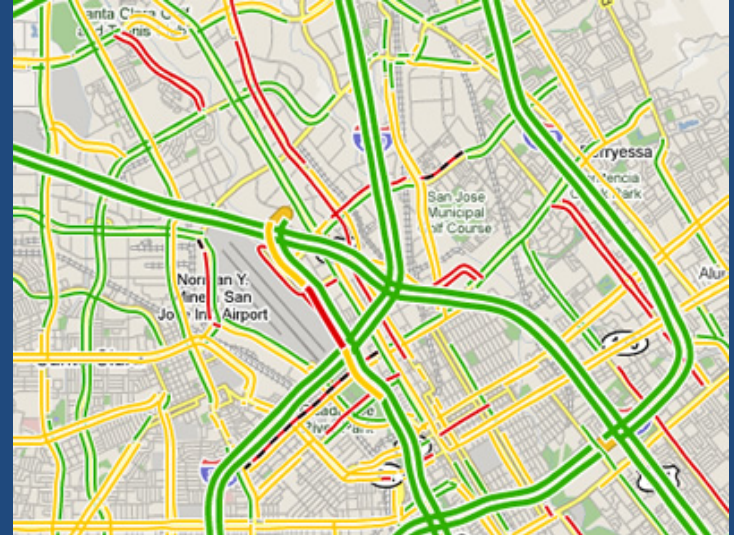


# Probe Data: Opening the Black Box



## Project Team

**Neil Ternowetsky**, B.Sc., Director of Information Systems

**Chao Bian**, B.Sc., M.Sc., System Developer

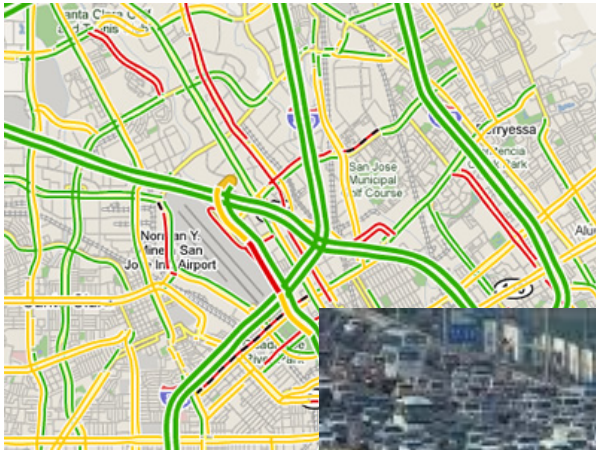
**Keenan Pattmore**, M.Sc., P.Eng, Transport Engineer

**Adam Rocan**, B.Sc., System Developer

NATMEC 2016  
Probe Data Collection and Accuracy



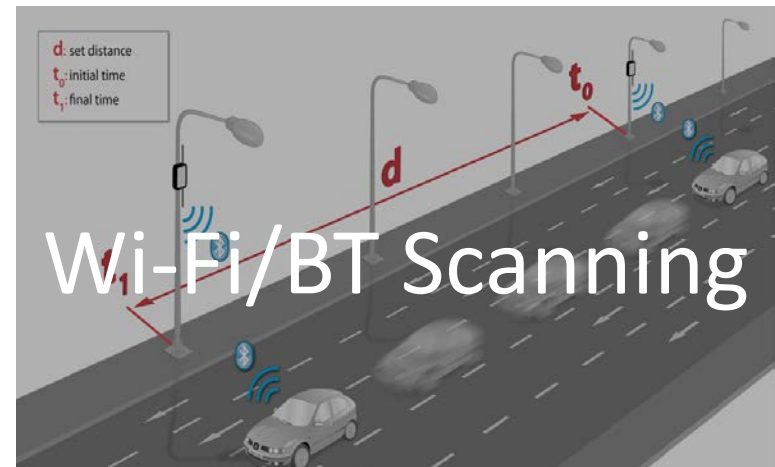
# What is Probe Data?



- Vehicle data that is used to determine traffic conditions.
- Used for:
  - Real-Time Traffic Monitoring
  - Congestion Studies
  - Assessing Infrastructure Investments

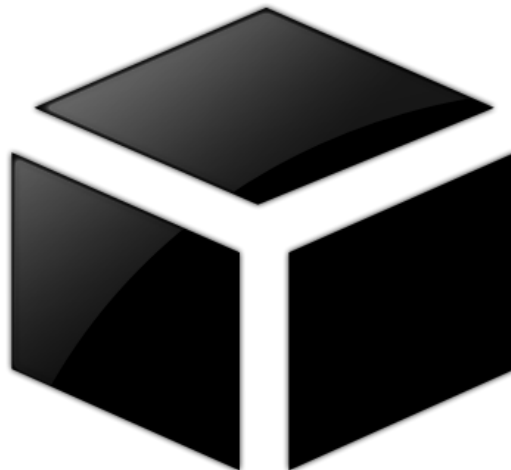
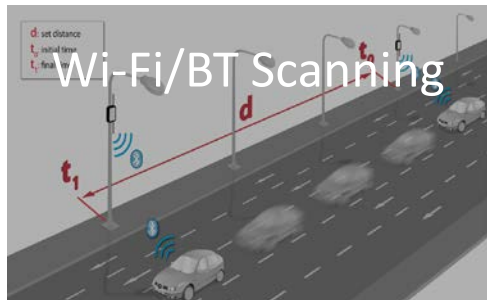
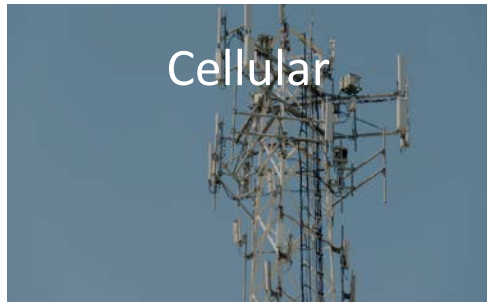
# What is Probe Data?

## Sources of Probe Data





# What is the Black Box?



# Outline of Presentation

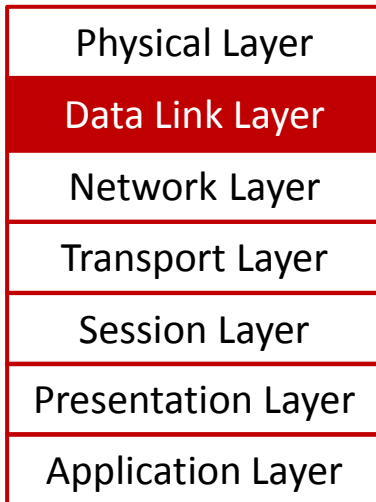
- Describe how Wi-Fi/BT scanning works
- Outline the architecture of a Wi-Fi/BT systems
- Details of unit and system testing
- Outline considerations for selecting a Wi-Fi/BT solution

# How Does Wi-Fi/Bluetooth Scanning Work?

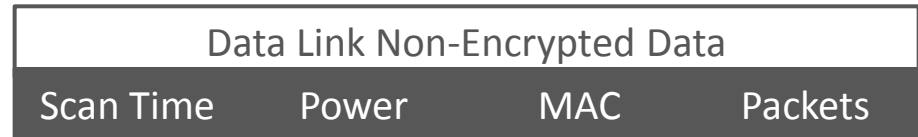
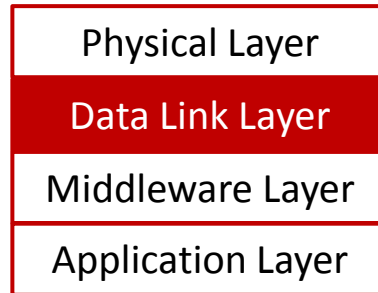


- Modem is setup in monitoring mode
- Performs cycles of scans for packets in the air
- Reads Data Link Layer of packets
- Extracts non-encrypted data

Wi-Fi Packet



Bluetooth Packet



# Architecture of a Wi-Fi/Bluetooth Scanning System

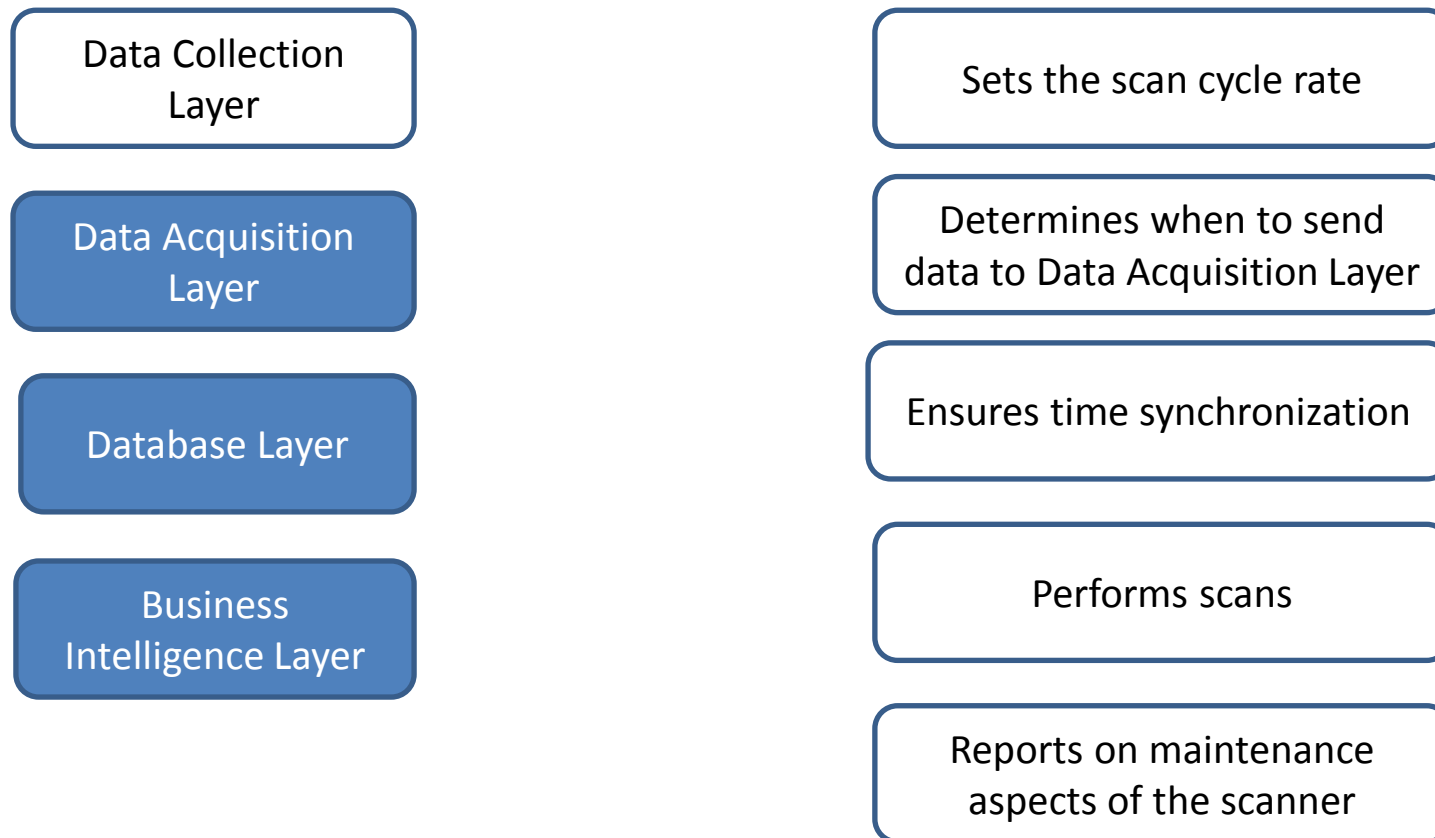
Data Collection  
Layer

Data Acquisition  
Layer

Database Layer

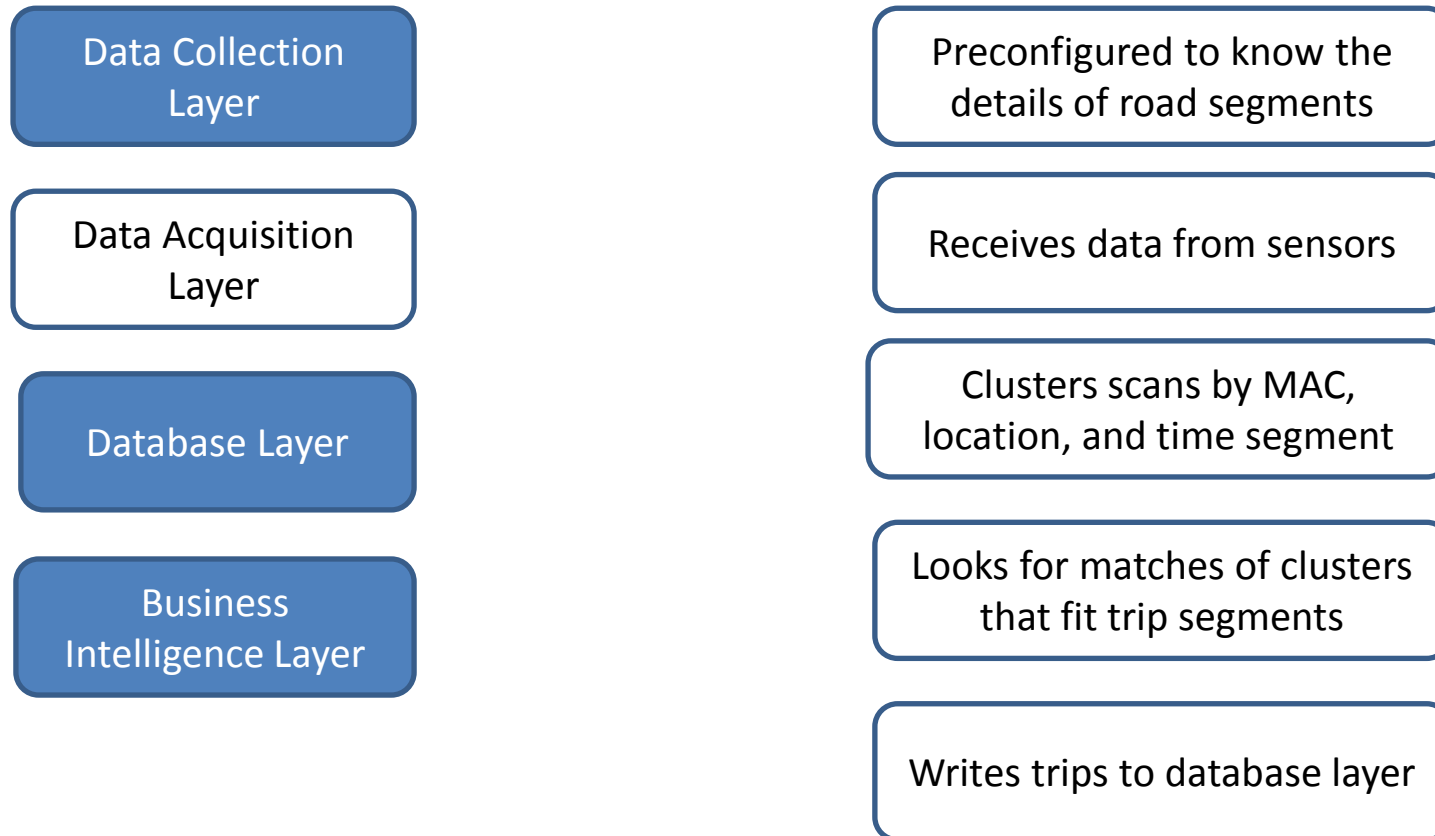
Business  
Intelligence Layer

# Architecture of a Wi-Fi/Bluetooth Scanning System

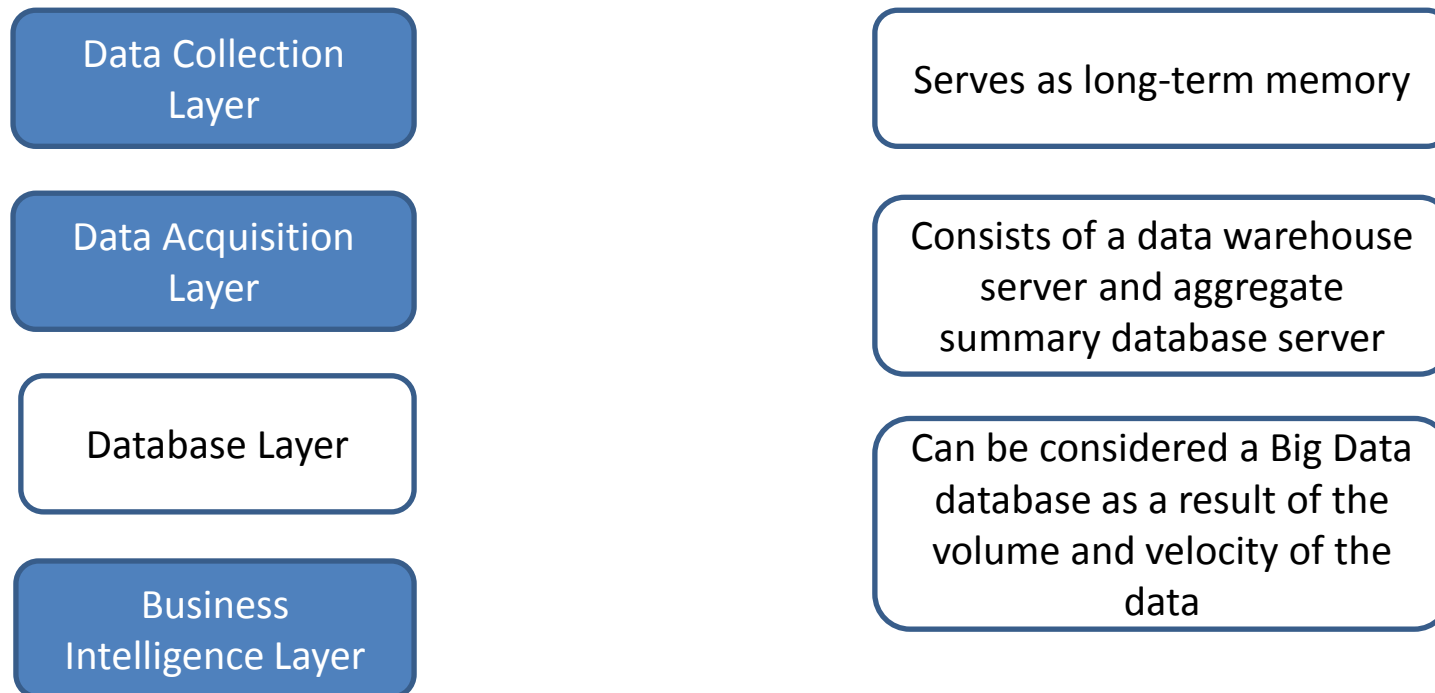




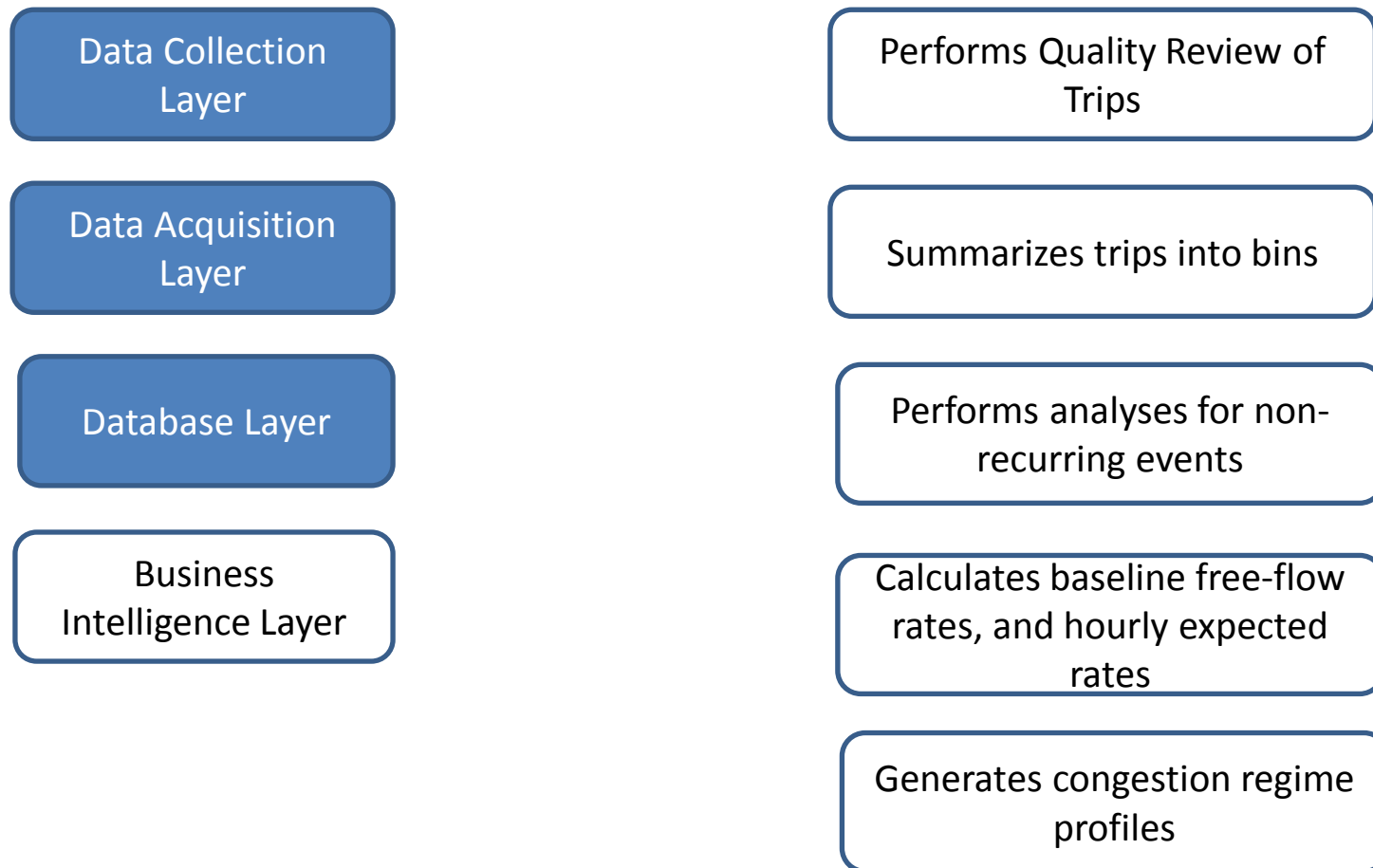
# Architecture of a Wi-Fi/Bluetooth Scanning System



# Architecture of a Wi-Fi/Bluetooth Scanning System



# Architecture of a Wi-Fi/Bluetooth Scanning System



# Wi-Fi/Bluetooth Scanner Testing

- **Unit Testing**
  - Device state
  - Antenna selection
  - Scan cycle
  - Scan selection

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing**
  - Outlier treatment
  - Data aggregation and reliability
  - System performance measures



# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment**
  - Sort by various modes
  - Distinguish between *Travel Times* and *Trip Times*
  - Treat extreme issues of scan radius

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment**
- **Treatment Methods**
  - Extreme limits
  - Hourly groups based on current conditions
  - Group of 10

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Extreme Limits**

- Eliminate based on an extreme max and min travel time limit

$$tt_{\max} = 1/3600 * d$$

$$tt_{\min} = 2 * PS / 3600 * d$$

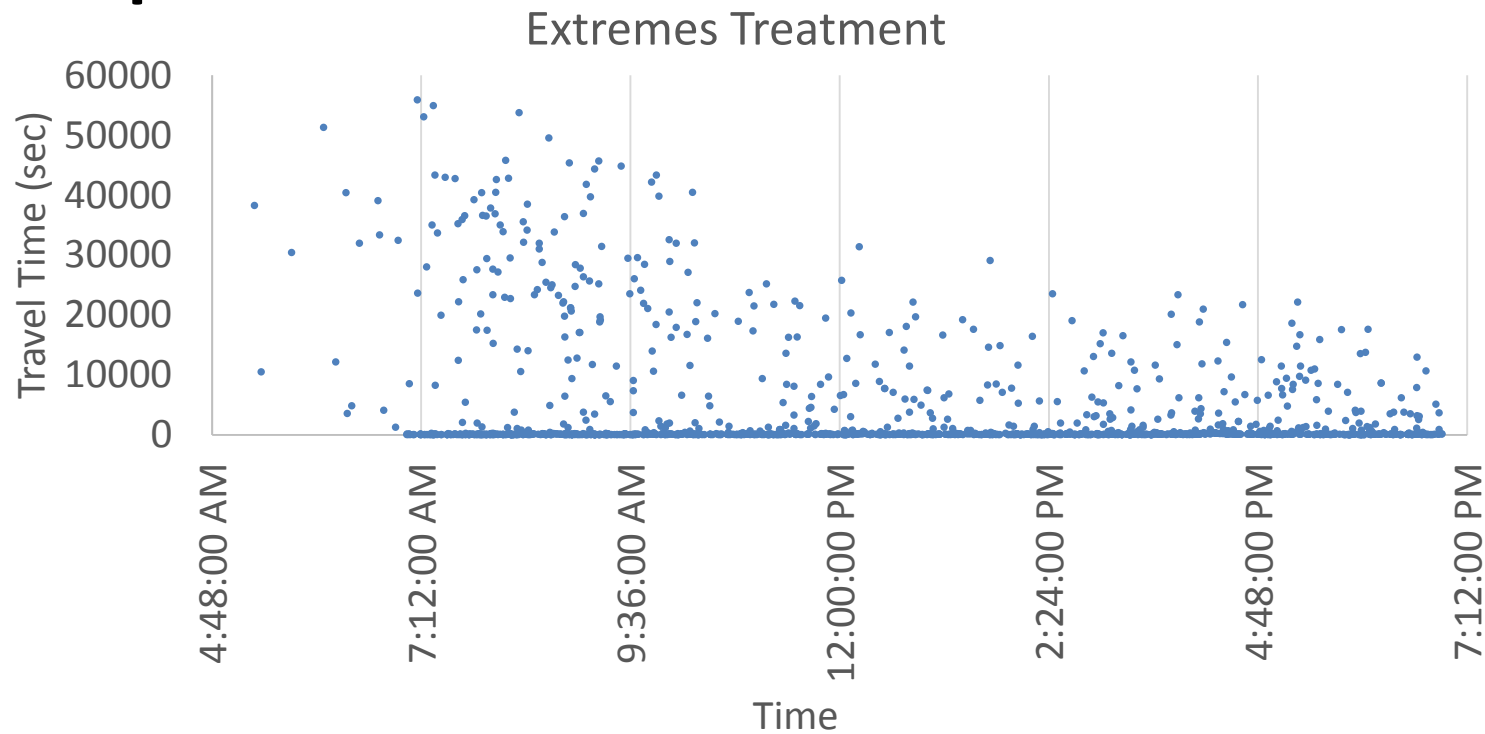
tt = Travel Time

d = Segment Length

PS = Posted Speed

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Hourly Groups**



# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Hourly Groups**

- Hourly processed based on current conditions

Current Time 1:00pm

11:00am	11:59am	12:00pm	12:59pm
$t_{1-123}$		$t_{124-234}$	
Median tt	110 sec	Median tt	116 sec

$$tt_{\max} = 5 * tt_{\text{PH}} / d$$

tt = Travel Time

$tt_{\text{PH}}$  = Travel Time Previous Hour

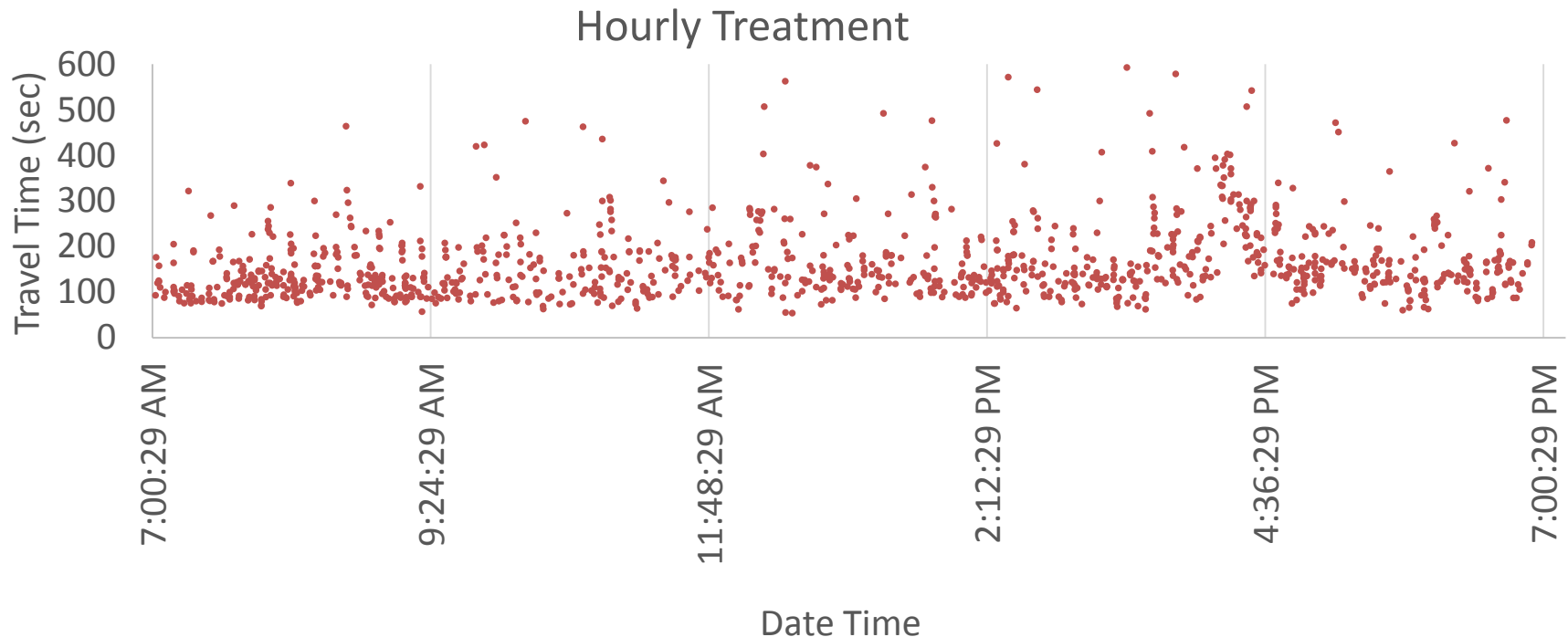
t = Trip Records

d = Segment Length



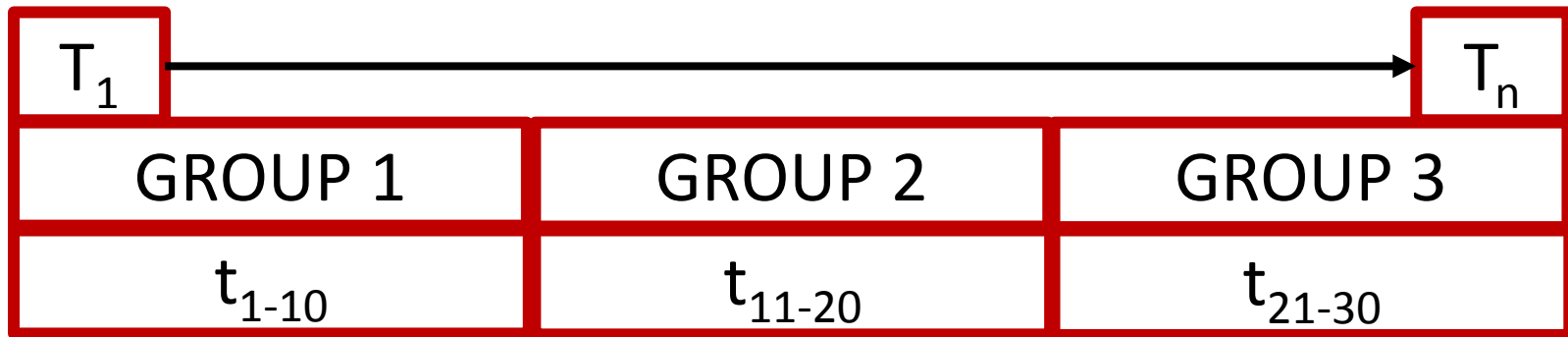
# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Hourly Groups**



# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Group of 10**



T = time  
t = trip record

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Group of 10**

GROUP 1
$t_{1-10}$

110	144	501	121	149	207	511	116	129	501
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Group of 10**

Define Allowable Difference

$$AD = SI * 60$$

Find min of  $t_{4-7}$

Define Initial Upper Bound

$$IUB = (\text{local}_{\min} + AD)$$

Define Initial Lower Bound

$$ILB = (\text{local}_{\min} - AD)$$

Are any trips within 10% IUB

If Yes

$$UB = IUB * 1.2$$

If No

$$UB = IUB * .9$$

110 144 501 121 149 207 511 116 129 501

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment>Group of 10**

Define Allowable Difference

$$AD = SI * 60$$

Find min of  $t_{4-7}$

Define Initial Upper Bound

$$IUB = (\text{local}_{\min} + AD)$$

Define Initial Lower Bound

$$ILB = (\text{local}_{\min} - AD)$$

Are any trips within 10% ILB

If Yes

$$UB = ILB * .9$$

If No

$$UB = ILB * 1.2$$

110 144

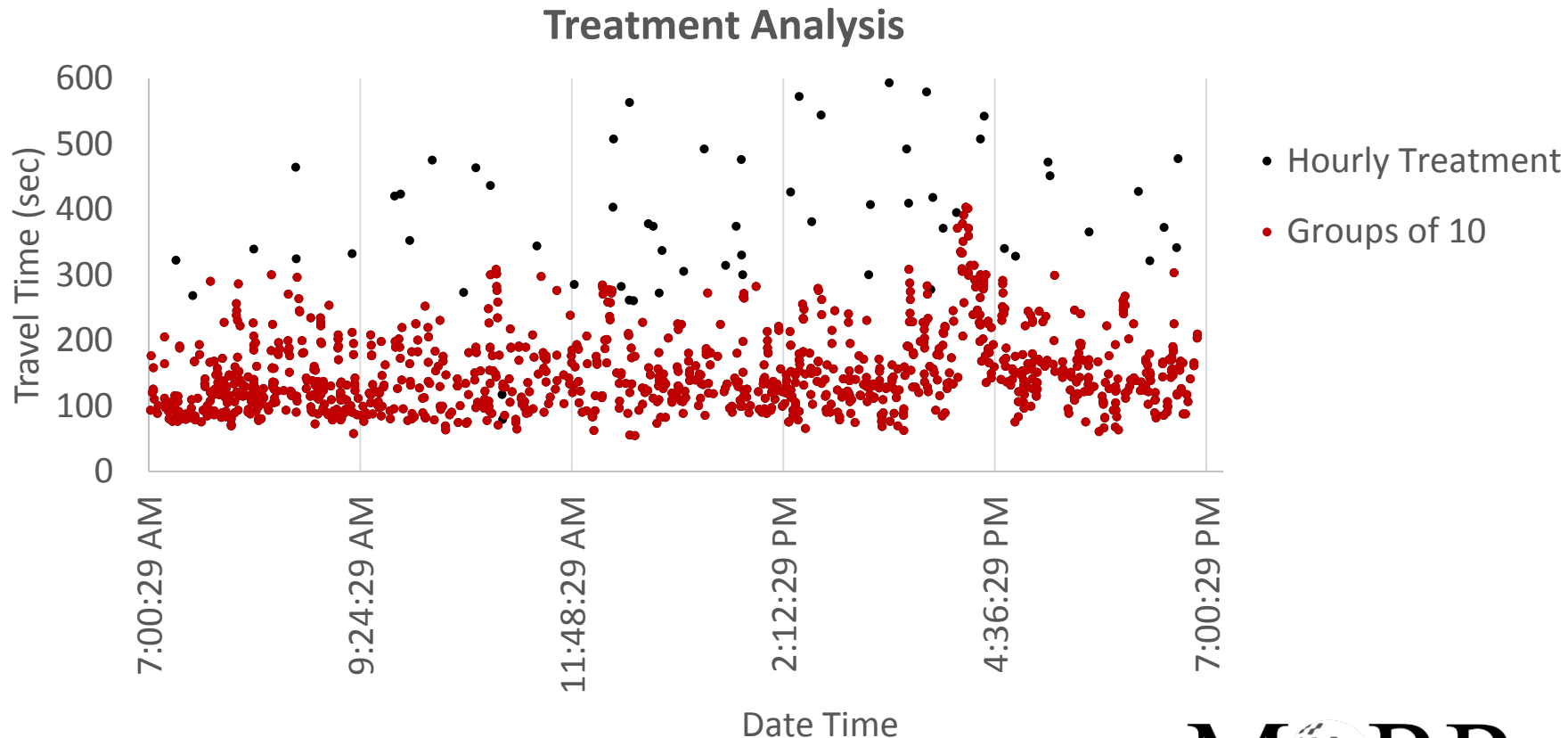
121 149 207

116 129



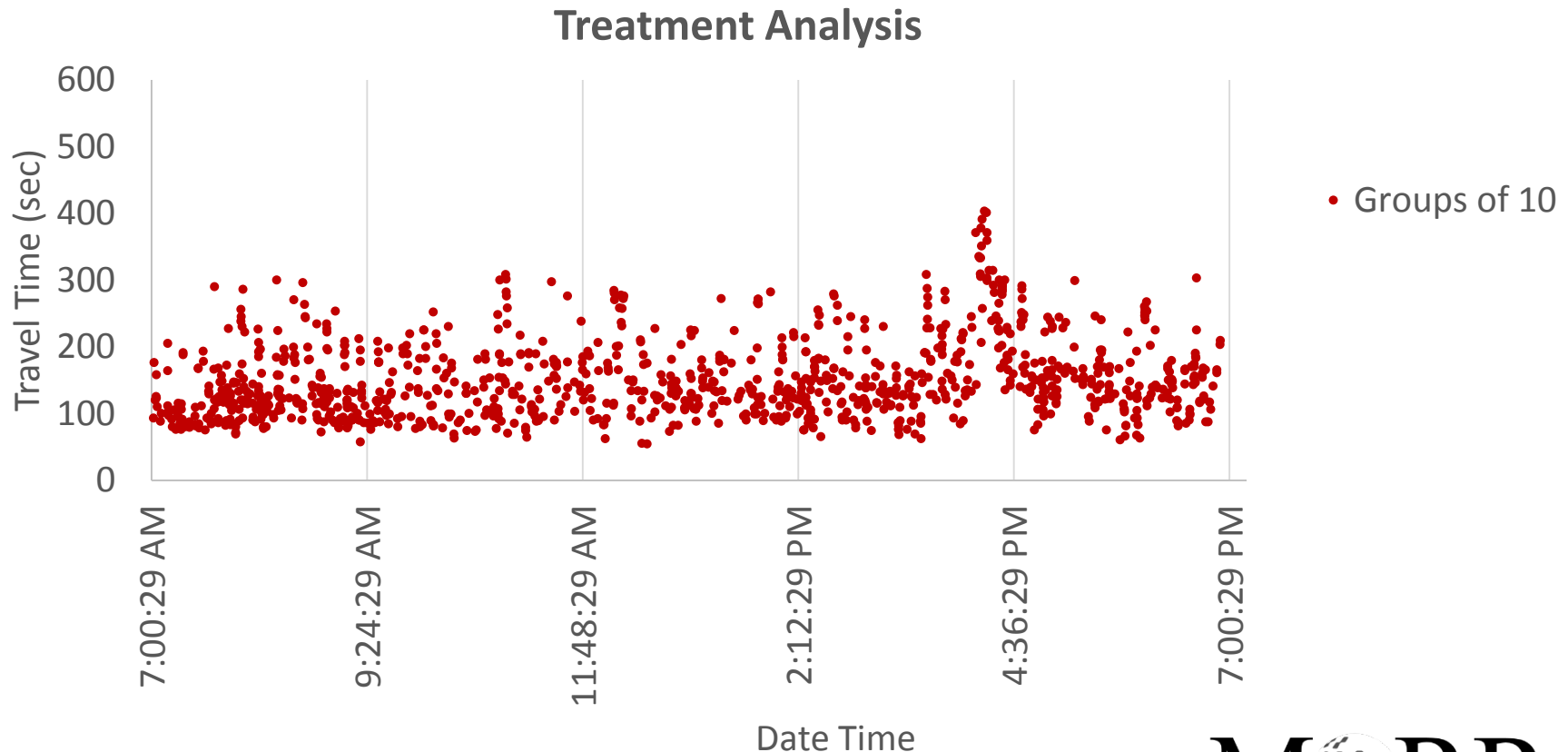
# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Outlier Treatment**



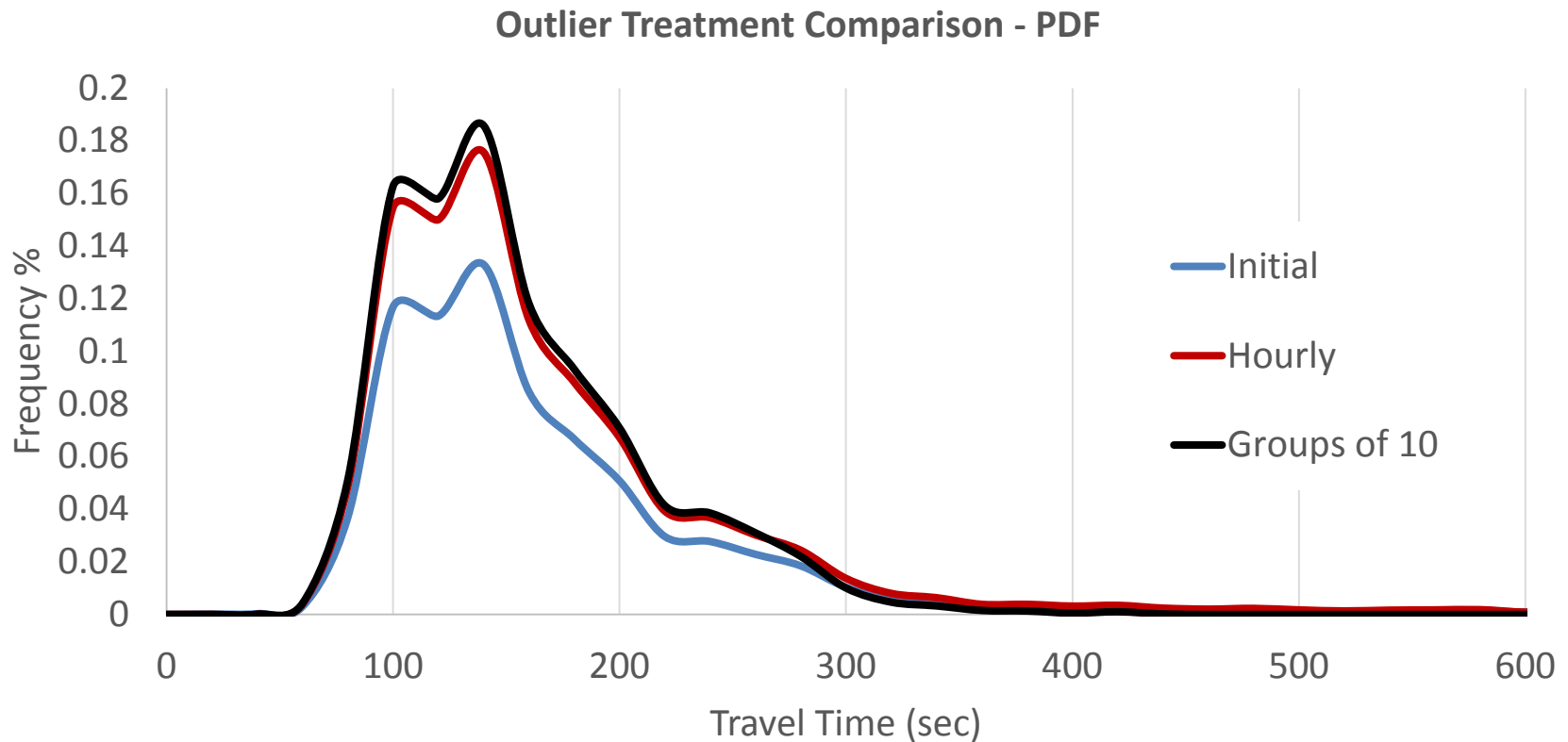
# Wi-Fi/Bluetooth Scanner Testing

- **System Testing>Outlier Treatment**



# Wi-Fi/Bluetooth Scanner Testing

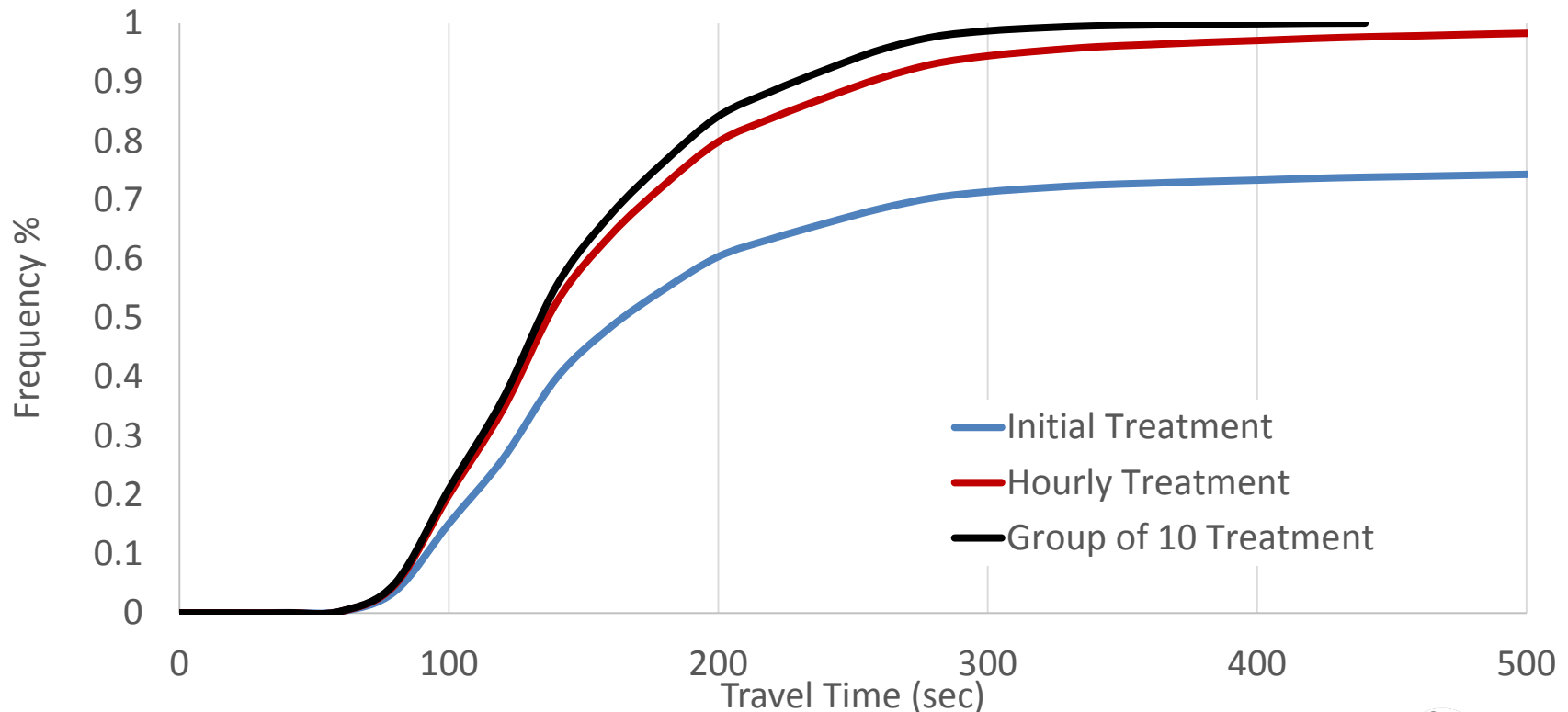
- **System Testing > Outlier Treatment**



# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Outlier Treatment**

Outlier Treatment Comparison - CDF



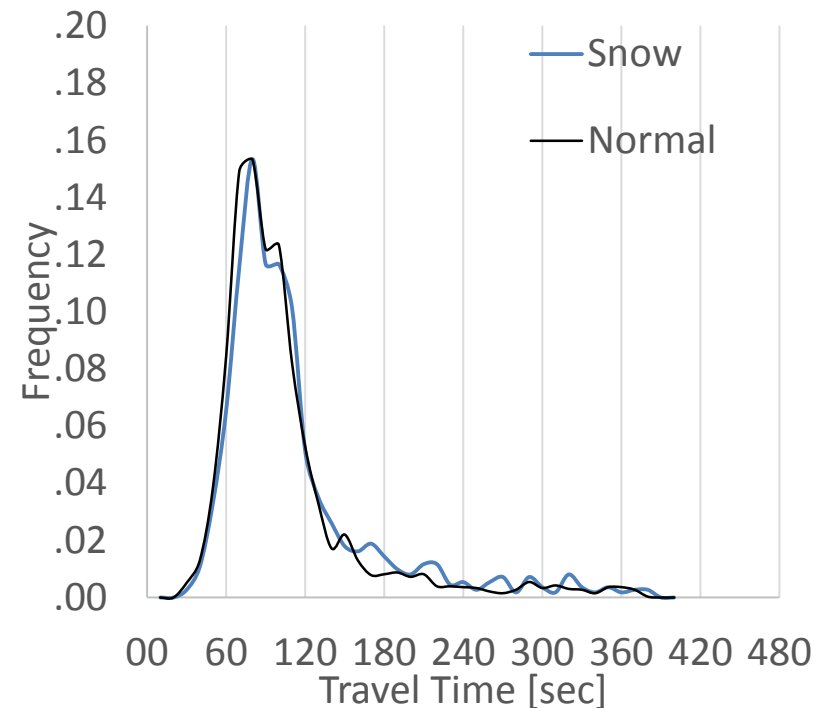
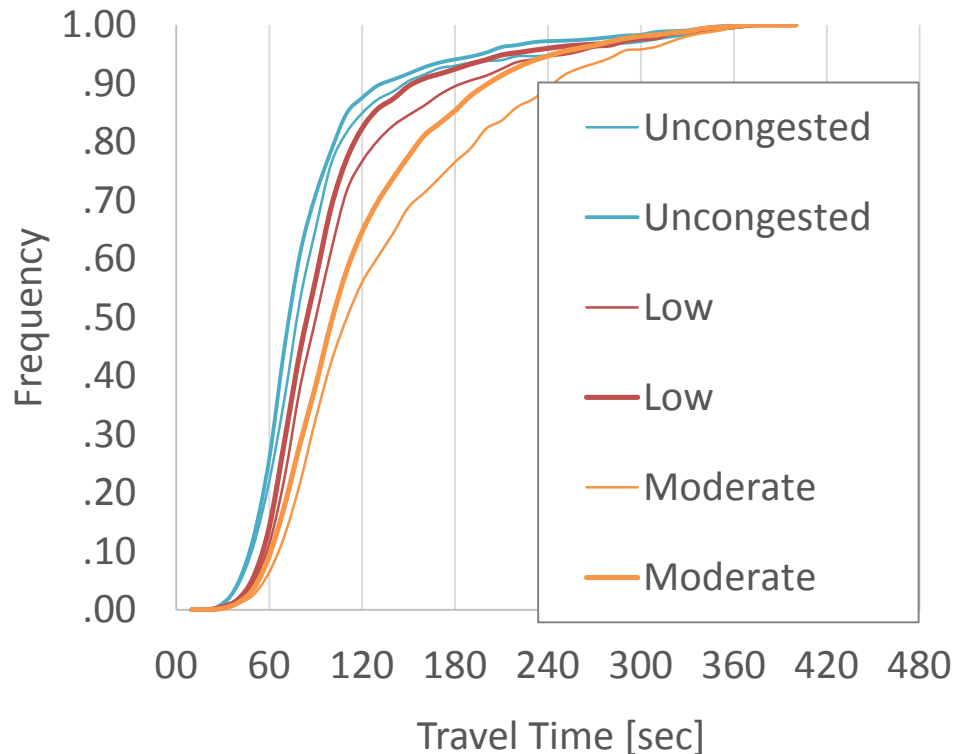
# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Data Aggregation and Reliability**
  - What data is available?
  - What is the reliability of the available data?

# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Data Aggregation and Reliability**

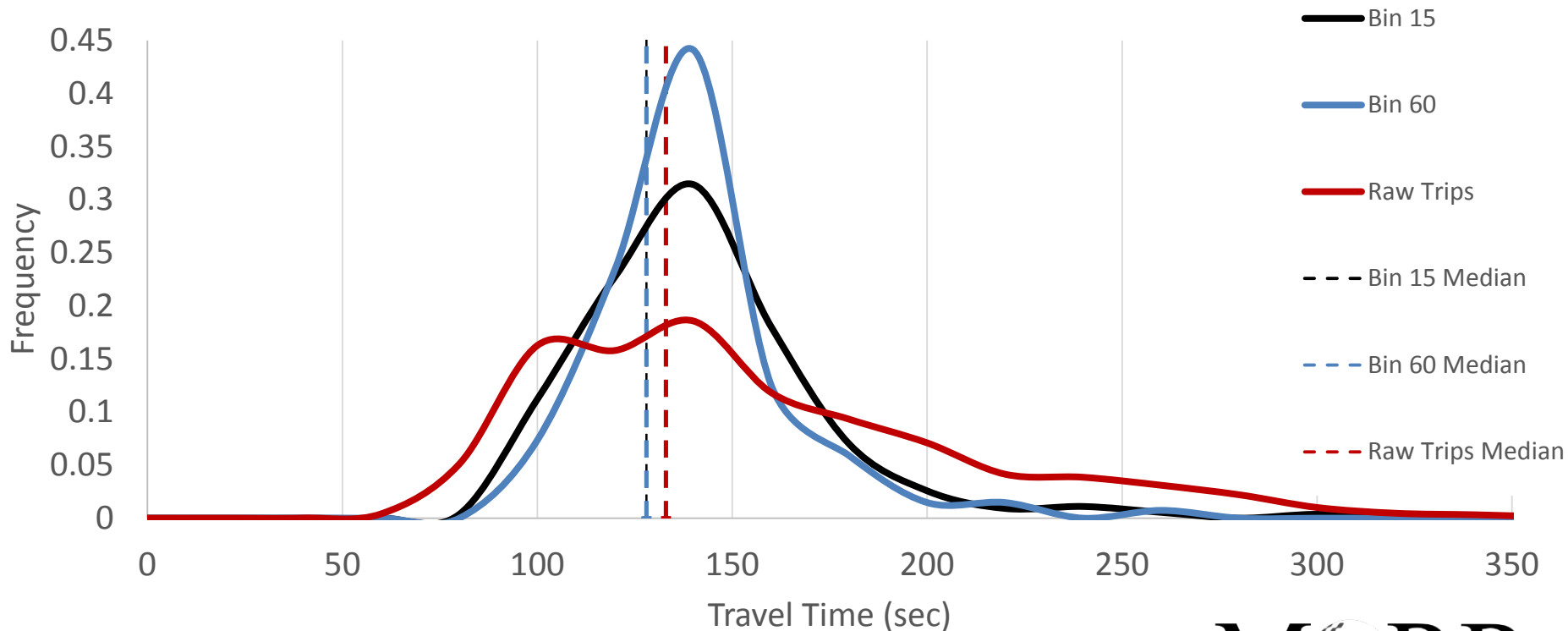
## Regime Analysis



# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Data Aggregation and Reliability**

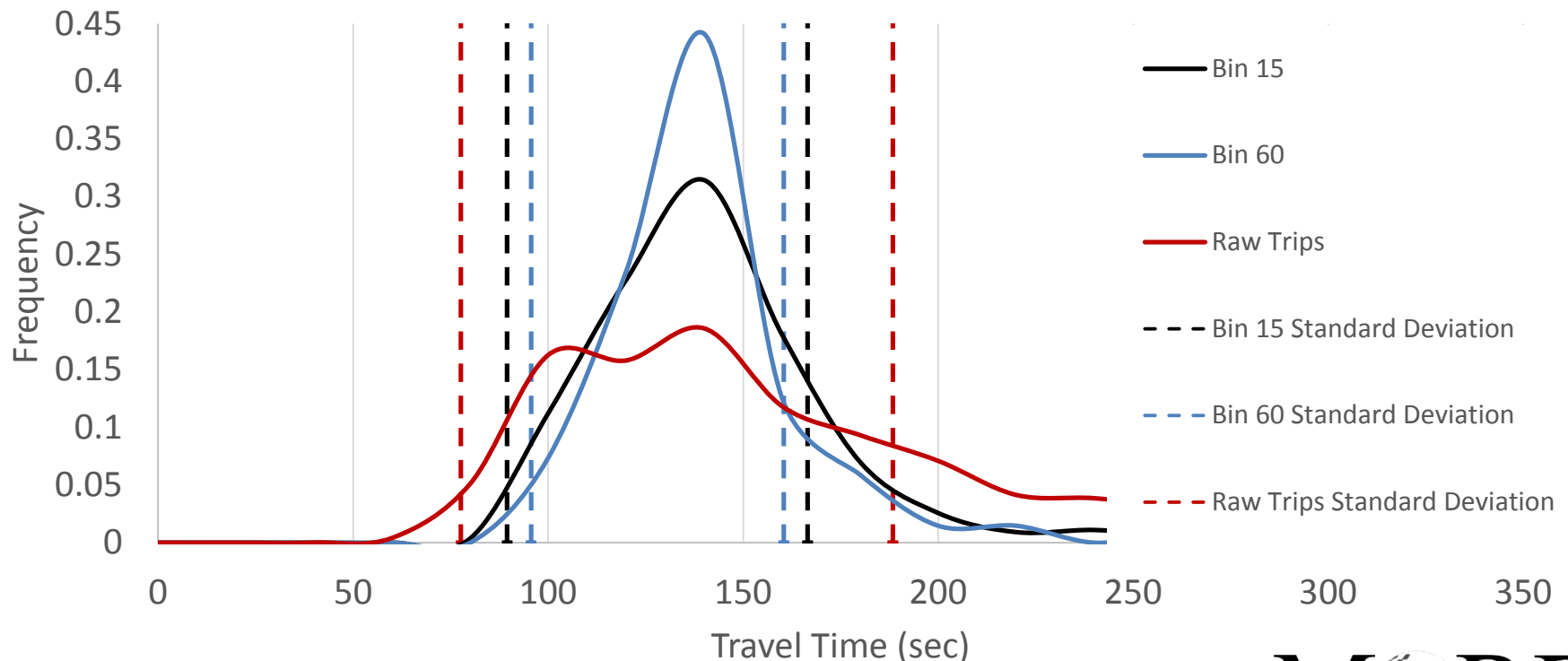
Data Reliability Comparison PDF & Median



# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Data Aggregation and Reliability**

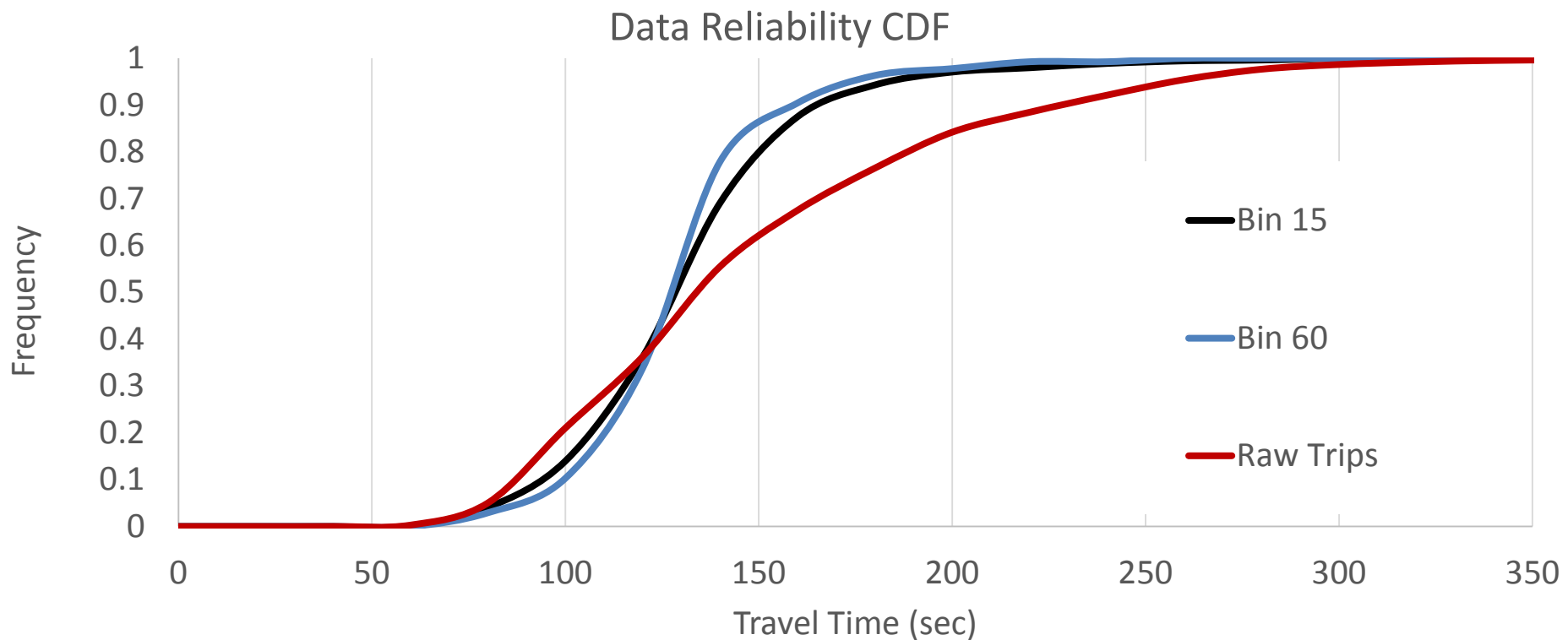
Data Reliability Comparison PDF & Standard Deviation





# Wi-Fi/Bluetooth Scanner Testing

- **System Testing > Data Aggregation and Reliability**



# Evaluating Wi-Fi/Bluetooth Scanning Systems

## Unit Considerations

Consideration	Requirement	Details
Scan cycle rate	High scan to sleep ration	Will allow for a higher sampling rate
Antenna type	Omni-directional(OD) antenna	A directional antenna will provide greater localisation, but can result in a sub-par sample. Localisation of scans using OD antennas can be performed using data processing methods. Considerations need to be made for Db level of antenna
Portability	System can be deployed and redeployed with minimal onsite effort.	Wi-Fi/Bluetooth sensors seem better suited for temporary studies and can be cost prohibitive for city wide deployment.
Antenna mounting	Internal mounting is recommended	Antennas can be fragile, and, if the scanners are being moved frequently, can be susceptible to damage.

# Evaluating Wi-Fi/Bluetooth Scanning Systems

## Unit Considerations

Consideration	Requirement	Details
Unit monitoring	Unit needs to frequently report on their current state.	Will allow for a higher sampling rate
Time synchronization	Scanners need to set their clocks to a centralized clock on a regular basis	Units can experience time drift. Relying solely on internal clocks can cause travel times with negative values to occur.
Scan cycle	Scan to rest ratio should be no less than 3:1	Scan ratios less than 3:1 will result in a lesser matching rate.
Scan selection	Not using a localisation based on signal strength	Measurements using signal strength are subject to a range of variables that are not consistent spatially or temporally

# Evaluating Wi-Fi/Bluetooth Scanning Systems

## System Considerations

Consideration	Requirement	Details
Outlier treatment	Use of group of 10 treatment is necessary to different between travel times and trip times	Will provide data that is a better representation of the segment performance.
Data aggregation	Clean trips is a required	Advanced analysis and regime development requires the least aggregated form of the trip data.
Real-time availability	Real-time reporting should be based on an index of the current condition against the expected rate	Reporting of travel-time and speed in real-time doesn't allow for simple indication of which road segments are performing well and which aren't.

# Evaluating Wi-Fi/Bluetooth Scanning Systems

## System Considerations

Consideration	Requirement	Details
Usage of median or mean	Median is required	Median is the better representation of central tendency, as the travel times are not a normal distribution
Performance measures	Focus on match rate, confidence intervals, and not sampling rate	A match rate of 4% of the segments AADT is suitable for a reliable representation. And, with the correct scan selection and outlier treatment methods, confidence intervals should be over 95% during non-free flow periods.

# Evaluating Wi-Fi/Bluetooth Scanning Systems



NATMEC 016  
Probe Data Collection and Accuracy