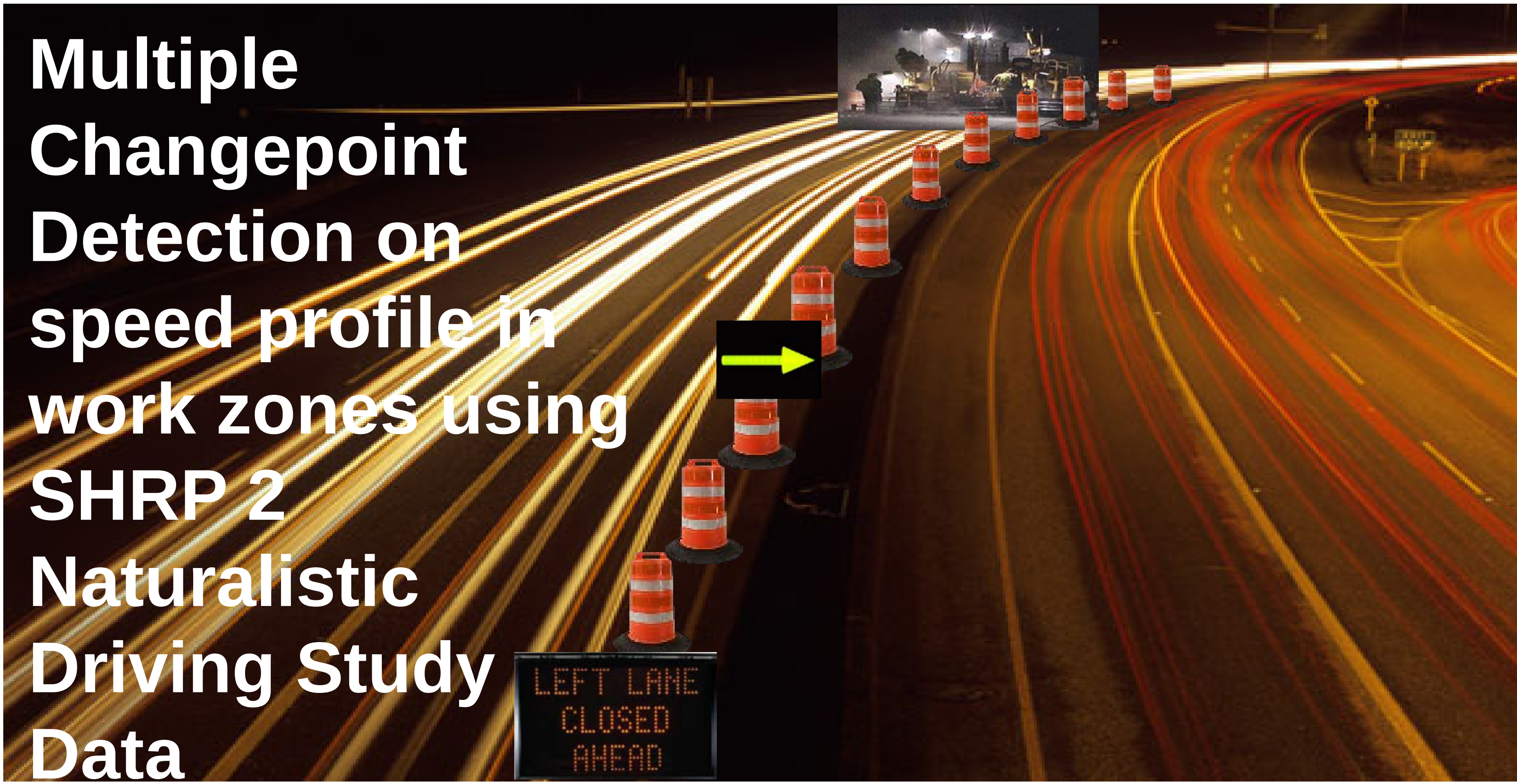
IOWA STATE UNIVERSITY

Institute for Transportation (InTrans)

Multiple Changepoint Detection on speed profile in work zones using SHRP 2 Naturalistic Driving Study Data



Hossein Naraghi

2017 Mid-Continent TRS

Background

- Presence of road constructions and maintenance
  - Increase safety risks
    - Disturbance to traffic flow
    - High cognitive work load for drivers
- **700-1000** fatalities per year in last 10 years
- **11%** increase in work zone fatalities (2011 – 2014)
  - Small decrease in non-work zone fatalities
- Limited details on available crash data
  - Officer interpretation, Witness testimony
- Unclear driver behavior and underlying causes of work zone crashes

Background

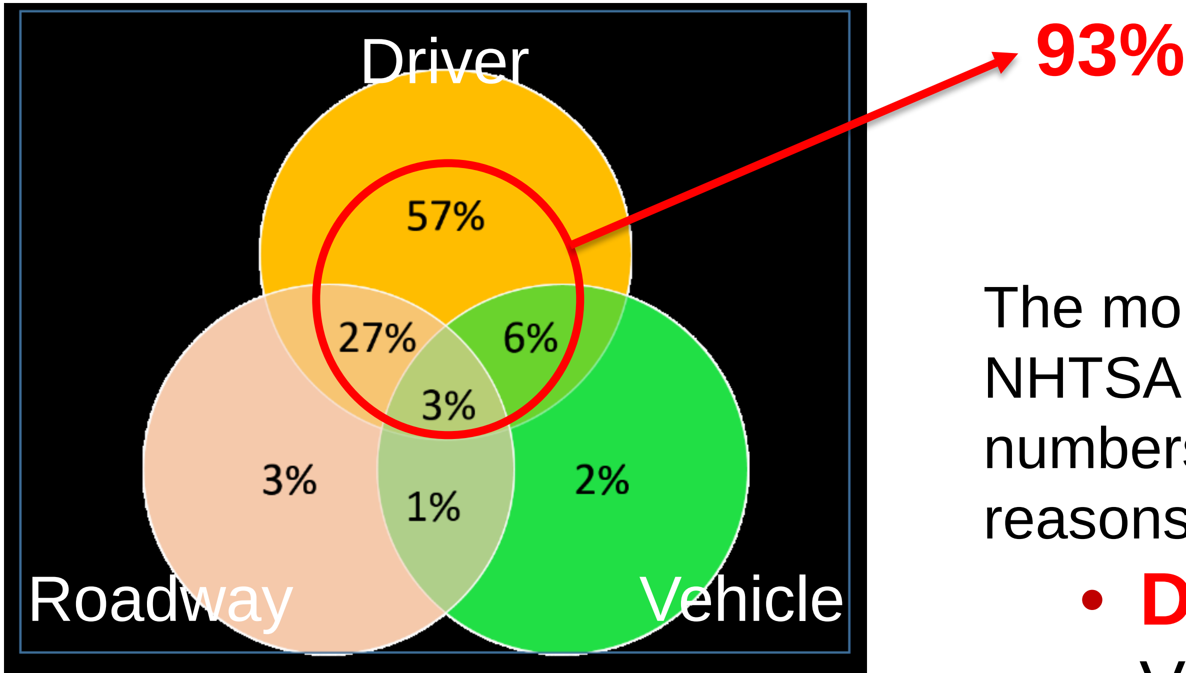
Speed management effectiveness

- Numerous countermeasures tends to get drivers attention to reduce speed
- How drivers react to measures like DMS, lane closure, merge sign, speed feedback sign, and ….?
- Are the safety measures effective?
- How effective are ITS strategies in getting driver attention?

- How to identify the most effective countermeasure if driver behavior is not well understood?

- SHRP2 NDS Time Series speed data collected at high frequency at 10HZ (every 0.1 second)
- Opportunity for a first hand observation of driver behavior

Background



Rumar et al.' study (1985)

The more recent study by NHTSA in 2015 confirms numbers and reveals critical reasons attributed to

- **Drivers 94%**
- Vehicles 2%
- Environment 2%
- Unknown 2%

NHTSA Traffic safety Facts 2015
(Data Source: NMVCCS 2005–2007)

Objectives

Utilize SHRP2 NDS time series data and forward video images

- Determine where drivers starts to react to different work zone features
- Identify the effectiveness of speed countermeasures
- Develop models to predict driver reactions to work zone features

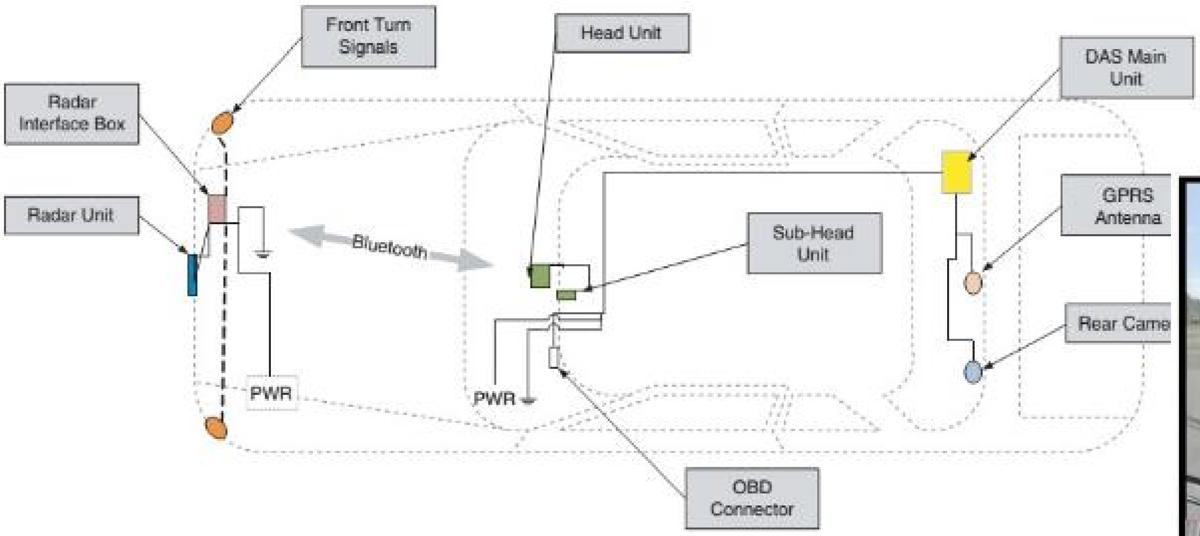
Data collection

- SHRP 2 Naturalistic Driving Study Data (NDS)
  - 3100 participant drivers
  - Oct 2010 – Nov 2013
  - 6 Sites
  - 5 M trips
  - 20 M data miles
  - 4 M GIGs data

Site	Primary Participant Count
Buffalo, NY	719
Tampa, FL	698
Seattle, WA	676
Durham, NC	504
Bloomington, IN	239
State College, PA	256

VTTI, 2013

Data Acquisition System



SHRP2 data acquisition system installation schematic (Antin et al., 2011)

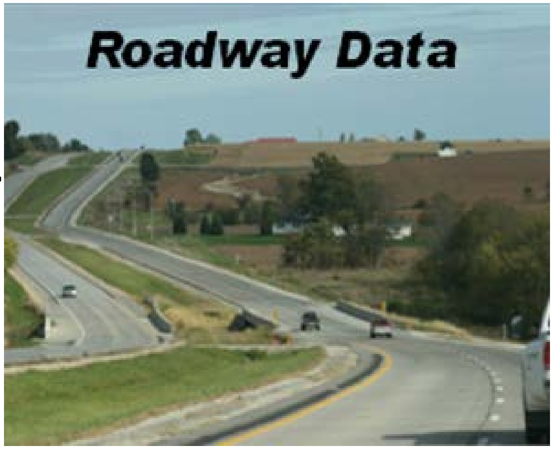
(The driver in the picture is not experimental participant and just for demonstration purpose only) (Dingus et al., 2014)

DAS Includes:

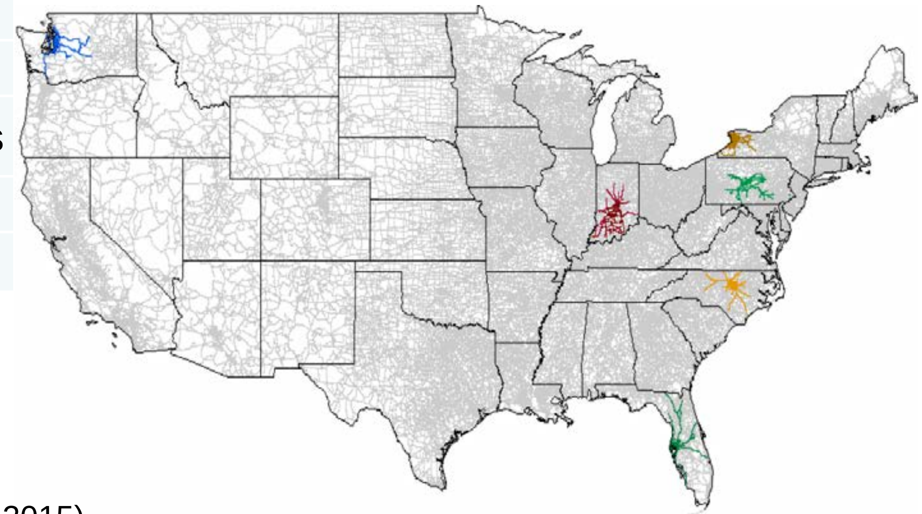
- Forward radar
- Four video cameras, including one forward-facing
- GPS
- Accelerometers
- Computer vision lane tracking
- DAS main unit

Data collection

- SHRP 2 Roadway Information Database (RID)
  - 25,000 miles of roadway data



Curve Radius	Lane Types
Curve Length	Shoulder Types
Curve Point of curvature	All Signs on MUTCD
Curve Point of tangency	Guardrails/Barriers
Curve Direction	Intersection Geometry
Grade	Median Types
Cross-slope/Super elevation	Presence of Rumble strips
Number of Lanes	Presence of Lighting
Lane Width	



A site map shows the RID database in the SHRP2 NDS (Smadi, 2015)

Data

Speed time series trajectories ¼ mile upstream of 1st WZ warning sign through work area

- 62 speed traces
- 29 male drivers
- 33 female drivers
- 35 Unique drivers
  - 18 male
  - 17 female
- 36 traces in open lane
- 26 traces in closed lane

WZ Configuration

- 4-lane divided Rural
  - Left lane closure
  - Cross-over
  - head to head traffic

WZ Features

- Identified in forward video
- Matched with nearest time stamp
- Vehicle position relative to the start of taper calculated for every 0.1 sec using GPS speed at location

Data quality assurance

- Forward video image quality
  - Excluded traces with very poor image quality
    - Identify the location of countermeasures
    - Detect the position of vehicle
    - Confirm active WZ
    - Verify free flow

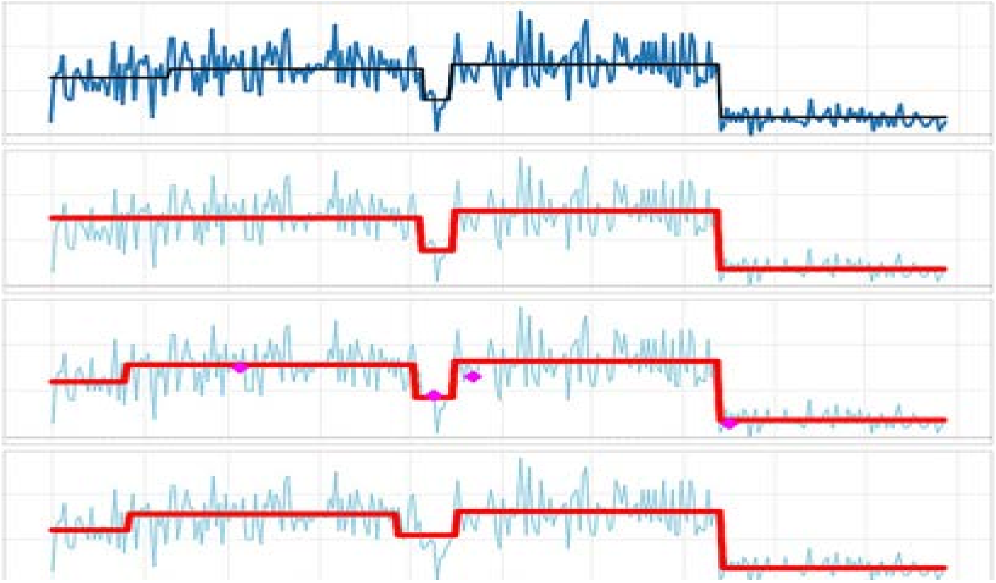
- Speed time series data issue was missing data
  - Excluded traces with more than 25% missing values
  - Utilize interpolation techniques to calculate missing values

Methodology

Multiple changepoint analysis

Used in different application areas

- Climatology
- Finance
- Medical imaging
- Oceanography



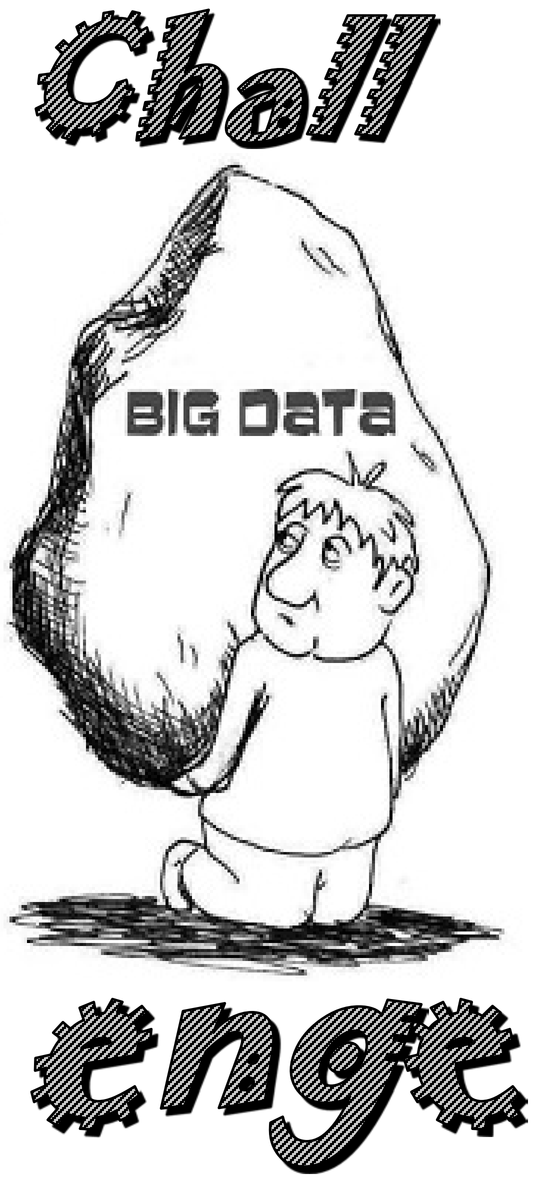
Multiple changepoint detection

Identify driver's reaction to series of countermeasures in WZ

Changepoint analysis or segmentation regression

- Aims to achieve homogeneity within time series data
- Method of time series analysis which divides data into sequence of isolated segments to show underlying properties of its source of isolation
- A time instance when statistical properties of data change
- Partition data into regions lies in between different important events
    - Series of speed countermeasures
        - DMS, merge sign, WZ speed limit, flashing arrow, speed feedback sign, …

- Use an efficient method for searching the solution space
- Choose the appropriate penalty which depends on the size of the changes and length of segments
  - Unknown prior to analysis
- Identifying the optimal number and location of changepoints



Multiple changepoint analysis

Pruned Exact Linear Time (PELT) package in R utilized to identify multiple changepoints by partitioning data into different regions based on different important events

Speed time series data is defined as:

$$V_{1:n} = (V_1, V_2, \ldots, V_n)$$

A changepoint may occur within this set when there exists a time

$$\tau \in \{1, \ldots, n-1\}$$

where statistical properties of

$$\{V_1, \ldots, V_\tau\} \text{ and } \{V_{\tau+1}, \ldots, V_n\}$$

are different in some way

Multiple changepoint analysis

For multiple changes, there are multiple changepoints, m, together with their positions, expressed as follows:

$$\tau_{1:m} = (\tau_1, \ldots, \tau_m)$$

Each changepoint is an integer between 1 and $n-1$

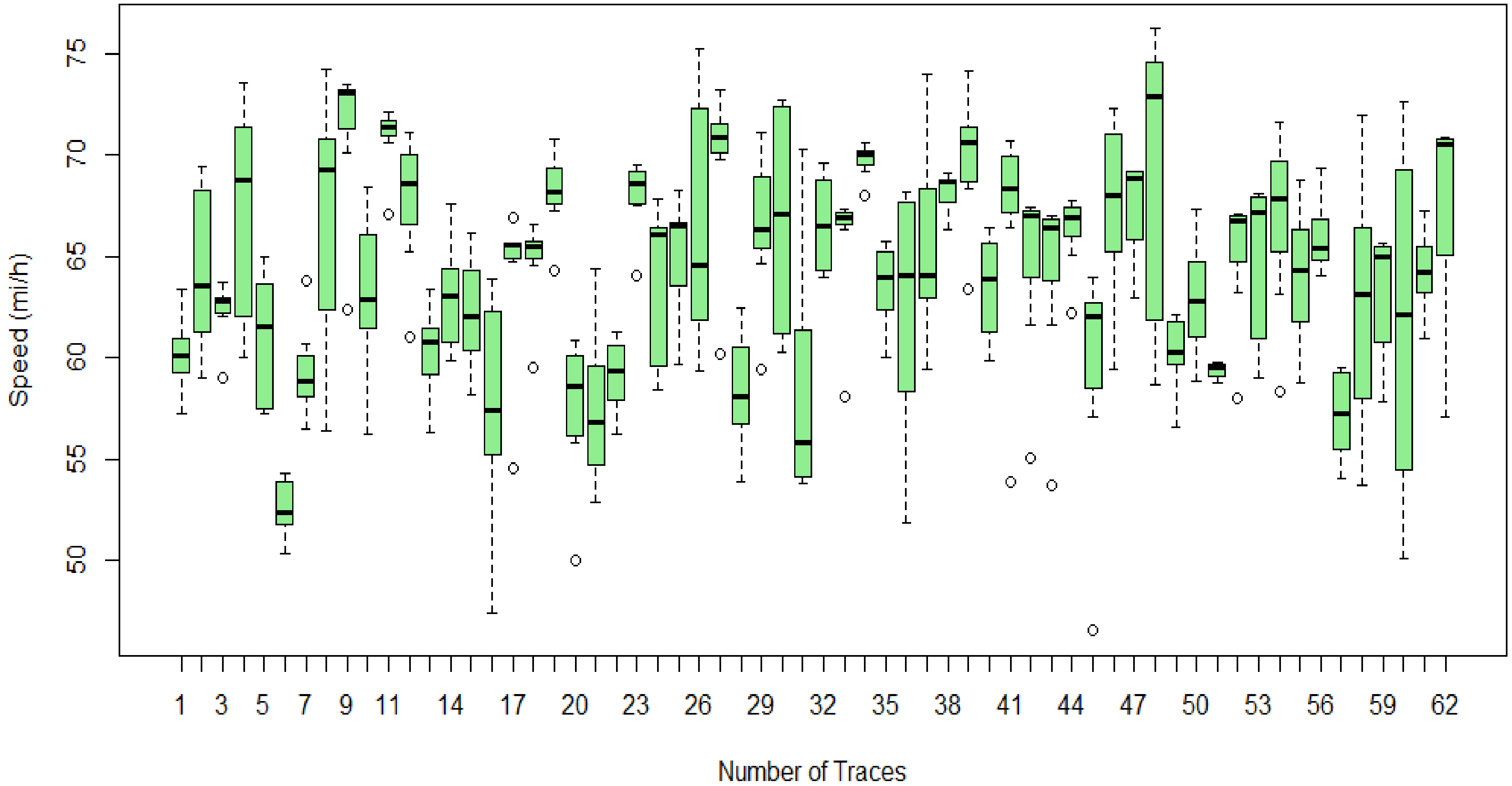
the m changepoints will split the data into $m + 1$ segments

The aim of the analysis is to efficiently and accurately estimate the location of multiple changepoints by minimizing the following equation:

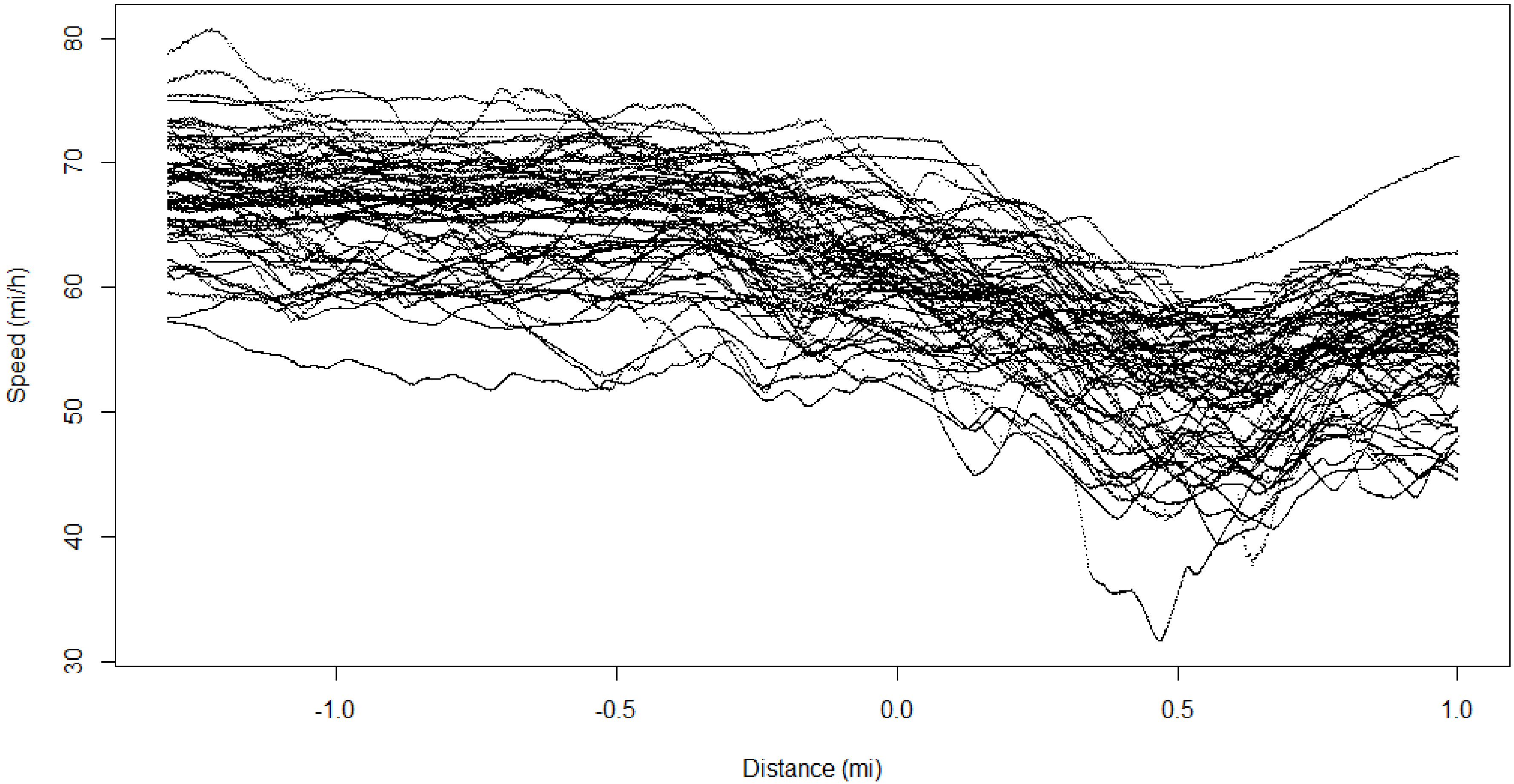
$$\sum_{i=1}^{m+1} \left[C\left(V_{(\tau_{i\_1}+1):\,\tau_i}\right) \right] + \beta f(m)$$

cost function penalty against overfitting

Speed traces distribution

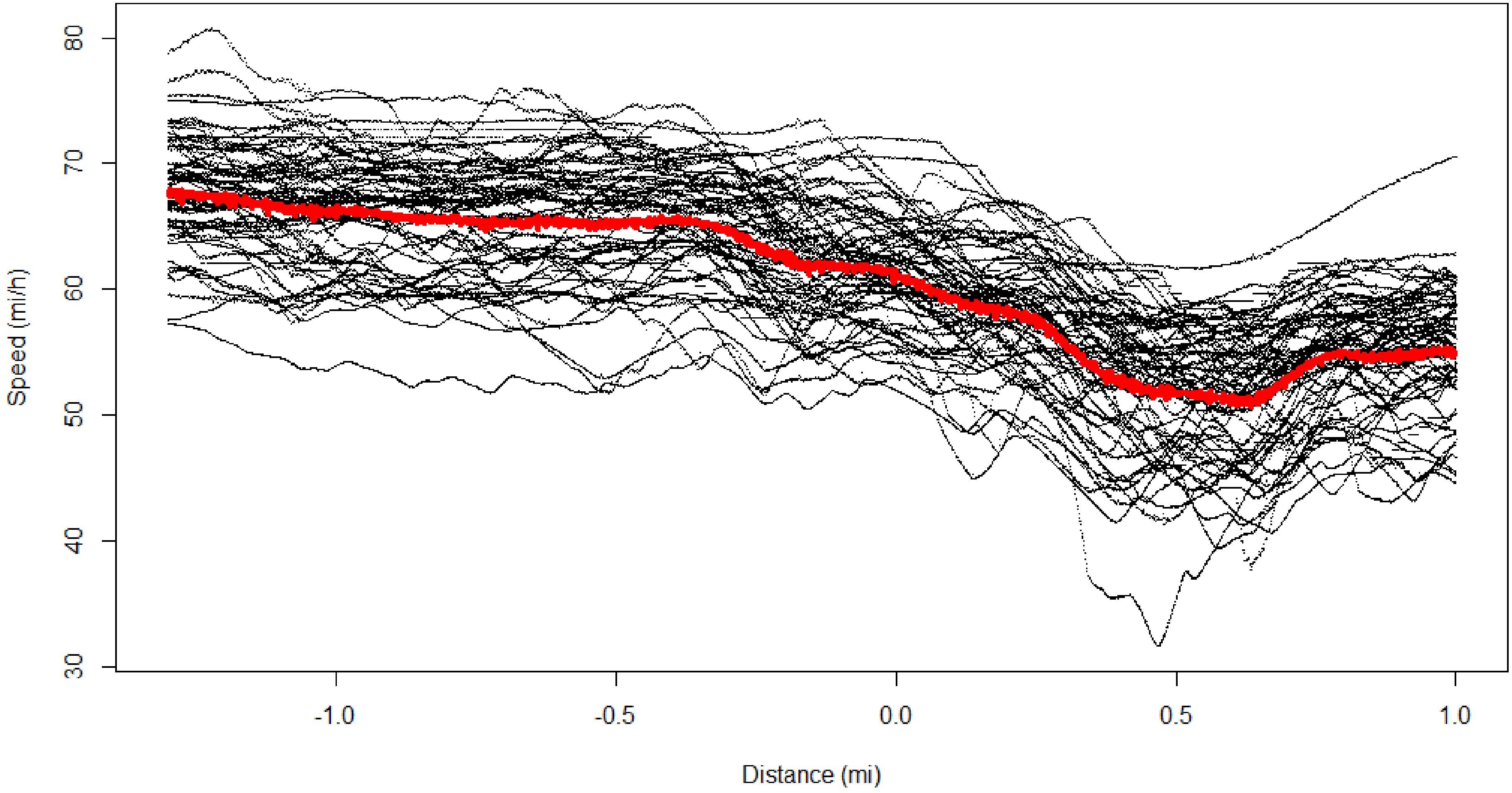


Time series speed traces

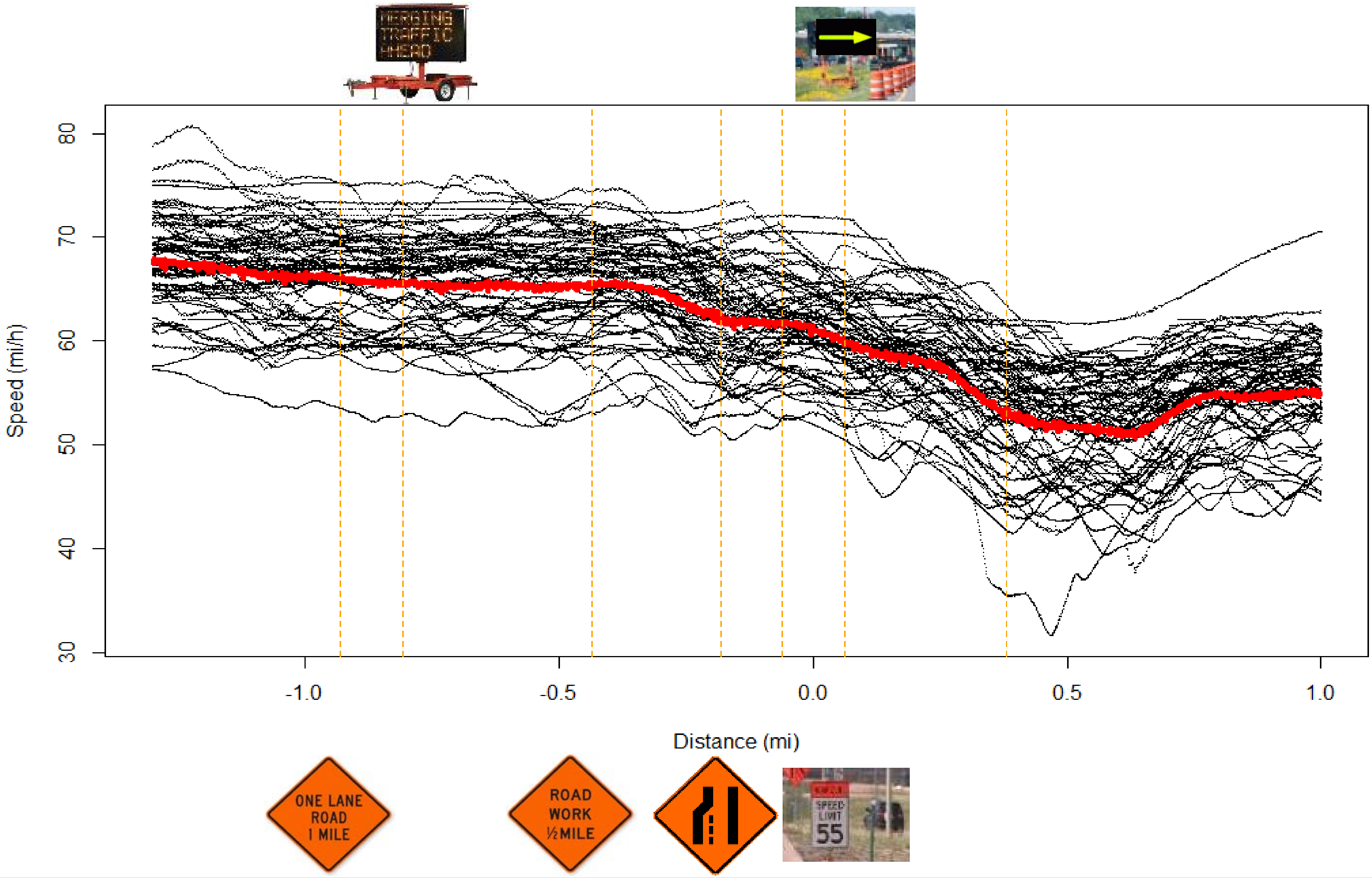


- 62 Time series speed traces
- 29 Male 33 Female
- 35 unique drivers
- 18 Male, 17 Female

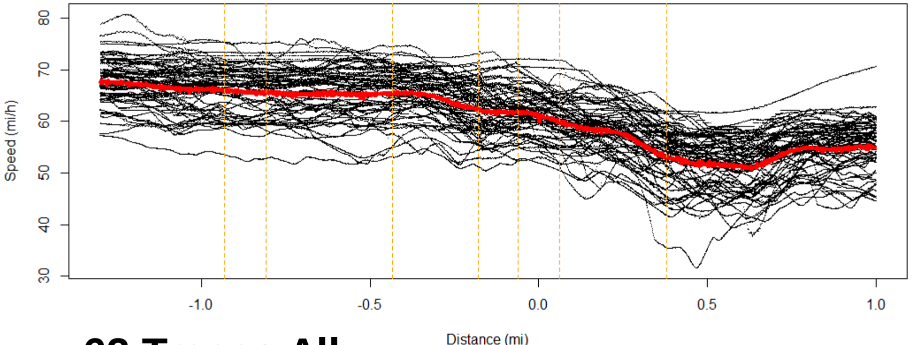
Mean of time series speed traces



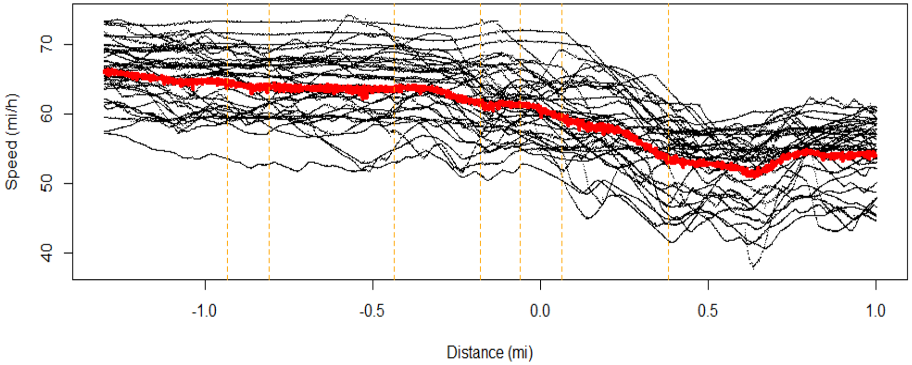
Countermeasure types and locations



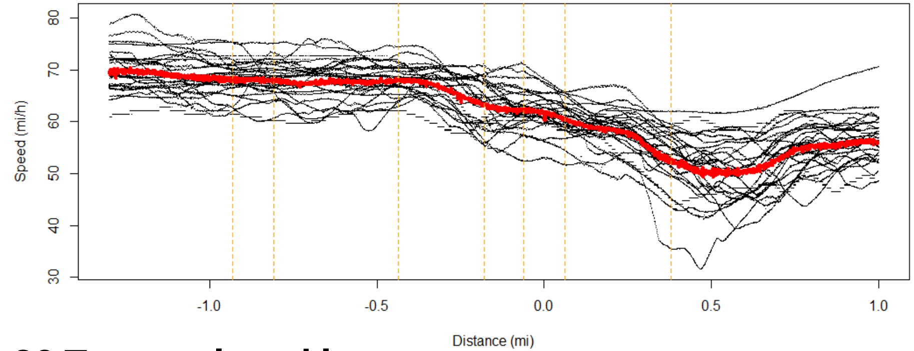
Mean of time series speed traces



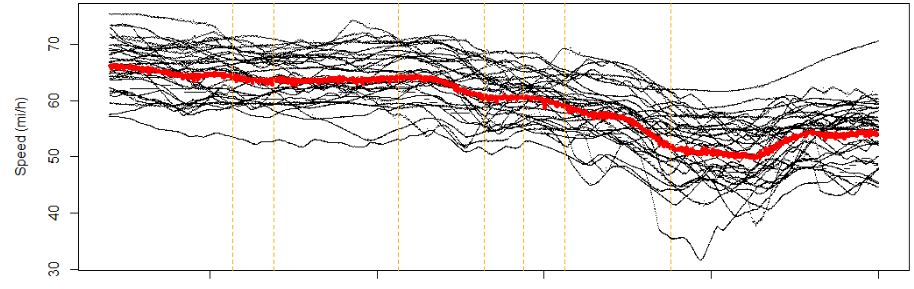
62 Traces All



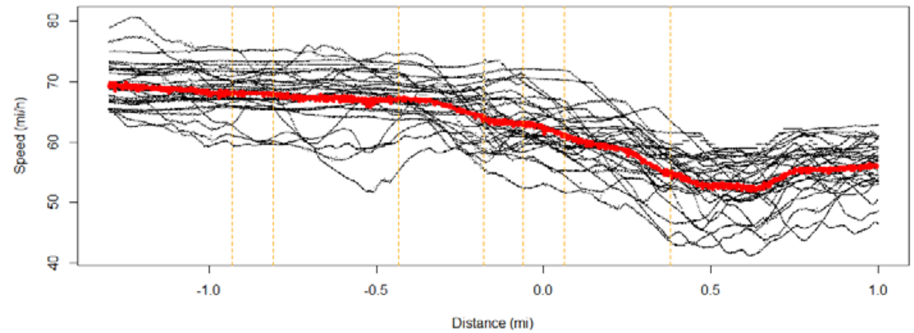
36 Traces open lane



26 Traces closed lane

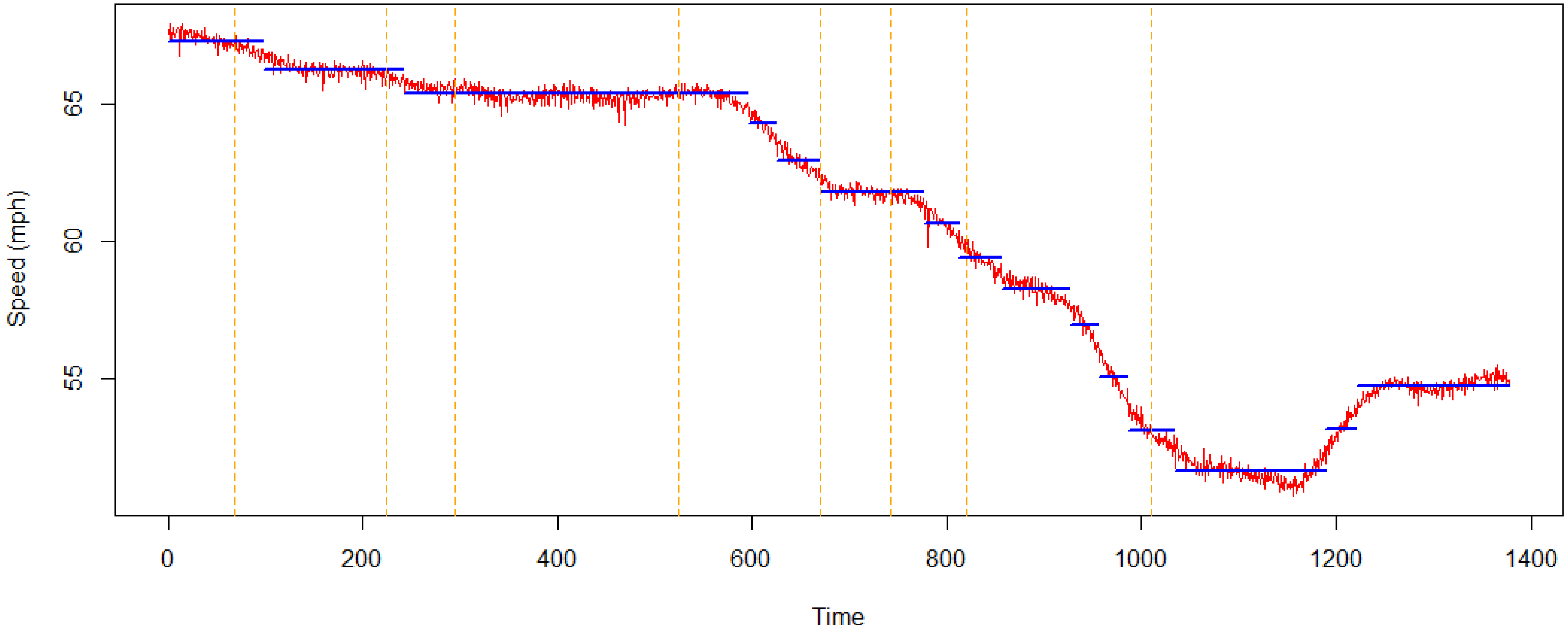


33 Traces female drivers



29 Traces male drivers

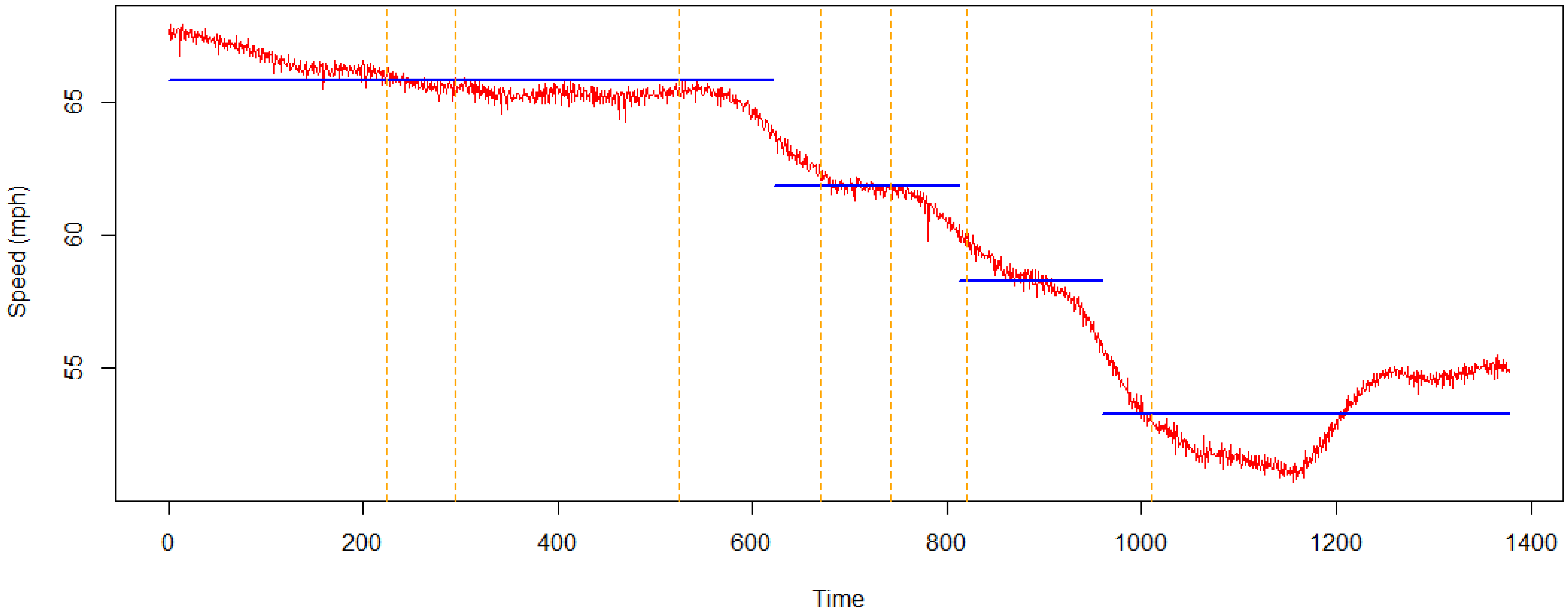
Mean Changepoints



Changepoint type : Change in mean
Method of analysis : PELT
Test Statistic : Normal
Type of penalty : **BIC** with value, 14.45678
Minimum Segment Length : 1
Maximum no. of cpts : Inf
Changepoint Locations : 97 241 595 624 669 776 812 856 926 955 986 1033 1189 1220

> ncpts(change_mean)
[1] 14
> seg.len(change_mean)
[1] 97 144 354 29 45 107 36 44 70 29 31 47 156 31 158
coef(change_mean)
$mean
[1] 67.3 66.3 65.4 64.3 63 61.8 60.7 59.4 58.2 56.9
[11] 55.1 53.1 51.6 53.2 54.7

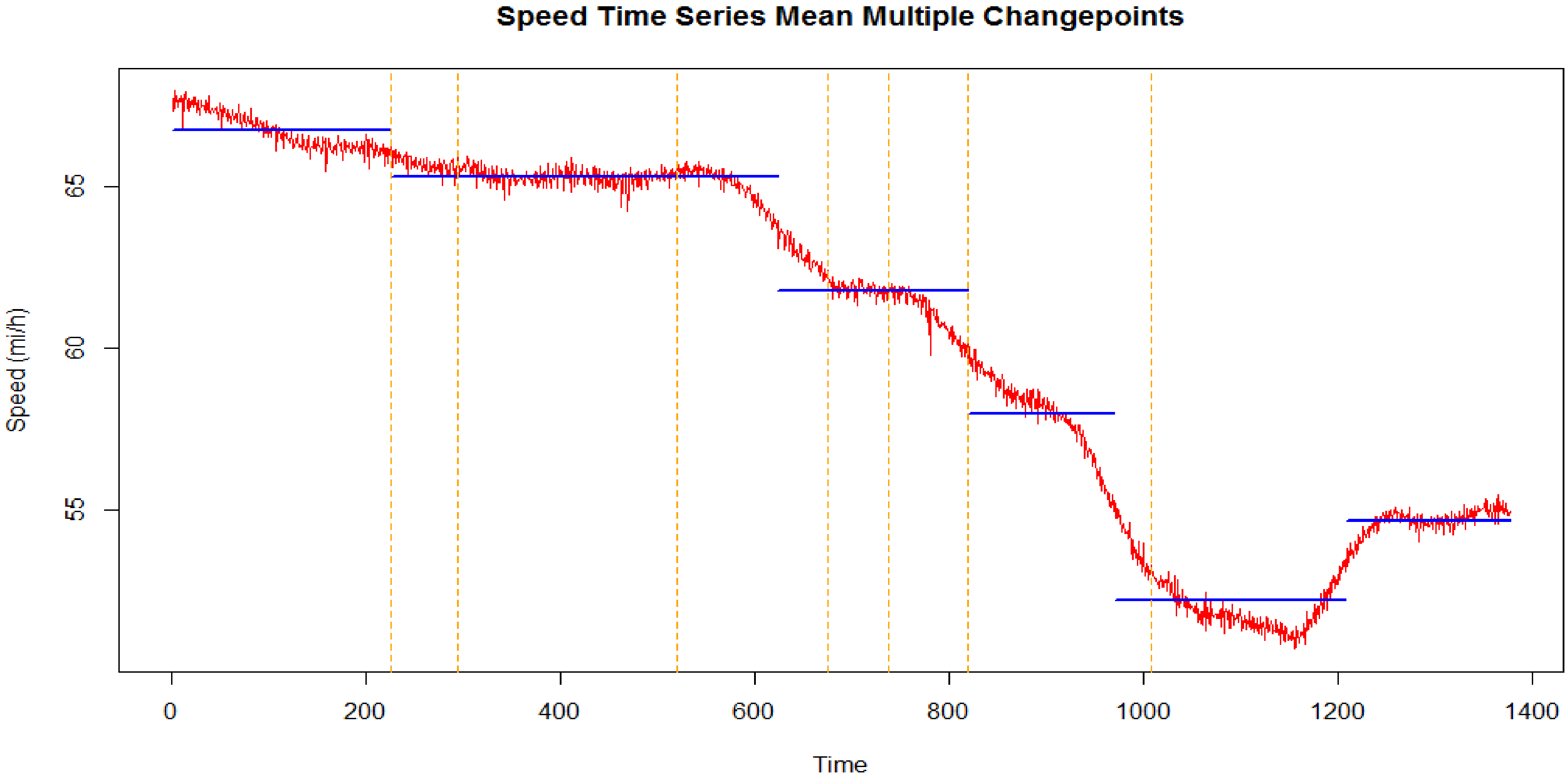
Mean Changepoints



Changepoint type : Change in mean
Method of analysis : PELT
Test Statistic : Normal
Type of penalty : **Manual** with value, 722.8388
Minimum Segment Length : 1
Maximum no. of cpts : Inf
Changepoint Locations : 621 812 959

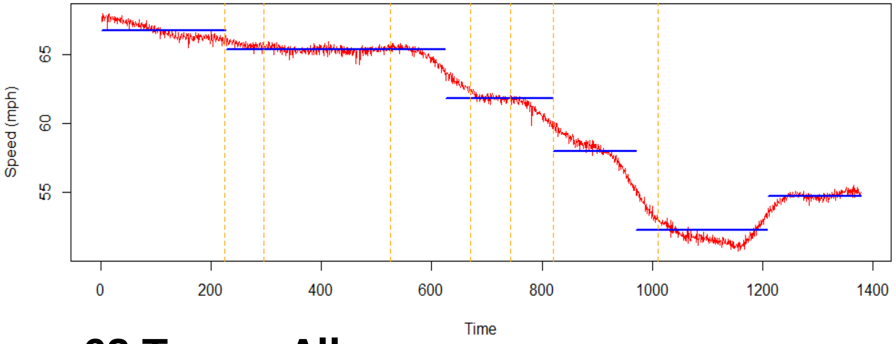
```
> ncpts(change_mean)
[1] 3
> seg.len(change_mean)
[1] 621 191 147 419
> coef(change_mean)
$mean
[1] 65.8  61.9  58.3  53.3
```

Mean Changepoints

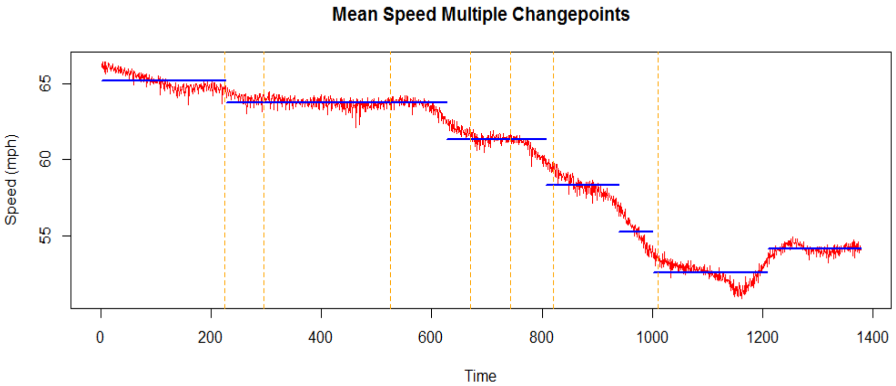


- Number of changepoints = 5
- Mean change coef: 66.8, 65.3, 61.8, 58, 52.2, 54.7
- Segment length: 226, 398, 196, 150, 239, 169

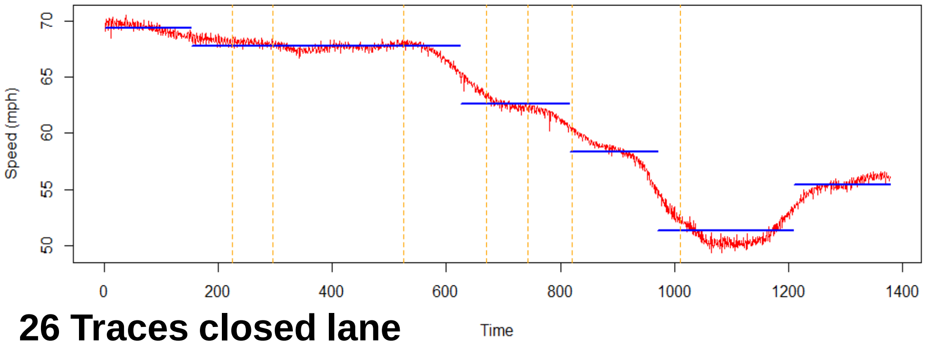
Mean Changepoints



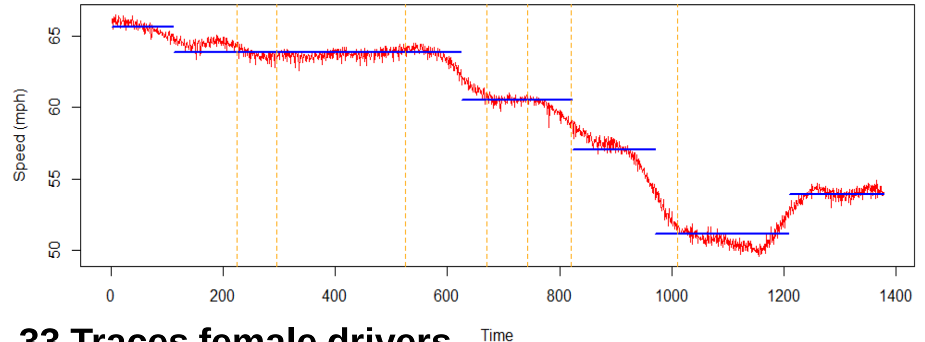
62 Traces All



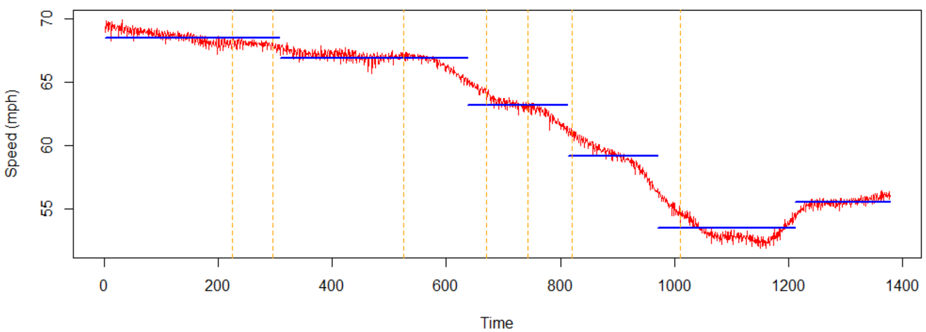
36 Traces open lane



26 Traces closed lane



33 Traces female drivers



29 Traces male drivers

Mean Changepoints Summary

Model Input Data	Changepoint PELT Model									
	Mean Speed (mi/h) Upstream	Change in Mean Speed (mi/h)	Mean Speed (mi/h) Reacting to first WZ sign and DMS	Change in Mean Speed (mi/h)	Mean Speed (mi/h) Reacting to Lane closure sign, WZ posted speed sign, start of taper and flashing arrow signs	Change in Mean Speed (mi/h)	Mean Speed (mi/h) Reacting to WZ posted speed sign, channelization	Change in Mean Speed (mi/h)	Mean Speed (mi/h) Reacting to start of work area	Upstream to work area speed reduction (mi/h)
All (62 Traces)	66.8	**-1.5**	65.3	**-3.5**	61.8	**-3.8**	58.0	**-5.8**	52.2	**-14.6**
Open lane (36 Traces)	65.2	**-1.4**	63.7	**-2.5**	61.2	**-3.3**	57.9	**-4.5**	53.4	**-11.8**
Closed lane (26 Traces)	**69.3**	**-1.6**	67.7	**-5.1**	62.6	**-4.2**	58.4	**-7.0**	**51.4**	**-17.9**
Female drivers (33 Traces)	65.6	**-1.7**	63.9	**-3.4**	60.5	**-3.5**	57.0	**-5.9**	**51.1**	**-14.5**
Male drivers (29 Traces)	**68.5**	**-1.6**	66.9	**-3.7**	63.2	**-4.0**	59.2	**-5.7**	53.5	**-15.0**

Conclusions

Changepoint PELT Model

- NDS data provides opportunity to understand and summarize driver behavior by analyzing time series data

- The analysis reveals a promising results on quantifying driver's speed behavior
  - Provide reaction location
  - Prime changepoint locations
  - Length of stable condition
  - Magnitude of mean speed reduction

Future Research

- Identify which work zone features are more likely to get driver's attention
  - DMS compared to static sign
  - DSFS vs. WZ speed limit sign
- Determine the effectiveness of the locations of the countermeasures
  - DMS located after 1st warning sign
  - DMS after the merge sign
- Investigate the effectiveness of WZ configurations
  - left lane vs. right lane closure
  - lane shift no shoulder vs. lane closure
- Look for roadway and environmental effects

Have patience

All things are difficult before they become easy