IV (--)

Figure 9. Diagram. Illustration of PC quadrants and boundaries.

ANALYSIS OF NEW DATA

New PC Fit Summary

For the new dataset, 1.5-s stretches were selected from each pass to train PC fit. Each datapoint consisted of a 1-s history of speed and longitudinal acceleration yielding a 20-component PC fit. One pass, a baseline, had no usable speed data and could not contribute to the analysis regardless of the interval selected. All other cases provided 15 datapoints from some point in the pass except where data were missing during the selected interval. Before fitting, the input variables were centered and scaled as in the previous analysis.

The loadings for the first three components from an example PC fit are shown in table 6 below. The first PC is an average of the 20 input variables weighted strongly toward the speed values. The second PC is effectively an average of the longitudinal acceleration values with a small differencing element from the speed. The third component has almost no contribution from speed and is the difference between the longitudinal acceleration in the early and late portion of the

1-s window, weighted toward the extremes. These results are consistent with the PC fit derived from the Phase 1 work.

Table 6. Variable loadings for the first three PCs.

Variable and Time Point	PC1	PC2	PC3
Longitudinal acceleration 0.0	-0.057	-0.306	-0.453
Longitudinal acceleration -0.1	-0.059	-0.309	-0.406
Longitudinal acceleration -0.2	-0.061	-0.312	-0.301
Longitudinal acceleration -0.3	-0.064	-0.313	-0.181
Longitudinal acceleration -0.4	-0.066	-0.313	-0.063
Longitudinal acceleration -0.5	-0.070	-0.312	0.060
Longitudinal acceleration -0.6	-0.073	-0.310	0.179
Longitudinal acceleration -0.7	-0.074	-0.308	0.303
Longitudinal acceleration -0.8	-0.076	-0.305	0.408
Longitudinal acceleration -0.9	-0.078	-0.300	0.456
Speed 0.0	-0.310	0.060	-0.006
Speed -0.1	-0.310	0.062	-0.005
Speed -0.2	-0.310	0.064	-0.005
Speed -0.3	-0.309	0.065	-0.005
Speed -0.4	-0.309	0.067	-0.006
Speed -0.5	-0.309	0.069	-0.007
Speed -0.6	-0.308	0.070	-0.008
Speed -0.7	-0.308	0.072	-0.009
Speed -0.8	-0.308	0.073	-0.011
Speed -0.9	-0.307	0.075	-0.013

The resulting PC fit was then applied to the full dataset. Due to missingness, the PCs could not be calculated for approximately 9.1 percent of the rows (dropping to 8.5 percent when omitting the baseline without speed data).

Ellipse Definition

Unlike the previous analysis, the normal driving ellipsoids in this analysis were defined for each work zone, rather than overall. To do this, the means and standard deviations of the first three PCs were calculated for each work zone using only the baseline passes through that work zone. The centers of the ellipsoids were determined using the PC means while the radii were defined as three times the standard deviation of the appropriate PC. For the three-dimensional ellipsoid, the formula for the boundary is:

$$\left(\frac{x-\mu_1}{3\sigma_1}\right)^2+\left(\frac{y-\mu_2}{3\sigma_2}\right)^2+\left(\frac{z-\mu_3}{3\sigma_3}\right)^2=1$$

Figure 10. Equation. Equation for three-dimensional ellipsoid for the first three PCs.

Where:

x, y, and z = the values of the PCs.
μ_i = the mean of the i^{th} PC.
σ_i = the standard deviation of the i^{th} PC.

Based on the results of the previous analysis, the first PC does not contribute strongly to the effectiveness of the ellipsoid, so reducing to an ellipse is reasonable. This simplifies the equation to:

$$\left(\frac{y-\mu_2}{3\sigma_2}\right)^2+\left(\frac{z-\mu_3}{3\sigma_3}\right)^2=1$$

Figure 11. Equation. Simplified equation for two-dimensional ellipsoid to represent the second and third PCs.

Where:

y and z = the values of the PCs.
μ_i = the mean of the i^{th} PC.
σ_i = the standard deviation of the i^{th} PC.

Using this equation, all datapoints can be characterized as inside or outside the ellipsoid.

Aligning to D26.5 Window

With the ellipses defined, the aligned passes were reduced from the D30 window to a D26.5 window where the locations of the vehicles were known and aligned within work zones. This removed the ends of some of the baseline passes, but since the ellipse work was all focused on the pre-event period, this portion of the data was not being used. By switching from temporal to spatial units, it was possible to compare the locations where the vehicles have shown out-of-ellipse patterns to examine how similar that behavior is across passes.

Three examples of such alignments are provided in figure 12 through figure 14. In each plot, the solid red line represents an SCE for a particular driver. Dotted blue lines represent other baseline passes for the same driver, while dashed green lines represent baseline passes for different drivers. Any point where a symbol occurs on the line is a location where the vehicle showed out-of-ellipse driving patterns. The shaded box, outlined in orange, shows the period in which the SCE occurred.

A number of specific observations can be drawn from these plots:

- SCE 10612575 (figure 12): most of the baseline passes appear to traverse this work zone with little difficulty, with only three displaying out-of-ellipse driving patterns. Of the three, two belong to the SCE driver and demonstrate very similar speed patterns through the work zone. The final has much more modest excursions and appears to be experiencing traffic queueing.
- SCE 136278188 (figure 13): this small work zone occurs between an interstate off-ramp and a stoplight, so most of the vehicles have to slow for the stoplight and/or traffic. As such, many of the passes exhibit ellipse departures in the same general area. In a situation like this, the ellipse may be less able to distinguish problematic driving because the space is not suited to smooth acceleration patterns.
- SCE 29877160 (figure 14): in this work zone, the SCE pass has no ellipse departures before the SCE occurs, while other drivers do. This seems to have happened because the subject driver was not at fault and the SCE occurred when the driver took evasive action due to another driver entering their lane with insufficient space. Other passes through the work zone demonstrate out-of-ellipse driving, but it does not seem to be more common for the subject driver than the others.

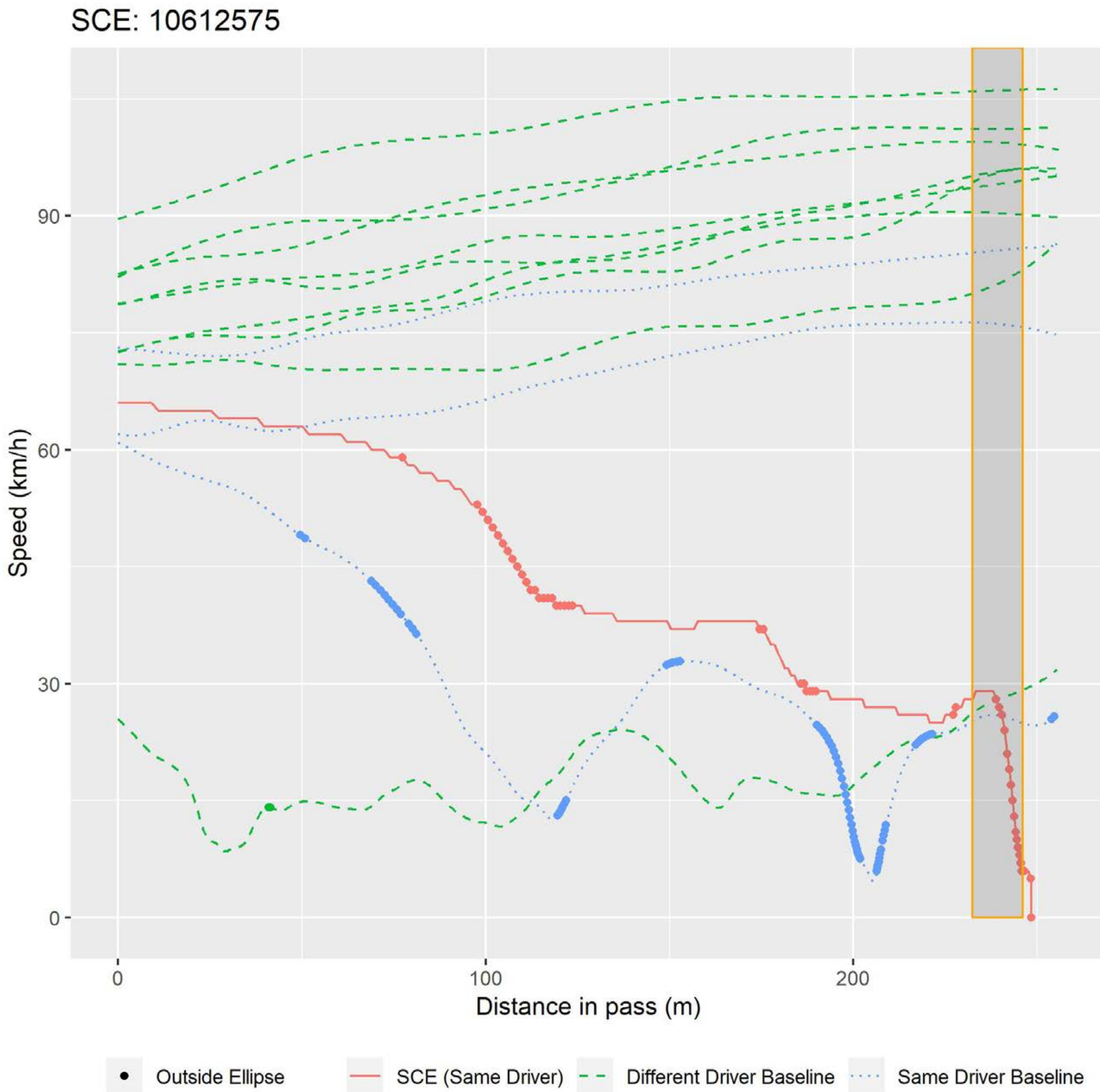


Figure 12. Graph. First example of work zones showing aligned speeds for baseline and SCE with ellipse excursions marked.

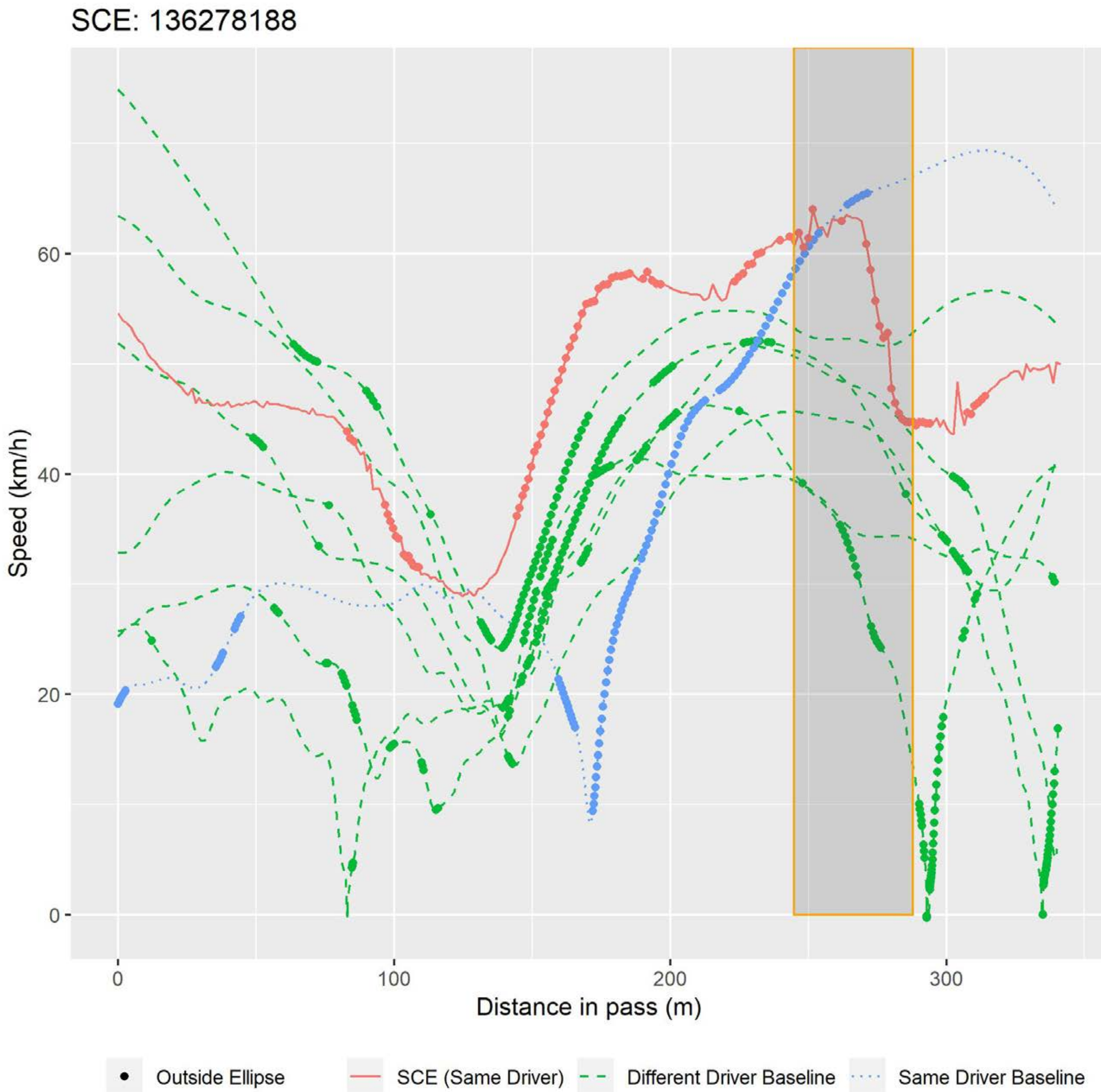


Figure 13. Graph. Second example of work zones showing aligned speeds for baseline and SCE with ellipse excursions marked.

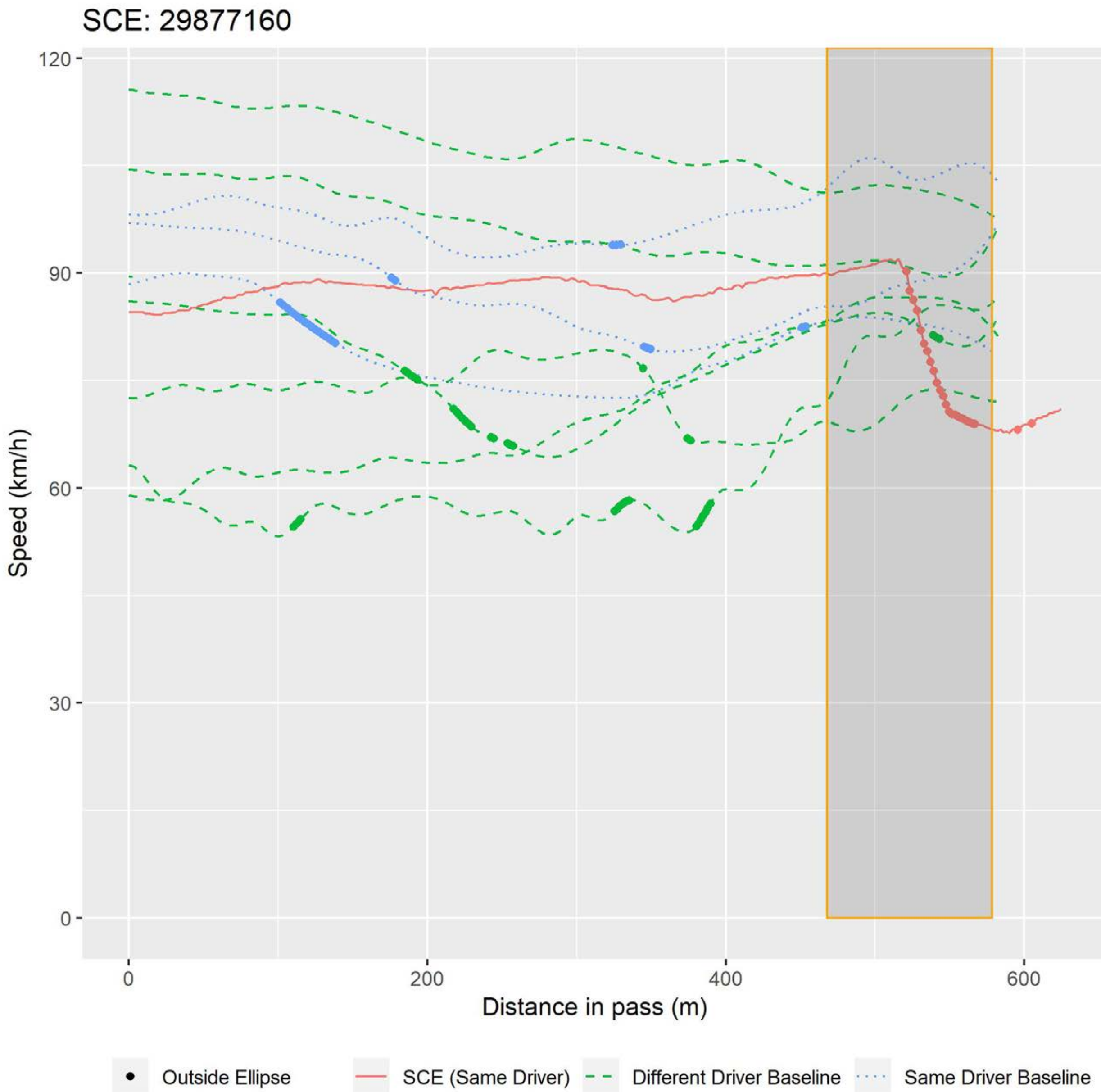


Figure 14. Graph. Third example of work zones showing aligned speeds for baseline and SCE with ellipse excursions marked.

These three sets of passes demonstrate a number of features regarding out-of-ellipse driving. First, when the structure of the work zone or the roadway in general requires substantial changes in speed or acceleration, the reliability of the ellipsoid as an indicator of problematic driving decreases. While the amount of out-of-ellipse driving may indicate poor behavior, it is not reasonable to expect driving in these areas to fall within the ellipse. Second, when drivers are not at fault, their driving may be sufficiently stable that they fall within the ellipse right up until the SCE. This is not required, and many of the not-at-fault crashes appear to depart from the normal driving ellipse during the pre-event period. In this dataset, at-fault crashes form the majority of

SCEs, with the subject driver determined to be at fault in 68 percent of cases. Finally, it appears that driver behavior is consistent, and individuals who frequently depart the ellipse will continue to do so on other passes through the work zone.

The relationship between out-of-ellipse driving and SCE status can also be examined using conditional logistic regression. Conditional logistic regression uses matched sets of data to control for confounding effects by allowing the risk to differ between strata and focusing on the effect of the variables of interest. In the context of this analysis, the work zones serve as strata with all passes through a given work zone having comparable experiences. The outcome of the regression is the presence of an SCE in the pass, with the time bin and ellipse departures serving as predictors. The driver effect cannot be used because all SCEs occur for the "same driver" level of the variable. As such, inclusion of the variable leads to perfect separation of a segment of the baselines leading to an unstable model fit. Similarly, the time variable must be changed from a binary same/different coding to the three-level categorical coding to avoid the same issue. Finally, since a binary coding of the ellipse departures is no longer required, the percentage of the pre-event out-of-ellipse driving time is used. The results of this regression are shown in table 7.

Table 7. Coefficients and odds ratios for model of SCE versus baseline.

Effect	Estimate	Standard Error	z Value	Pr (>|z|)	Odds Ratio
Percent out-of-ellipse	0.0603	0.0181	3.33	0.0009	1.06
Weekday off-peak	-0.5753	0.4140	-1.39	0.1646	0.56
Weekend/holiday	-2.6526	1.0381	-2.56	0.0106	0.07

Pr = probability under the null hypothesis of equal odds; z = z-score.

Based on these data, there is a significant difference between the time bins ($X^2 = 3.97$, $p = 0.0009$), but that difference seems to be predominantly for the weekend/holiday bin, which has only 2 of the 40 SCEs. The percent of time with out-of-ellipse driving is also significant, with each percentage of the pre-event driving occurring outside the normal driving ellipse, increasing the odds that an SCE occurred by 6 percent. Out-of-ellipse driving exceeds 1 percent of total driving in 60 percent of SCE passes, and 39 percent of baseline passes, with approximately 30 percent of the latter group being the drivers experiencing SCEs. Based on these results, it appears that out-of-ellipse driving is associated with the presence of SCEs, but it does not guarantee poor outcomes for the pass.

Driver, Time, and SCE Effects

To examine the effect of SCE status, driver, and time of day on out-of-ellipse driving, the passes were limited to the pre-event period, which were determined based on travel distance into the work zone. After aligning the data using the odometry and limiting to the pre-event period, only 310 baseline passes remained, covering 40 of the SCEs. Table 8 shows the relative proportions of cases that depart the ellipse prior to the SCE location split on four different factors, which are denoted by color in the table. The odds ratios and upper and lower bounds of the 95 percent confidence intervals are given for each comparison.

Table 8. Ellipse excursions as a function of different factors.

Variable	Never Outside (Count)	Never Outside (Percent)	Outside 1+ Steps (Count)	Outside 1+ Steps (Percent)	Total	Odds Ratio	2.5 Percent Odds Ratio Lower Bound	97.5 Percent Odds Ratio Upper Bound
Baseline	174	56.13	136	43.87	310	2.38	1.19	4.72
SCE	14	35.00	26	65.00	40			
At fault (SCE only)	12	40.00	18	60.00	30	2.67	0.48	14.79
Not at fault (SCE only)	2	20.00	8	80.00	10			
Different driver	149	60.57	97	39.43	246	2.56	1.60	4.11
Same driver	39	37.50	65	62.50	104			
Different time bin (wrt SCE)	124	61.39	78	38.61	202	2.09	1.36	3.21
Same time bin (wrt SCE)	64	43.24	84	56.76	148			

wrt = with respect to.

The following observations can be made from table 8:

- SCE versus baseline: driving outside of the normal driving ellipse is more strongly associated with SCEs than baselines. The odds of a pass containing an SCE including problematic driving are 2.38 times higher than that of a baseline pass.
- Fault: as mentioned above, while some not-at-fault SCE passes do not demonstrate any out-of-ellipse driving, the majority do. Based on the available data indicating nonsignificant *p*-value and wide confidence intervals, there is not sufficient evidence to conclude that there is a difference between the odds of problematic driving when the subject vehicle was at fault compared to when it was not.
- Driver effect: the odds ratio for the driver effect is more extreme than the baseline versus SCE effect, with drivers involved in an SCE having odds 2.56 times than those of non-SCE-involved drivers. This may indicate that certain drivers behave in a way that produces risky situations and may be prone to SCEs when other factors occur on the roadway.
- Time effect: passing through a work zone during a time similar to when an SCE has occurred appears to increase the odds of abnormal driving by a factor of 2.09. This indicates that certain elements of those timeframes (such as increased congestion) may increase the frequency of ellipse excursions regardless of driver behavior.

While the driver, time, and SCE effects appear to be significant, it is reasonable to hypothesize that not all of these effects will be significant when adjusting for the others. One way to address this concern is by modeling the variables together. To examine this possibility, logistic regressions for the probability of ellipse excursions based on combinations of driver, time bin, and SCE status were constructed. When all three effects are included in the model, the SCE effect is nonsignificant ($p = 0.66$). Additionally, the driver and time bin effects appear to be additive, with a proposed interaction term being nonsignificant in the model ($p = 0.81$). This is consistent with the supposition that SCE-prone drivers engage in risky driving at a relatively fixed rate and do not correct that behavior when in difficult driving conditions. Under this interpretation, SCEs would occur when those drivers encounter an unobserved trigger, possibly related to the other drivers on the road. As such, the model fit and odds ratio estimates for each effect, adjusted for the others, are provided in table 9.

Table 9. Coefficients and odds ratios for model of ellipse excursions (present or absent).

Effect	Estimate	Standard Error	z Value	Pr (>|z|)	Odds Ratio
(Intercept)	-0.6245	0.1552	-4.02	0.0001	N/A
Same driver	0.7941	0.2490	3.19	0.0014	2.21
Same time bin	0.5594	0.2291	2.44	0.0146	1.75

N/A = not applicable; Pr = probability under the null hypothesis of equal odds; z = z-score.

These effect sizes are consistent with those from the individual odds ratios shown above, each showing about a 15 percent decrease in magnitude after adjusting for the others. The driver effect remains larger than the time effect by about 25 percent, supporting the conclusion that the behavioral element is a larger contributor than circumstances.

Video Annotation versus Normal Driving Ellipse

Using the video data available, a set of annotations was compiled for driving inside work zones. Annotations included the presence of barrels/cones and barriers, lane closures, and workers. All elements were coded on an s-by-s basis, so each unit of time an element was present in a video converts to 10 observations in the data. As such, it is possible to compare the proportion of out-of-ellipse driving when the work zone elements are present or absent. Table 10 provides these values for driving in the D26.5 period described above.

Table 10. Relationship between ellipse excursions and work zone elements.

Work Zone Element	Observed Time (Deciseconds)	Proportion Outside: Absent	Proportion Outside: Present	Ratio (Present/Absent)
General proportion outside	N/A	0.094	N/A	N/A
Text sign	10	0.094	0.000	0.00
Warning lights	63	0.094	0.000	0.00
Police vehicle	10	0.094	0.000	0.00
Crossroad	10	0.094	0.000	0.00
Merging	472	0.095	0.052	0.55
Barrier	7,819	0.099	0.058	0.58
Traffic signs	3,100	0.096	0.060	0.63
Parked cars	150	0.094	0.085	0.90
Equipment	1,617	0.094	0.093	0.99
Exit lane	82	0.094	0.103	1.09
Turn arrow sign	90	0.094	0.111	1.18
Digital sign	308	0.094	0.113	1.20
Work zone signs	2,234	0.093	0.128	1.38
Barrels or cones	18,062	0.080	0.136	1.71
Queue forming	3,194	0.092	0.149	1.63
TMA	40	0.094	0.176	1.87
Workers present	60	0.094	0.183	1.95
Closed lane	3,630	0.089	0.189	2.13
Entering the motorway	40	0.094	0.450	4.79
Exiting the motorway	20	0.105	0.450	4.28
Decreased lane	20	0.094	0.500	5.32
Lane opening	10	0.094	0.700	7.44
Chicane	10	0.094	0.800	8.50

N/A = not applicable; TMA = truck-mounted attenuator.

Many of the recorded work zone elements are rare and only occur for a few s, but a number produce substantial changes in the rate of abnormal driving. Several elements, such as workers present, queue forming, and lane opening, produce the expected increase in the rate of out-of-ellipse behavior. Others, like warning lights, crossroads, and merging, actually decrease the rate at which driving out-of-ellipse occurs. While this may be a real effect, it is also possible that some of these elements propagate up the traffic stream so that their effect is reduced before they appear in the annotation. For instance, while crossroads frequently have stoplights or other traffic controls that could lead to driving patterns that depart the ellipse (see SCE 136278188 in figure 13), those ellipse excursions may not occur at exactly the point that the crossroad is encountered. Similarly, when merging is required, it can cause flow problems upstream from the merging location before the vehicle actually merges. Conversely, lane openings and queues should have an impact much closer to where the subject vehicle encounters them.

CHAPTER 4. STM APPROACH

OVERVIEW OF THE TOPIC MODELING APPROACH

Phase 1 explored how topic modeling, which was developed to analyze text data, can be applied to naturalistic driving data (NDD). This approach has previously been used for driving data by McLaurin et al. (2018). The goal of topic modeling is to use co-occurrences of words in documents to define topics, which are essentially collections of words that "go together." The approach is probabilistic (specific words may or may not appear in a document), in contrast to hierarchical clustering, where specific words must be present to be categorized in a particular group. Instead, in topic modeling, topics are groups of words that are more likely to be seen together than expected by chance. This probabilistic quality is consistent with the use of language.

In Phase 1, the specific topic modelling algorithm LDA was applied, and analysis incorporated both kinematic and categorical variables. In Phase 2, further exploration of methods (discussed in the *Methods Interim Report: Reporting Task D3.1 for Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones* led to focusing only on categorical variables and to using the STM topic modeling algorithm (Kovaceva et al. 2017).

The use of STM (and other topic models) differs from the traditional critical-events approach to analyzing NDD to identify trip segment patterns in the behavior of drivers (in this case, passing through work zones). The STM approach describes all the data from a driver's trip segments instead of just the short periods of the trip where a critical event happened (e.g., 6 s). The approach also enables the discovery of unknown driving patterns by not requiring the analyses to have predetermined features that will be extracted from the raw data (e.g., predetermine to extract all instances where acceleration exceeded a threshold).

To discover long-range patterns in large volumes of highly variable driving data, the data need to be transformed in a way that reduces the volume and provides structure without subsetting the dataset (Wang et al. 2013). One common approach for such a data transformation is to summarize the distribution of the data using measures of central tendency. This approach provides simple metrics for comparison, but masks the subtle variability that constitutes the patterns of interest. An alternative approach is to use linear or nonlinear dimensional reduction methods to represent the data using fewer parameters (Van Der Maaten et al. 2009). Many of the driving patterns identified in controlled settings represent nonlinear changes in vehicle control (McLaurin et al. 2018). As a result, nonlinear dimensional reduction methods may be well suited to represent the driving data in a way that preserves the patterns of interest.

One of the nonlinear dimensional reduction methods that can produce interpretable results is the topic modeling method (Blei 2012). Topic modeling is an unsupervised dimensional reduction method designed to identify word co-occurrence patterns called latent topics (Steyvers and Griffiths 2007). In text applications, topics provide a meaningful overall description of a body of documents as well as a description of the individual documents. For example, topic modeling has been applied to the thousands of papers published in the journal *Science* to describe the papers according to their topic areas and to identify the domains of investigation that have defined science over the last 100 years (Blei and Lafferty 2007). Topic modeling has also been applied to population genetics, computer vision, and human activity recognition (Blei 2012, Farrahi and Gatica-Perez 2011, and Huynh et al. 2008). For the study's particular analysis, it is important to observe co-occurrences of work-zone-related (WZR) labels together with secondary task labels.

NDD contains categorical text elements in the form of annotated information taken from the recorded driving records. Weather conditions, traffic information, and secondary tasks are typical data that can be represented as text elements and analyzed with topic modeling. Even kinematic data transformed into discrete information can be included in these analyses (e.g., out-of-ellipse driving developed in PCA). The number of unique "words" will be limited to the coding schemes applied in the annotation process, but the frequencies of these elements are unique to each trip in the database.

STM

Phase 1 of this project evaluated LDA, the simplest implementation of probabilistic topic modelling (Blei et al. 2003). Phase 2 involved exploratory analyses of driving-data-as-text data describing the actual events with STM (Roberts et al. 2014).

Understanding the application of topic modeling to driving requires definitions of terms that have been adopted in the text-processing context:

- A word is the basic unit of discrete data, defined as an item from a vocabulary (i.e., a list of possible words). In this context, a word is defined as any actual word from the free-text narratives in the dataset, or any level of a categorical variable describing any portion of a driving epoch (SCE or baseline).
- A document is a sequence of words. In the driving context, a document is an event, and the sequence of words is the collection of all descriptors of a particular document (i.e., event). This study used only the D30 portion of each event in STM analyses and considered two segments of data within that portion of the event as described below.
- A topic is a probability distribution over the vocabulary of words. The topics represent collections of words that frequently occur together in documents. This is the key element to be implemented in the driving context and the key goal of STM. The question is whether certain words or groups of words occur with notably different probability in crash-relevant versus baseline events. If so, this identifies a potential mechanism for distinguishing between the two events.

Extending LDA, STM is a general framework for topic modeling that accounts for document-level covariate information (e.g., event types: baseline or SCE) and allows for correlations among topics. The covariates can improve inference and qualitative interpretability and are allowed to affect the proportion of the topics, topical content, or both. STM can estimate the effect of covariates on topic proportions. In practice, it means that STM can estimate whether a particular topic (e.g., secondary tasks in active work zone) is more prevalent in a certain event type (e.g., SCE). Furthermore, time-related covariates can be included to understand how the words used within a topic change over time; such covariates are called content covariates.

However, topic modeling is a form of unsupervised learning. This means that there are no ground truth data, making it challenging to validate the output of a topic model. Instead, topics are interpreted based on their most associated words.

DETAILS OF STM ANALYSES

Two approaches were considered to choose the segment of data to be used for STM analyses:

- Event-based approach (i.e., with reference to the location of SCEs).
- Inside/outside work zone approach (i.e., only include parts of data that are inside work zone or outside work zone, and compare if there are differences between inside and outside work zone).

Approach 1 (event-based) has been the common approach used in many analyses of naturalistic driving studies. The raw data were requested following this approach.

To decide whether approach 2 (inside/outside work zone) could be considered further, the number of events where the work zone start could be identified was checked (table 11). The results suggest that this approach is not viable since there are so many baselines without data on work zone start. The inside/outside work zone approach might be used in a future study with even longer events requested.

Table 11. Number of events with identified work zone start and end.

Event Type	SCEs	Baselines
Have identified work zone start	50	192
Do not have identified work zone start	2	251
Have identified work zone end	44	137
Do not have identified work zone end	8	306

The spatially aligned road segments covering vehicle motion for 21.5 s prior to the time of minimum acceleration to the end of the SCE (i.e., the D30 zone shown in figure 3) were used in the analysis of each event. The length (in time) of each baseline event is not necessarily the same as the length (in time) of its matching SCE; however, the spatial length (in distance) would be the same. This is due to different travel speeds through the same location. Furthermore, the lengths (in time) of different SCEs are not the same either, since each SCE may have a different travel speed. The expectation was that STM could provide a clearer picture of what conditions/activities co-occur that can differentiate SCEs from their matching baselines in the same exact location.

From 444 total events, only 382 (335 baselines and 47 SCEs) could be spatially aligned. Therefore, only 382 events were used for this experiment. The analysis was done with the STM package developed by Roberts et al. (2014). Annotation data were used as the main input data for this experiment. Kinematic data were used by including the variable "out-of-ellipse," which was derived from the PCA analysis. Excursions from the normal driving ellipse were considered "abnormal" or potentially unsafe driving. Finally, punctuation and whitespace were first removed from the input data. The event type was selected as the covariate to be studied.

The number of topics must be defined as an input to analysis, and identifying the ideal number of topics is challenging since there is no single correct way to address this issue. The selection of the number of topics to use is normally done by first running the model with several different numbers of topics and then inspecting the values of the STM diagnostic parameters from the resulting models, taking into account the total number of events (fewer than 400 events can be used). Several models were tested with different numbers of topics (2, 5, 10, 15, and 20) and were compared in terms of exclusivity and coherence. Exclusivity is a measure of the extent to which topics are composed of terms that are not shared with other topics, and coherence describes the similarity between words in each topic (Roberts et al. 2014). The models with larger exclusivity and coherence are more desirable. However, as the number of topics increases, exclusivity generally increases and coherence generally decreases. Based on a balance between these two measures, 10 topics were selected.

After applying STM, specific words were linked to specific topics. The topics were labelled as topic 1 through topic 10. Intuitive or meaningful topic labels were not automatically generated. The analyst determined what the topics represent based on the frequency of the different words within the topics. This required examining each topic with a collection of words that are associated with the topic. If, for example, the words asleep, emotion, and angry were linked to a topic, the topic label of impairments is appropriate because these are several categories of impairments in the data. In addition to labeling the topics, suitable experts reviewed the topic compositions and interpreted their implications.

The determination of the words associated with each topic can be done in several ways. Because of this, the STM package provided the words associated with each topic separately based on four different measures. These four measures are:

- Highest probability. This straightforward approach selects the words that have the highest probability of occurring in a particular topic. The problem here is that certain words such as event and vehicle will show up as high-probability words for many topics, and words such as have or get that are universal add relatively little value when separated from the words around them (in text).
- Frequency and exclusivity (FREX). Defined as the ratio of word frequency in a topic to word-topic exclusivity (Bischof and Airoldi 2012). FREX considers how often a term occurs in a topic, but then adjusts this based on the degree to which the term is exclusive to that topic.
- Lift. Refers to the probability of word occurrence in a topic divided by the probability of word occurrence across the corpus. Lift weights words more heavily if they occur infrequently in other topics. The problem with this metric is that words that appear infrequently are likely to score well.
- Score. Logarithm of word frequency within a topic divided by the logarithm of word frequency in other topics.

FREX, lift, and, score all incorporate word frequency within a topic and adjust by their frequency in other topics. The mechanisms of the adjustments are different, which results in different sets of words used to define topics (Roberts et al. 2014). In this report, all metrics are included since a single best statistic was not identified.

To clarify how topics are derived from the variables available in the data, figure 15 depicts two example D30 passes through work zones (labeled $D30_x$ and $D30_z$). Each work zone pass contained common words that were together defined as a topic (topic 1). It is important to note that STM does not take into account the order in which these words occur in the dataset. As shown in figure 15, the three annotated variables "out-of-ellipse," "cellphone dialing (handheld)," and "biting nails/cuticles" occurred within the D30 zone at different locations and could be observed in any order within the D30 zone. Thus, the words in a topic cannot be used to determine causality patterns among the words.

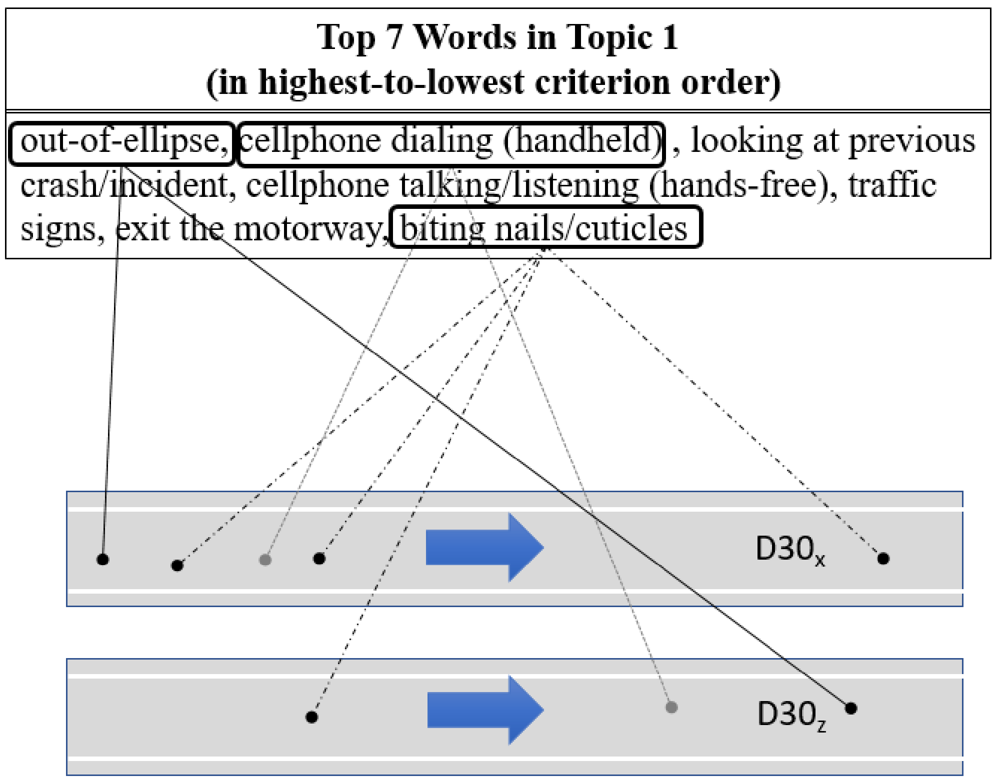


Figure 15. Graph. Variable spatial arrangement in topics.

RESULTS

The 10 topics from the analysis are shown in table 12. Each topic is shown here with the three words that have the highest probability in that topic and the expected proportion of that topic in the whole set of events included for the analysis. Topic 5 was the most prominent, while topic 2 was the least prominent.

Table 12. Expected topic proportions (three words with highest probability are shown with the topic number).

Topic Number	Three Key Words	Probability
5	Barrels/cones start, talking/singing audience unknown, biting nails/cuticles.	0.161
9	Traffic signs, crossroad, sign text.	0.140
8	Both hands off, queue forming, cellphone browsing.	0.122
1	Out-of-ellipse, cellphone dialing (handheld), looking at previous crash/incident.	0.112
10	Equipment, barrels/cones start, passenger in rear seat interaction.	0.095
4	Barrier start, speed limit, free flow (no lead traffic).	0.090
7	Other personal hygiene, other external distraction, other nonspecific internal eye glance.	0.085
3	Passenger in adjacent seat interaction, digital sign, cellphone talking/listening (handheld).	0.081
6	Adjusting/monitoring radio, speed limit in work zone, dancing.	0.061
2	Closed lanes, parked car, flow with some restrictions.	0.053

The 10 topics shown in table 12 are also listed in table 13, together with a designation of whether topics are associated with an SCE, baseline, or both. Table 13 also includes intuitive topic labels, developed using judgement from the research team. Table 14 through table 23 list the 7 highest ranked words based on the 4 measures for each of the 10 topics.

Table 13. Identified topics, event types, and labels.

Topic Number	Event Type	Label
1	SCE	Out-of-ellipse accompanied by both internal and external distraction.
2	SCE	Internal distraction in active work zone.
3	both	In-vehicle interaction with passengers (auditory engagement).
4	baseline	Free flow with secondary tasks related to food.
5	both	Mild distraction.
6	baseline	Free flow in work zone.
7	baseline	Free flow with nonspecific secondary tasks.
8	SCE	Secondary tasks in conjunction with queue forming.
9	baseline	External distraction related to signs.
10	baseline	Interaction with passengers/child in rear seat.

Table 14. Criterion and top words for topic 1: out-of-ellipse accompanied by both internal and external distraction.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Out-of-ellipse, cellphone dialing (handheld), looking at previous crash/incident, cellphone talking/listening (hands-free), traffic signs, exit the motorway, biting nails/cuticles.
Lift	Cellphone talking/listening (hands-free), cellphone dialing (handheld), out-of-ellipse, looking at previous crash/incident, exit the motorway, traffic signs, biting nails/cuticles.
FREX	Cellphone dialing (handheld), out-of-ellipse, looking at previous crash/incident, cellphone talking/listening (hands-free), exit the motorway, barrels/cones start, traffic signs.
Score	Out-of-ellipse, cellphone talking/listening (hands-free), cellphone dialing (handheld), looking at previous crash/incident, traffic signs, exit the motorway, barrels/cones start.

FREX = frequency and exclusivity.

Table 15. Criterion and top words for topic 2: internal distraction in active work zone.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Closed lanes, parked car, flow with some restrictions, workers present, enter the motorway, moving object in vehicle, decrease lane.
Lift	Chicane, closed lanes, drowsy/sleepy/asleep/fatigued, other known secondary task, enter the motorway, moving object in vehicle, parked car.
FREX	Closed lanes, parked car, flow with some restrictions, workers present, enter the motorway, moving object in vehicle, TMA.
Score	Closed lanes, parked car, enter the motorway, other known secondary task, flow with some restrictions, workers present, moving object in vehicle.

FREX = frequency and exclusivity.

Table 16. Criterion and top words for topic 3: in-vehicle interaction with passengers (auditory engagement).

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Passenger in adjacent seat interaction, digital sign, cellphone talking/listening (handheld), work zone sign start, road lanes, exit the motorway, adjusting/monitoring climate control.
Lift	Digital sign, road lanes, work zone sign start, passenger in adjacent seat interaction, exit the motorway, cellphone talking/listening (handheld), adjusting/monitoring climate control.
FREX	Digital sign, passenger in adjacent seat interaction, work zone sign start, cellphone talking/listening (handheld), road lanes, exit the motorway, adjusting/monitoring climate control.
Score	Passenger in adjacent seat interaction, digital sign, road lanes, cellphone talking/listening (handheld), work zone sign start, exit the motorway, adjusting/monitoring climate control.

FREX = frequency and exclusivity.

Table 17. Criterion and top words for topic 4: free flow with secondary tasks related to food.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Barrier start, speed limit, free flow (no lead traffic), eating without utensils, drinking from open container, smoking cigar/cigarette, unknown type secondary task present.
Lift	Eating without utensils, drinking from open container, smoking cigar/cigarette, speed limit, free flow (no lead traffic), unknown type secondary task present, barrier start.
FREX	Speed limit, free flow (no lead traffic), eating without utensils, barrier start, drinking from open container, smoking cigar/cigarette, unknown type secondary task present.
Score	Barrier start, speed limit, smoking cigar/cigarette, free flow (no lead traffic), eating without utensils, drinking from open container, unknown type secondary task present.

FREX = frequency and exclusivity.

Table 18. Criterion and top words for topic 5: mild distraction.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Barrels/cones start, talking/singing audience unknown, biting nails/cuticles, reaching for food-related or drink-related item, distracted by construction, traffic signs, looking at previous crash/incident.
Lift	Reaching for food-related or drink-related item, distracted by construction, biting nails/cuticles, talking/singing audience unknown, barrels/cones start, looking at previous crash/incident, traffic signs.
FREX	Reaching for food-related or drink-related item, biting nails/cuticles, talking/singing audience unknown, barrels/cones start, distracted by construction, looking at previous crash/incident, traffic signs.
Score	Barrels/cones start, reaching for food-related or drink-related item, talking/singing audience unknown, biting nails/cuticles, distracted by construction, traffic signs, looking at previous crash/incident.

FREX = frequency and exclusivity.

Table 19. Criterion and top words for topic 6: free flow in work zone.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Adjusting/monitoring radio, speed limit in work zone, dancing, sign/arrow for turning, free flow, pet in vehicle, exit lane.
Lift	Free flow, sign/arrow for turning, exit lane, pet in vehicle, removing/adjusting jewelry, speed limit in work zone, adjusting/monitoring radio.
FREX	Speed limit in work zone, sign/arrow for turning, adjusting/monitoring radio, free flow, pet in vehicle, exit lane, removing/adjusting jewelry.
Score	Speed limit in work zone, adjusting/monitoring radio, removing/adjusting jewelry, dancing, free flow, sign/arrow for turning, pet in vehicle.

FREX = frequency and exclusivity.

Table 20. Criterion and top words for topic 7: free flow with nonspecific secondary tasks.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Other personal hygiene, other external distraction, other nonspecific internal eye glance, free flow (leading traffic present), none apparent, object in vehicle other, adjusting/monitoring other devices integral to vehicle.
Lift	Flow unstable/vehicles unable to pass/temporary stoppages, lighting cigar/cigarette, other external distraction, other nonspecific internal eye glance, other personal hygiene, drinking with lid and straw, no additional.
FREX	Other personal hygiene, other external distraction, other nonspecific internal eye glance, object in vehicle other, free flow (leading traffic present), drinking with lid and straw, no additional.
Score	Other external distraction, other personal hygiene, other nonspecific internal eye glance, adjusting/monitoring other devices integral to vehicle, object in vehicle other, none apparent, no additional.

FREX = frequency and exclusivity.

Table 21. Criterion and top words for topic 8: secondary tasks in conjunction with queue forming.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Both hands off, queue forming, cellphone browsing, cellphone holding, cellphone locating/reaching/answering, combing/brushing/fixing hair, removing/inserting/adjusting contact lenses or glasses.
Lift	Cellphone holding, cellphone other, cellphone texting, police vehicle, queue forming, cellphone locating/reaching/answering, reaching for object other.
FREX	Queue forming, cellphone holding, both hands off, cellphone locating/reaching/answering, cellphone texting, reaching for object other, combing/brushing/fixing hair.
Score	Both hands off, queue forming, cellphone locating/reaching/answering, cellphone holding, cellphone browsing, removing/inserting/adjusting contact lenses or glasses, reaching for object other.

FREX = frequency and exclusivity.

Table 22. Criterion and top words for topic 9: external distraction related to signs.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Traffic signs, crossroad, sign text, barrels/cones start, cellphone browsing, looking at previous crash/incident, out-of-ellipse.
Lift	Sign text, crossroad, traffic signs, looking at previous crash/incident, cellphone browsing, barrels/cones start, out-of-ellipse.
FREX	Crossroad, traffic signs, sign text, barrels/cones start, looking at previous crash/incident, cellphone browsing, adjusting/monitoring climate control.
Score	Traffic signs, sign text, crossroad, barrels/cones start, cellphone browsing, looking at previous crash/incident, out-of-ellipse.

FREX = frequency and exclusivity.

Table 23. Criterion and top words for topic 10: interaction with passengers/child in rear seat.

Criterion	Top Seven Words (In Highest-to-Lowest Criterion Order)
Probability	Equipment, barrels/cones start, passenger in rear seat interaction, child in rear seat interaction, warning lights, unknown, traffic signs.
Lift	Passenger in rear seat interaction, child in rear seat interaction, equipment, warning lights, unknown, barrels/cones start, adjusting/monitoring climate control.
FREX	Equipment, passenger in rear seat interaction, child in rear seat interaction, warning lights, barrels/cones start, unknown, adjusting/monitoring climate control.
Score	Passenger in rear seat interaction, equipment, barrels/cones start, child in rear seat interaction, warning lights, unknown, traffic signs.

FREX = frequency and exclusivity.

An estimation of whether each topic was more strongly associated with baseline or SCE conditions was made, and the results are shown in table 24 and figure 16. The positive values of the event estimates for topics 1, 2, and 8 indicate that SCEs are more prominent ($p < 0.001$); for topics 4, 6, 7, 9, and 10, the negative values of the event estimates indicate that baselines are more prominent ($p < 0.05$). Topics 3 and 5 are neither baseline nor SCE because the mean difference in topic proportions between baseline and SCE for these two topics is not significant ($p > 0.05$).

Table 24. Model to estimate marginal effects between baseline and SCE for each topic.

Topic	Intercept	Intercept Standard Error	Intercept *p*-Value	SCE Estimate	SCE Standard Error	SCE *p*-Value
1	0.094422	0.001923	<0.001	0.059524	0.006080	**<0.001**
2	0.050715	0.004081	<0.001	0.081777	0.014941	**<0.001**
3	0.097910	0.003362	<0.001	0.001313	0.009799	0.893
4	0.097910	0.00387	<0.001	-0.05143	0.010160	**<0.001**
5	0.145490	0.002545	<0.001	0.009647	0.007126	0.177
6	0.070463	0.003604	<0.001	-0.042753	0.007884	**<0.001**
7	0.100479	0.005175	<0.001	-0.074830	0.011877	**<0.001**
8	0.125945	0.009222	<0.001	0.075745	0.029484	**0.0106**
9	0.132837	0.003286	<0.001	-0.023881	0.008111	**0.00344**
10	0.101252	0.003899	<0.001	-0.035239	0.010824	**0.00123**

Note: Values in bold are significant under the significant level $\alpha = 0.05$.

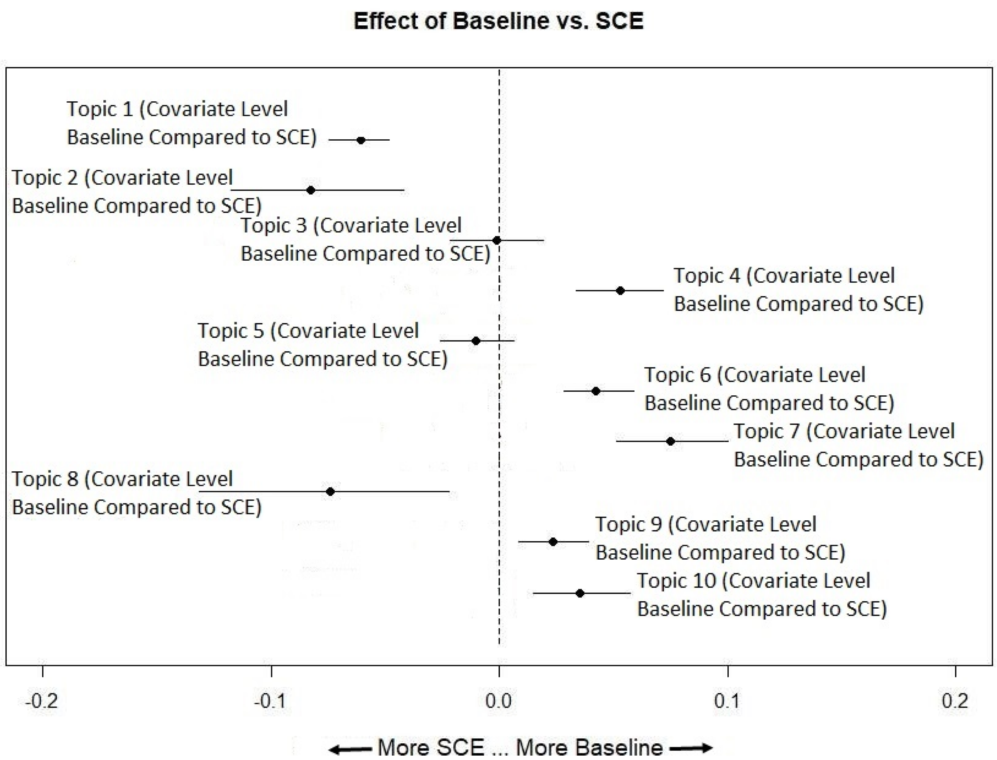
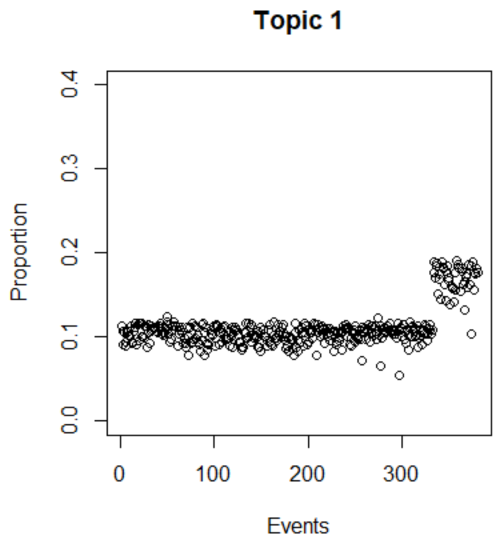


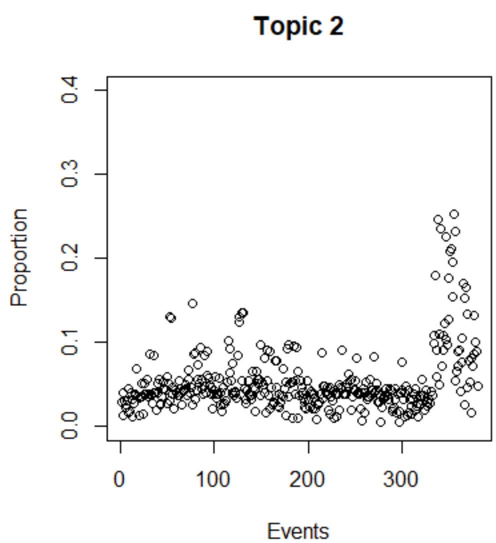
Figure 16. Graph. Mean and confidence interval of the difference in topic proportions between baseline and SCE for each topic.

Another method to present how prevalent the different topics are across the different events is shown in figure 17 and figure 18. The events are sorted such that the first 335 events are baselines and the last 47 events are SCEs. This highlights the proportions of the topics in each event and highlights the difference between the proportion of topic 1 and topic 2 in baselines and in SCEs, respectively. Similar figures for topics 3 through 10 can be found in appendix C.



© 2019 University of Michigan.
Note: The first 335 events are baselines and the rest are SCEs.

Figure 17. Graph. Proportion of topic 1 across the different events.



© 2019 University of Michigan.
Note: The first 335 events are baselines and the rest are SCEs.

Figure 18. Graph. Proportion of topic 2 across the different events.

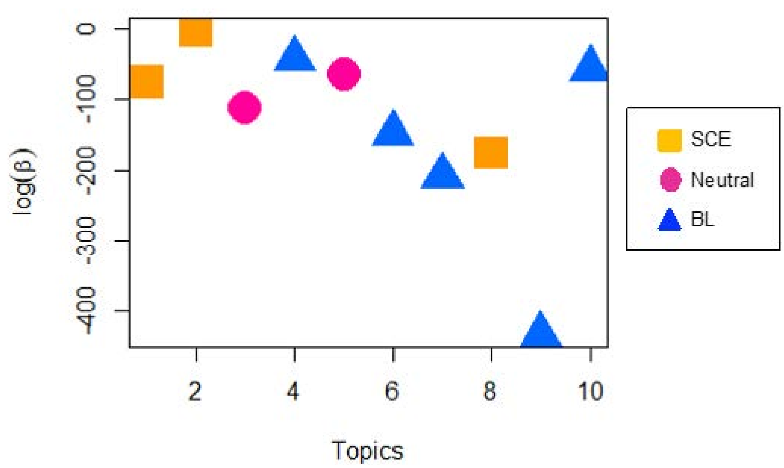
The following observations and comments can be made about the topics (discussed in the order of their expected proportions listed in table 12):

- Topic 5 describes mild distraction in work zones, as characterized by barrels and cones, with drivers reaching for food, distraction by a previous incident, or distraction by construction. Given the project focus on work zone behavior, it is not surprising that this topic is the most common topic represented in the dataset. This topic might be expected to characterize more risky behavior since the drivers are engaging in secondary tasks and may not pay attention to the work zone ahead. However, this topic is almost equally represented in baselines and SCEs in this dataset.
- Topic 9 includes work zone items such as barrels, cones, and traffic signs, but also driver-related secondary tasks such as distraction by a previous incident and cellphone browsing. This topic is the second most common topic and is associated more strongly with baselines.
- Topic 8 shows events where a queue is forming and the drivers are reaching for something, holding a cellphone, or texting on a cellphone. It is possible that when the drivers are queuing, they may decide to do something in the vehicle out of boredom. Alternatively, an already-distracted driver may come upon a queue, potentially resulting in an SCE. This topic is the third most common topic and is associated more strongly with SCEs. This suggests that queue forming in conjunction with other behaviors (resulting either from the queue or prior to encountering the queue) could lead to a risky situation. Avoiding queue forming in work zones could potentially help to increase the safety level in work zones.
- Topic 1 is related to out-of-ellipse excursions and internal and external secondary tasks such as cellphone (handheld) or talking and looking at a previous incident. This topic is more related to SCEs than baselines. This is consistent with the PCA results where out-of-ellipse behavior indicates potential safety-relevant behavior. Together with internal and external secondary tasks, the out-of-ellipse excursions may particularly result in SCEs.
- Topic 10 is related to work zone equipment and warning lights and interaction with passengers. This topic is about interaction with passengers when passing through a work zone. It is associated with baselines.
- Topic 4 is characterized with notations of barrier start, free flow, and food-related secondary tasks. This free flow condition may lead the driver to feel safe enough to start engaging in secondary tasks. This topic is associated with baselines.
- Topic 7 is also related to free flow (following vehicle at a long distance) and to nonspecific secondary tasks such as other external distractions. This topic is not related to work zone keywords and is more specific to baselines.
- Topic 3 is related to in-vehicle interaction with passengers (as characterized by passenger interaction and cellphone talking/listening) in work zones (as characterized by the words digital sign and work zone sign start). The topic is neither related to baselines nor to SCEs.
- Topic 6 is related to secondary tasks during free flow in work zones as characterized by adjusting/monitoring radio, dancing, free flow, and speed limit in work zone. The topic is related to baselines. This topic is the second least prominent in the dataset. Generally, it is not common to see free flow in work zones.

- Topic 2 is the least prominent of all the topics in this dataset. This topic is related mainly to work zone items such as closed lanes, parked cars, workers present, and truck-mounted attenuator (TMA) (a protective crash cushion mounted on a truck when traffic can be exposed to work zone objects). The flow is with some restrictions. This topic is related to SCEs. This suggests that one should be careful when passing in highly active work zone, particularly when there is internal distraction (characterized here by a moving object in a vehicle).

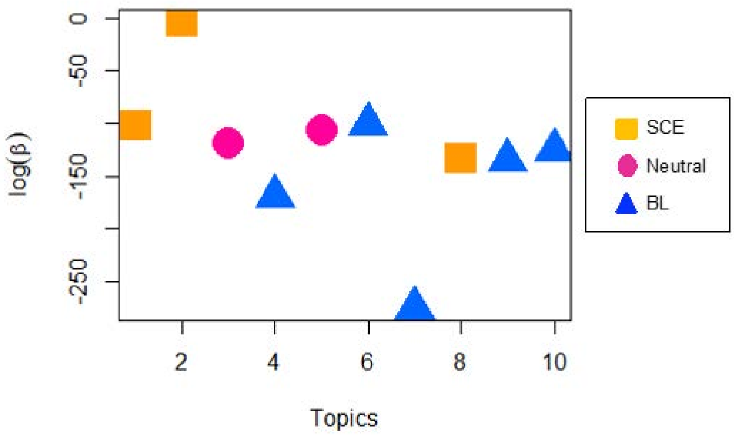
In general, all free-flow-related topics (topics 4, 6, and 7) are more represented in baselines, while the queue-forming-related topic (topic 8) is more represented in SCEs. Therefore, avoiding queue forming in work zones could be an important step toward improving safety levels in work zones. Related to secondary tasks, analysis suggests that while conducting secondary tasks alone may not lead to an SCE, when done in combination with out-of-ellipse behavior or a highly active work zone, this could lead to an SCE.

To look more closely at the association of specific words with baseline and SCE topics, figure 19 through figure 38 show the natural log of the probability of a particular word for each topic. Higher values indicate greater association (i.e., the value 0 on the Y-axis means that the probability of the word (e.g., workers present) is 1, and as the value of probability of the word approaches 0, the value of the Y-axis goes to negative infinity). The topics are coded as baseline-related (blue triangles), SCE-related (orange squares), or neutral (pink circles). Each topic is categorized as WZR, context-related (CR), or secondary-task-related (STR).



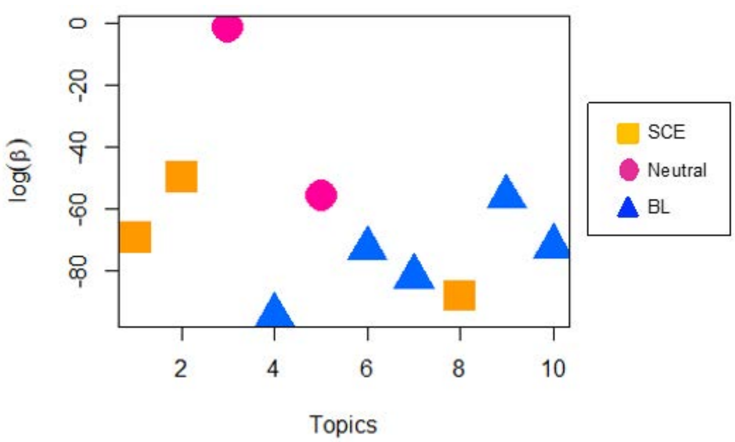
BL = baseline.

Figure 19. Graph. Probability of workers present (WZR) with baseline, SCE, or neutral.



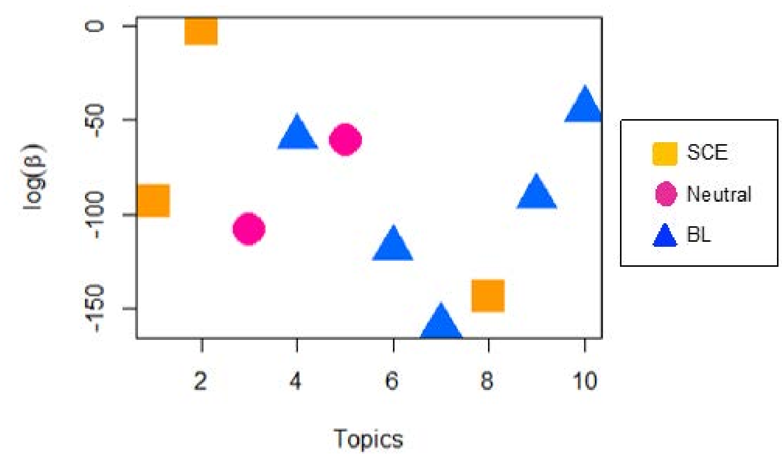
BL = baseline.

Figure 20. Graph. Probability of TMA (WZR) with baseline, SCE, or neutral.



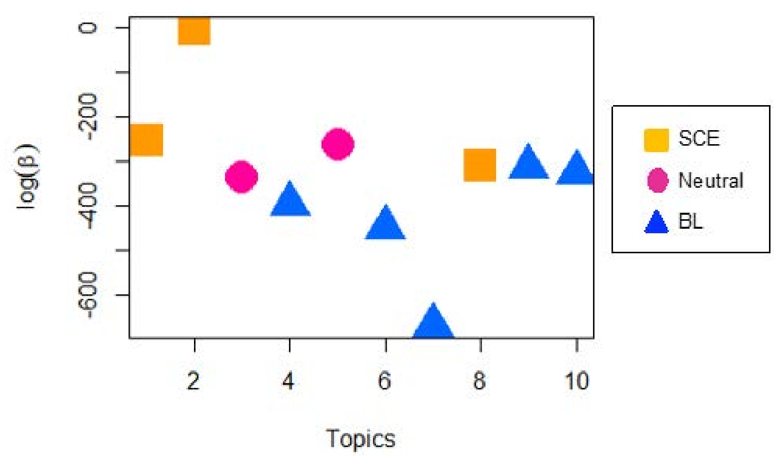
BL = baseline.

Figure 21. Graph. Probability of digital sign (WZR) with baseline, SCE, or neutral.



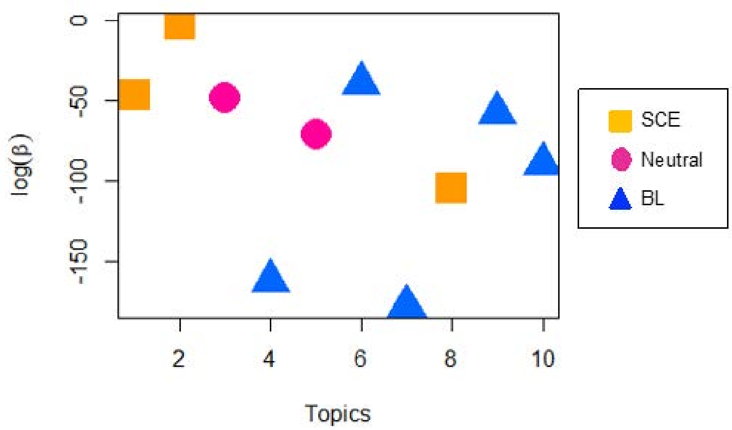
BL = baseline.

Figure 22. Graph. Probability of parked car (WZR) with baseline, SCE, or neutral.



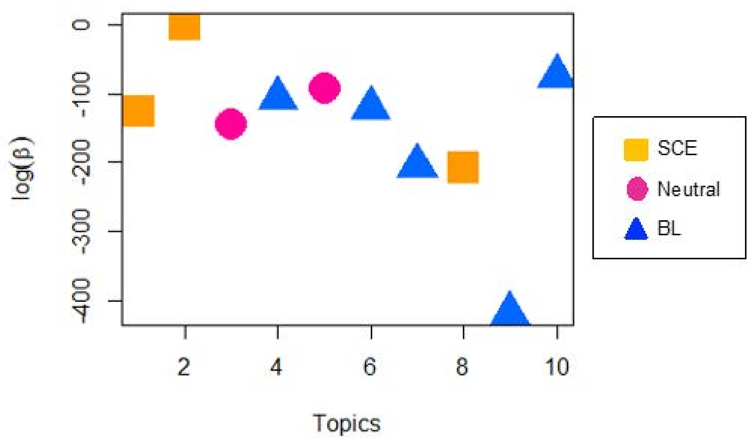
BL = baseline.

Figure 23. Graph. Probability of chicane (WZR) with baseline, SCE, or neutral.



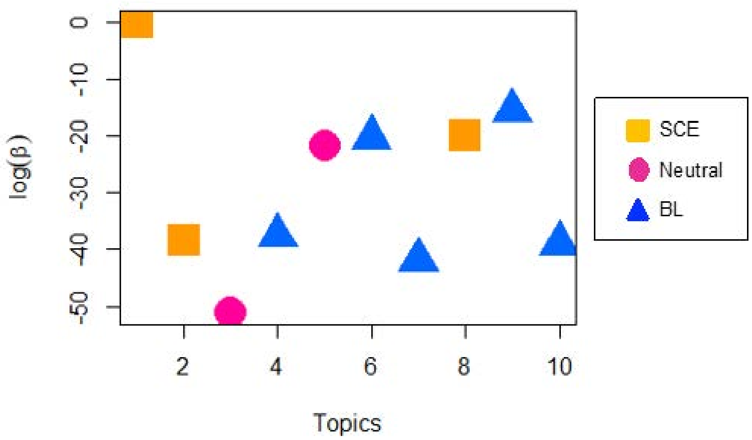
BL = baseline.

Figure 24. Graph. Probability of decrease lane (WZR) with baseline, SCE, or neutral.



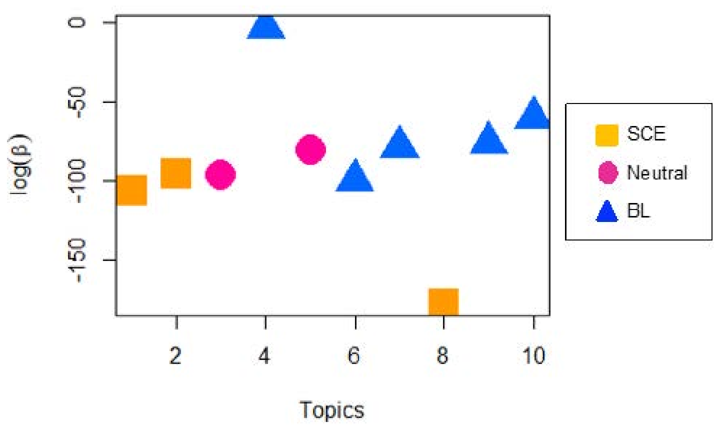
BL = baseline.

Figure 25. Graph. Probability of closed lanes (WZR) with baseline, SCE, or neutral.



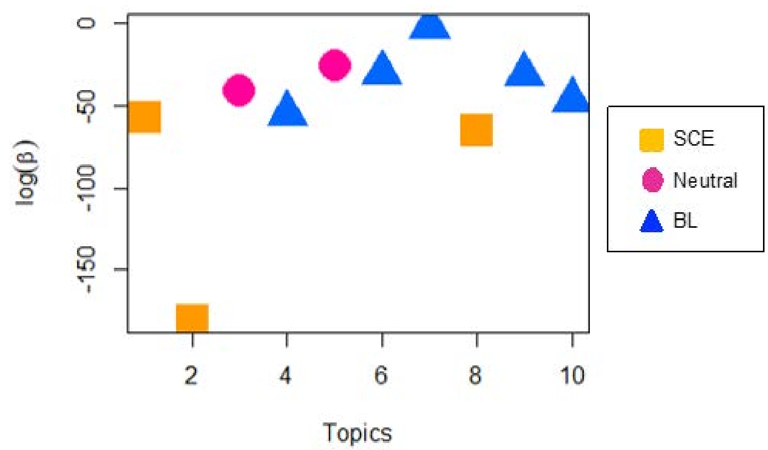
BL = baseline.

Figure 26. Graph. Probability of out-of-ellipse (CR) with baseline, SCE, or neutral.



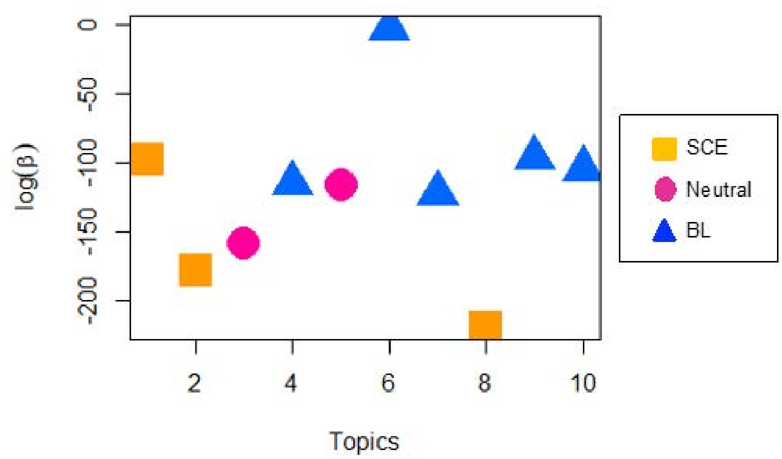
BL = baseline.

Figure 27. Graph. Probability of free flow no lead traffic (CR) with baseline, SCE, or neutral.



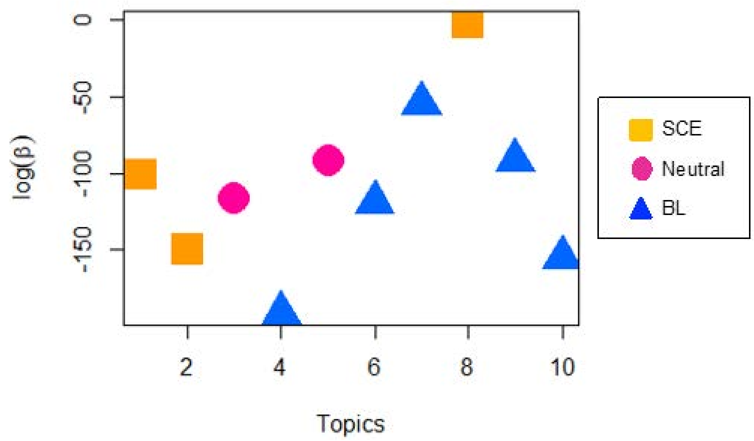
BL = baseline.

Figure 28. Graph. Probability of free flow leading traffic present (CR) with baseline, SCE, or neutral.



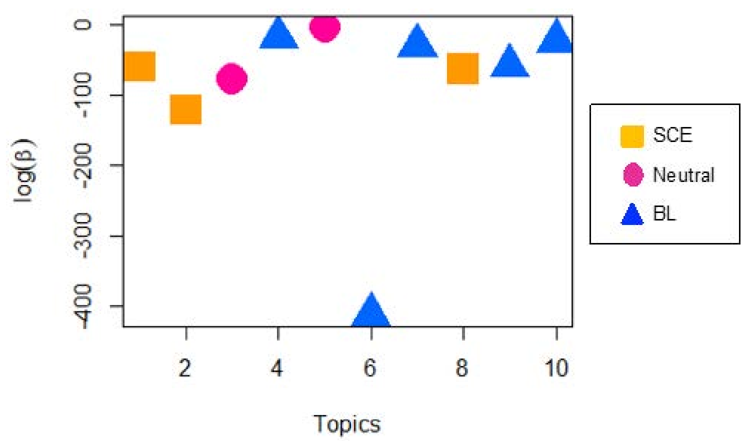
BL = baseline.

Figure 29. Graph. Probability of free flow (CR) with baseline, SCE, or neutral.



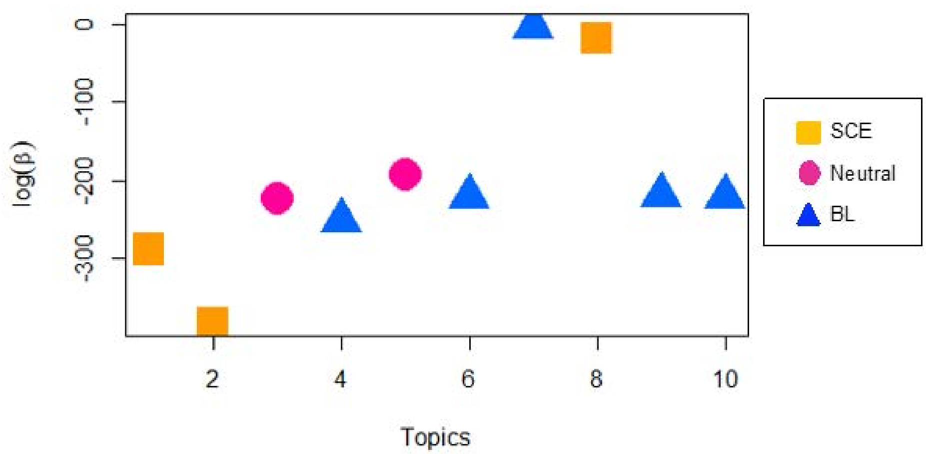
© 2019 University of Michigan.
BL = baseline.

Figure 30. Graph. Probability of queue forming (CR) with baseline, SCE, or neutral.



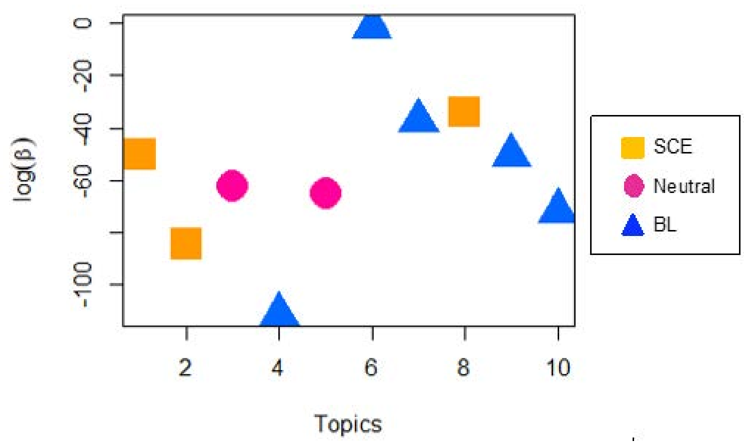
© 2019 University of Michigan.
BL = baseline.

Figure 31. Graph. Probability of distracted by construction (STR) with baseline, SCE, or neutral.



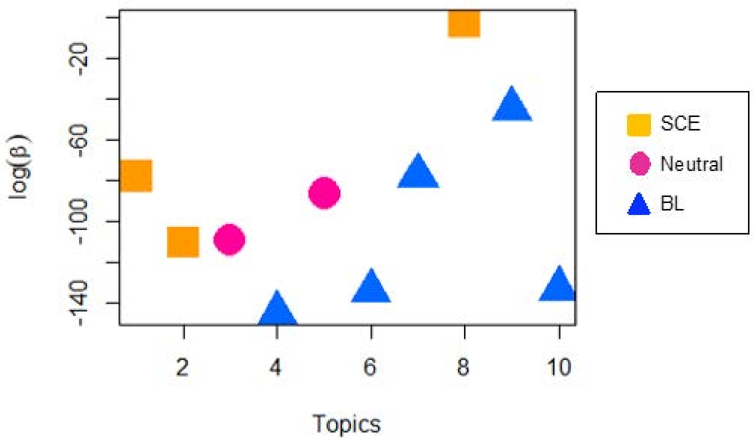
BL = baseline.

Figure 32. Graph. Probability of adjusting/monitoring other devices integral to vehicle (STR) with baseline, SCE, or neutral.



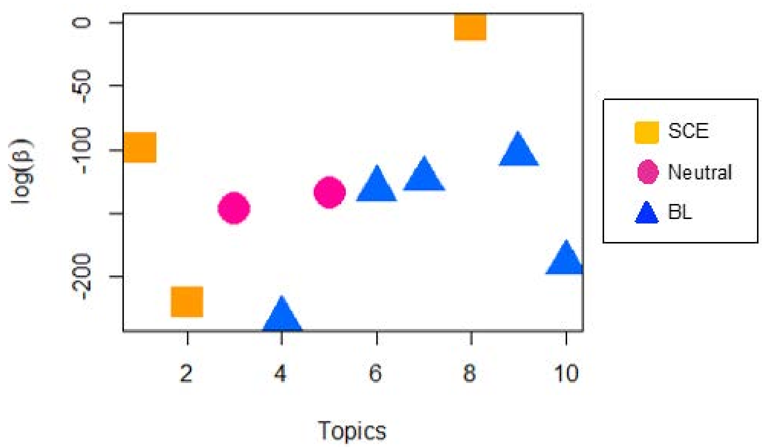
BL = baseline.

Figure 33. Graph. Probability of adjusting/monitoring radio (STR) with baseline, SCE, or neutral.



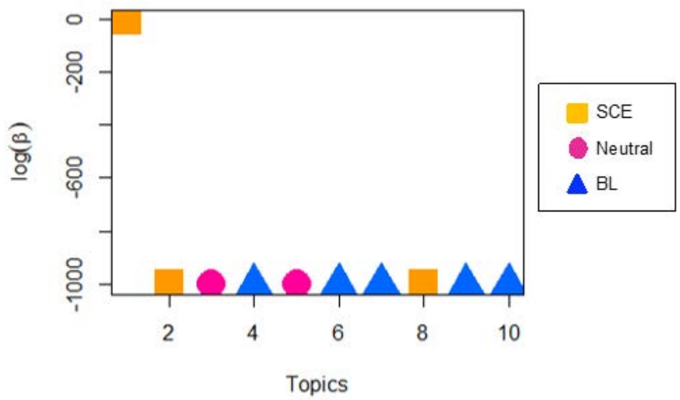
BL = baseline.

Figure 34. Graph. Probability of cellphone holding (STR) with baseline, SCE, or neutral.



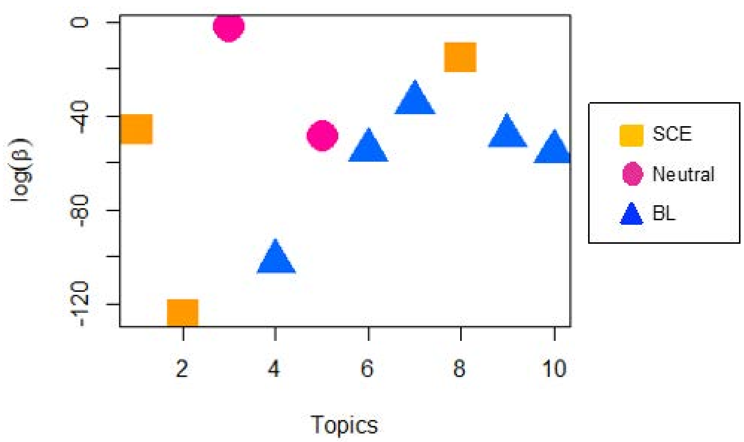
BL = baseline.

Figure 35. Graph. Probability of cellphone locating/reaching/answering (STR) with baseline, SCE, or neutral.



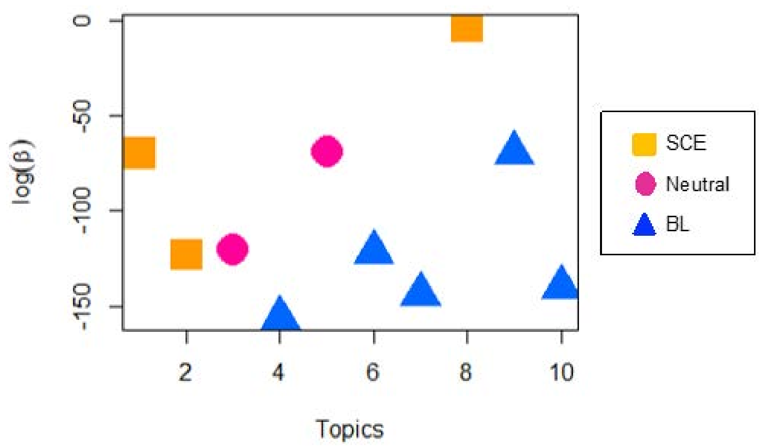
BL = baseline.

Figure 36. Graph. Probability of cellphone talking/listening hands-free (STR) with baseline, SCE, or neutral.



BL = baseline.

Figure 37. Graph. Probability of cellphone talking/listening handheld (STR) with baseline, SCE, or neutral.



BL = baseline.

Figure 38. Graph. Probability of cellphone texting (STR) with baseline, SCE, or neutral.

The following observations were made for the selected words related to work zone, context, and secondary tasks (figure 19 through figure 38):

- Among work zone variables, most (workers present, TMA, parked cars, chicanes, decreased lane, and closed lane in figure 19, figure 20, and figure 22 through figure 25) are associated with specific SCE topics (often topic 2), with the exception of digital signs (topic 3, neutral, as can be seen in figure 21).
- Free-flow-related words are associated with single baseline topics (figure 27 through figure 29), whereas out-of-ellipse and queue forming are associated with single SCE topics (figure 26 and figure 30).
- The highest probability of seeing queue forming is in topic 8, which is related to SCE. The queue forming might be one of the causation mechanisms/processes leading to SCEs.
- Internal distractions like cellphone holding, locating, reaching, answering, and texting (figure 34 through figure 38) have the highest probability to appear in topic 8, which is related to SCEs. In this dataset, cellphone texting is more common than hands-free talking.
- Unlike the other cellphone activities, talking/listening handheld (figure 37) is strongly associated with topic 3, neutral.

Based on these observations, while the sole presence of workers may not increase the risk of an SCE happening, the co-occurrence of some or all of workers' presence, TMA, parked car, chicane, closed lane, and/or decreasing lane together may increase the risk of an SCE happening. Similarly, the observations also suggest that the co-occurrence of queue forming with internal distraction activities may increase the risk of an SCE happening.

Finally, STM permits correlations between topics, as shown in table 25. Positive correlations between topics indicate that both topics are likely to be present within a document. Negative correlations indicate that the topics are less likely than chance to both be present in a single document. Interestingly, baseline topics are not necessarily positively correlated with each other, nor are SCE topics.

Table 25. Correlations among topics.

Topics	1 (SCE)	2 (SCE)	3 (Neutral)	4 (Baseline)	5 (Neutral)	6 (Baseline)	7 (Baseline)	8 (SCE)	9 (Baseline)	10 (Baseline)
1 (SCE)	**1.00**	R	R	R	R	R	R	R	R	R
2 (SCE)	0.46	**1.00**	R	R	R	R	R	R	R	R
3 (Neutral)	0.22	0.04	**1.00**	R	R	R	R	R	R	R
4 (Baseline)	*-0.54*	0.06	0.35	**1.00**	R	R	R	R	R	R
5 (Neutral)	0.29	0.54	0.42	0.46	**1.00**	R	R	R	R	R
6 (Baseline)	*-0.43*	*-0.40*	0.13	0.22	0.17	**1.00**	R	R	R	R
7 (Baseline)	*-0.50*	*-0.63*	*-0.42*	*-0.24*	*-0.58*	0.15	**1.00**	R	R	R
8 (SCE)	0.18	*-0.31*	*-0.62*	*-0.79*	*-0.79*	*-0.35*	0.37	**1.00**	R	R
9 (Baseline)	*-0.08*	*-0.10*	0.66	0.41	0.52	0.57	*-0.36*	*-0.68*	**1.00**	R
10 (Baseline)	*-0.37*	0.24	0.51	0.93	0.51	0.12	*-0.44*	*-0.84*	0.51	**1.00**

R = redundant; values match those below the diagonal.
Note: Positive correlations are bolded, and negative correlations are italicized and red.

SUMMARY

The experiment with STM using the Phase 2 dataset suggests that all free-flow-related topics (topics 4, 6, and 7) are more represented in baselines, while the queue-forming-related topic (topic 8) is more represented in SCEs. Although the occurrence of queue forming by itself may not lead to an SCE, the co-occurrence of queue forming with internal distraction activities may increase the risk of an SCE happening. Therefore, avoiding queue forming in work zones is an important step toward improving safety levels in work zones.

Internal distraction and secondary tasks appear in several topics. While conducting secondary tasks alone may not lead to an SCE, secondary tasks together with out-of-ellipse driving behavior, or performing secondary tasks in a highly active work zone, may lead to an SCE.

CHAPTER 5. APPLICATION TO WORK ZONE SPEED LIMITS

CONTROL THEORY PAPERS WITH HUMAN BEHAVIOR COMPONENTS

From the perspective of applied control theory, the predefined section of the road (i.e., work zone) can be taken as a “system” within a larger system (i.e., transportation network). Several studies cover macroscopic traffic flow models for diverse work zone activities (Bharadwaj et al. 2018). Although this type of model is very useful in understanding (and managing) how work zones and even specific work zone activities affect traffic flow on a macro scale, they have no explicit parameters on driver behavior. The driver-specific microscopic parameters, such as characteristic speed profiles or the pattern of acceleration/deceleration, also affect the work zone conditions (i.e., safety, congestion).

Taking a control theoretic approach, the variable to be regulated in the work zone system can be considered as the vehicle density per section (traffic flow rate) subject to constraints of safety (the control aim is to have zero SCEs, or the relaxed control aim is to reduce the probability of having an SCE). It is very common to apply VSL control algorithms for traffic flow stabilization as well as guaranteeing safety (Yu and Abdel-Aty 2014). An alternative approach might involve applying delay balancing, especially in scenarios where the controlled congestion area has off-ramp and on-ramp artillery branches (Iordanidou et al. 2016). The application of VSLs involves adaptively changing the posted speed limits in the upstream traffic according to changes in the monitored section of the work zone downstream. In other words, the VSL control algorithms use the measurements from the work zone or hazard zone to change the posted speed limits in the buffer/prework zone part of the road. Thus, when the vehicles arrive at the work zone, their speed profiles are already corrected and shaped to stay in a safe boundary to satisfy a certain throughput or reduced risk of an SCE. The second strategy, “delay balancing,” can be an application of a VSL control at a larger scale in the traffic network. Instead of using only posted speed limits, it can also use on-ramp and off-ramp branches in the traffic network to adjust the speed profile of the road segment that is being regulated by controlling the rate of traffic joining the traffic stream. Delay balancing algorithms can handle multiple bottleneck situations since their scale of application extends to larger portions of the traffic network.

One of the most interesting applications for regulating the flow rate through influencing individual driver behavior can be found in the work by Davis (2016). In that study, the problem involves a two-to-one lane reduction case where the traffic flow rate in the merged lane has to be controlled to avoid congestion. Employing connected vehicles featuring ACC systems, Davis (2016) suggests reducing the speed of the soon-to-be-closed lane, forcing/nudging the drivers to perform early lane changes well before the hazard zone. Although only 30 percent of the fleet had connected vehicle features, this level of fleet penetration improved the situation and reduced congestion. This type of early and precautionary behavior shaping can be also adopted as a VSL treatment in the work zone scenarios.

In the study by Yu and Abdel-Aty (2014), the compliance of the drivers to the VSL application was of particular interest. This study demonstrates that any active or passive treatment in a work zone scenario involving speed limits/regulations should have compliance from drivers; otherwise, it may be less efficient or fail altogether.

ROLE OF PCA/NORMAL DRIVING IN FEEDING CONTROL THEORY COMPUTATION OF VSLs

Introduction

Some specific qualities of the PCA results are important in envisioning how the method might be used in a practical application. First, excursions from the ellipse are indicators of increased crash risk, whether that risk is introduced by the driver or the situation. However, the prediction performance of the metric is not good enough to support strong interventions such as automatic braking. Instead, it makes more sense to use the PCA metric to:

1. Identify impending issues in real time as they arise.
2. Identify specific drivers who may be engaging in driving behavior that could lead to issues within the work zone.
3. Characterize work zones broadly as involving relatively more or less problematic driving.

To implement the PCA approach, some type of technology must track a vehicle's time-history of speed, acceleration, or position. Vehicle-to-infrastructure (V2I) communication is most promising because vehicles can send their own time-histories of speed and position, but that solution requires reasonable fleet penetration of equipped vehicles. In contrast, speed sensors are currently deployable. The disadvantage of speed sensors is that they provide a snapshot of speeds at one location rather than a history, so additional work would be needed to translate the PCA approach to this different type of data. For this report, communication is assumed to be possible.

Given communicated speed histories, real-time calculation of PC2 and PC3 values is straightforward, and traffic can be monitored for excursions from specific limits of the ellipse. Ellipse boundaries can be determined based on past examples and/or adjusted for a specific work zone as data are collected.

Identifying Impending Issues

Impending safety issues could be identified by increases in the number of vehicles making ellipse excursions. This will often happen because of increases in traffic density, but it can also be caused by specific driver actions. Either way, recalculation of speed limits upstream may be warranted when excursions increase. In particular, the approach to identifying VSLs described by Yu and Abdel-Aty (2014) would work well in conjunction with PCA.

As described above, Yu and Abdel-Aty (2014) detail a method in which they add a crash risk model to the standard set of optimization equations associated with calculating VSLs that maximize throughput while minimizing crash risk. As congestion increases, slower speeds are warranted, but noncompliance can influence the success of VSLs. The PCA results provide two potential inputs to this model. First, Yu and Abdel-Aty (2014) use a logistic regression model in which mean speed is used to predict crash risk. However, the PCA out-of-ellipse model may provide a more nuanced crash risk model for this purpose since it incorporates acceleration patterns as well as speed. Second, the PCA approach identifies noncompliant vehicles in terms of more than just speed noncompliance, although speed noncompliance is clearly important.

An alternative and interesting potential countermeasure is described by Davis (2016), who looks at the possibility of using I2V and ACC-equipped vehicles to essentially operate as “pilot cars” to induce more appropriate deceleration and speed-keeping. With as little as 30 percent fleet penetration, the influence of these vehicles helps improve throughput and compliance in other vehicles. Here again, PCA has the potential to identify situations when VSLs should be implemented (i.e., lowered limits) and vehicle decelerations should be monitored (or even controlled using the pilot ACC-equipped vehicles).

Identifying Specific Drivers

In the same way that PCA can identify impending issues with general driving in a work zone, it can also identify specific drivers who may be contributing to problems. Results show that the increased crash risk predicted by out-of-ellipse excursions is about one-half due to drivers and one-half due to environmental conditions. Thus, another potential use case for improving work zone safety is to identify specific drivers for targeted countermeasures.

A key aspect of the PCA approach that is consistent with the control theory literature is that both approaches indicate the benefit of smoother, gentler deceleration and acceleration. Control theory indicates that to maximize throughput under congested conditions, deceleration to optimal (variable) speed limits should occur well before the point where lower speeds are needed. Out-of-ellipse excursions indicate that this is not occurring, and countermeasures to improve deceleration to appropriate speeds are warranted.

When driver behavior is the issue, countermeasures that are specific to a given driver may be more effective than general warnings or notices. I2V communication provides a potential way to provide warnings or indications directly to drivers who may be inattentive or aggressive. While this type of countermeasure is still relatively far in the future, it is worth considering how personal notifications could work in this situation (and other similar situations) to try to elicit improved behavior.

Characterize Work Zones

A third potential use of PCA is simply to characterize driving in work zones broadly. The analyses process generated ellipse boundaries specific to each work zone based on the baseline behavior observed in those work zones. However, one can turn this logic around and ask why some work zones have larger ellipse boundaries (i.e., more variability in the PCs) than others.

As shown in table 8, certain aspects of work zones tend to produce more out-of-ellipse excursions. These include barrels/cones, closed lanes, and queues forming, but other characteristics such as barriers, merging, and equipment are associated with reduced or equal likelihood of excursions. These results are consistent with many observations from work zones indicating that elements requiring change by the driver can lead to increased risk of crashing.

For work zones that have a number of elements producing out-of-ellipse excursions at a higher-than-average rate, it would be reasonable to implement countermeasures specifically aimed at helping drivers anticipate those elements and manage them smoothly. VSLs are one path to improving throughput, but an essential component is to ensure that drivers manage speed smoothly with reasonable spacing between them and nearby vehicles (either ahead or to the side). This can be difficult to achieve, but Davis's (2016) ideas for using a subset of vehicles to control the general flow might be promising.

CHAPTER 6. APPLICATION TO WORK ZONE TREATMENT GUIDANCE

WORK ZONE TREATMENT BACKGROUND

Introduction

Any maintenance or construction activity on or near the road introduces a disturbance to the road user. The challenge for transportation professionals is to determine the most effective design layout that has proven performance characteristics. Each construction activity has unique qualities that need to be considered. Work zones have many distinguishing characteristics defined by their location, road type, duration, and size, and in some situations they may be mobile (snow clearing, mowing, or road marking activities).

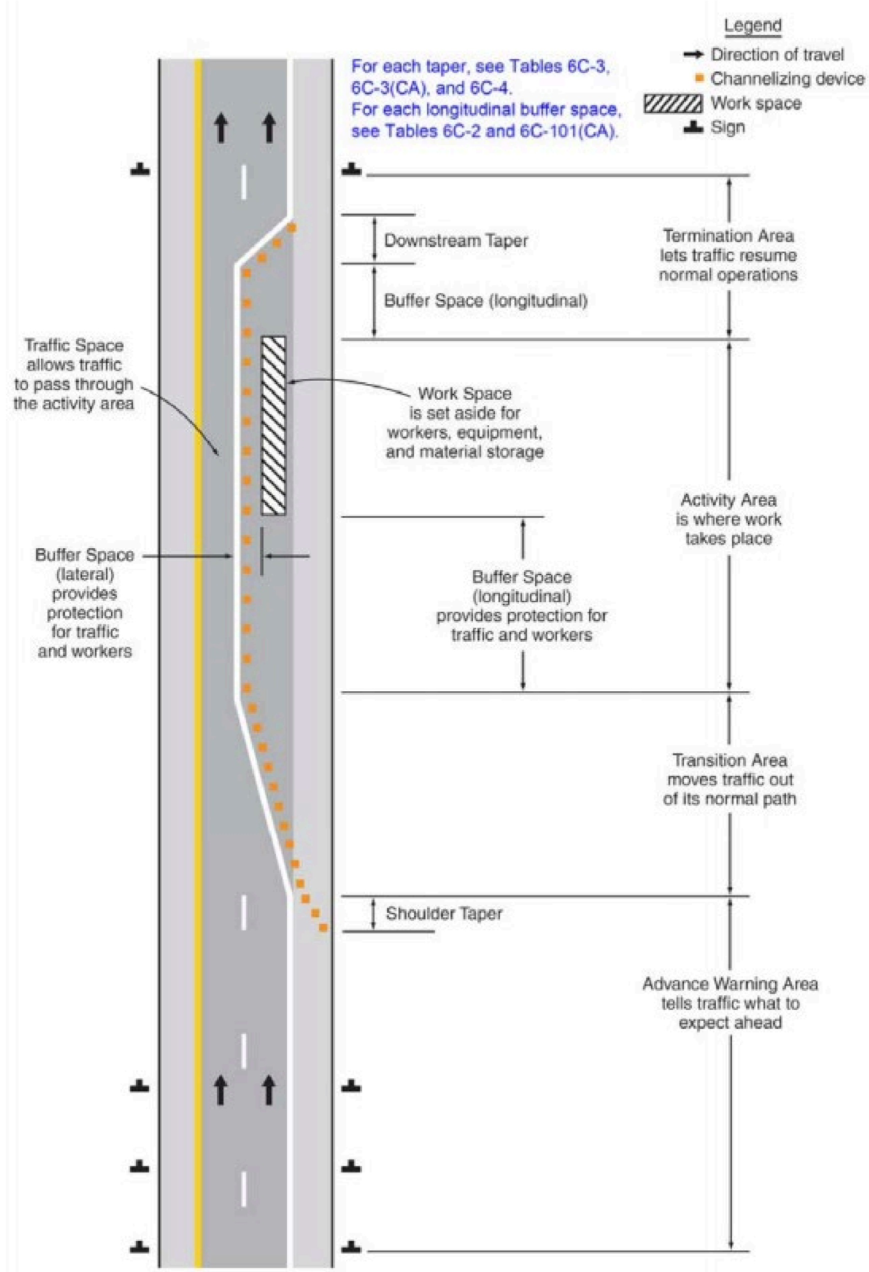
The information collected from the SHRP2 database and presented in the previous chapters is essentially restricted to somewhat higher speed roads and do not include mobile work zone conditions. This results in an analysis knowledge base that may not represent all work zone types. However, the results are presented in the context of a generic work zone as much as possible.

Strategy and Current Practice

The United States developed the *Manual on Uniform Traffic Control Devices (MUTCD)* for streets and highways as a national standard on the design of traffic control devices and provides guidance on their use and placement for normal and temporary traffic control (FHWA 2009). The latter is relevant for this report, as work zones require temporary situations. The *MUTCD* is founded on the principle that consistent design and use of traffic control devices reduce the uncertainty for interpretation by road users and should improve safety.

Figure 39 is an example provided in the *MUTCD* on work zone layout and is an important reference for the subsequent discussion. Independent of the work zone type, it is important to consider the four main divisions of the road segment related to the site:

- Advance area. The region upstream of the work zone informing traffic of the upcoming construction activities.
- Transition area. Corrects traffic placement or conditions for the downstream regions where the maintenance or construction activities occur. Speed and lane layouts are modified to accommodate the constraints on the road due to the activities.
- Activity area. Area where construction activities take place.
- Termination area. Downstream area adjacent to area where traffic can return to normal operations and speed and lane layouts are no longer modified.



Source: FHWA.

Figure 39. Diagram. Concepts for work zone layout (USDOT 2012).

Safety Issues

Although there is no single consideration that can be used to quantify the safety level of a road network or even an isolated work zone location, speed has been one of the main indicators of safety. More difficult is quantifying the lateral motion of vehicles, in particular the merging behavior of vehicles that is required in work zones where transitions between different travel lane layouts are common. The combination of traffic flow and travel speed also provide important clues to safety performance.

The use of instrumented vehicles in NDSs has allowed for the observation of driver behavior in situ. The SHRP2 dataset allows for the observation of driver activities, such as secondary tasks or direction of attention away from the forward view, to be annotated and correlated with the surrounding road environment. The known increased risk of crashes due to driver distraction can thus be studied in the context of work zone design. The design of safer work zones requires an understanding of the propensity of drivers to direct their focus to construction or traffic control equipment lateral to the travel lanes, which leads to riskier traffic conditions. Similar distractions inside the vehicle can be studied in the context of the work zone, particularly if travel speeds are reduced and monotony or irritation causes the driver to divert their attention to cellphone use.

The earlier analyses of the PCA and STM methods allow for discrete work zone elements to be studied in combination with driver behavior parameters. The PCA method allows for the driving style of the driver to be monitored and identify the situations when these changes occur from normal to potentially unsafe conditions. The STM method allows for the unbiased observations of the work zone, traffic, and driver behavior factors to be studied to observe conditions that are more or less favorable for safety.

Safety Issues Identified

Figure 3 provides the framework for the analysis with regard to the work zone layout that is relevant for the discussion of the PCA and STM analysis results in the context of work zone treatments. All the analyses from the STM method use information collected in the D30 zone. This means that data for safe or baseline trips through the work zone are compared to trips with an SCE (crash or near crash). All the data are aligned to compare the same road mile post sections for trips occurring within a 2-week window surrounding an SCE.

It is important to consider the information in the context of the results presented in figure 15. Topics can be assigned to general sections of the work zone based on associated words. One must look beyond the topic definition to understand the importance of the words such as the "lift" or "FREX" ratings. In figure 15, the "biting nails/cuticles" could occur three times in one event, but only once in another. In another situation, the presence of barrels and signs may indicate topics connected to the whole work zone. The results in table 10 can also be useful to interpret the presence of words in topics. Table 10 provides insight into how often a work zone feature (barrel, barrier, etc.) appeared in the studied work zones and how often these features are associated with out-of-ellipse events. Some of these features are intentionally introduced to the work zone to produce a change in traffic conditions and the PCA results can be used to identify borders where traffic conditions change more abruptly, leading to SCEs in some cases.

Using the words observed in the topics, it is possible to locate their most likely position in the work zone layout, as shown in figure 38. For example, chicanes and lane endings are indicators that the topics are occurring in or around the transition area. Words like "workers present," "equipment," and "construction activities" indicate that the topic is relevant for the activity area. Unfortunately, the STM method does not contain a time-history element that could discriminate between the advance warning area and termination area.

The four important zones defined by the *MUTCD* in figure 39 are identified in relation to a generic work zone in figure 40. While the specific zone location was difficult to pinpoint for topics, figure 40 does agree with earlier work zone safety data that were summarized in the Appropriate Speed Saves All People (ASAP) project where more crashes were observed in the advance warning or transition area than in the activity and termination zones (Nocientini et al. 2013). This is intuitively valid as this is where the most dynamic changes in traffic occur as the traffic flow must adapt to the work zone speed and lane structure from the upstream conditions. This is typified by the key words "queue forming," "out-of-ellipse," and "flow unstable."

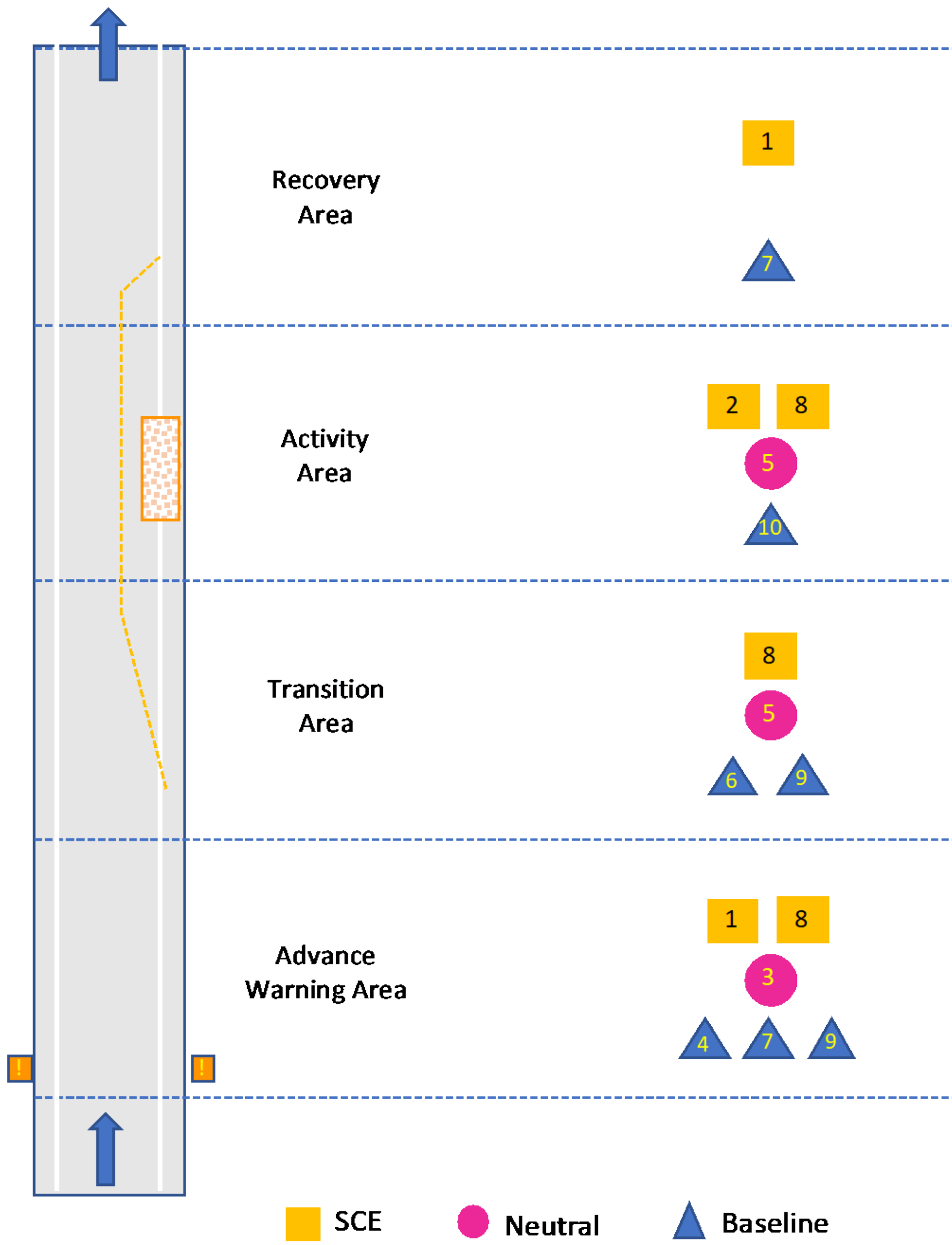


Figure 40. Diagram. Likely position of topic in work zone based on topic content.

The SCE-related topics (1, 2, and 8) occur in all portions of the work zone in figure 40 according to the words related to the types of work zone elements included in the topic. Topic 8, which includes queue forming, might occur anywhere prior to the termination area. Topic 1, which is out-of-ellipse driving plus a variety of secondary tasks, tends to occur at the beginning or end of the work zone. This is potentially a consequence of the definition of out-of-ellipse driving that is associated with strong braking or accelerations. These speed changes would tend to occur as the situation changes in the transition and recovery areas. Finally, topic 2 is specific to the active work zone and is also characterized by distractions.

The results of the STM analysis highlight that the driver is often involved in secondary tasks while negotiating the work zone. In fact, all 10 topics listed in table 13 through table 23 , whether they are more aligned with baselines or SCEs, included one or more secondary tasks in the top words. For all but topic 6, secondary tasks are notable enough to be included in the summary labels (interpretation). The STM identifies the presence of work zone traffic control devices (signs, barrels) and drivers distracted by the work zone (topic 5), observing a previous crash (topic 9), or out-of-ellipse driving patterns (topic 9). Topic 5 and topic 9 are the two most common topics observed in all of the cases studied with STM. Topic 5 is neutral, but on the borderline of being more associated with SCEs, while topic 9 is more clearly associated with SCEs (figure 14).

The distribution of the topics in the different zones identified in figure 40 highlights how the types of topics vary between the zones. It is important to note the number and diversity of topics in the transition area and particularly the advance warning area. These highlight the complex driving conditions. There may be only small changes in the driving context that change a safe driving condition to an unsafe situation.

An additional information source from the STM that allows interpretation of work zone treatment is shown in figure 19 through figure 38. These figures indicate the frequency of specific words in a topic and analysis of key words related to work zone treatment and traffic control devices that provide guidance for traffic engineers. Table 10 complements this information with the relationship between work zone features and out-of-ellipse driving, an indicator of deviations from safe driving conditions. Chicanes, closed or decreased lanes, parked cars (an indication of work zone activity), and TMAs are most strongly related to at least one topic associated with SCEs. The presence of workers is associated with several topics, but most strongly with SCE topic 2, and least associated with the most common baseline topic (topic 9).

Digital signs, in contrast, are seen to be less associated with safety-critical topics and most connected to a neutral topic (topic 3). While the number of digital signs in the annotated data is relatively few, this could reflect their positive influence on speed reductions. The benefit of digital signs is reported in the ASAP project and is consistent with observations about the need for communication with drivers.

Topic 9, the most common baseline topic, is interesting to review as well. It includes work zone elements at the start of the work zone, driver distraction elements, and driving the vehicle under out-of-ellipse conditions. Since this topic is associated with safer baseline conditions, it suggests that otherwise unsafe conditions may not lead to crashes unless a potential conflict or threat is present. This situation is observed in many topics where there are combinations of “unsafe” driver behavior (e.g., distractions, out-of-ellipse) with driving conditions (e.g., construction activities) and perceived risk conditions (e.g., equipment or workers present). These topics could reflect safer work zone passages when traffic conditions were less congested (i.e., free flowing).

Distractions due to construction or activities outside the vehicle seemed to be quite common in many of the top 10 topics, for both safe and safety-critical situations. Because the study topic is work zones, distractions are present; but, the association of distractions to both baselines and SCEs suggests that it is not a strong risk factor among the many secondary tasks in which drivers engaged in work zones.

There is no information in this study to determine if the tendency for secondary tasks is greater in work zones than in normal conditions. In the ASAP project, evidence shows that the crash modification factor (CMF) for work zones is over 2 (i.e., crashes are twice as common in a work zone than on "normal" roads) and highlights the need to reduce any work zone or traffic feature that will divert the driver away from the traffic task (Sorensen et al. 2015). The work zone analyses with STM and PCA also do not allow investigation of the role of absolute speed on driver engagement in secondary tasks. The current analysis only considers the PCA ellipse analysis that addresses sudden changes from a baseline condition. The influence of reduced work zone traffic speeds on driver secondary tasks and distraction toward work zone construction was not possible in this analysis.

CHAPTER 7. CONCLUSIONS

SUMMARY

The analyses in this project focused on investigating SHRP2 work zone data. The SHRP2 samples were selected to include SCEs that occurred in work zones and a set of baselines from the same work zones.

The general goal was to try to define “normal” driving behavior within work zones as a way of identifying deviations from “normal” as “abnormal.” With baselines as the example (or reference) set of normal driving and SCEs as cases of abnormal driving, the developed methods were reasonably successful at distinguishing between the two conditions.

Two methods were developed, PCA and STM, each of which has different strengths. PCA is a dimension-reduction method that works with continuous measures such as the kinematic speed and acceleration signals. PCA constructs linear combinations of a large number of variables so that only a few such linear combinations (or PCs) are required to reconstruct nearly all of the information in the original data. In addition, each PC is orthogonal to the others, making PCs easy to use in analysis. In this case, the variables were speed and acceleration at every 0.1 s of driving over a 1.5-s interval.

STM is best suited for categorical variables, and identifies groups of categories or qualities of work zone driving that tend to co-occur. The combinations of these characteristics could shed light on work zone and driver behavior characteristics that are more associated with SCEs or with baseline driving.

CONTRIBUTION OF INDIVIDUAL DRIVERS

A key result of this study is an emphasis on the contribution of individual drivers to risk in work zones. That is, the work zone treatments observed may be safe for undistracted drivers, but drivers are nonetheless often engaged in secondary tasks while driving in a work zone. The PCA analysis shows that the driver and the time of day made similar contributions to risk in SCEs. Time of day is related to the tendency for queueing to occur (i.e., when traffic volume is higher). Queueing is expected to contribute to risk, especially when secondary tasks are involved. Indeed, queuing produced out-of-ellipse driving and is also a key word in the SCE-associated topic 8, along with several kinds of secondary tasks. Topic 8 is associated with SCEs, further supporting the idea that the combination of distraction and queues forming is a particular mechanism for SCEs in work zones.

Individual drivers, however, make a similar contribution to risk (as defined by ellipse excursions) as time of day. That is, some drivers tend to have greater accelerations and decelerations in the same work zone at the same time of day as compared other drivers, contributing to the risk of SCEs.

The individual drivers observe and experience the changes and instabilities in the traffic flow in the forms of queuing, slowing down, conflicts (increased risk of rear-end crashes) during car following, lane changes, and takeover maneuvers. Previous work has shown that individual behavior can contribute to or mitigate this unstable traffic flow, even under the same traffic density conditions (Davis 2016). Thus, if individual drivers can be "nudged" toward apparently altruistic behavior (i.e., slowing down or speeding up for the sake of sustaining a stable traffic flow), they can improve both safety and throughput for the overall road network.

A potential area of development for work zone treatments is to identify more effective ways of nudging. For example, digital message signs, continuous posted speed limits, and audio broadcasts throughout the work zone (similar to information broadcasts in tunnels) all provide the same information to the whole driver population in a given work zone. However, the STM analysis highlighted how driver distraction and noncompliant behavior is often unrelated to the work zone equipment or treatments. Therefore, a more effective approach may require active guidance of individual drivers through the work zone while simultaneously measuring and monitoring the traffic flow in macroscopic scale to form the information needed for "nudging" purposes.

Individual guidance could be accomplished two broad ways: either targeted messaging to individual drivers, or controlling a subset of vehicles in the traffic stream that would then influence the compliance of the remainder of the vehicles. Individual messaging could be accomplished using I2V communication combined with the identification of "problem" individuals via the PCA safety ellipse algorithm (implemented on V2I messages). While broad implementation of this technology is some years away, in principle, the behavior of individuals could also be sensed from the roadside and targeted messages or enforcement could be implemented.

For nudging without individual messaging, the ideas of Davis (2016) have real potential. He simulated the potential of I2V controlling ACC-equipped vehicles such that they complied with VSLs. Interestingly, he found that when only one-third of the vehicles in the traffic stream were compliant (keeping both targeted speed and headway), the whole traffic stream was much closer to compliance and both throughput and safety were improved. Automated vehicle driving may be the most complex but best approach to this issue. Increased penetration of active safety systems like lane keeping and ACC may help actively guide vehicles through work zones if the vehicle and infrastructure handshaking can be achieved. This handshaking can be in the form of V2I communication where onboard systems react to infrastructure monitoring systems. ACC systems could be remotely controlled to introduce better vehicle speed distributions up and downstream from a work zone. Lane keeping systems need cues for the lane edges, and this requires that marking requirements for these systems are known and employed. Future systems could benefit from work-zone-specific markings that signal the driver when they are encouraged or discouraged to change lanes, reducing unstable traffic conditions.

Technological control of ACC settings is some years away, but the idea of placing compliant vehicles in the stream to improve the behavior of all vehicles could be achieved by other means. For example, heavy truck companies might be engaged to help by encouraging truck drivers with ACC to select particular settings to maintain headways. The desired set speed could be conveyed by digital message signs, thus requiring no new technologies to be in place. A relatively low level of compliance with suggested speeds and headways might improve traffic throughput and safety in general.

The results of the STM and PCA analyses agree that driver behavior dominates the safety performance of work zones. Distraction is a significant problem, and while free flow (or steady flow) conditions are associated with safer driving, they cannot always be achieved. Thus, additional means of encouraging drivers to be attentive are needed.

INDICATIONS FOR WORK ZONE TREATMENT CHOICE

The analyses of the PCA and STM techniques provide insight into work zone layout and operations. The results highlight the importance of smooth flowing traffic and reduced excursions of vehicles from a safe behavior ellipse defined in the PCA analysis. The STM analysis highlights the tendency of drivers to be distracted by construction activities, and this can be attributed to cases with conditions associated with dangerous (safety critical) conditions.

The ASAP project identified the desire by many transportation authorities to minimize the reduction of speeds from regular posted speeds (Nocientini et al. 2013). This strategy, when construction activities allow, needs to be reviewed in the context of potential external distractions. The frequency of distractions to construction activities outside the vehicle may be problematic if they occur at higher operational speeds where less maneuverability is offered in the transition or activity zone. The material available does not allow for a rigorous analysis of the influence of distraction sources (worker present, parked car, equipment, etc.). However, the common occurrence in the different topics, associated with baseline or safety-critical conditions, indicates that work zones have triggers for driver distraction that must be eliminated to reduce the crash risk.

The analyses in this project do not indicate whether specific observed work zone treatments are themselves causes of increased or decreased crash risk. Interestingly, digital message signs are the only work zone treatment variable associated strongly with baseline cases. This could indicate their benefit in communicating with drivers, a key theme of the project’s conclusions.

In general, the message of the SHRP2 data is that work zone treatments designed to get the attention and compliance of drivers have the best potential to improve safety. In other words, the results suggest that the recommendations for physical treatments in the form of barrels, cones, chicanes, etc., are likely to be reasonable. The key addition should be frequent and clear messaging to help drivers anticipate issues and pay attention to traffic. The more specific and individual the information, the better it may work.

INPUT FROM MICHIGAN DEPARTMENT OF TRANSPORTATION (MDOT) ON CURRENT CHALLENGES AND PLANS

Project participants met with representatives of MDOT. There were a few key themes for this discussion.

First, MDOT has embraced communication technology and has been installing the necessary infrastructure to enable I2V or V2I communication in the not-too-distant future. They envision work zones as an early use case, but primarily as information for equipped vehicles. By populating and updating the traffic flow restrictions (TFR) database, work zone information can be obtained via messaging in real time.

One major challenge for the implementation of I2V/V2I communication, however, is the establishment of standards, especially for WZR messages. Another is reaching a critical mass of equipped vehicles to make use of messages. Many of the Detroit-based automakers have pledged to install communication capability in their new vehicles, so there is some hope that the critical mass will be reached sooner rather than later. However, the value of MDOT's improved communication and information infrastructure can only be realized if manufacturers make use of it. Thus, MDOT must rely on manufacturers and app developers to see potential benefits from these systems.

A second theme of discussion revolved around the significant challenges MDOT faces in just knowing where construction workers, signs, and equipment are at any given moment. MDOT has begun to invest in equipment that can be attached to signs to track them and is evaluating vests and other devices that can be worn by workers to enable tracking them as well. This is critical to any work zone implementation because work zones change and move over time, signs move or are covered or knocked over, and information given to drivers needs to be consistent across physical signs, digital message boards, and I2V messaging.

Third, MDOT is also evaluating and beginning to implement technology to measure speeds in work zones and use them to identify the back of queues. When speeds go below a certain level, this is considered the back of the queue. This information is, in turn, posted on digital message boards to inform drivers more accurately of when they will see slowdowns. Devices that are typically used for traffic counting (which includes lane-specific counts and vehicle type identification) can be used for speed estimation, but these are generally expensive and can measure more than is needed for this purpose. MDOT is evaluating inexpensive roadside sensors that only measure speeds. At this time, only 10 percent of work zones where queues develop have speed sensors.

Finally, MDOT has been considering working with heavy truck companies to make use of the communication around work zones. Heavy trucks are a potential partner in maintaining VSLs and tempered acceleration and deceleration in the same way that Davis envisioned using ACC-equipped vehicles to achieve the same goal (Davis 2016).

LIMITATIONS AND FUTURE WORK

During the analysis of work zone driver behavior, a shortcoming in the established SHRP2 data analysis was observed. The original annotation by the subcontractor that archives the SHRP2 dataset of SCEs was conducted after the initial data collection period based on kinematic triggers (e.g., hard braking or crashes) of the subject vehicle. These events were coded using a standard codebook to describe the traffic and driver situations observed in the video records. Driver behavior was only available in the archiving subcontractor annotations that have predominantly focused on the subject vehicle's kinematics and driver behavior.

Review of the work zone SCEs indicates that many work zone SCEs are observed after lane changes of a surrounding vehicle occurred. While this behavior may be indirectly observed in "queue forming" or "unstable flow" variables, there is no reliable real-time information on the relative velocity of surrounding traffic. As a result, observable traffic parameters like lane changes are not explicitly available for analysis.

An exploratory investigation was made to determine if the kinematics of neighboring traffic could be automatically processed from the existing forward video data from the subject vehicle. Figure 41 shows vectors representing the relative motion of image elements such as signs and neighboring vehicles. The size and direction of the vectors indicate not only the subject vehicle's own relative motion to the environment, but more importantly the motion of other vehicles that could be processed to quantify lane changes downstream from the subject vehicle, which can result in changing traffic flows and potential conflicts for the subject vehicle's forward motion. This video processing requires more development, but indicates the need to extend the current SHRP2 analysis techniques to address traffic behavior downstream, independent from the subject vehicle's own kinematics.

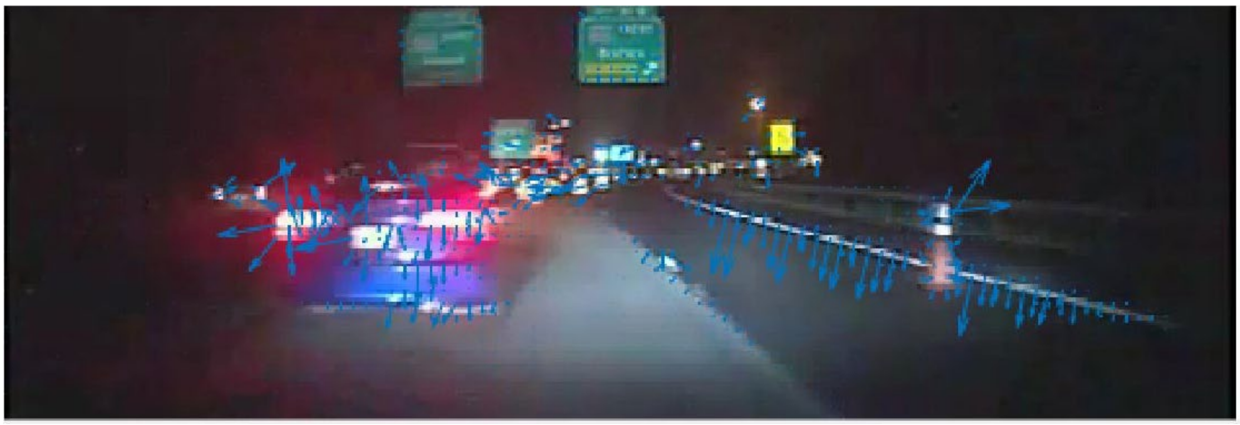


Figure 41. Photo. Extraction of relative velocity of surrounding objects.

Another limitation for application of this work is that the data were continuous and longitudinal measurements of an individual vehicle's kinematic data (velocity and position). In contrast, general traffic flow monitoring and management applications of highways are observed using cameras or loop-counters located in a fixed position and covering only a short segment of the road network. Thus, the analysis focused on longitudinal data, while a typical implementation would be based on cross-sectional measurements (other than via V2I). A focused future analysis of the vehicle kinematic data in SHRP2 could allow both the long-term longitudinal measurements along the road network and at each location to be combined to estimate cross-sectional traffic information. Such data would allow for the development of the relationship between drivers' longitudinal movements and cross-sectional observations of several vehicles in the traffic flow.

APPENDIX A. VIDEO CODING

FORWARD VIDEO ANNOTATION (FORWARD VIDEO CODING)

Additional annotation was implemented to annotate the refined dataset of only interstate events by coding the forward video. The coding protocol included the following variables to identify and categorize the work zone treatments with regard to control, activity, and type:

- Control:
    - Traffic signs (billboard style with text).
    - Barrels or cones.
    - Sign and text for a fine (special sign).
    - Warning lights on the sign (flashing lights on).
    - Chicane (S-shape).
    - Digital sign.
    - TMA (truck with obvious sign for work zone).
- Activity:
    - Workers present.
    - Equipment (on the road or on the side).
    - Cars parked along the road.
- Type:
    - Divert traffic.
    - Decrease lane (width or number).
    - Closed lane.
    - Mobile work zone (vehicle with flashing lights).
- Time of "control" (traffic sign, digital sign, warning light).
- Event nature (only for baseline events):
    - Event with a following vehicle.
    - Event with a lead vehicle.
    - Event with animal.
    - Event with merging vehicle.
    - Event with obstacle/object in roadway.
    - Event with oncoming traffic.
    - Event with pedestrian.
    - Event with vehicle in adjacent lane.
    - Event in intersection.
    - Other.
    - Single vehicle.
    - Unknown conflict.

Narrative—text comment from the annotator.

FACE VIDEO ANNOTATION

Secondary Tasks

If any of the following secondary tasks were observed, they were coded, along with start and end times of engagement in the secondary task:

- Talking/singing, audience unknown.
- Dancing.
- Reading.
- Writing.
- Passenger in adjacent seat, interaction.
- Passenger in rear seat, interaction.
- Child in adjacent seat, interaction.
- Child in rear seat, interaction.
- Moving object in vehicle.
- Insect in vehicle.
- Pet in vehicle.
- Object dropped by driver.
- Reaching for object, other.
- Object in vehicle, other.
- Cellphone, holding.
- Cellphone, talking/listening, handheld.
- Cellphone, talking/listening, hands-free.
- Cellphone, texting.
- Cellphone, browsing.
- Cellphone, dialing handheld.
- Cellphone, dialing handheld using quick keys.
- Cellphone, dialing hands-free using voice-activated software.
- Cellphone, locating/reaching/answering.
- Cellphone, other.
- Tablet device, locating/reaching.
- Tablet device, operating.
- Tablet device, viewing.
- Tablet device, other.
- Adjusting/monitoring climate control.
- Adjusting/monitoring radio.
- Inserting/retrieving CD (or similar).
- Adjusting/monitoring other devices integral to vehicle.
- Looking at previous crash or incident.
- Looking at pedestrian.
- Looking at animal.
- Looking at an object external to the vehicle.
- Distracted by construction.

- Other external distraction.
- Reaching for food-related or drink-related item.
- Eating with utensils.
- Eating without utensils.
- Drinking with lid and straw.
- Drinking with lid, no straw.
- Drinking with straw, no lid.
- Drinking from open container.
- Reaching for cigar/cigarette.
- Lighting cigar/cigarette.
- Smoking cigar/cigarette.
- Extinguishing cigar/cigarette.
- Reaching for personal body-related item.
- Combing/brushing/fixing hair.
- Applying makeup.
- Shaving.
- Brushing/flossing teeth.
- Biting nails/cuticles.
- Removing/adjusting clothing.
- Removing/adjusting jewelry.
- Removing/inserting/adjusting contact lenses or glasses.
- Other personal hygiene.
- Other nonspecific internal eye glance.
- Other known secondary task.
- Unknown type (secondary task present).
- Unknown.

Impairments

Impairments were coded based on the variable driver impairments used in SHRP2 dictionary (Hankey et al. 2016):

- None apparent.
- Drowsy, sleepy, asleep, fatigued.
- Blackout.
- Angry.
- Other emotional state.
- Drugs, medication.
- Drugs, alcohol.
- Other illicit drugs.
- Restricted to wheelchair.
- Impaired due to previous injury.
- Deaf.

- Other.
- Unknown.

Hands Ever Off (Wheel)

This variable could have values of “both hands off,” “unknown,” or “not applicable” (at least one hand was always on the wheel). When the value was “both hands off” or “unknown,” the timestamps at start and end were coded for each event where this was observed.

APPENDIX B. PCA BACKGROUND

The following text is largely taken from the *Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones: Phase I Final Report* for this project (Flannagan et al. 2016). It is included to provide general background on PCA. It has been edited to improve accessibility.

GENERAL DESCRIPTION OF PCA

The goal of the PCA approach is to create an efficient description of any driving epoch (e.g., a 1-s sample). Using this description, investigation can focus on the distribution of descriptors for normal driving and identify occasions when driving falls outside of this distribution. The inclusion of crash-related events (CREs) in the sample allows for the exploration of how such events might differ from baseline on the developed description.

PCA is a data reduction method that was developed over 100 yr ago (Pearson 1901). It is used with datasets that have a large number of correlated variables for each observation. These datasets can be difficult to use in analysis. PCA serves two key purposes:

- Reduces the total number of variables needed to describe each observation.
- Creates orthogonal variables, making it possible to use these variables as predictors without concern about collinearity problems.

The starting point for PCA is a set of variables that are observed for a reasonably large number of cases. PCA creates a set of new variables, or dimensions, each of which is a linear combination of the original variables. These new variables are called PCs. Each PC is also orthogonal to all other PCs—a quality that is not true of the original variables. Finally, PCs are selected such that the first PC is the linear combination of variables that accounts for the most variance in the dataset, followed by the linear combination of variables that is 1) orthogonal to the first PC, and 2) accounts for the most variance among all possible orthogonal PCs. The third must be orthogonal to the first two, and so on. The last PC is determined because there will be only one remaining linear combination of variables that is orthogonal to all of the others.

Once the full set of PCs is determined, the linear combinations associated with each PC can be used to calculate a value, or "score," for each original point in the dataset. For example, suppose a dataset has 20 variables for each observation. PCA can produce 20 new scores (also "dimensions" or "variables") that retain all of the original information (i.e., the 20 PC scores can be used to reconstruct the original variables for every observation in the dataset). However, because PCs are created in order of the amount of variance accounted for, the information in each dataset is typically described with almost equal precision by using only the first few PCs. Thus, by dropping most of the PCs, the same data that were originally characterized by 20 variables can be characterized with many fewer (often 3–5).

A simple example is shown in figure 42 and figure 43. Here, two variables, *X* and *Y*, are highly correlated, as shown in figure 42. Triangles in red are used to illustrate a representation of the first PC, as shown in figure 43. The line shows the first PC, which accounts for 95.2 percent of the total variance in the data. The score for each point is the location along the PC line at the intersection of the perpendicular line segment that passes through the point being scored. This score, in arbitrary units, is the distance along the PC line from the origin.

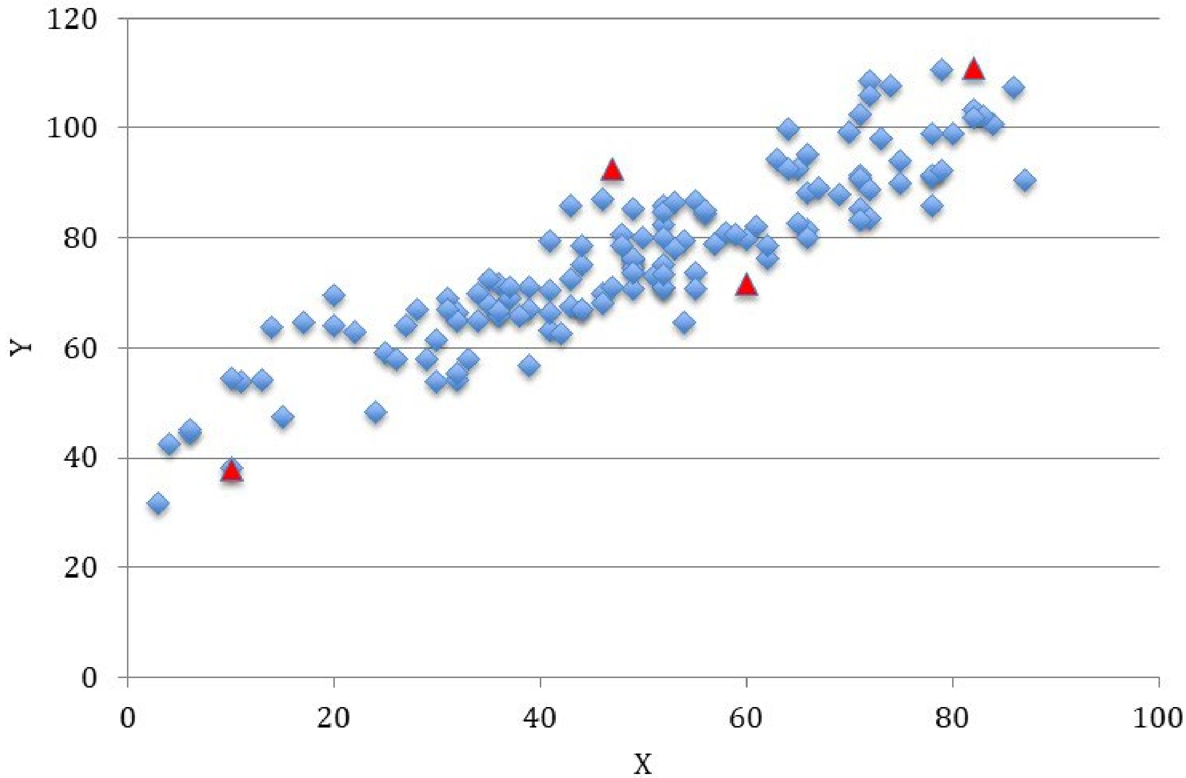


Figure 42. Graph. PCA illustration. Triangles and diamonds represent points plotted on two hypothetical dimensions, X and Y (modified from Flannagan et al. 2016).

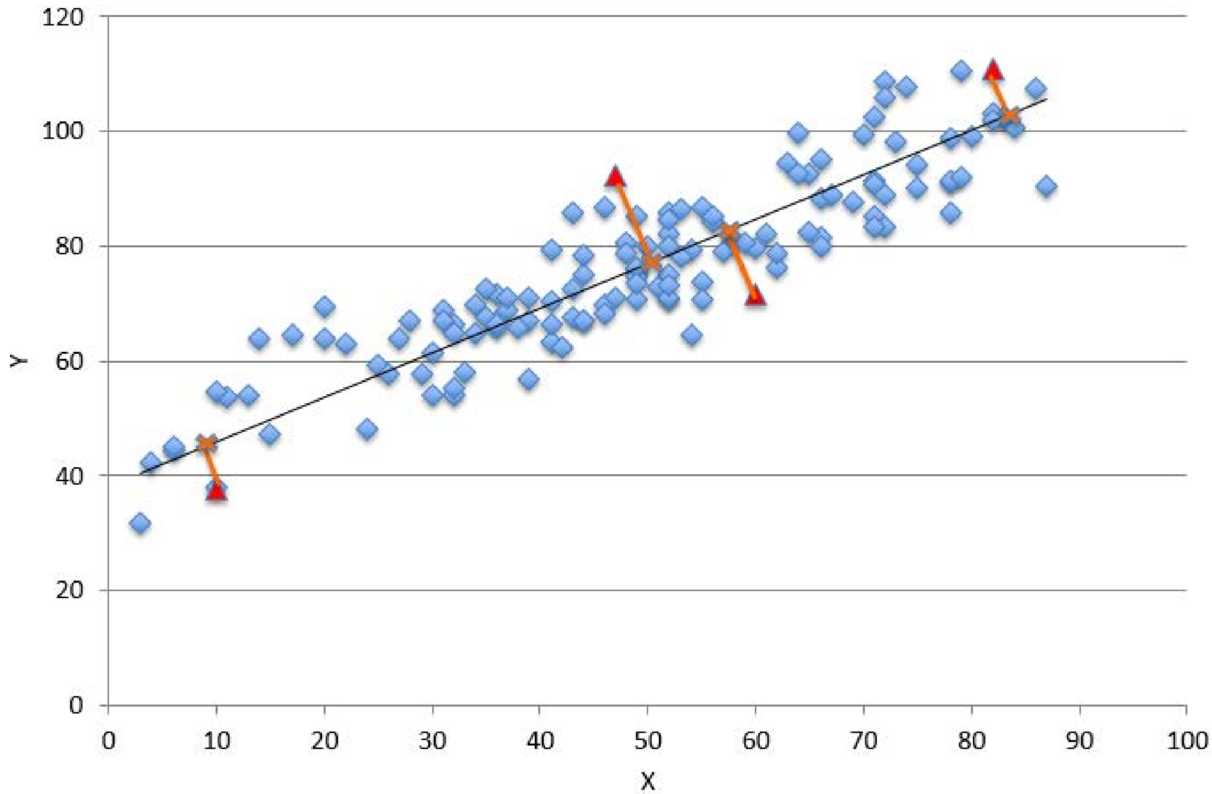


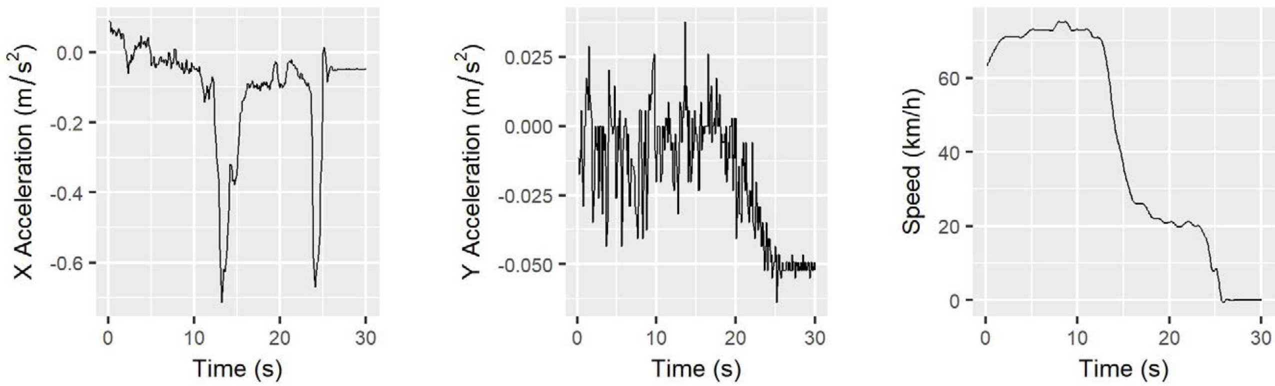
Figure 43. Graph. Illustration of first PC and scoring of each point (modified from Flannagan et al. 2016).

Although the variance associated with the length of each of the short line segments in figure 38 is lost by using only the first PC score to describe each point, the description is very efficient and captures the majority of the variance in the whole dataset. When each point has 20 dimensions or more, the reduction to a few dimensions that capture the majority of variance is efficient and helpful for analysis. Thus, PCA is a dimension-reduction method that also serves to orthogonalize the retained dimensions. (Note that the orthogonal nature of each subsequent PC was not illustrated.) A challenge of using PCA is that the PC scores are often difficult to interpret. Interpretation is sometimes possible, but in general, PCA results can be useful for computation and prediction but not necessarily for interpretation.

Once PCs were constructed, the goal was to use them to distinguish between baselines (normal) driving and CREs. This was done using a set of candidate methods including Linear Discriminant Analysis, random forests, and BART. Attempts at using Quadratic Discriminant Analysis showed that it performed consistently less well than the other methods, so it was dropped. These methods are described in more detail below.

Construction of PCs

In this application, each event contains data at 0.1-s intervals for certain kinematic measures. This is illustrated in figure 44 through figure 46 in which time series for longitudinal (x) acceleration, lateral (y) acceleration, and speed are shown for the entire event.



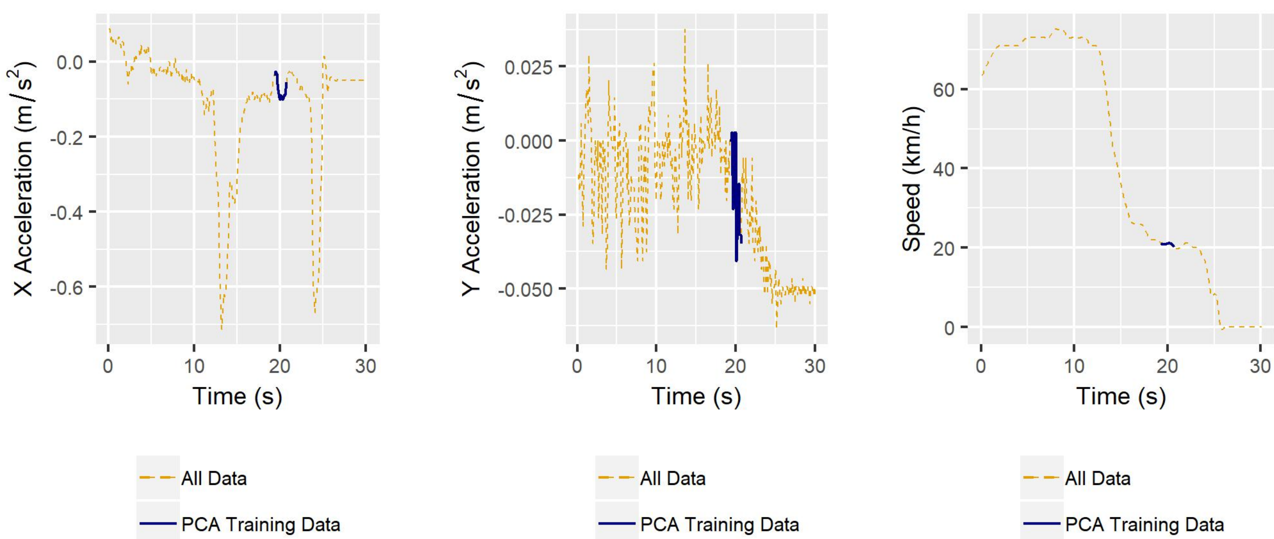
A. Example of time series x-acceleration for one crash event.

B. Example of time series y-acceleration for one crash event.

C. Example of time series speed for one crash event.

Figure 44. Graph. Example time series for one crash event: (left to right) x-acceleration, y-acceleration, speed.

Figure 45 shows a closer view of 1.5 s taken from the time series in figure 44. Figure 45 shows the individual datapoints for each of the three signals for the 1.5-s interval.



A. Example of time series x-acceleration for one crash event.

B. Example of time series y-acceleration for one crash event.

C. Example of time series speed for one crash event.

D. Closeup of 1.5-s sample x-acceleration.

E. Closeup of 1.5-s sample y-acceleration.

F. Closeup of 1.5-s sample speed.

Note: Subfigure A through subfigure C show a random 1.5-s sample, and subfigure D through subfigure F show a closer view with individual points that make up that sample.

Figure 45. Graph. Closeup of data for 1 s of the example event (lines and points in graphs).

Each 1-s sample of data can thus be described by 10 times the number of measurements (k) taken at each sampling timepoint (i.e., *10k* variables describe what occurred during that 1 s of driving). This dataset (illustrated in figure 42) has 3 kinematic measures, resulting in 30 variables for each sample (10 for longitudinal acceleration (a_{long1}, a_{long2}, etc.), 10 for lateral acceleration (a_{lat1}, a_{lat2}, etc.), and 10 for velocity (v_1, v_2, etc.)). Within these 30 variables, values are highly correlated. For example, longitudinal acceleration at time 0.1 s is highly correlated with longitudinal acceleration at time 0.2 s, making this an ideal application for PCA. PCA should produce a small set of linear combinations of these variables (i.e., PCs) that can describe the same 1 s of data, but far more efficiently than the original 30 variables.

Since the goal is to develop a method that can be used in applications, identifying PCs specific to this sample were not of interest. Instead, using PCs on future driving data was the target of interest. Thus, the first step was to establish whether a given set of linear combinations (PC scores) would be stable as a way of describing many different driving epochs.

Evaluating the stability of PC scores across different driving samples was done as follows. A short interval from each event was randomly selected on which to build the PC scores. Events have different lengths, but are typically about 30 s long. The short interval length was also varied to identify the event length that was most stable. For this purpose, 1.5 s proved to be most stable.

Next, the randomly selected 1.5-s intervals were used to develop PCs. Interestingly, lateral acceleration did not improve either the development of PCs or (later) the prediction of abnormal driving. Thus, three PCs based only on longitudinal acceleration and speed were sufficient to capture most of the variance in each 1.5-s period of driving. As described above, each PC is a linear combination of the original variable values (i.e., longitudinal acceleration and speed), and thus it is simply a set of coefficients that can be applied to any 1.5-s interval of driving for which those same variables are observed.

Using the PC score coefficients developed on the random sample, the three PCs were calculated for every 1.5-s interval of driving in the dataset. The amount of variance in the 1.5-s intervals was accounted for by PC scores that were developed on the random sample that was then calculated. In the original (training) sample, 98 percent of the variance was accounted for, and in the test sample (i.e., all other driving), 96 percent of the variance in the original measures was accounted for.

Finally, this process was repeated 100 times using different random samples to look at the overall stability of the PC scores and subsequent analysis using those scores. PC scores were quite stable across different random samples and different 1.5-s intervals of driving. Thus, analyses concluded that using three PC scores could efficiently describe any 1.5-s period of driving (at least in this sample).

To illustrate, figure 1 shows a set of these 1.5-s intervals, plotted in PC space, rather than the original time series (as in figure 40). Black dots are 1.5-s intervals of baseline driving, and the dots connected by the red line are a series of 1.5-s intervals from a CRE.

There are several notable features of figure 1. First, most of the black dots cluster in the center of the space. This represents "normal" driving, where all three PC scores for a large number of 1.5-s samples of driving are like a lot of other 1.5-s samples of driving. Second, the set of points connected by the red lines is a sequence of 1.5-s samples from a CRE. Note that during driving, when three PCs describe each 1.5-s sample, the points can move in PC space over the course of the drive. The connected dots show how one sample progresses over time outside of the "normal" driving cloud of points. These points move into a part of the space where "abnormal" driving is suspected. Finally, the axes of this space (i.e., the three PCs) are somewhat difficult to interpret. However, they can be loosely thought of as follows:

PC 1: [Average speed] + [average x acceleration]; weighted toward speed.
Interpretation: forward momentum; PC1 increases as speed or acceleration increases.
PC 2: [Average x acceleration] – [average speed]; weighted toward x acceleration.
Interpretation: changing speed; PC2 has high magnitude when changing speeds (slow + high acceleration or fast + negative acceleration).
PC 3: [Early x acceleration] – [late x acceleration]; weighted toward start and end; mostly ignores speed.
Interpretation: changing acceleration; PC3 has high magnitude when acceleration changes sign over the timeframe (especially in the middle).

APPENDIX C. TOPIC ASSOCIATIONS WITH EVENT TYPE

In figure 46 through figure 53, the left side shows the association of each event with the topic being graphed. Events are sorted by baseline first, followed by SCEs, which make up the last 47 events. Data for topic 1 are in figure 14, while data for topic 2 are in figure 15.

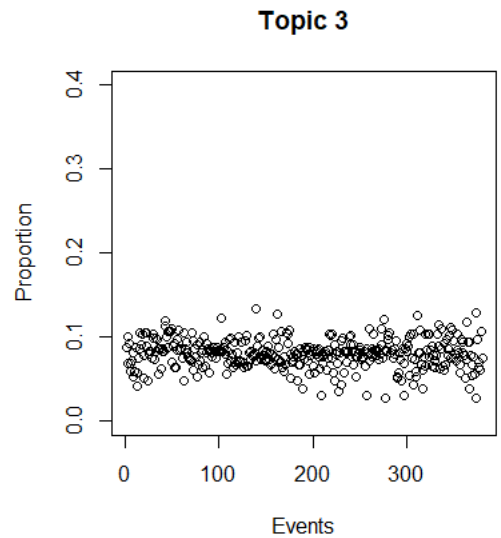


Figure 46. Graph. Association of each event with topic 3.

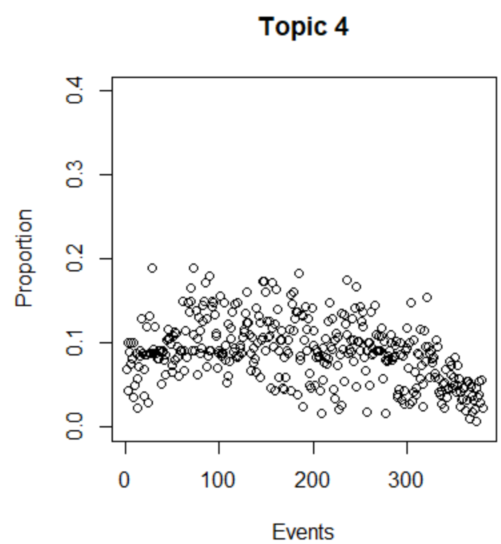


Figure 47. Graph. Association of each event with topic 4.

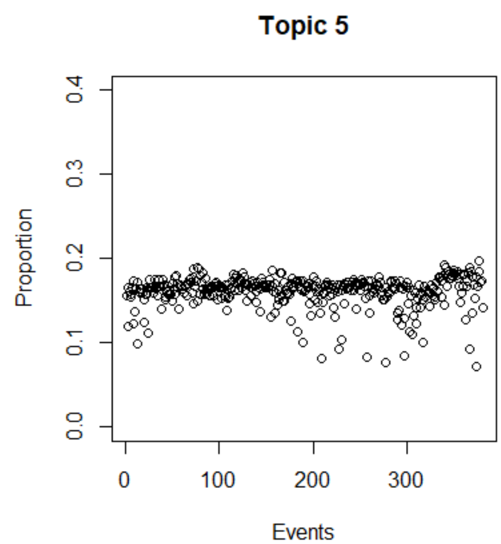


Figure 48. Graph. Association of each event with topic 5.

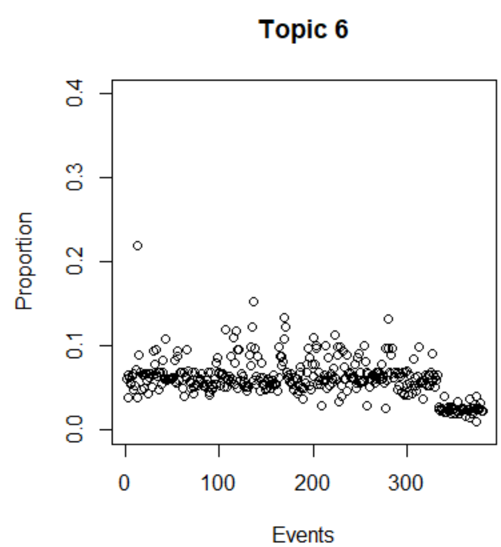


Figure 49. Graph. Association of each event with topic 6.

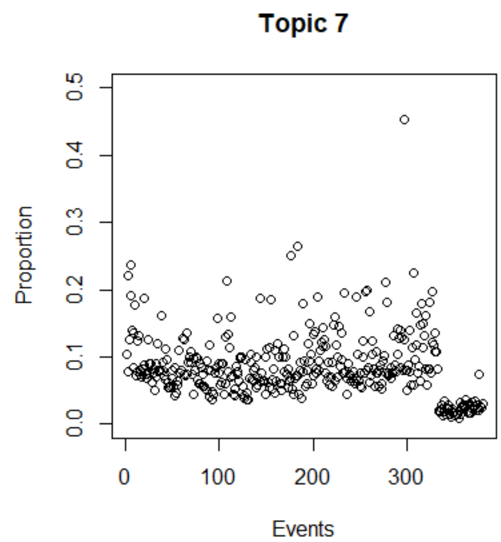


Figure 50. Graph. Association of each event with topic 7.

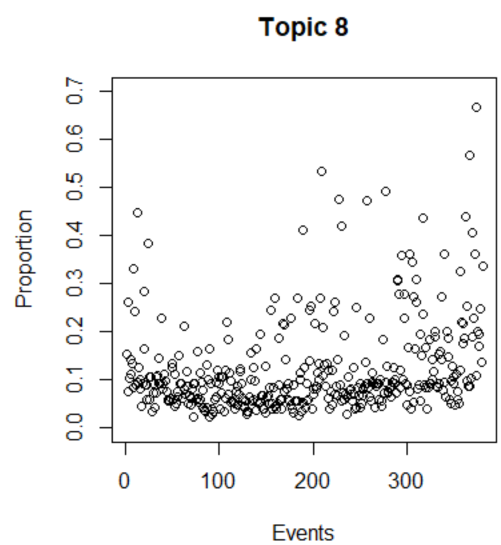


Figure 51. Graph. Association of each event with topic 8.

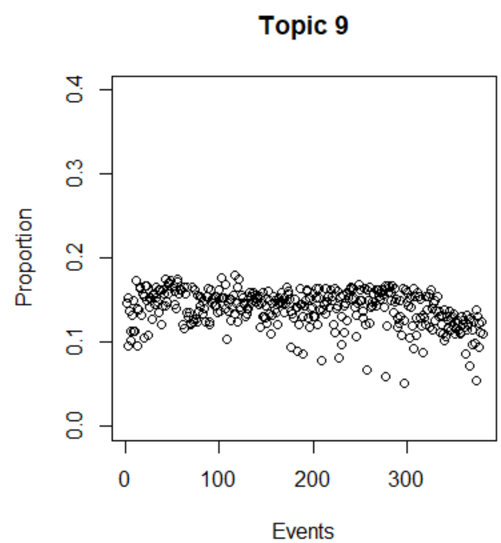


Figure 52. Graph. Association of each event with topic 9.

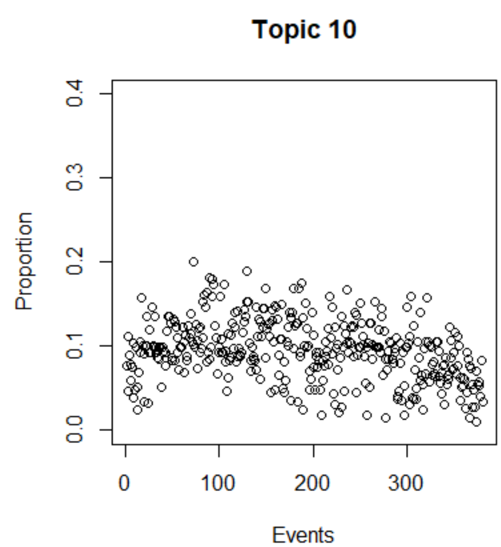


Figure 53. Graph. Association of each event with topic 10.

REFERENCES

Bharadwaj, N., Edara, P., Sun, C., Brown, H., and Chang, Y. (2018). “Traffic Flow Modeling of Diverse Work Zone Activities.” *Transportation Research Record*, *2672*(16), pp. 23–34.

Bischof, J. and Airoldi, E. (2012). “Summarizing Topical Content with Word Frequency and Exclusivity.” *Proc., 29th International Conference on Machine Learning*, J. Langford and J. Pineau, eds., Vol. 2, Omnipress, Madison, WI, pp. 201–208.

Blei, D.M. (2012). “Probabilistic Topic Models.” *Communications of the ACM*, *55*(4), pp. 77–84.

Blei, D.M. and Lafferty, J.D. (2007). “A Correlated Topic Model of Science.” *The Annals of Applied Statistics*, *1*(1), pp.17–35.

Blei, D.M., Ng, A.Y., and Jordan, M.I. (2003). “Latent Dirichlet Allocation.” *Journal of Machine Learning Research*, *3*, 993–1022.

Box, G.E.P., Jenkins. G.M., and Reinsel, G.C. (1994). *Time Series Analysis: Forecasting and Control*, Prentice Hall, Upper Saddle River, NJ.

Cheng, Y. and Langbort, C. (2016). “Model of Informational Nudging in Transportation Networks.” *Proc., 55th IEEE Conference on Decision and Control (CDC)*, IEEE Control Systems Society, New York, pp. 7598–7604.

Chipman, H.A., George, E.I., and McCulloch, R.E. (2010). “BART: Bayesian Additive Regression Trees.” *The Annals of Applied Statistics*, *4*(1), 266–298.

Davis, L.C. (2016). “Improving Traffic Flow at a 2-to-1 Lane Reduction with Wirelessly Connected, Adaptive Cruise Control Vehicles.” *Physica A: Statistical Mechanics and its Applications*, *451*, pp. 320–332.

Farrahi, K. and Gatica-Perez, D. (2011). “Discovering Routines from Large-Scale Human Locations using Probabilistic Topic Models.” *ACM Transactions on Intelligent Systems and Technology (TIST)*, *2*(1).

Flannagan, C., Kovaceva, J., Selpi, Leslie, A.J., and Thomson, R. (2016). “Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones: Phase I Final Report.” *University of Michigan Transportation Research Institute*.

Hankey, J.M., McClafferty, J.A., and Perez, M.A. (2016). “Description of the SHRP 2 Naturalistic Database and the Crash, Near-Crash, and Baseline Data Sets.” *Virginia Tech Transportation Institute*.

Huynh, T., Fritz, M., and Schiele, B. (2008). “Discovery of Activity Patterns using Topic Models.” *Proc., 10th International Conference on Ubiquitous Computing*, ACM, New York, pp. 10–19.

Iordanidou, G.R., Papamichail, I., Roncoli, C., and Papageorgiou, M. (2016). "Integrated Motorway Traffic Flow Control with Delay Balancing." *IFAC-Papers Online*, *49*(3), pp. 315–322.

Kovaceva, J., Tan, V., Selpi, Leslie, A.J., and Flannagan, C. (2018). "Methods Interim Report: Reporting Task D3.1 for Analysis of SHRP2 Data to Understand Normal and Abnormal Driving Behavior in Work Zones." *University of Michigan Transportation Research Institute*.

McLaurin, E.J., Lee, J.D., McDonald, A.D., Aksan, N., Dawson, J., Tippin, J., and Rizzo, M. (2018). "Using Topic Modeling to Develop Multi-Level Descriptions of Naturalistic Driving Data from Drivers with and without Sleep Apnea." *Transportation Research Part F: Traffic Psychology and Behavior*, *58*, pp. 25–38.

Nocientini, A., La Torre, F., Thomson, R. (2013). "Deliverable 2.1, State of the Art Report on Speed Management Methods." *Appropriate Speed Saves All People (ASAP)*.

Pearson, K. (1901). "LIII. On Lines and Planes of Closest Fit to Systems of Points in Space." *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, *2*(11), pp. 559–572.

Roberts, M.E., Stewart, B.M., and Tingley, D. (2014). "STM: R Package for Structural Topic Models." *Journal of Statistical Software*.

Sorensen, G., Bolling, A., Vadeby, A., Cocu, X., Nocentini, A., Aleksa, M., and Saleh, P. (2015). "Deliverable 5.1, Final Report." *Appropriate Speed Saves All People (ASAP)*.

Steyvers, M. and Griffiths, T. (2007). "Probabilistic Topic Models." *Handbook of Latent Semantic Analysis*, Laurence Erlbaum, Mahwah, NJ, pp. 427–448.

Transportation Research Board. (2013). "The 2nd Strategic Highway Research Program Naturalistic Driving Study Dataset." *The National Academies of Sciences, Engineering, and Medicine*.

Federal Highway Administration. (2009). "*Manual on Uniform Traffic Control Devices*," *U.S. Government Publishing Office, Washington, DC*.

van der Maaten, L., Postma, E., and van den Herik, J. (2009). "Dimensionality Reduction: A Comparative Review." *Journal of Machine Learning Research*, *10*, pp. 66–71.

Wang, X., Mueen, A., Ding, H., Trajcevski, G., Scheuermann, P., and Keogh, E. (2013). "Experimental Comparison of Representation Methods and Distance Measures for Time Series Data." *Data Mining and Knowledge Discovery*, *26*(2), pp. 275–309.

Yu, R. and Abdel-Aty, M. (2014). "An Optimal Variable Speed Limits System to Ameliorate Traffic Safety Risk." *Transportation Research Part C: Emerging Technologies*, *46*, pp. 235–246.

Recycled
Recyclable

HRDS-20/12-19(WEB)E