Using Big Data Safety Analytics for Proactive Traffic Management

Dr. Mohamed Abdel-Aty, PE
Trustee Chair
Pegasus Professor and Chair
Department of Civil, Environmental & Construction Engineering

UCF

University of Central Florida, USA

Editor-in-Chief, Accident Analysis & Prevention

Purpose

- **Challenges** in the transportation system
  - Increase in travel demand
  - Growth in congestion
  - Need to improve safety
  - Reality of limited resources
- **Solution**: (Pro)Active Traffic Management
  - Dynamically manage recurrent and non-recurrent (incident) congestion based on prevailing traffic conditions
- **Benefits**
  - Maximize the efficiency of the facility
  - Increase safety

Proactive Perspective

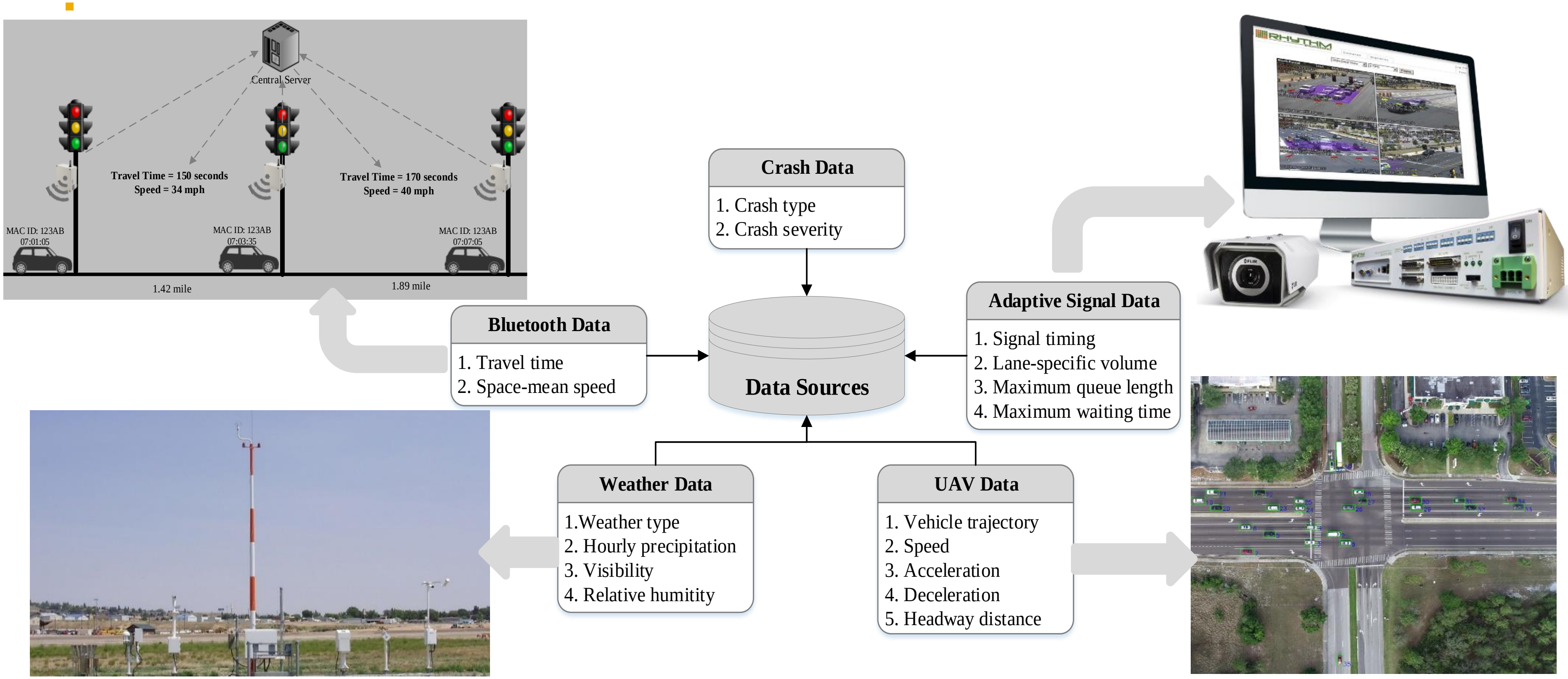
- Traditional Approach
  - Where congestion/queues have formed
  - Where the incident has occurred
  - Where inclement weather has been detected
- ***Proactive*** Perspective
  - Where the congestion/queues are <u>about to form</u>
  - Where a crash is <u>more likely to occur</u>
  - Where inclement weather is <u>about to begin</u>

❖ **Key: prediction in real-time**

Data and Monitoring

- Application of Intelligent Transportation Systems (ITS)
  - **Traffic Detection Systems**
    - Automatic Vehicle Identification (AVI) Systems
    - Microwave Vehicle Detection Systems (MVDS)
  - **Weather Detection Systems**
    - Weather sensors (e.g. temperature, precipitation,
  - **Countermeasures**
    - Variable Speed Limit
    - Ramp Metering strategies
    - Queue Warning
    - Dynamic Rerouting and Traveler Information
    - Adaptive Signals

Data Availability



Data Collection

- Currently Available Data Source Characteristics

Data	Availability		Contents				Granu-larity	Source Agency
	Freeway & Expressway	Arterial	TMS[1]	SMS[2]	Volume	Incident		
MVDS	√		√		√		30s-60s	CFX, RITIS
AVI(Toll tag)	√			√			Individual	CFX
BlueMAC[3]		√		√			Individual	Orange County
Iteris[3]		√		√			Individual	Orange County
BlueTOAD[3]		√		√			Individual	Seminole County, FTE
InSync		√			√		15min	Orange County
SPM		√			√		Individual[4]	Seminole County
HERE	√	√		√			1min	RITIS
INRIX				√			1min	RITIS
Twitter	√	√	Congestion Only			√	–	Twitter Mining

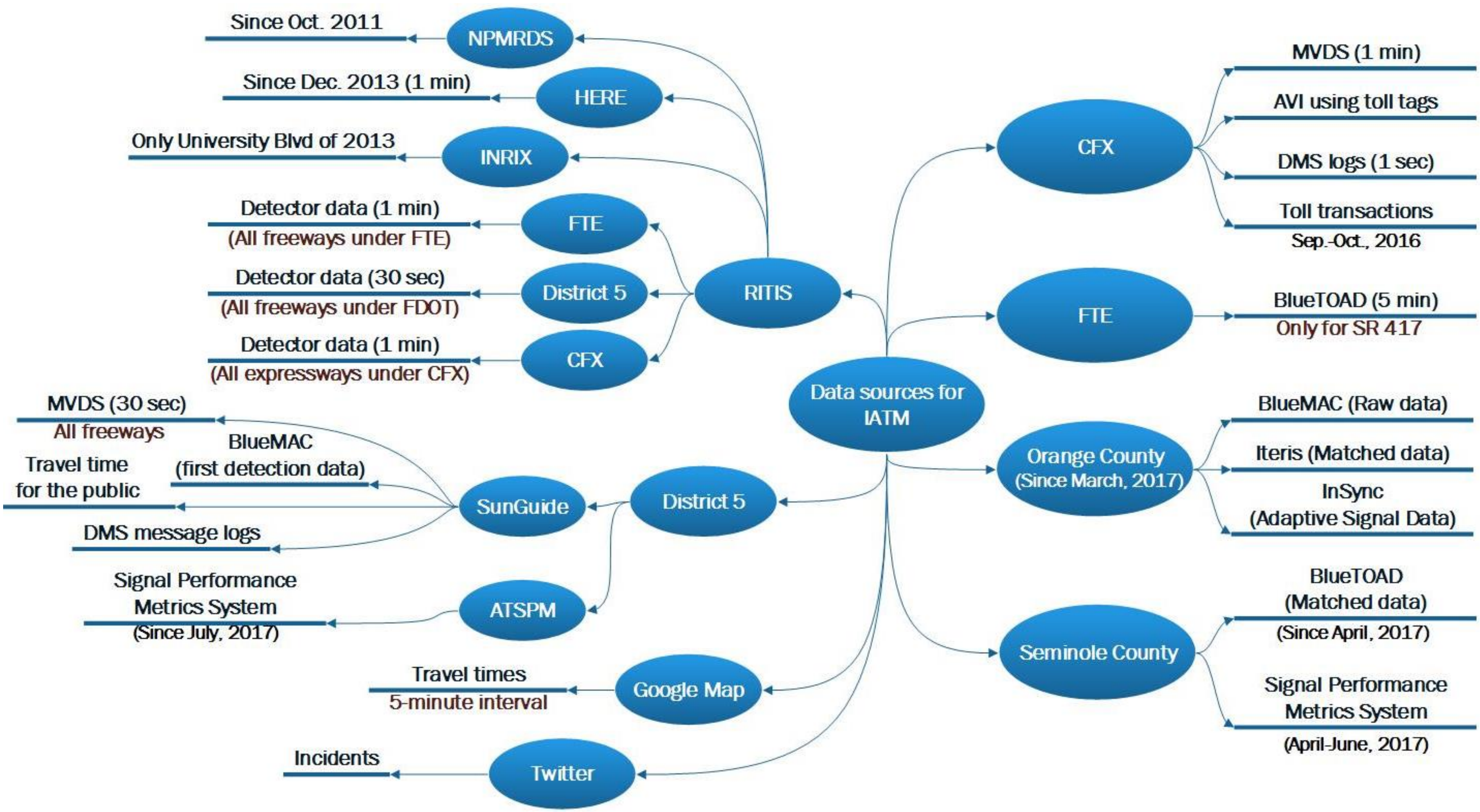
1 Time Mean Speed;
2 Space Mean Speed;
3 All are Bluetooth system, only BlueMAC provides raw log data;
4 SPM records individual count.

Data Collection Technologies

- Size of Current Datasets
  - In order to manage, store, and utilize the collected big data efficiently, the UCF research team has purchased two high-end computing severs with in-house funding

Source	Name of Dataset	Monthly raw data size (Gbytes)	Expected periods (Years)	Expected Size (GB, as of 2017)	Format	Data collection period
CFX	MVDS	11	5	660.00	dat	From 2013 to 2017
	AVI	10	5	600.00	csv	From 2013 to 2017
	DMS	0.34	5	20.51	csv	From 2013 to 2017
Seminole County	SPM	125	3	4500	MS SQL Server backup database	From January 2015
	BlueTOAD	0.3	4	14.4	csv	From 2014 to 2017
Orange County	InSync	0.35	2	8.44	csv	From 2015 to 2017
	BlueMAC	2	1	24.00	csv	Detector deployed since Dec 2016. Still under deployment.
	Iteris	0.25	3	9.00	csv	From 2015 to 2017
RITIS	MVDS	6.1	4	292.80	csv	From 2014 to 2017
	HERE	8.20	4	393.60	csv	From Oct, 2013 till now
Total		164		6523		

Summary of Data Collection

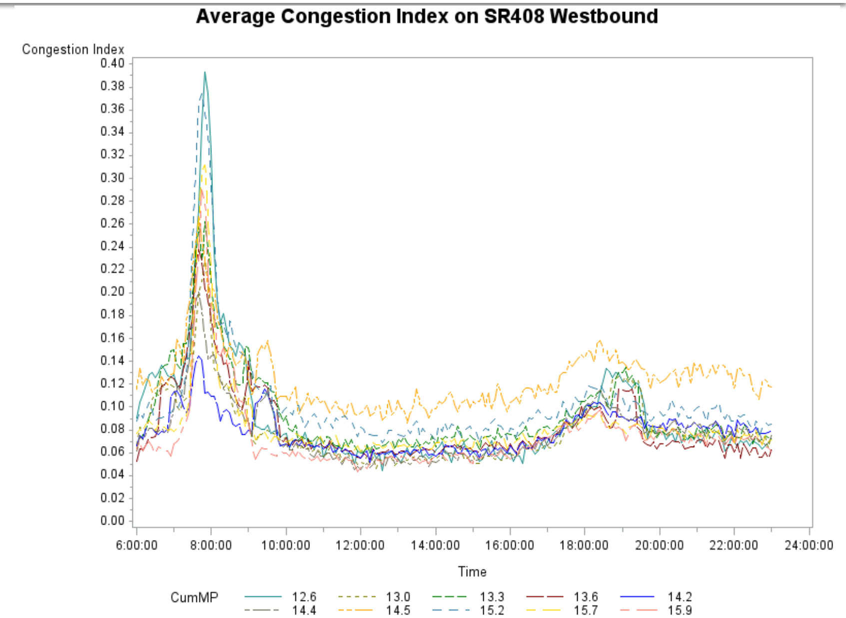


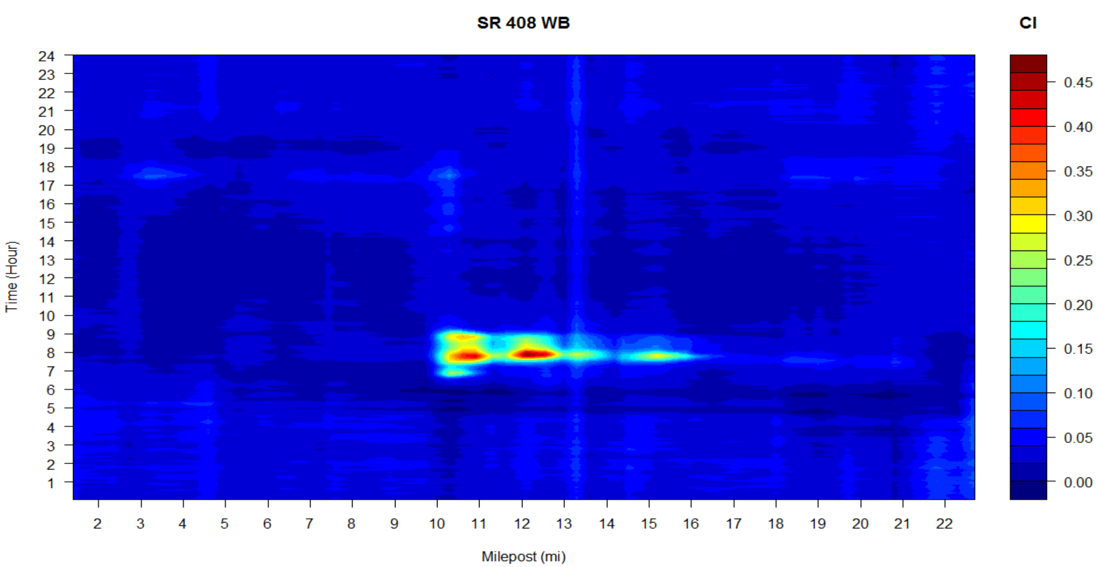
Application of Big Data in Operation: Congestion Measurement

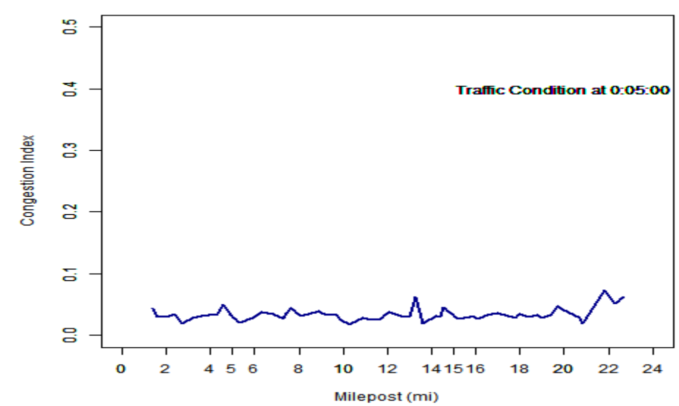
- MVDS Congestion Index
  - Congestion Index (CI) =

$$\frac{\text{Free Flow Speed} - \text{Speed}}{\text{Free Flow Speed}}$$

  - Free Flow Speed = 85th percentile speed

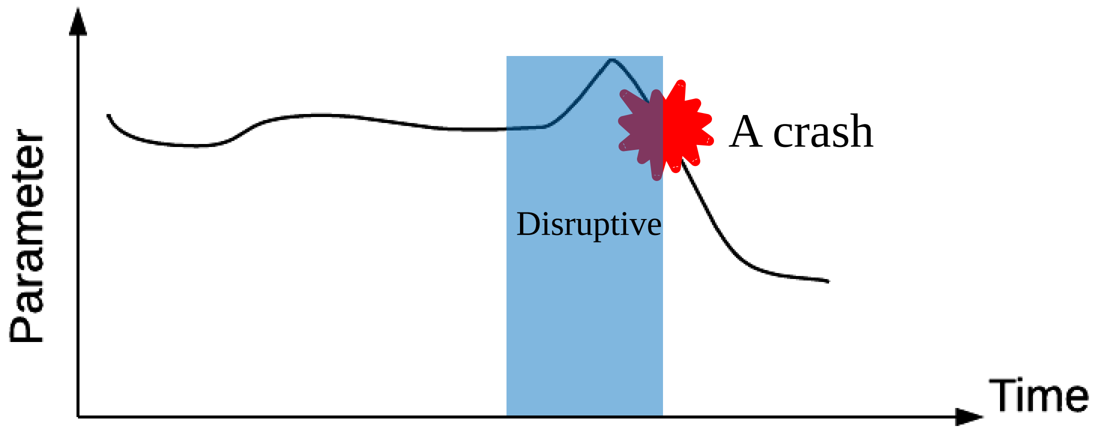






Real-Time Safety

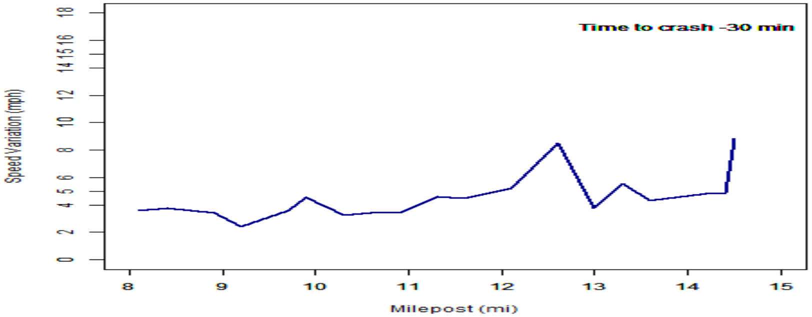
- Real-time safety analysis

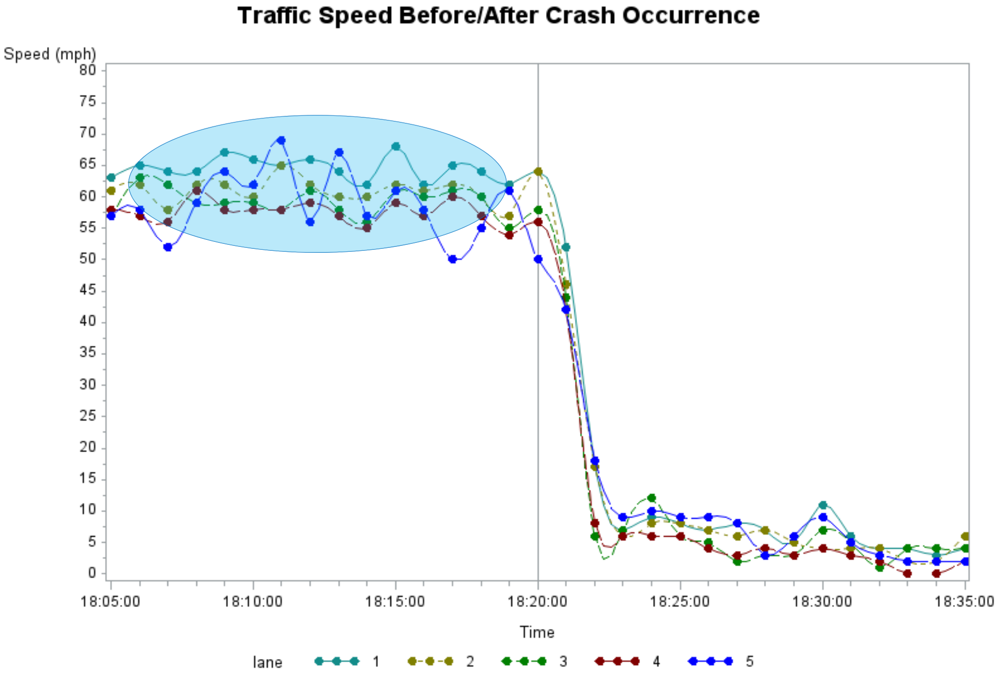


- Implementations
  - Understanding the microscopic crash mechanisms
  - Estimating crash likelihood in real-time
  - Improving traffic safety in real-time

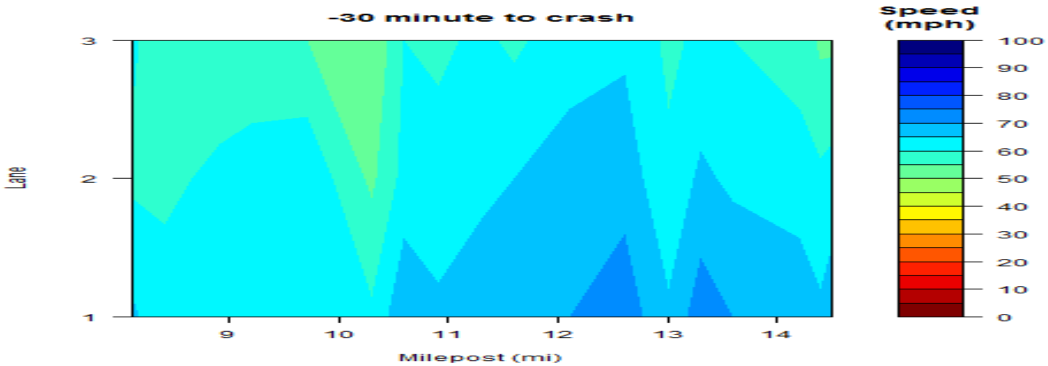
Application of Big Data in Safety: Traffic Safety in Real-Time

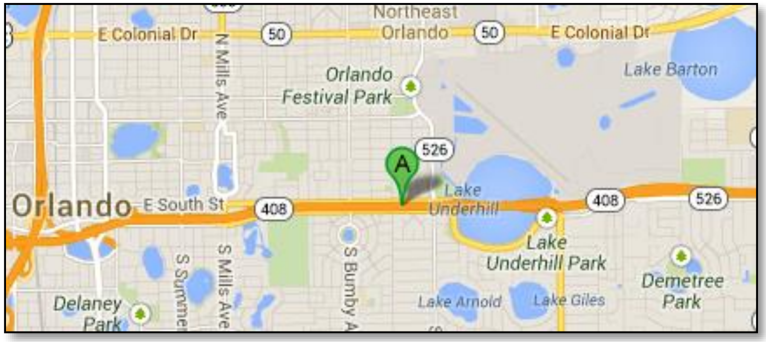
- Effects of traffic safety on operation
  - Crash-leading congestion
  - Real-time speed profile by lane
  - Real-time speed variation





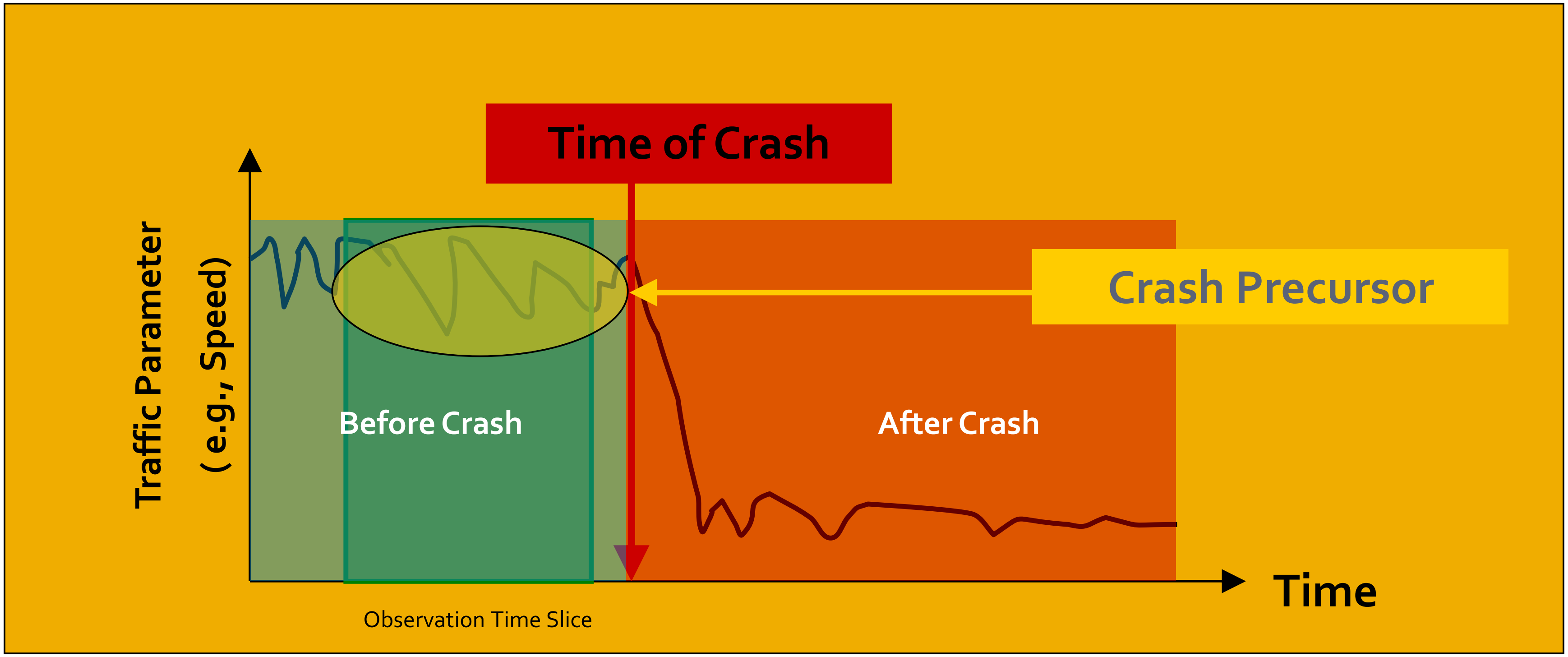
Direction: WB, @MM:12.1, Type: Rear-end, Vehicles: 4, Injuries: 3





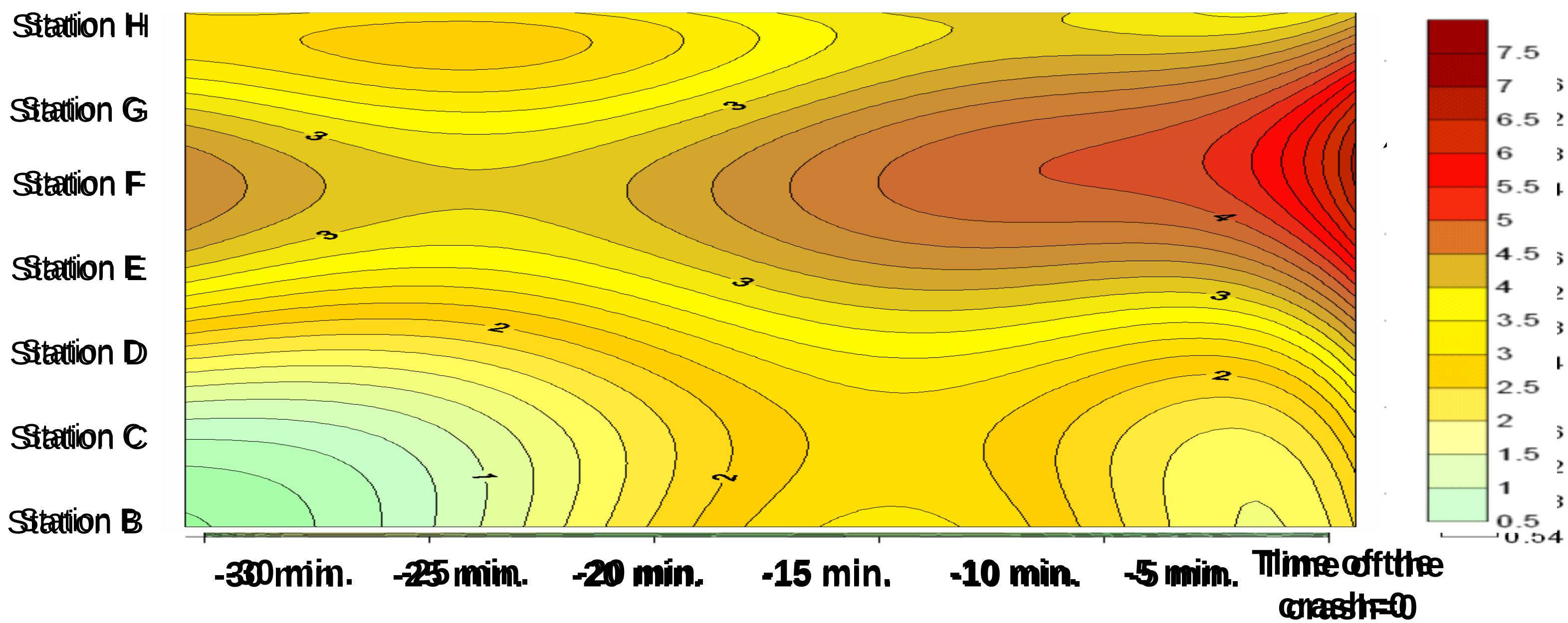
Proactive Perspective of Traffic Safety (1)

- What patterns are we looking for?
  - Developing a Hybrid Detailed Crash Prediction System Using ITS Data on I-4 and Evaluating the Application Strategies
  - Speed profile before/after crash on I-4



Proactive Perspectives of Traffic Safety (2)

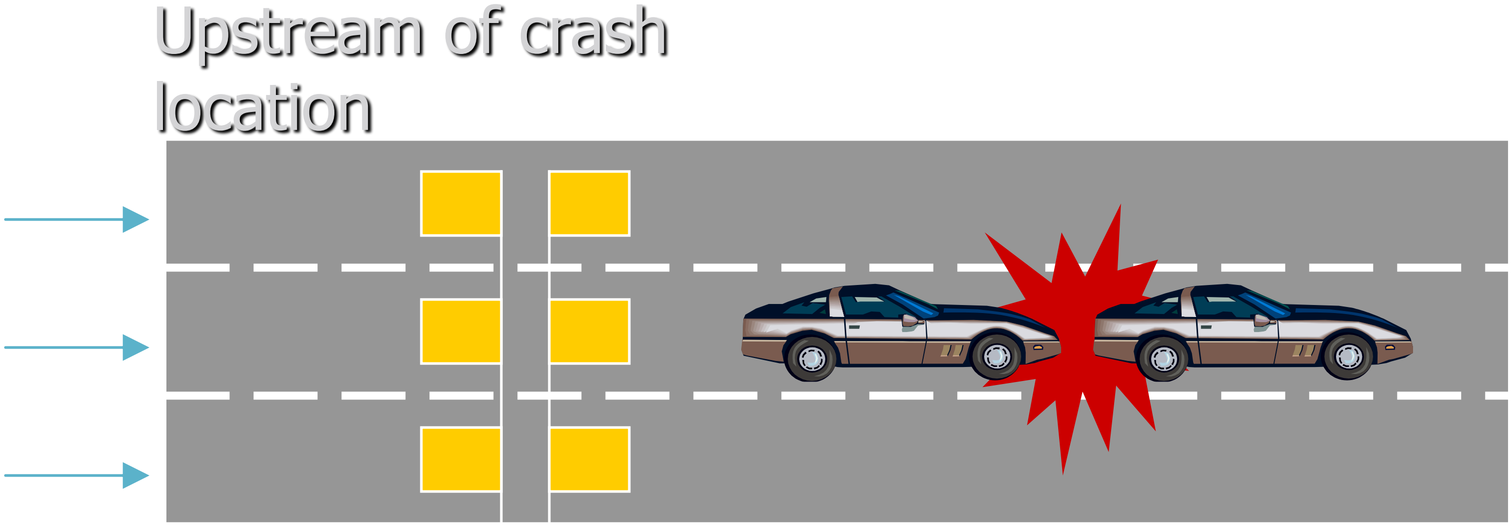
- Real-time crash risk for I-4



Hazard Ratio: Contour plots of hazard ratios corresponding to coefficient of variation in speed (42 model outputs)

Crash Prediction Model
Crash Precursors

Crash Precursor #1:
Variation in speed upstream of crash location

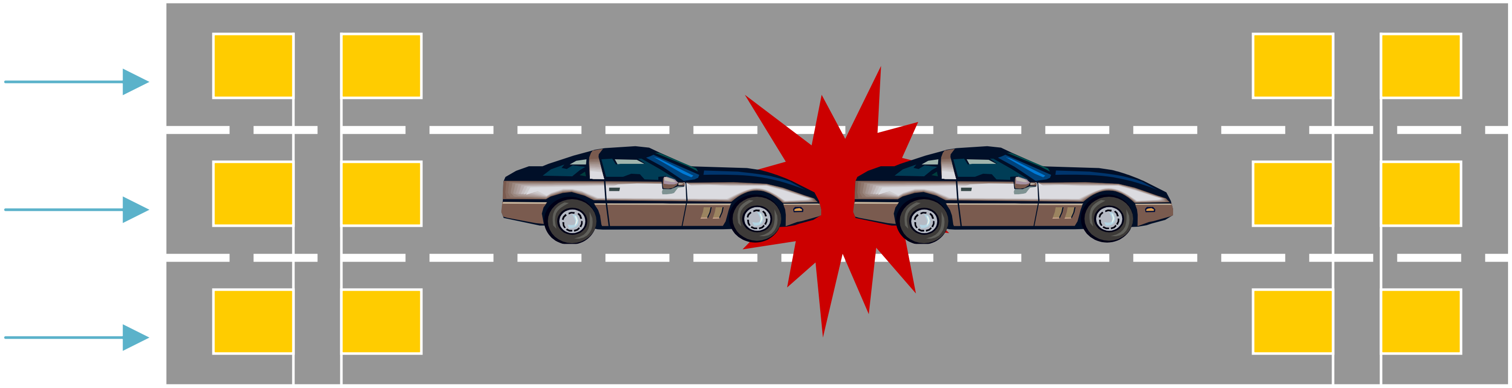


Crash Prediction Model
Crash Precursors

Crash Precursor #2:
Speed upstream differential of crash location with speed downstream

Upstream of crash location

Downstream of crash location



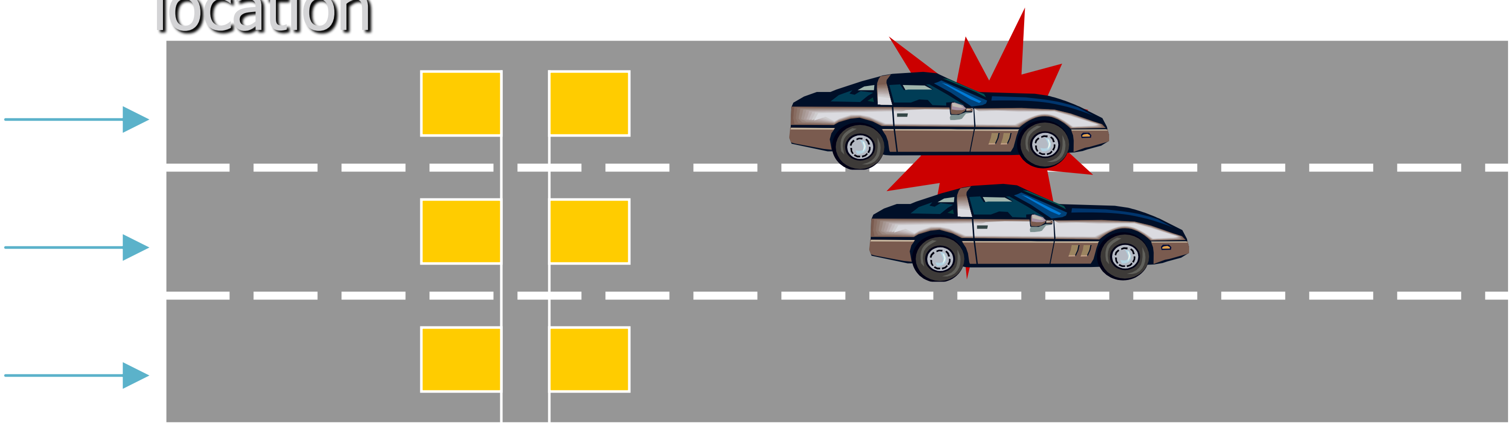
Crash Prediction Model

Crash Precursors

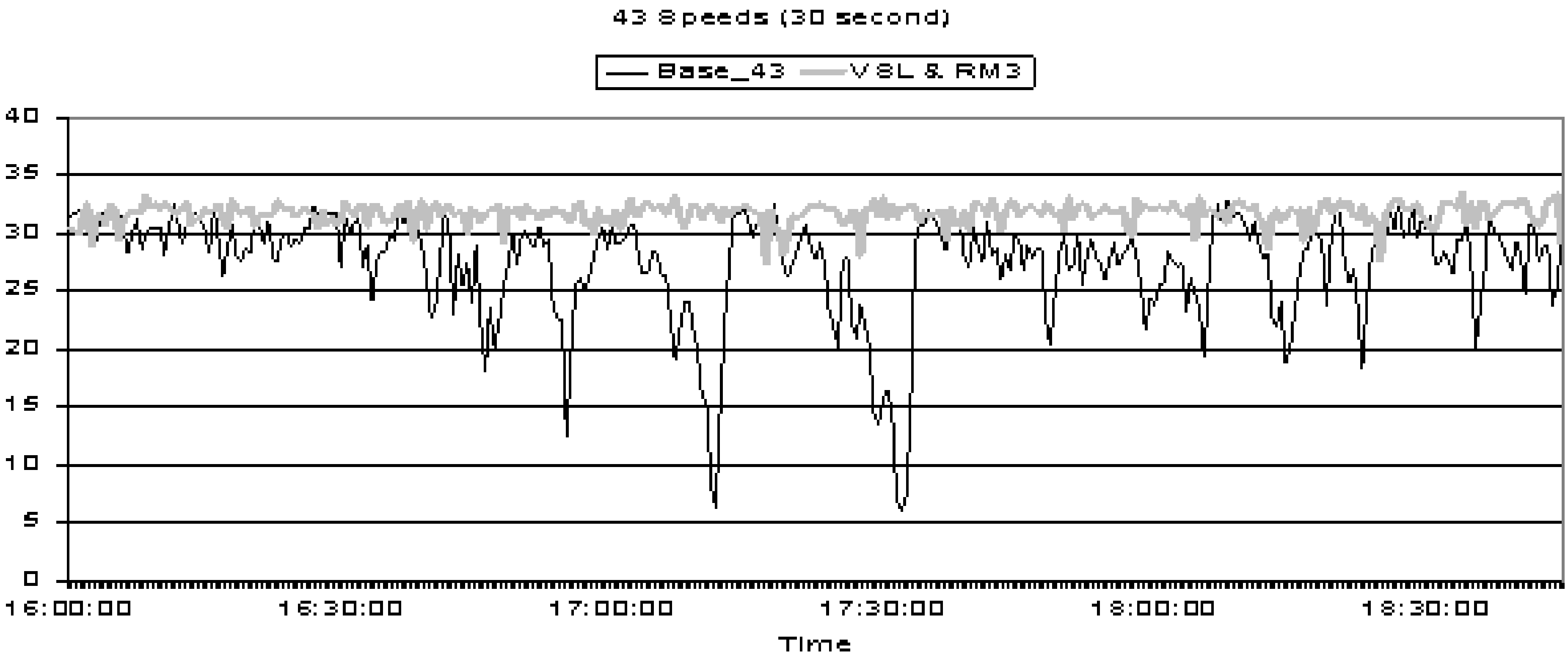
Crash Precursor #3:

Covariance of volume across adjacent lanes upstream of crash location

Upstream of crash location



Micro-Simulation: Speed Harmonization



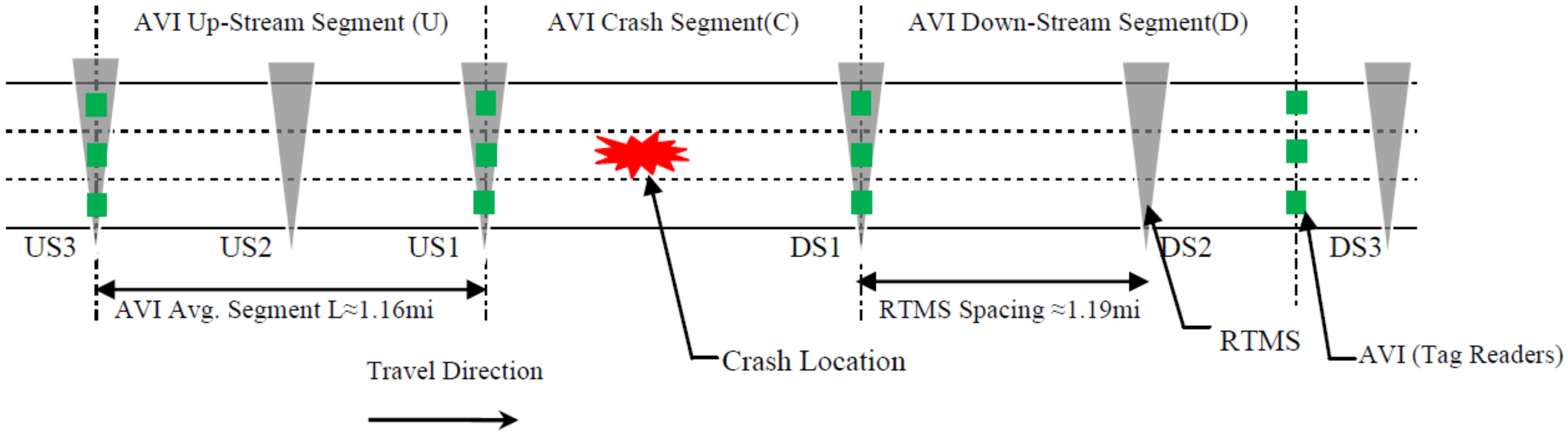
Stabilization in 30-second speed profiles following ITS strategies implementation

VSL & VMS in Driving Simulator

Warning Message

Variable Speed Limit

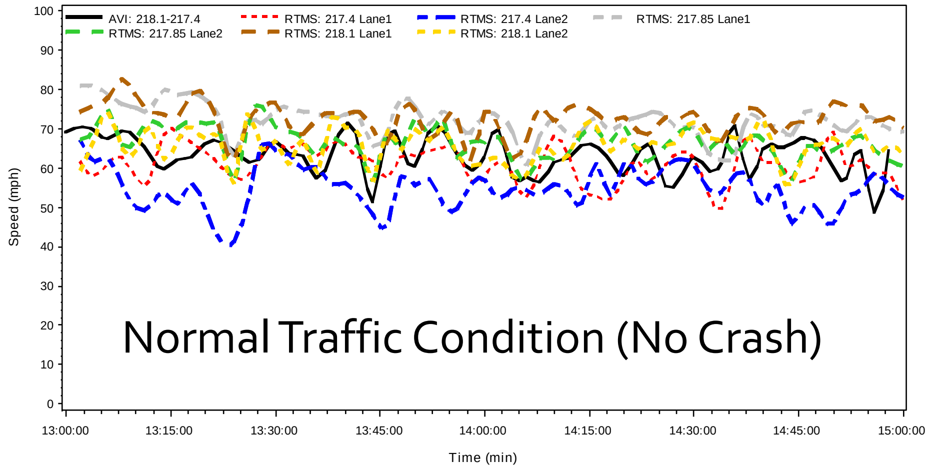
A DATA FUSION FRAMEWORK FOR REAL-TIME RISK ASSESSMENT ON FREEWAYS

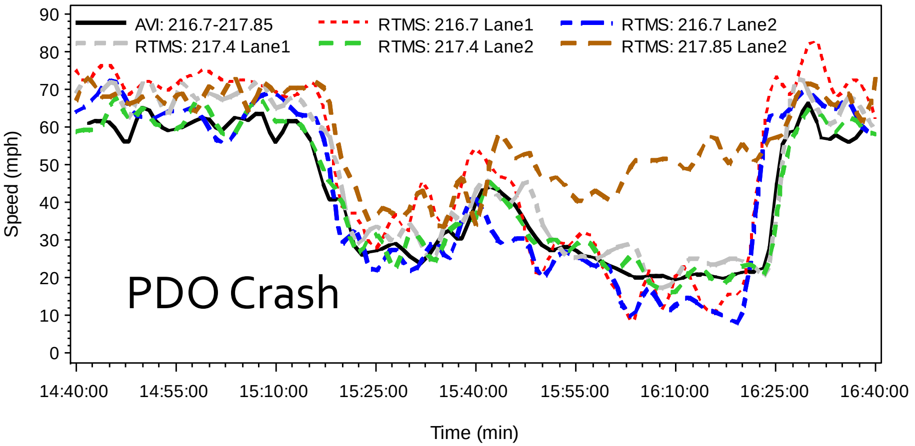


Arrangement of RTMS and AVI Segments

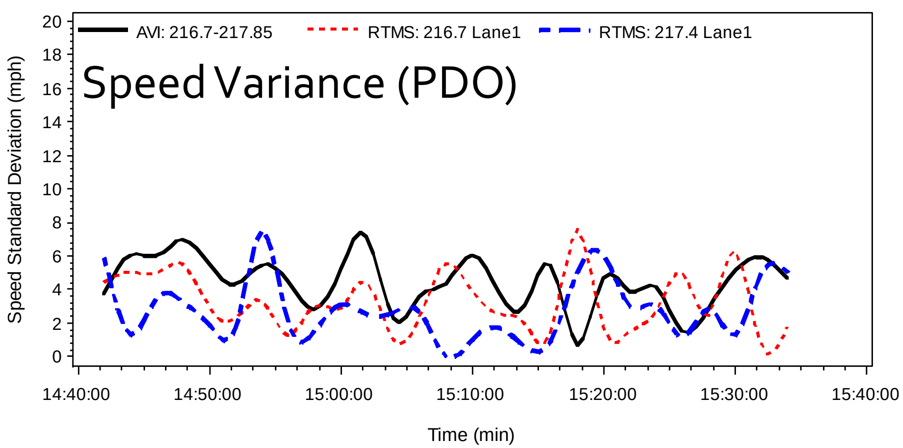
- The 15-mile on I-70 in Colorado is equipped with AVI, RTMS, and Weather Stations.
- There were five sets of data used in this study; roadway geometry data, crash data, and the corresponding AVI, RTMS and weather data.
- The crash data were obtained from CDOT for a 15-mile segment on I-70 for 13 months (from October 2010 to October 2011).
- Traffic data consists of space mean speed captured by 12 and 15 AVI detectors located on each east and west bounds, respectively along I-70. Volume, occupancy and time mean speed are collected by 15 RTMSs on each direction.
- AVI estimates SMS every 2-minute while RTMS provides traffic flow parameters every 30-second. Weather data were recorded by three automated weather stations along the roadway section for the same time period.
- The roadway data were extracted from Roadway Characteristics Inventory (RCI) and Single Line Diagrams (SLD).

Explanatory Comparison between AVI and RTMS Data

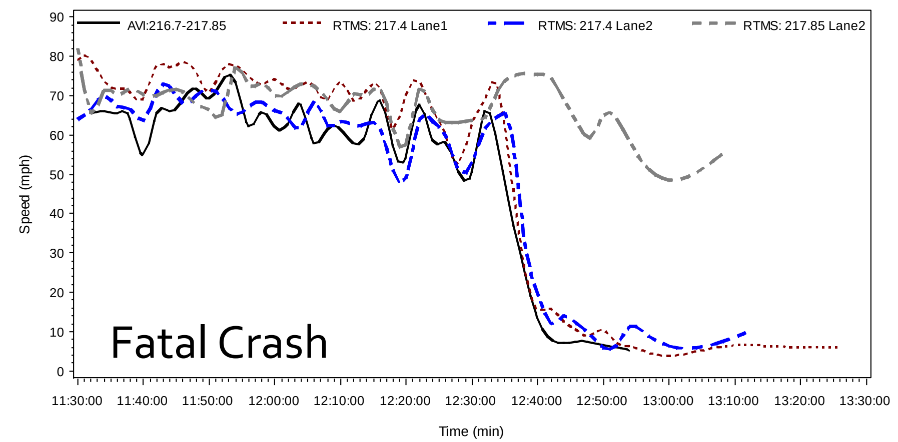




Crash Location: Milepost 217.7



Crash Location: Milepost 217.7

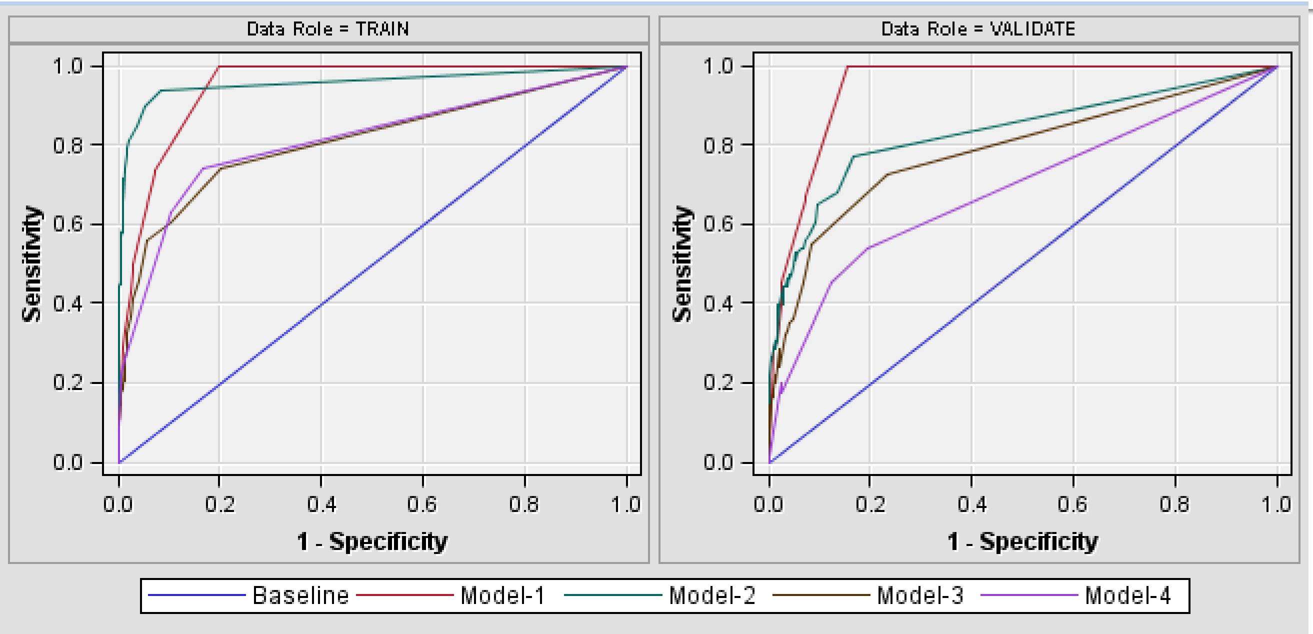


Crash Location: Milepost 217.5

Variable Importance

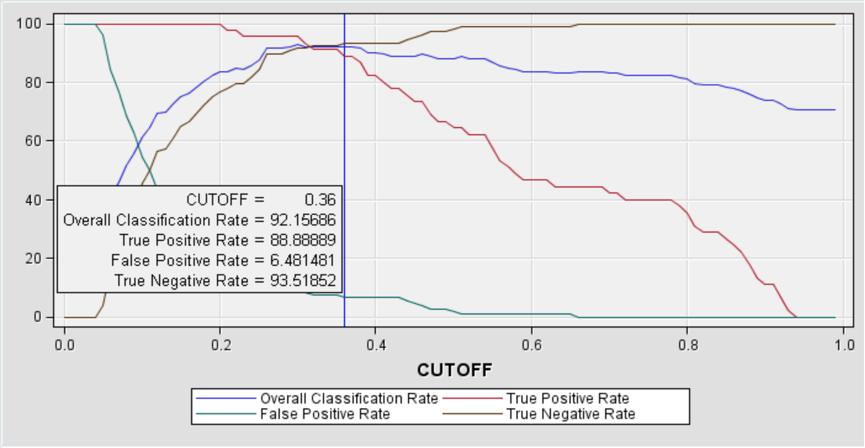
Model-1 (All Data)		Model-2 (RTMS)		Model-3 (AVI)		Model-4 (Weather)	
Variables	**Var. Import.**	**Variables**	**Var. Import.**	**Variables**	**Var. Import.**	**Variables**	**Var. Import.**
Avg. Occ. Upstream1_Time Slice _2	1.000	Avg. Occ. Upstream 2_Time slice_3	1.000	Log. Coef. of Var. of Speed Crash Segment Time Slice_2	1.000	1-Hour Visibility	1.000
Avg. Occ. Upstream 2_Time slice_3	0.887	Log. Coef. of Var. of Speed Upstream 1_Time Slice_2	0.997	Avg. Speed Downstream Segment Time Slice_2	0.899	10-Minute Precipitation	0.459
Log. Coef. of Var. of Speed Crash Segment Time Slice_2	0.798	Avg. Speed Upstream 2_Time Slice_2	0.804	Avg. Speed Downstream Segment Time Slice_3	0.741	1-Hour Precipitation	0.324
Avg. Speed Downstream Segment Time Slice_2	0.742	S.D. Occ. Upstream 2_Time Slice 2	0.541	Avg. Speed upstream Segment Time Slice_2	0.537		
1-Hour Visibility	0.684	Avg. Speed Downstream 1_Time Slice_2	0.457				
Grade	0.661	Avg. Speed Downstream 2_Time Slice_2	0.391				
S.D. Occ. Upstream 3_Time Slice 2	0.642	Avg. Occ. Upstream1_Time Slice _2	0.374				
No. of Lanes	0.521	Avg. Occ. Upstream2_Time Slice _2	0.348				
Avg. Speed Upstream 1_Time Slice_2	0.519	Log. Coef. of Var. of Volume Downstream 2_Time Slice_2	0.249				
Avg. Speed Downstream Segment, Time Slice_3	0.431						
Abs. Deg. of Curve	0.337						
10-Minute Precipitation	0.335						
Log. Coef. of Var. of Volume Downstream 2_Time Slice 3	0.334						
Log. Coef. of Var. of Speed Upstream Segment, Time Slice 3	0.329						

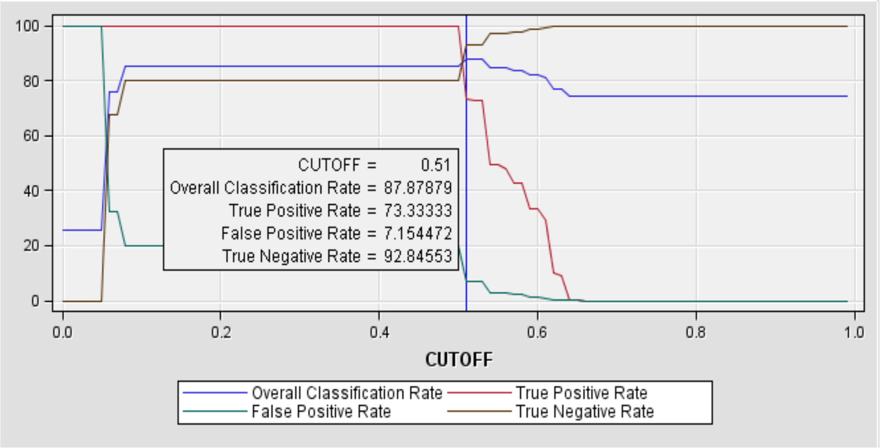
Models Comparison

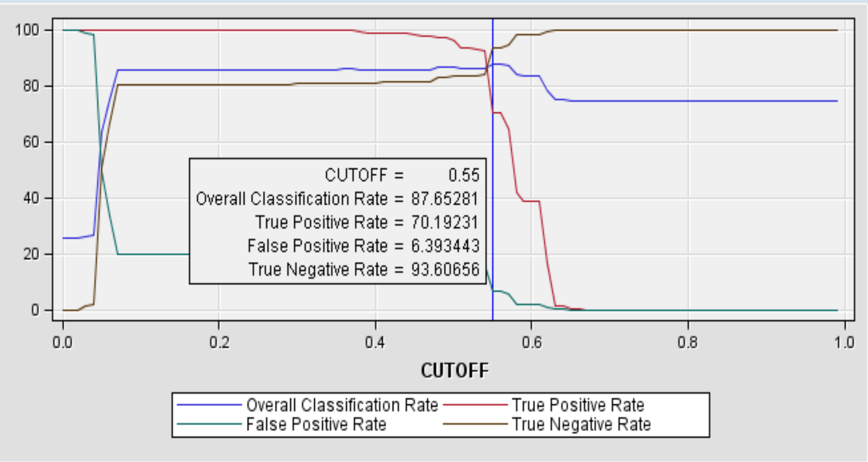


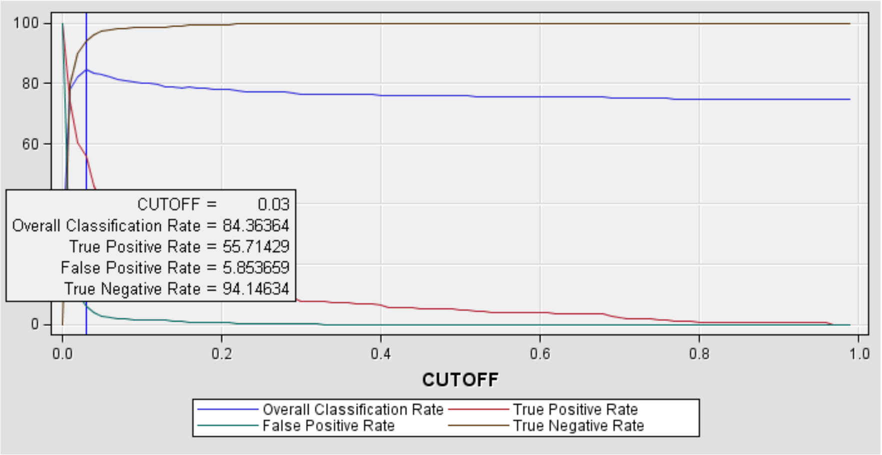
Model	Model Description	Overall Classification Rate	True Positive Rate	False Positive Rate	True Negative Rate	ROC Index
Model-1	All Data	92.157%	88.889%	6.481%	93.519%	0.946
Model-2	RTMS	87.879%	73.333%	7.154%	92.845%	0.762
Model-3	AVI	87.653%	70.192%	6.393%	93.607%	0.721
Model-4	Weather	84.364%	55.714%	5.854%	94.146%	0.675

Classification Rates

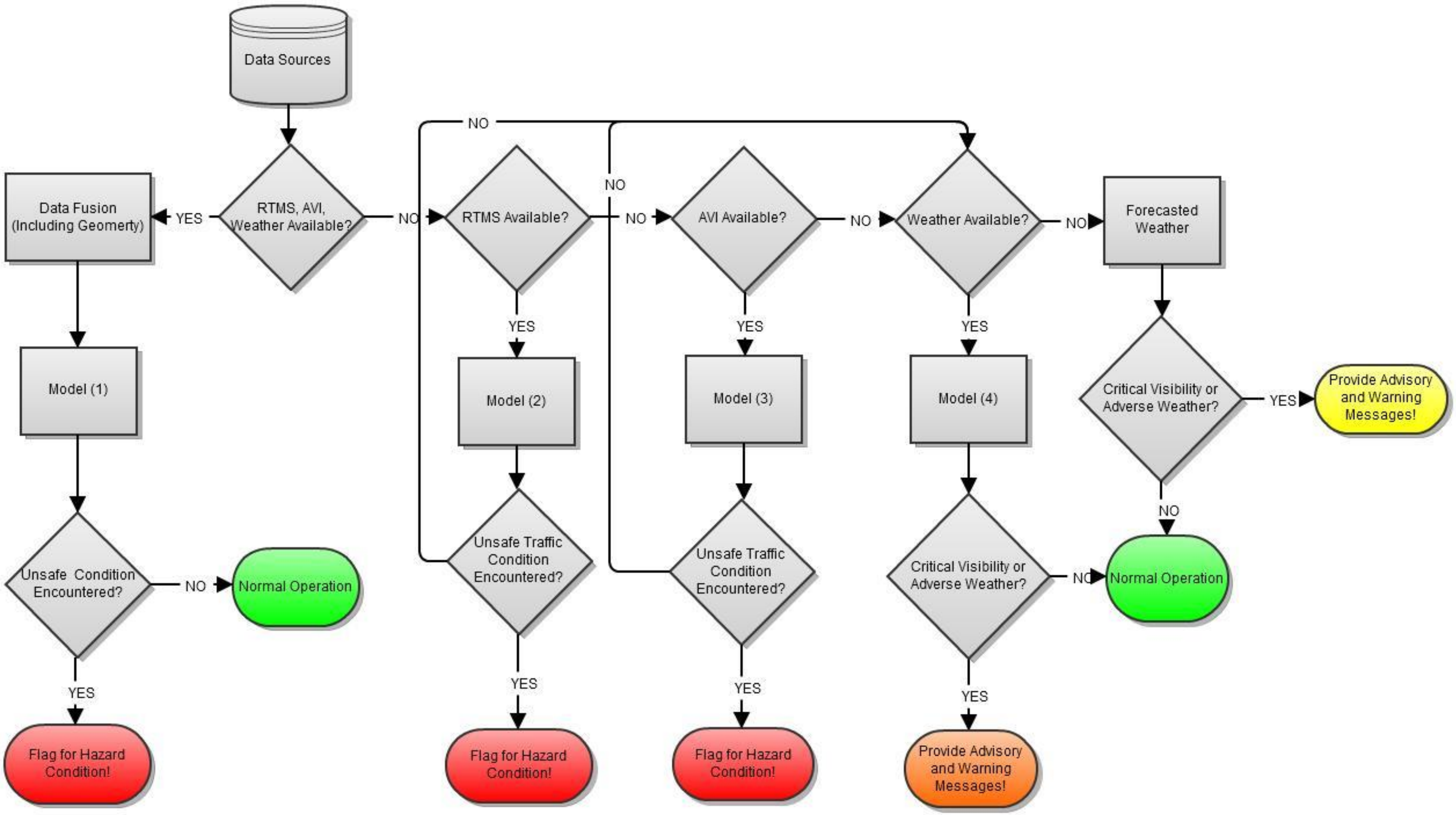






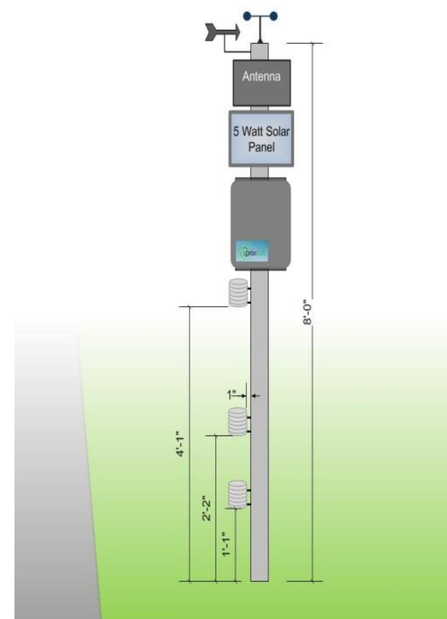


Framework of the Real-Time Risk Assessment



Investigation into Real-time Weather & Traffic Data

- Real-Time weather and traffic monitoring system on I-4



Weaving Sections Real-time safety analysis

- Real-time crash model for weaving segments

Variables	Mean	Std.	p-value
Intercept	-7.86	0.79	0.00
Speed difference (Spd_{dif})	0.11	0.03	0.00
Log(Vehcnt)	0.65	0.12	0.00
Weaving configuration (1=no lane change)	0.57	0.20	0.01
Weaving influence length (Lmax)	0.21	0.07	0.00
Road surface condition (1=Wet)	1.22	0.24	0.00
Training ROC*		0.716	
Validation ROC		0.704	

Speed difference: 0 if Bm_spd < Em_spd;
otherwise $\boxed{\text{Bm\_spd- Em\_spd}}$

$$L_{max}=5728(1+\boxed{VR})^{1.6}-1566N_{WL}$$

Using VSL to reduce

Using RM to reduce

Then, improve safety of weaving segments in real-time

Pro-Active Traffic Management Algorithm (1)

- Ramp metering (RM)
  - Updated every 5 minutes based on occupancy and crash risk

$$r(k) = r(k-1) + K_R(\hat{o} - o_{k-1}) \boxed{+ K_s(\hat{p} - p_{k-1})}$$

- Green-phase duration, $g(k)$, is calculated as follows,

$$g(k) = \left(\frac{r(k)}{r_{sat}}\right).C \qquad g_{\min} \le g(k) \le g_{\max}$$

- Queue Control

$$g' = \begin{cases} g & Queue \le 10 \\ g+1 & 10 < Queue \le 20 \\ g+2 & 20 < Queue \le 30 \\ g+3 & Queue > 30 \end{cases}$$

Active Traffic Management Algorithm (2)

- Variable speed limit (VSL)

  When crash risk is higher than critical crash risk, VSLs at the upstream and the downstream of the congested weaving segment are activated

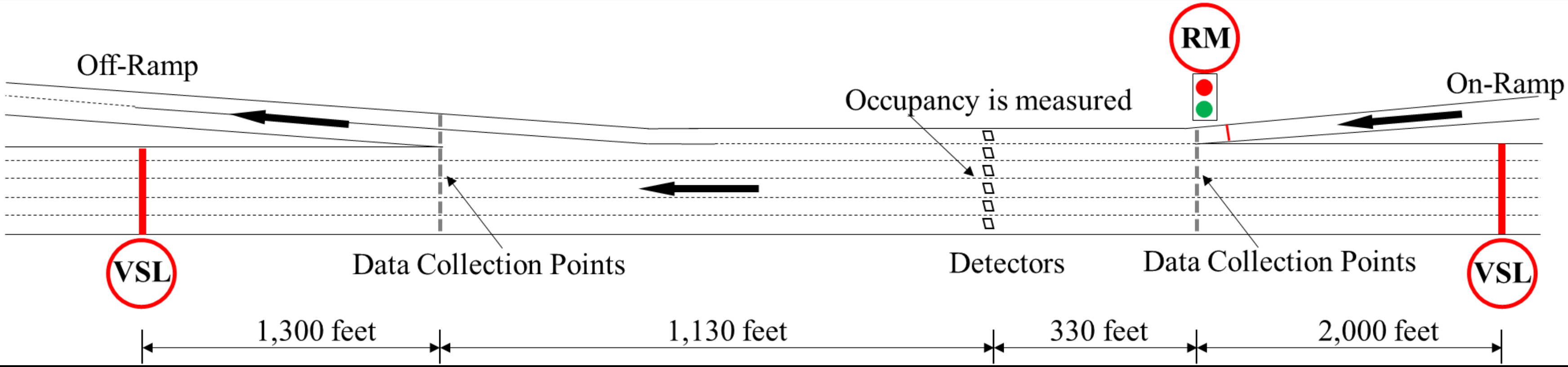
- RM-VSL

  RM is always active;

  When queue is more than 10 vehicles, VSL is activated to reduce speed limit on mainline to provide more gap for ramp vehicles

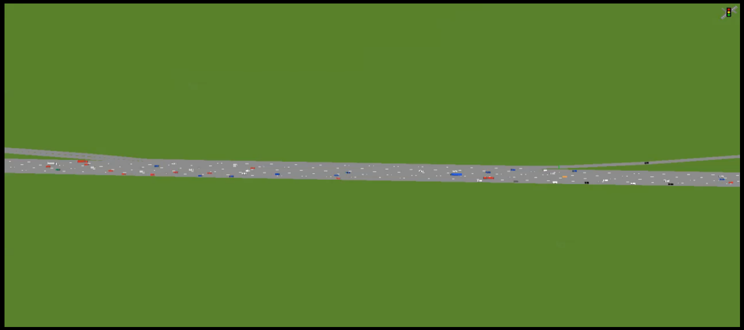
In Microscopic Simulation VISSIM
Through Component Object Model (COM) interface
Coded by Visual Basic for Application

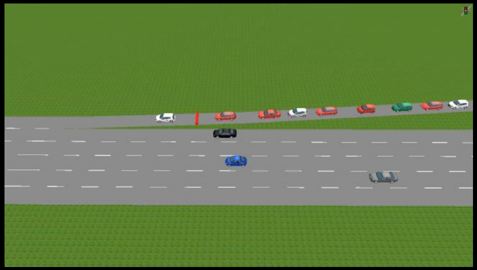
Experiment design

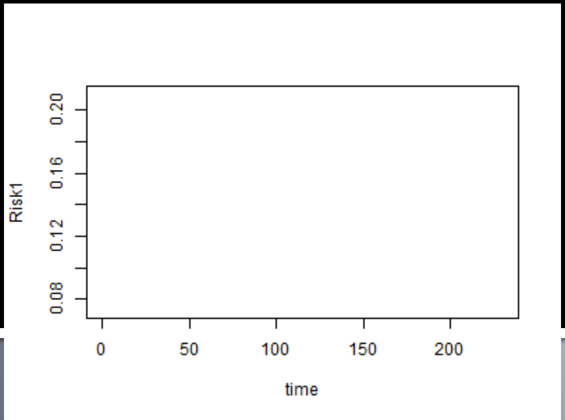


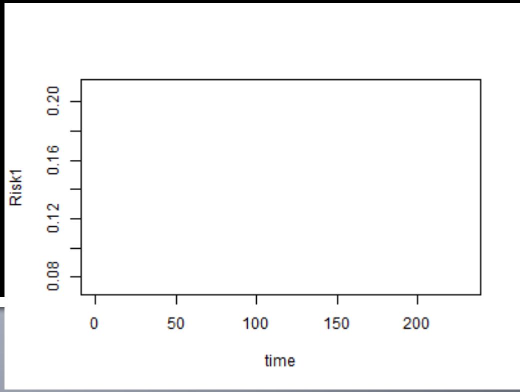
Case	VSL	RM
1	N/A#	NA
2	N/A	Ks=0
3	N/A	$Ks=2.5\times10^6$ (no Queue Control)
4	N/A	$Ks=2.5\times10^6$ (Queue Control)
5	Upstream 50 mph, Downstream remains	N/A
6	Upstream 45 mph, Downstream remains	N/A
7	Upstream remains, Downstream 60 mph	N/A
8	Upstream remains, Downstream 65 mph	N/A
9	Upstream 50 mph, Downstream 60 mph	N/A
10	Upstream 45 mph, Downstream 60 mph	N/A
11	Upstream 50 mph, Downstream 65 mph	N/A
12	Upstream 45 mph, Downstream 65 mph	N/A
13	Upstream 45 mph, Downstream 55 mph	$Ks=2.5\times10^6$ (Queue Control)

Impact of RM on real time Crash Odds

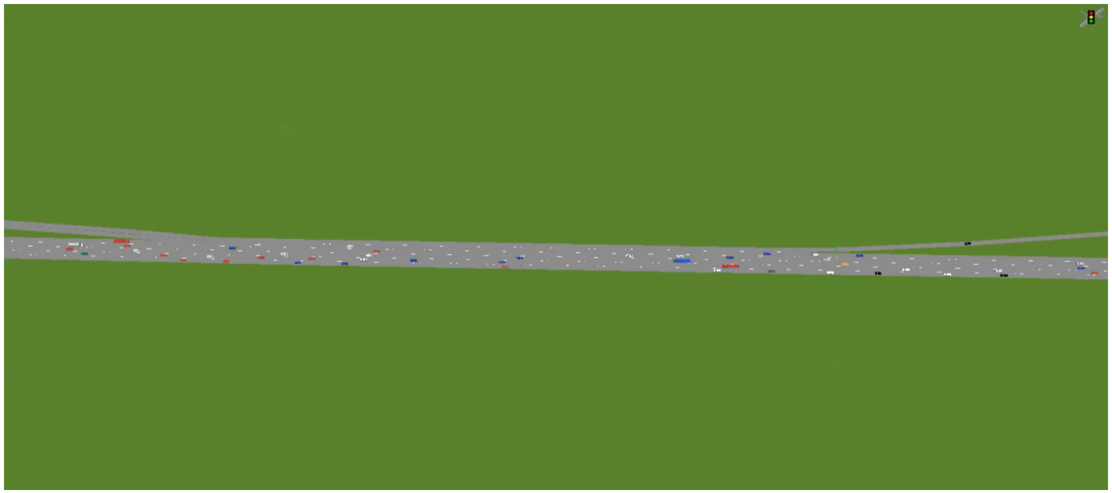


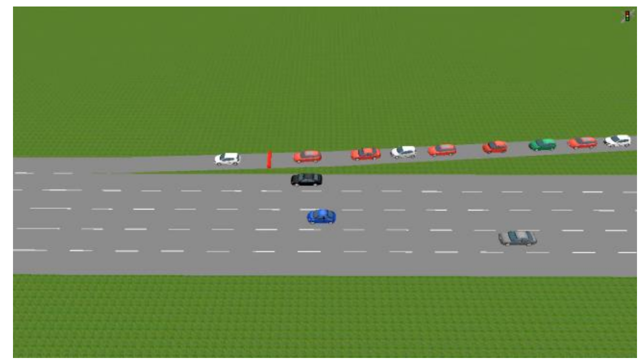


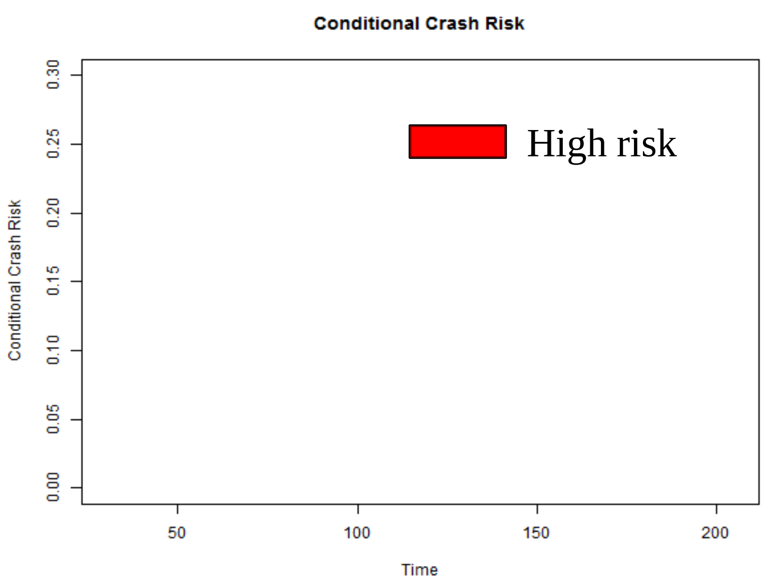


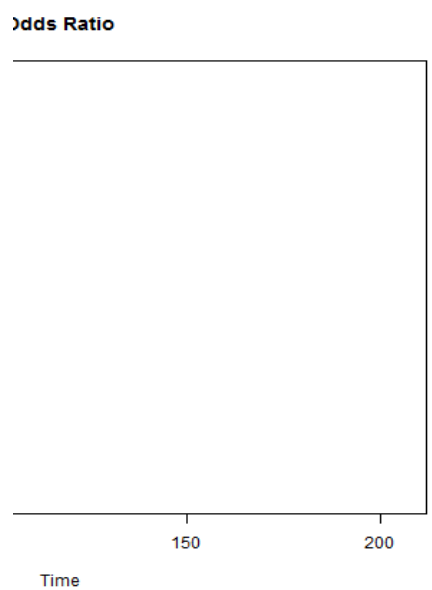


Impact of VSL and RM









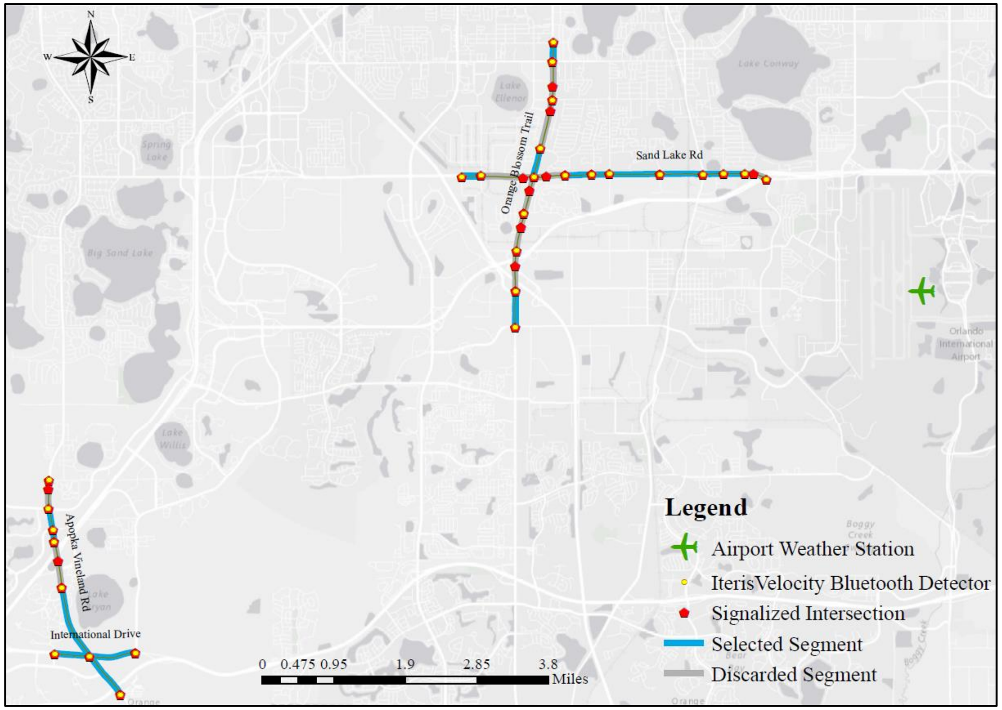
Impact of P-ATM

ATM results

Description		Case	Weaving			Non-weaving		Average travel time for total
			Conflict	Conflict Change %	OR	Conflict	Conflict Change %	
N/A	No ATM	1	705	N/A	1.00	59	N/A	98.3
RM	Traditional	2	653	-7.3	1.01	38	-35.6	97.9
	No Queue	3	555	-21.2	0.95	41	-30.5	113.7
	Queue control	4	621	-11.9	0.92	40	-31.7	101.4
VSL	Up 50	5	639	-9.3	0.88	62	5.8	100.1
	Up 45	6	575	-18.4	0.82	43	-26.9	101.3
	DW 60	7	705	0.1	1.00	59	-0.3	97.7
	DW 65	8	705	0.0	1.00	60	1.4	97.4
	Up 50, Dw 60	9	639	-9.3	0.88	63	7.7	99.8
	Up 45, Dw 60	10	575	-18.4	0.82	44	-25.2	101.1
	Up 50, Dw 65	11	639	-9.3	0.88	63	7.8	99.6
	Up 45, Dw 65	12	575	-18.4	0.82	43	-26.1	101.0
RM-VSL	Queue control &Up 45	13	586	-16.8	0.94	43	-27.6	105.0

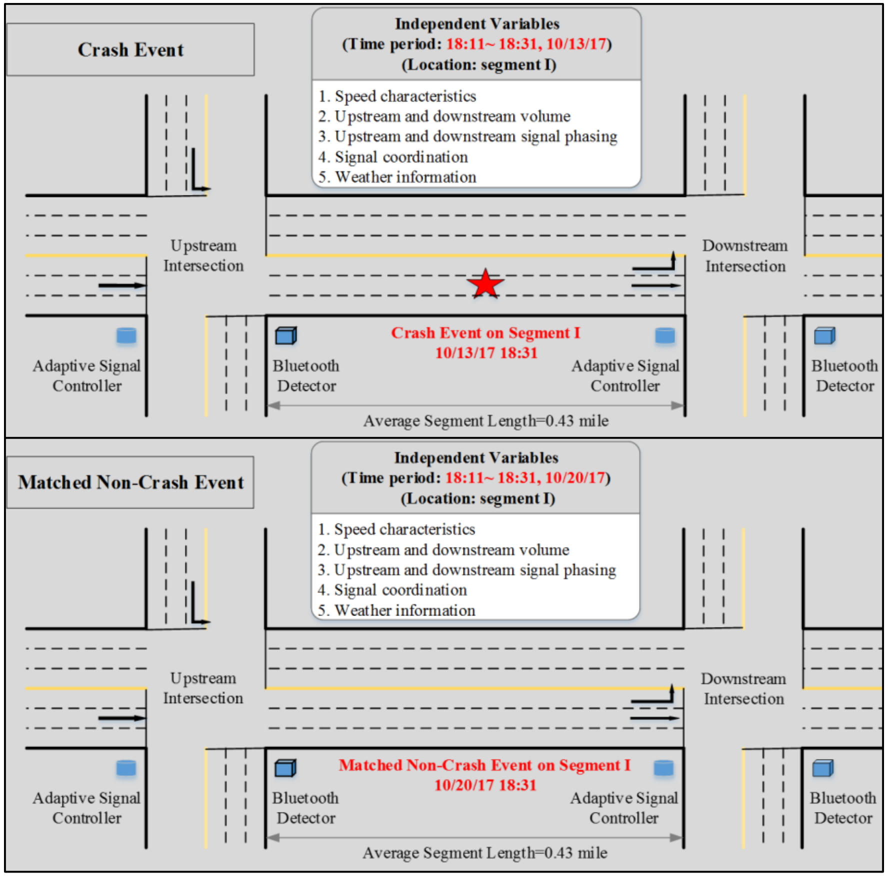
Arterials Real-Time Safety

- Four urban arterials in Orlando, Florida were chosen;
- Crash data were collected from March, 2017 to December, 2017;
- Space-mean speed data collected by 23 IterisVelocity Bluetooth detectors;
- Signal timing and traffic volume provided by 23 adaptive signal controllers;
- Weather characteristics collected from MCO



Data Preparation

- Matched case-control design with a control-to-case ratio of 4:1 was employed to select the corresponding non-crash events for each crash event;
- Three confounding factors, i.e., location, time of day, and day of the week, were selected as matching factors;
- The real-time travel speed data were extracted for a period of 20 minutes (divided into four 5-minute time slices) prior to crash occurrence



Results

- **Bayesian Conditional Logistic Models**

Parameter	Slice 1		Slice 2		Slice 3		Slice 4	
	Mean (95% BCI)	Hazard Ratio	Mean (95% BCI)	Hazard Ratio	Mean (95% BCI)	Hazard Ratio	Mean (95% BCI)	Hazard Ratio
Avg_speed	-0.049 (-0.071, -0.029)	0.952	-0.025 (-0.048, -0.004)	0.975	-	-	-	-
Up_Vol_LT	0.024 (0.007, 0.044)	1.024	0.024 (0.005, 0.044)	1.024	0.024 (0.006, 0.045)	1.024	0.036 (0.014, 0.06)	1.037
Down_GreenRatio	-	-	-0.042 (-0.075, -0.011)	0.959	-	-	-	-
Rainy	0.551 (0.02374, 1.065)*	1.735	0.667 (0.055, 1.274)	1.948	0.682 (0.037, 1.322)	1.978	0.72 (0.078, 1.341)	2.054

Implementation

UCF team has implemented the Real-Time risk estimation in the following:

- I-4, I-95 and CFX network in Orlando
- I-70 in Colorado
- Motorways in the Netherlands
- City streets in Cyprus
- Expressways in China
- Currently Orange county arterials
- 30 Km of an Expressway in Stockholm is currently operational

Vision of Big Data for Transportation

- **More Proactive (but data intensive) approaches / Real-Time**
- **Ever richer information**
  - Smartphones, sensors, onboard vehicle hardware, provide continuous data
  - Traffic status, weather conditions in real-time
- **Better operation and safety**
  - Bottleneck detection in real-time
  - Crash risk evaluation and prediction in real-time
- **More accurate prediction**
  - Formation of congestion, queue length, congestion duration
  - Crash-prone conditions: unstable traffic flow, adverse weather
- **Timely communication**
  - Media: smartphone, webpage, DMS, radio
  - Suggested countermeasures: trip planning, route choice, travel time calculation, VSL, speed advice, RM, etc.

Thank You

Dr. Mohamed Abdel-Aty, PE

Center for Advanced Transportation Systems Simulation
Department of Civil, Environmental & Construction Engineering
University of Central Florida
Orlando, FL 32816-2450