Electronic Data Patterns in Vehicles Involved in Unsafe Outcomes

David-Erick Lafontant, Kean University, NJ

&

Yao A. Tapoayi, Black Hawk College, IL

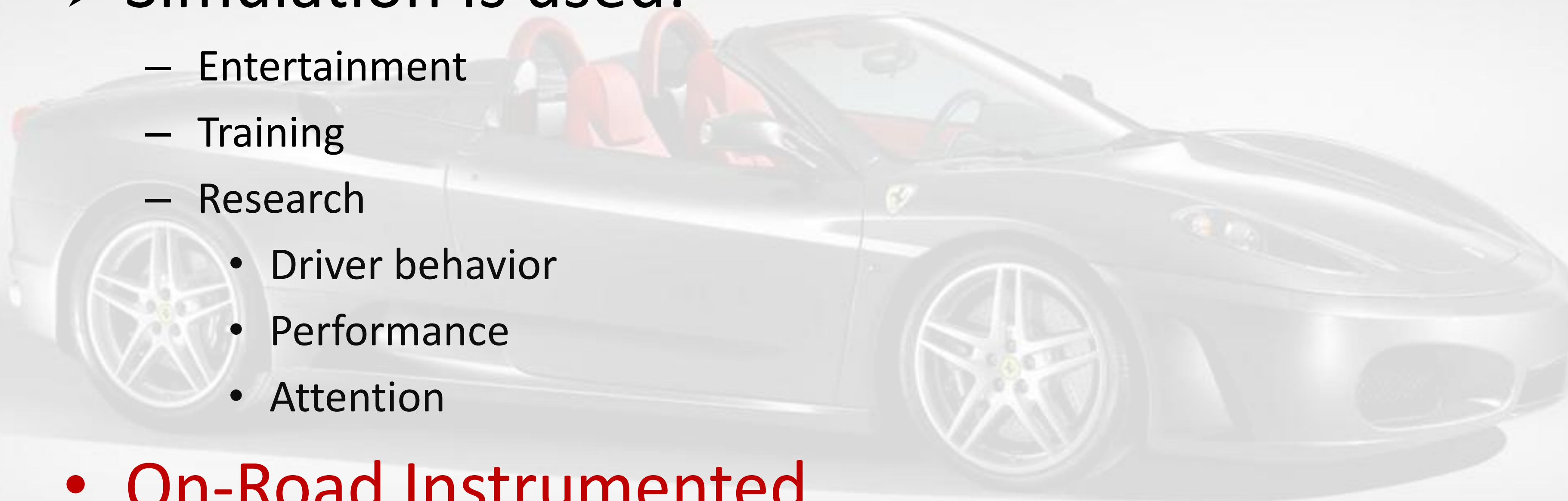
Mentor: Dr. Jeffrey Dawson

ISIB-University of Iowa

Outline

- Importance of Driving Simulation & Research
- 100-car- study
- Data availability + Definitions of Variables
- Statistical methods
- Results
- Comments

Driving Simulation VS. On-Road Instrumented Vehicles



- Simulation Is used:
  - Entertainment
  - Training
  - Research
    - Driver behavior
    - Performance
    - Attention
- On-Road Instrumented
  - Training
  - Research

Importance of Driving Research

- Public Health Issue
- National Highway Traffic Safety Administration (NHTSA)
  - 43,000 deaths/year (117/day) in US
  - Driver inattention

Previous and Ongoing Research

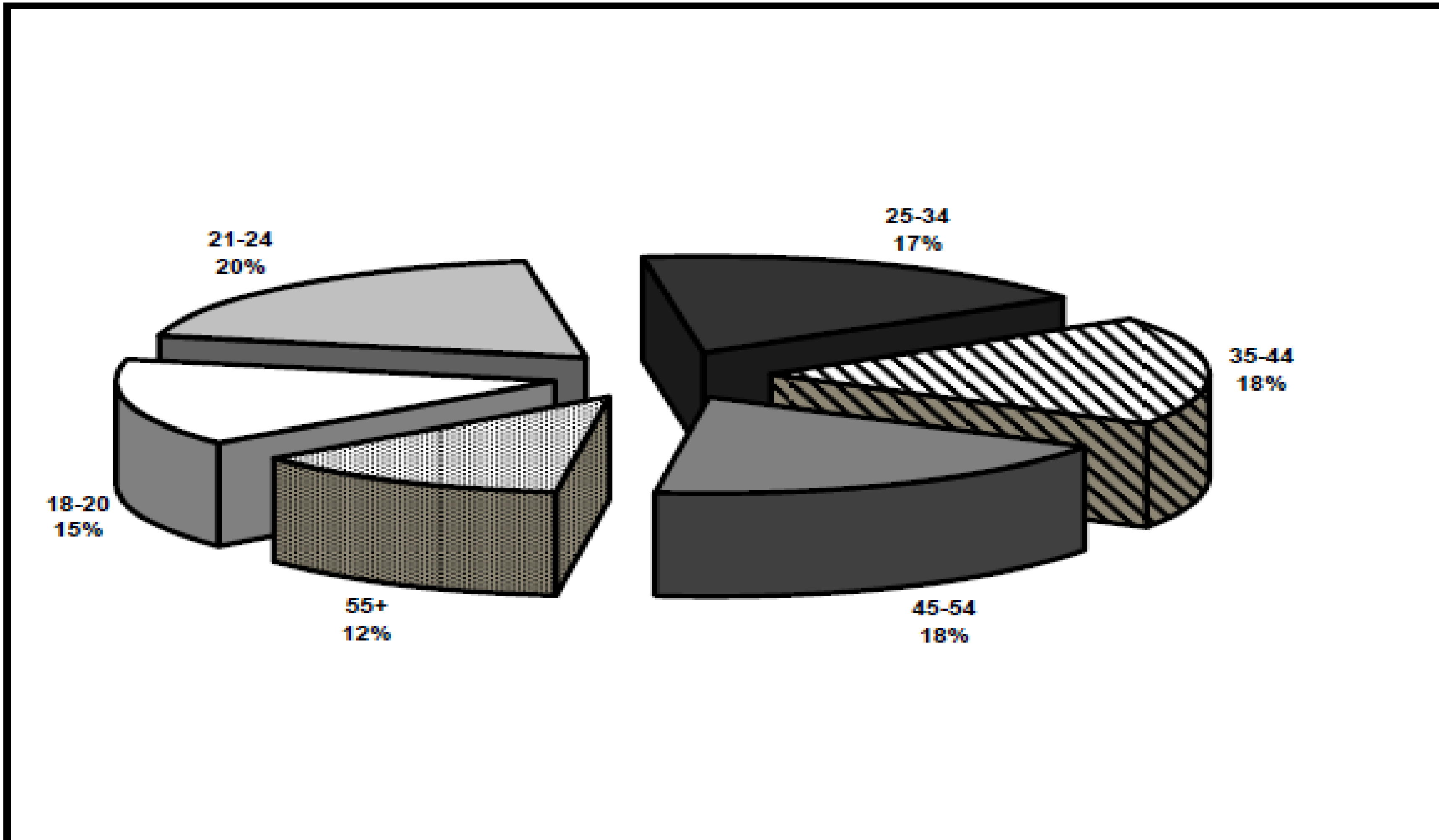
- University of Iowa
  - Public Policy Center
    - Teen Drivers
  - Departments of Neurology & Biostatistics
    - Elderly Drivers
- Virginia Tech Transportation Institute(VTTI)
  - 100-Car Naturalistic Study

100-Car Naturalistic Study

- 12 – 13 months of data collection
- 2,000,000 vehicle miles of driving
- 42,300 hrs of driving
- 241 total drivers
- 15 police-reported & 67 non-police reported crashes
- 761 near-crashes
- Contains many extreme driving cases
  - severe drowsiness, impairment, judgment error, risk taking
  - secondary task engagement, aggressive driving, traffic violations
- Crash – any physical contact between the subject vehicle and another vehicle, fixed object, pedestrian, pedal cyclist or animal
- Near-Crash – situations requiring a rapid, severe evasive maneuver to avoid a crash

Age Distribution Pie Chart of Drivers

(VTTI/100CarMain/pg94)



100-Car Naturalistic Study
(cont'd)

- *Data Collection Instrumentation*
  - Five channels of digital, compressed video
  - Front and rear radar sensors
  - Accelerometers
  - Machine vision-based lane tracker
  - GPS
  - Vehicle speed sensor
- Nearly 80 % crashes & 65 % of all near-crashes
  (due to distraction, fatigue, or just looking away)

TOOLS

VTTI/100CarMain/pg73,75,76

The Data

- 12 out of 213 data for ID 8302

	A	B	C	D	E	F	G	H	I	J
1	webfile_i	sync	time	gas_pedal	composite	gps_spee	yaw_rate	gps_headi	lateral_ac	longitudir
2	8302	1	42.486	0	0	0	-1.30229	0	0.058063	0.039847
3	8302	2	42.537	0	-1	0	-1.30229	0	0.067664	0.041357
4	8302	3	42.598	0	-1	0	-1.30229	0	0.059437	0.037433
5	8302	4	42.635	0	-1	0	-1.13953	0	0.056912	0.033135
6	8302	5	42.661	0	-1	0	-1.30229	0	0.060079	0.024917
7	8302	6	42.689	0	-1	0	-1.13953	0	0.054122	0.018462
8	8302	7	42.715	0	-1	0	-1.13953	0	0.052102	0.014417
9	8302	8	42.738	0.392	-1	0	-1.30229	0	0.056416	0.005566
10	8302	9	42.764	0.392	-1	0	-1.13953	0	0.064504	-0.00354
11	8302	10	42.792	0.392	-1	0	-1.30229	0	0.071205	-0.01252
12	8302	11	42.818	0.392	-1	0	-1.30229	0	0.079185	-0.0215

- Up to 28963 pieces of data

	A	B	C	D	E	F	G	H	I	J
28478	9119	9366	985.528	0	39.14639	39.2	-0.16297	178	0.03101	-0.03675
28479	9119	9367	985.628	0	39.14639	39.2	-0.32573	178	0.004021	-0.03595
28480	9119	9368	985.728	0	39.14639	39.2	-0.32573	178	0.037205	0.004823
28481	9119	9369	985.828	0	39.76776	39.2	-0.48849	178	0.047888	-0.03272
28482	9119	9370	985.928	0	39.76776	39.2	-0.16297	178	0.054713	-0.03681
28483	9119	9371	986.028	0	39.76776	39.2	-0.97677	178	0.028934	-0.02273
28484	9119	9372	986.128	0	39.76776	41	-0.32573	178.1	0.036491	-0.03511
28485	9119	9373	986.228	0	40.38913	41	-0.00021	178.1	0.022137	-0.02787
28486	9119	9374	986.328	0	40.38913	41	-0.16297	178.1	0.062905	-0.0466
28487	9119	9375	986.428	0	40.38913	41	0.325307	178.1	0.063496	0.001952
28488	9119	9376	986.528	0	41.0105	41	-0.65125	178.1	0.030752	-0.02432
28489	9119	9377	986.628	0	41.0105	41	-0.16297	178.1	0.039122	-0.06138
28490	9123	2965	344.118	28.616	16.15565	17.6	-1.13953	66.5	0.036918	0.233203
28491	9123	2966	344.218	28.616	16.15565	17.6	-1.46505	66.5	0.023842	0.237047
28492	9123	2967	344.318	28.616	16.15565	17.6	-1.13953	66.5	0.042308	0.212916
28493	9123	2968	344.417	28.616	16.15565	17.6	-1.30229	66.5	0.025906	0.22398
28494	9123	2969	344.518	28.616	18.64114	17.6	-1.62781	66.5	0.02098	0.218394
28495	9123	2970	344.618	28.616	18.64114	17.6	-1.46505	66.5	0.043797	0.216923
28496	9123	2971	344.718	28.616	18.64114	17.6	-1.46505	66.5	0.015131	0.210683
28497	9123	2972	344.818	28.616	18.64114	17.6	-1.13953	66.5	0.028556	0.217552
28498	9123	2973	344.918	28.616	18.64114	17.6	-1.62781	66.5	0.05148	0.219795
28499	9123	2974	345.018	28.616	18.64114	17.6	-1.62781	66.5	0.040809	0.220783
28500	9123	2975	345.118	28.616	18.64114	22.5	-1.30229	65	0.042429	0.225446
28501	9123	2976	345.218	28.616	19.88388	22.5	-1.79057	65	0.047345	0.228241
28502	9123	2977	345.318	28.616	19.88388	22.5	-1.13953	65	0.021552	0.230655
28503	9123	2978	345.418	43.512	19.88388	22.5	-1.30229	65	0.023846	0.231814

- 342000 pieces of data for Near-Crash

Our Goals

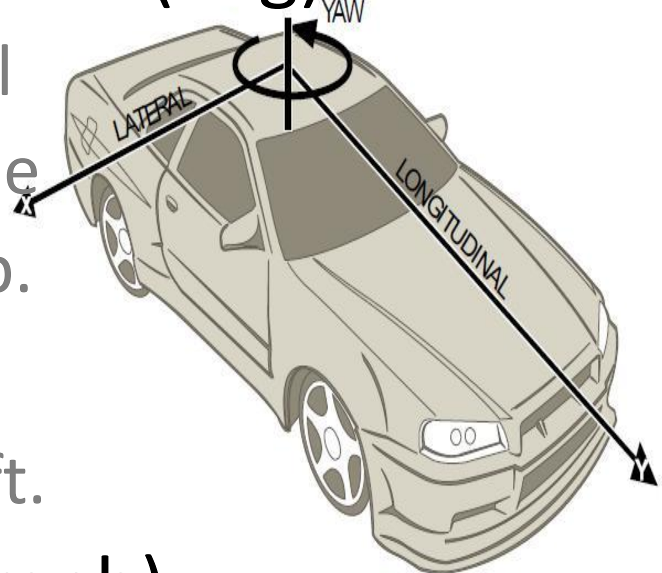
- Better Understanding of crashes and near-crashes
  - Analyze Electronic data
  - Compare the Crashes and Near-Crashes with respect to data patterns prior to these events
    - (drivers could be warned of high risk situations)

Data Used

- 68 crashes & 761 near crashes
  - www.vtti.vt.edu
- Several measures captured
- 30 seconds before the event
- 10 seconds after the event (≈ 10 Hz)
- Narrative descriptions

Definitions of Variables of Interest

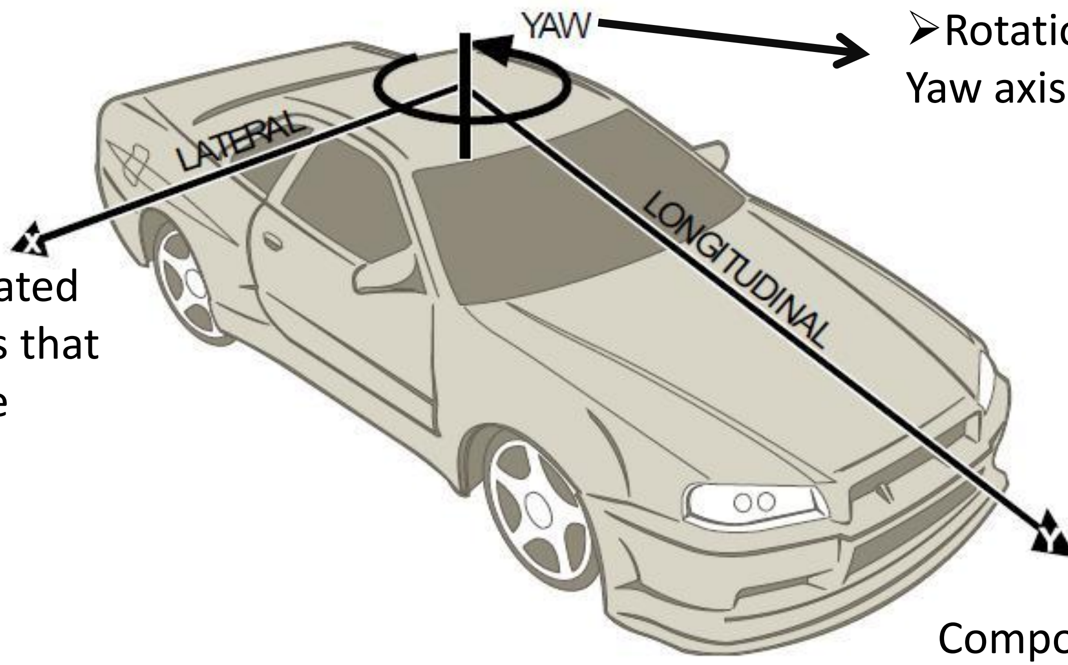
- SDs of Lateral Acceleration(in g)
  - Positive indicates lateral acceleration as generated by the vehicle turning to left.
- SDs of Longitudinal Acceleration(in g)
  - Positive indicates longitudinal acceleration as generated by the vehicle accelerating from a stop.



- SDs of Yaw Rate(in deg/s)
  - Positive in vehicle turns to left.
- SDs of Composite Speed(in mph)
  - Forward and reverse motion is positive

Definitions of Variables of Interest (cont'd)

➢Acceleration: rate of change of velocity as a function of time



➢Rotation around Yaw axis

➢The acceleration created when a vehicle corners that tends to push a vehicle sideways.

Component of linear acceleration along the x axis

Statistical Methods

- Time Series Plots
  - Look for data quality, outliers
  - Compare with narrative descriptions
- Data reduction of 4 variables
  - 5-15 second before event
  - Mean of SD of each variables
  - Mean of Means of Speed

2 Sample T-test

- Welch Test
  - Independent samples
  - Unequal variances
  - $$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left(\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}\right)}}$$
  - 4 SDs & 1 Mean
- Crash Vs. Near-Crash

2 Sample T-test (suite)

- Welch Test
- Hypothesis :

$$H_0: \mu_{sdcrash} = \mu_{sdnearcrash}$$

$$H_A: \mu_{sdcrash} \neq \mu_{sdnearcrash}$$

$$H_0: \mu_{meanspdcrash} = \mu_{meanspdnearcrash}$$

$$H_A: \mu_{meanspdcrash} \neq \mu_{meanspdnearcrash}$$

- $\alpha = .05$
- $$t = \frac{(\bar{X}_1 - \bar{X}_2)}{\sqrt{\left(\frac{{S_1}^2}{n_1} + \frac{{S_2}^2}{n_2}\right)}}$$

The R Code (1/2)

```r
tscrash <- read.table("F:\\ISIB DRVING SIMILATION\\vtech\\ts_crash.txt", header = T)
tsncrash <- read.table("F:\\ISIB DRVING SIMILATION\\vtech\\ts_ncrash.txt", header = T)

##############################################
dat <- tscrash[, c(1, 3, 5, 9, 10, 7)]
datn <- tsncrash[, c(1, 3, 5, 9, 10, 7)]
Lgt <- table(dat$webfile_id)
Lgt.c <- cumsum(Lgt)
ids <- dat$webfile_id[Lgt.c]
Lgtn <- table(datn$webfile_id)
Lgt.cn <- cumsum(Lgtn)
idsn <- datn$webfile_id[Lgt.cn]
##############################################



########################################################################
Z <- dat[dat$webfile_id == 0, ]
window <- rep(NA, length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25]))
laacc <- rep(NA, length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25]))
loacc <- rep(NA, length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25]))
yawrate <- rep(NA, length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25]))
speed <- rep(NA, length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25]))

getwindow <- function(id, whichdata, whatdoyouwant) {
  Z <- whichdata[whichdata$webfile_id == id, ]
  Z[, 2] <- Z[, 2] - Z[1, 2]

  for(i in 1:length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25])){
    window[i] <- Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25][i]
    laacc[i] <- Z[, 4][Z[, 2] == window[i]]
    loacc[i] <- Z[, 5][Z[, 2] == window[i]]
    yawrate[i] <- Z[, 6][Z[, 2] == window[i]]
    speed[i] <- Z[, 3][Z[, 2] == window[i]]
  }
    if(whatdoyouwant == "sdlaacc"){
    return(sd(laacc))
  }
  if(whatdoyouwant == "sdloacc"){
    return(sd(loacc))
  }
  if(whatdoyouwant == "sdyawrate"){
    return(sd(yawrate))
  }
  if(whatdoyouwant == "sdspeed"){
    return(sd(speed))
  }
  if(whatdoyouwant == "meanspeed"){
    return(mean(speed))
  }

}
###########################################################



###########################################################
getwhatyouwant <- function(all_ids, whichdat, whatdoyouwant){
  D <- NULL
  for(i in 1: length(all_ids)){
    D  <- c(D,getwindow(all_ids[i], whichdat, whatdoyouwant))
  }
return(D)
}
######
sdlac<- na.omit(getwhatyouwant(ids, dat, "sdlaacc"))
```

The R Code (2/2)

```
png(filename = "F:\\ISIB DRVING SIMILATION\\vtech\\crashfullplot.png",
    height = 1000, width = 1500)
    A <- NULL
par(mfrow=c(8,9))
for(i in 1: length(ids)){
  A  <- c(A,getwindow(ids[i], dat, "sdspeed"))

}
dev.off()

png(filename = "F:\\ISIB DRVING SIMILATION\\vtech\\nearcrashfullplot.png",
    height = 1000, width = 1500)
  B <- NULL
for(i in 1: length(idsn)){
         B <- c(B,getwindow(idsn[i], datn))
}
dev.off()
A <- na.omit(A)
B <- na.omit(B)
############################################################################


############################################################################
jraphsvs1525 <- function(id, whichdata) {

  M <- whichdata[whichdata$webfile_id == id, ]
  M[, 2] <- M[, 2] - M[1, 2]
for(i in 1:length(M[, 2][M[, 2]>= 15 & M[, 2]< 25])){
  window[i] <- M[, 2][M[, 2]>= 15 & M[, 2]< 25][i]
  laacc[i] <- M[, 4][M[, 2] == window[i]]
  loacc[i] <- M[, 5][M[, 2] == window[i]]
  yawrate[i] <- M[, 6][M[, 2] == window[i]]
  speed[i] <- M[, 3][M[, 2] == window[i]]
}
```

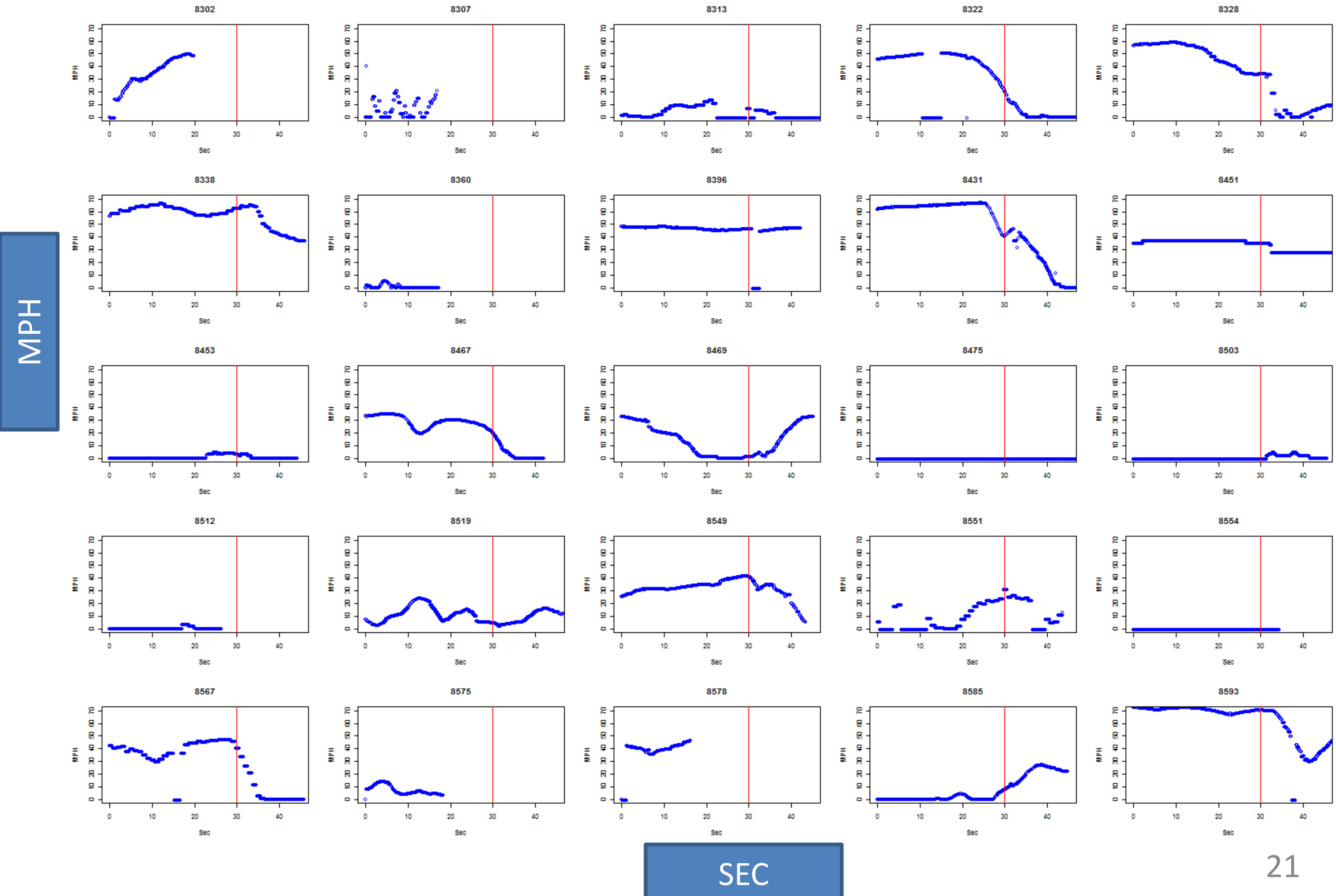
```
saveplotforwhich <- function(whichids, whichdata){
png(filename = "F:\\ISIB DRVING SIMILATION\\vtech\\graphs3.png",
      height = 1000, width = 1500)

  par(col = "blue")
  par(mfrow=c(8,9))
  for(i in 1: length(whichids)){
      graphspeedvs1525(whichids[i],whichdata)
    if(i == 72)
      break
    }

dev.off()
}
####
saveplotforwhich(ids, dat)
saveplotforwhich(idsn, datn)
###################################################################


###################################################################
t.test(A, B, var.equal=TRUE)
###
t.test(A, B, var.equal=FALSE)
t.test(A, B, paired=FALSE)
###
summary(A)
summary(B)
t.test(A, B, var.equal=TRUE, paired=FALSE)
###################################################################
t.test(msc, mscn, var.equal=FALSE)
summary(msc)
summary(mscn)
```

The R Code Summary

```
tscrash <- read.table("F:\\ISIB DRVING SIMILATION\\vtech\\ts_(

dat <- tscrash[, c(1, 3, 5, 9, 10, 7)]

for(i in 1:length(Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25])){
  window[i] <- Z[, 2][Z[, 2]>= 15 & Z[, 2]< 25][i]
  laacc[i] <- Z[, 4][Z[, 2] == window[i]]
  loacc[i] <- Z[, 5][Z[, 2] == window[i]]
  yawrate[i] <- Z[, 6][Z[, 2] == window[i]]
  speed[i] <- Z[, 3][Z[, 2] == window[i]]
}
   if(whatdoyouwant == "sdyawrate"){
     return(sd(yawrate))
   }
   if(whatdoyouwant == "sdspeed"){
     return(sd(speed))
   }

  t.test(A, B, var.equal=FALSE)
```

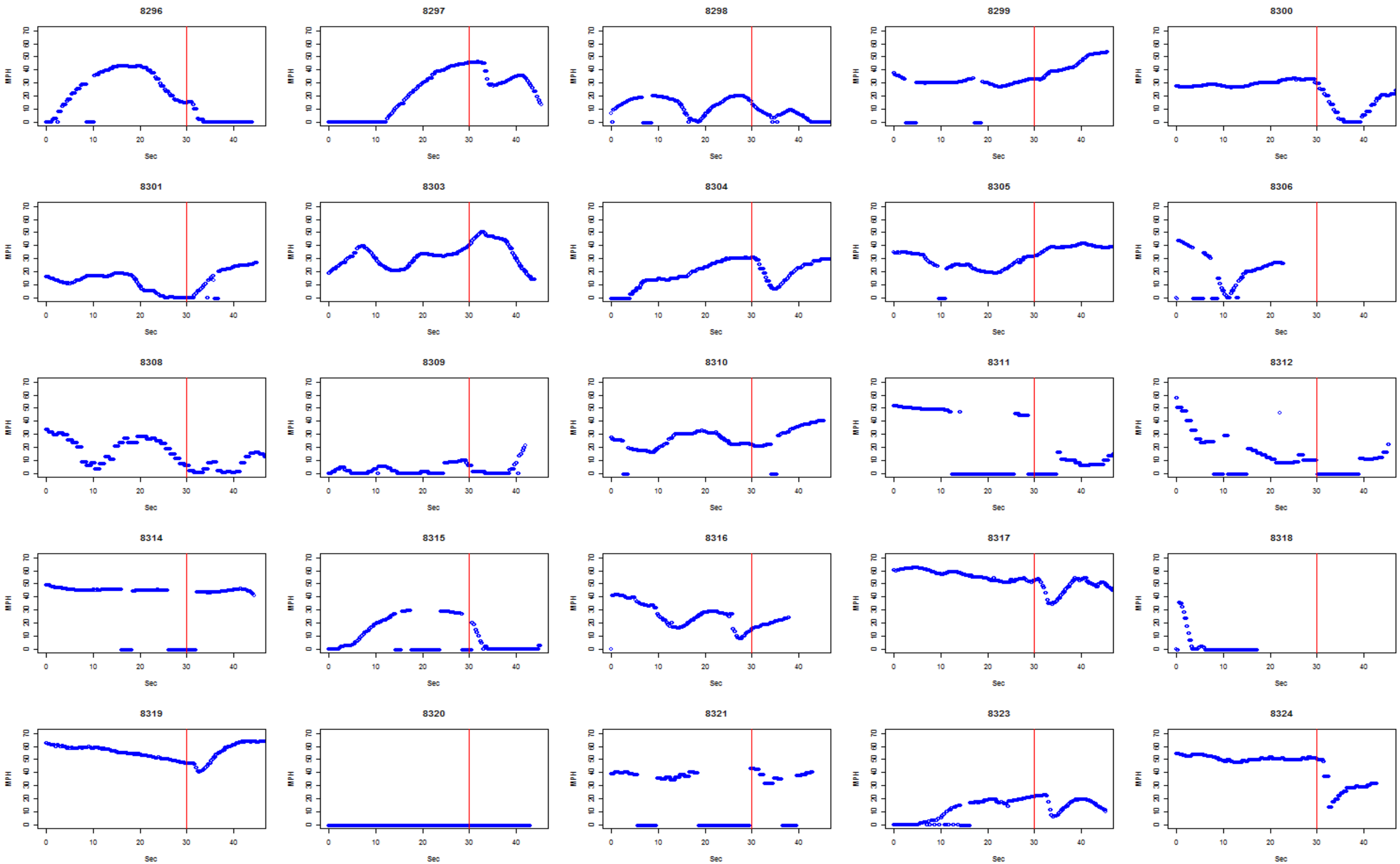
- Read the data
- Take wanted variables
- Find window in wanted variables
- Return them
- Graph
- T-test

Time series Plots for 25/68 Crash



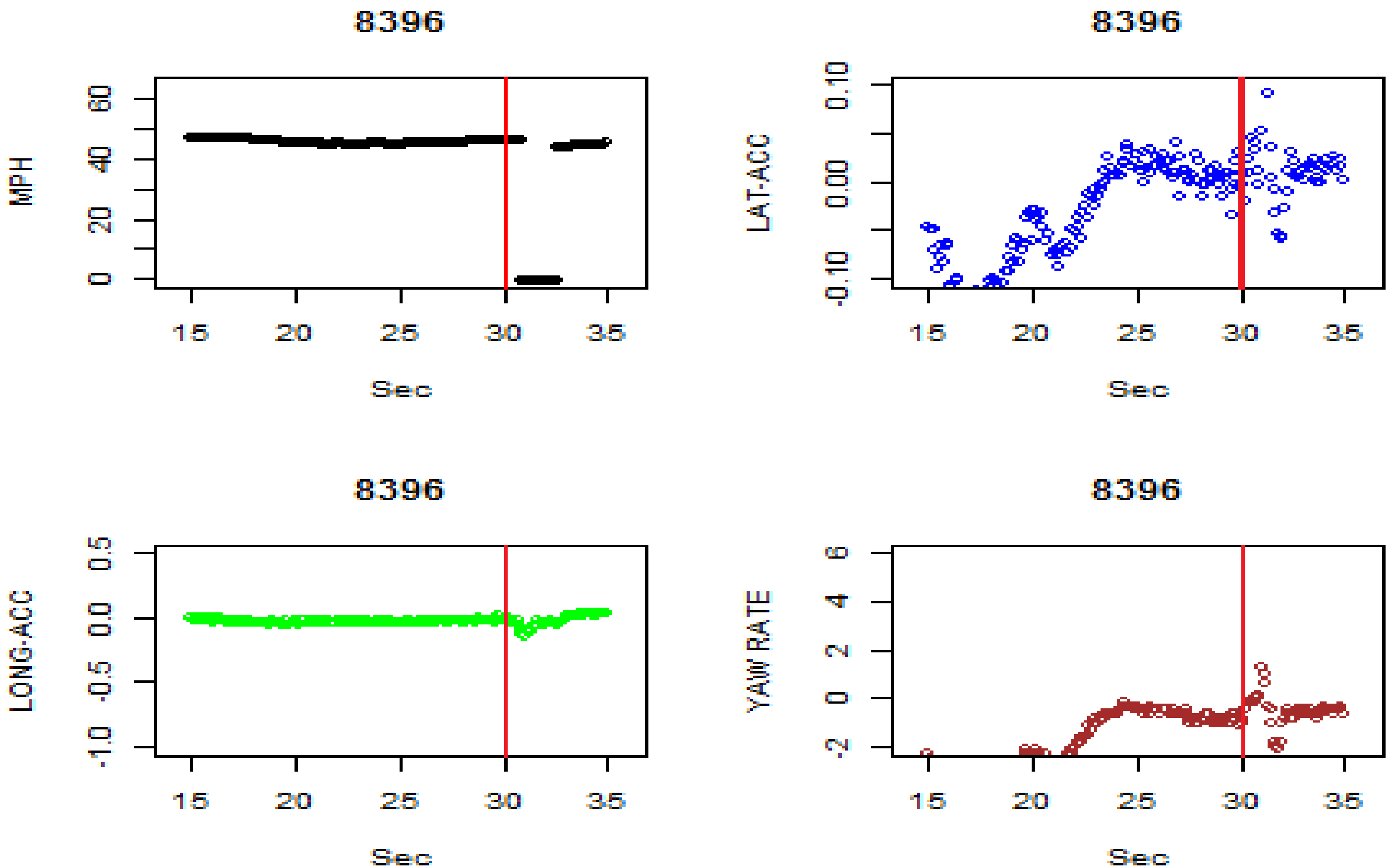
Time series Plots for 25/761 for Near-Crash events

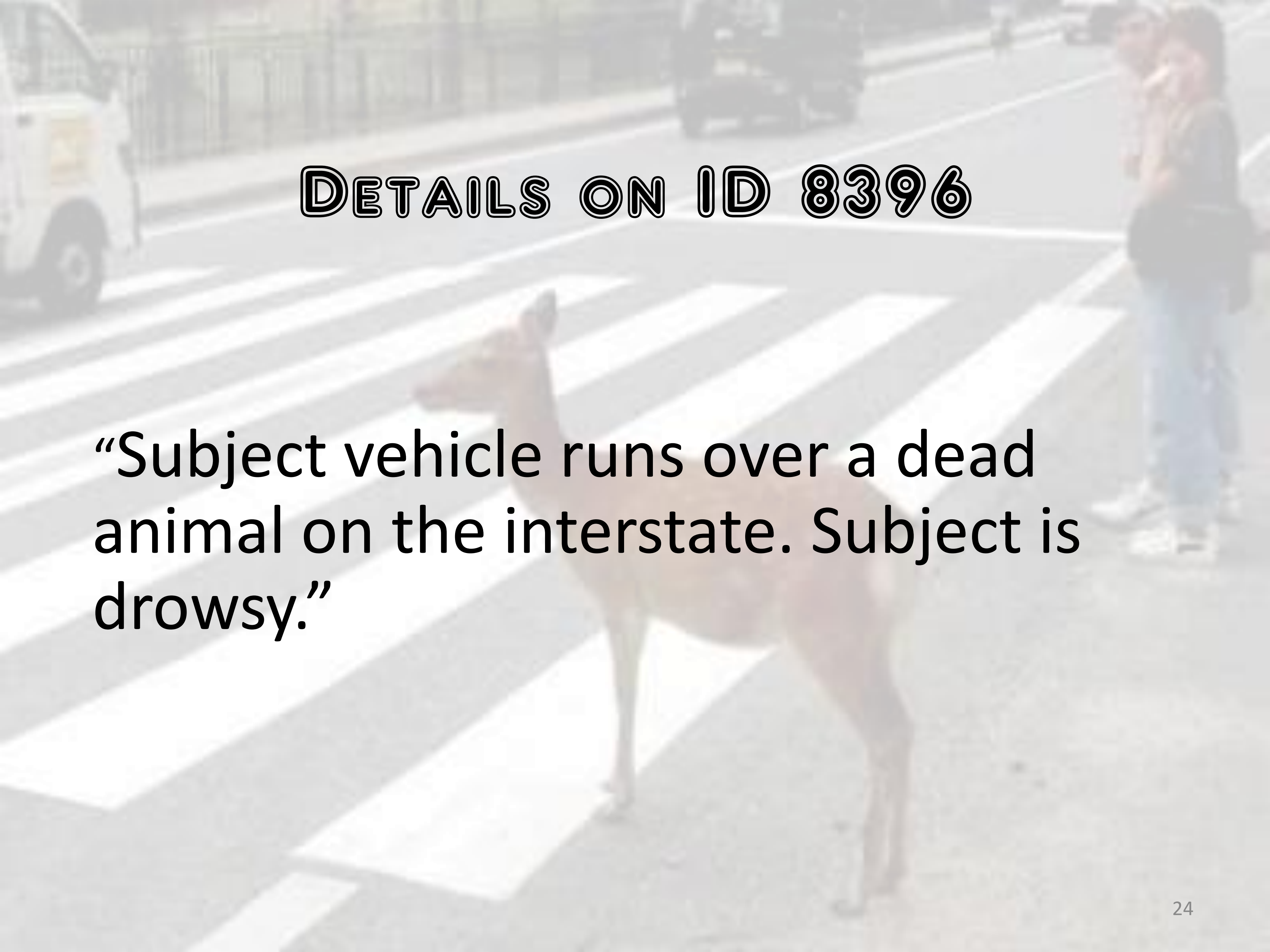
MPH



SEC

Graph crash 8396

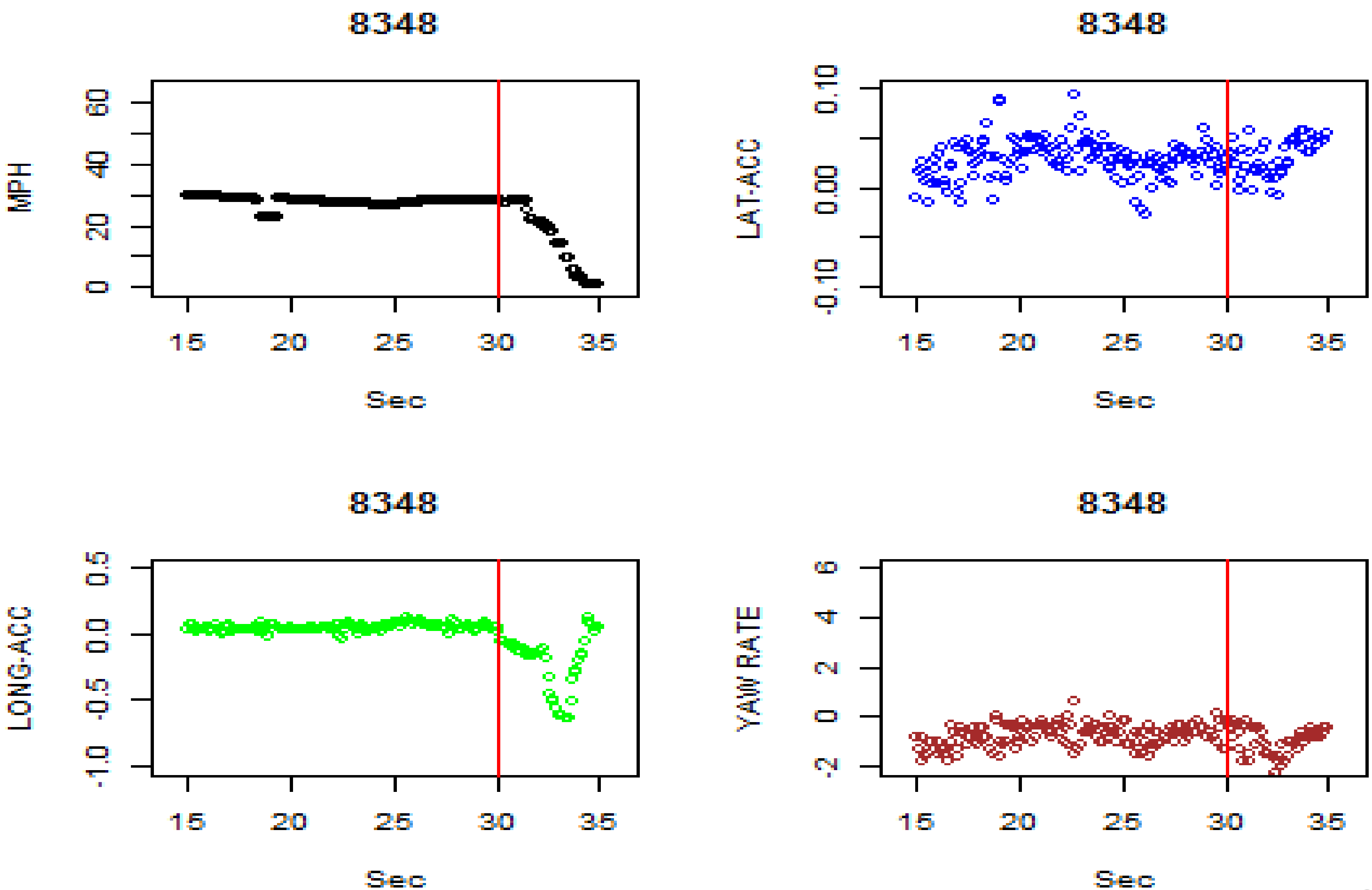




Details on ID 8396

“Subject vehicle runs over a dead animal on the interstate. Subject is drowsy.”

Graph near-crash 8348



Details on ID 8348

"Subject is distracted by eating in the car when lead vehicle stops suddenly because an oncoming vehicle is turning to its left, across the path of the lead and subject vehicles, to enter a parking lot. Subject must brake hard to avoid hitting the lead vehicle in the rear, and car following subject must also brake to avoid hitting subject vehicle in the rear."

Results for Welch test

MEANS → ↓ RESULTS	SD of Lat-Acc	SD of Lon-Acc	SD of Yaw-rate	SD of Speed	Mean of Speed
t	1.893	0.944	2.297	-1.392	-2.881
p-value	0.062	0.348	0.024	0.167	0.005
Mean of … crash	0.050	0.064	2.718	2.861	21.511
Mean of … near-crash	0.039	0.057	1.653	3.395	28.754

Discussion

- Incomplete data observed
  - Difficult to implement reliable warning device
- Crashes lower speed than Near-Crashes
- Many Crashes were mild in nature
  - Hitting stationary object in parking lot lightly
- Higher Variability in Yaw-rate & Lat-Acc for Crashes
  - Indicates loss of vehicle control

Limitation

- Non-Event Data needed
- Had to treat all events as independent
  - Drivers ID’s
    - Drivers were repeated
    - Results may be over-studying significance

Future Research

- Obtain data on non-event segments
- Electronic data
  - Prediction
  - Crash & Near-Crash
  - Real time

References

- Virginia Tech Transportation Institute (VTTI)
- National Highway Traffic Safety Administration (NHTSA)

Acknowledgment

- Jeremy Sudweeks, VTTI
  - Provided data
  - Dictionary
- Ming Yang & Mitch Thomann & Yu-Hui Huang
  - Help with R Codes
- ISIB
  - Food
  - Dorm
  - Activities
  - Course & Info (Dr. Zamba, Dr. Kate, Dr. Oleson, Dr. Chaloner, Dr. Brian, everyone else)