Identification and Analysis of Composite Travel Time Distributions in a Traffic Stream Utilizing Probe Vehicle Data

Angshuman Guin, PhD
Michael Hunter, PhD
Michael Rodgers, PhD
James Anderson

Georgia Institute of Technology

Overview

Objective: Identify multiple distributions in travel time datasets

- **Purpose**
  - Identify heterogeneity in the dataset (different subgroups of traffic which may experience different service quality along the same roadway)
  - Improve accuracy of before-after analysis (impact of roadway improvements on different road user subgroups)
  - Separate lanes biases in dataset

Data Collection Sites (Full Extent)

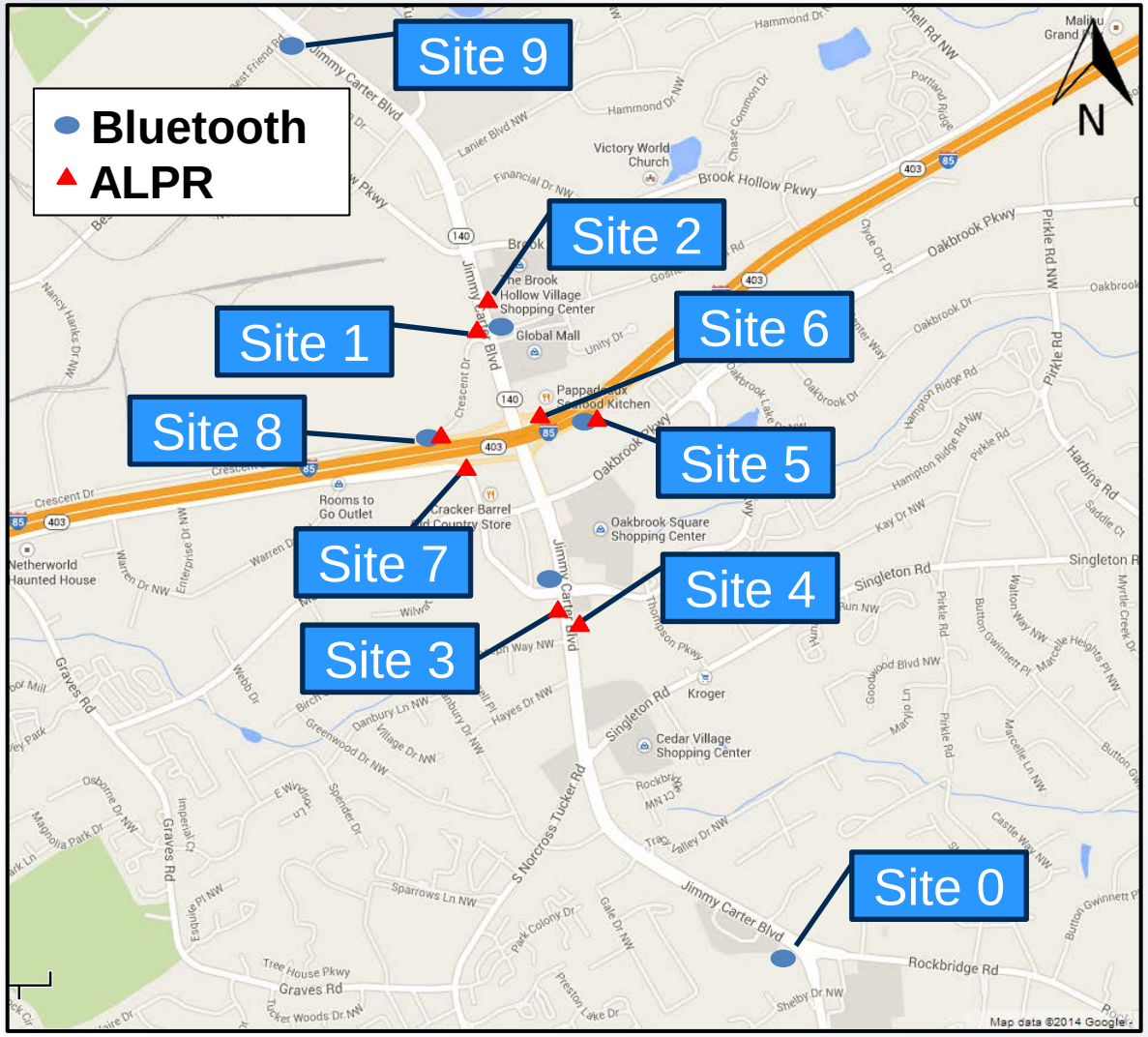
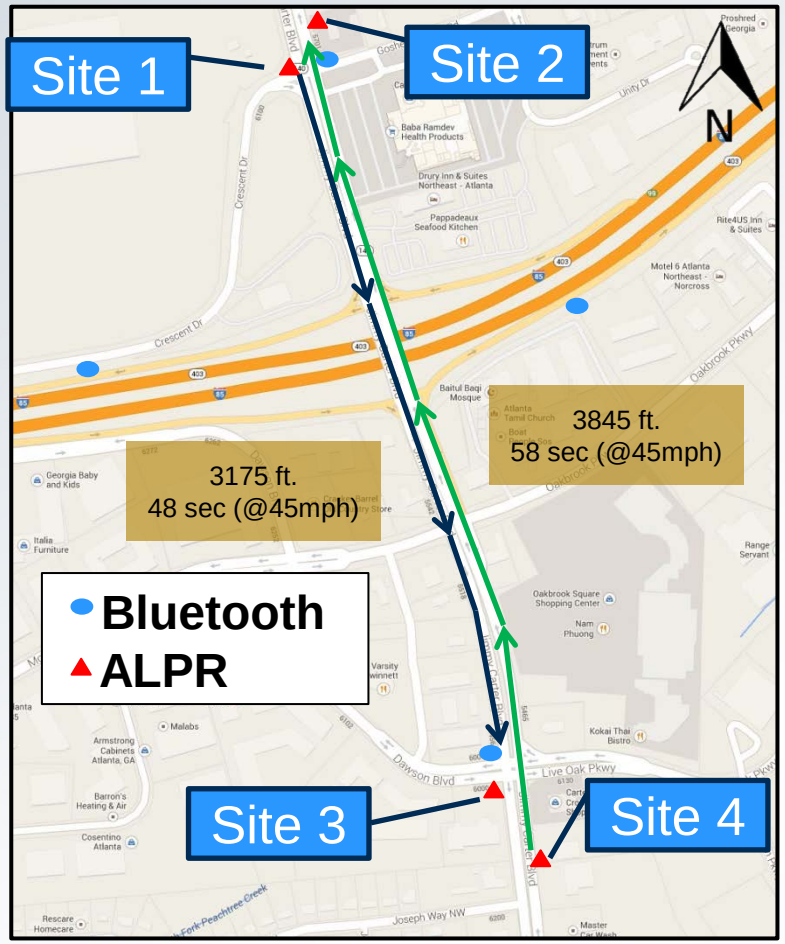


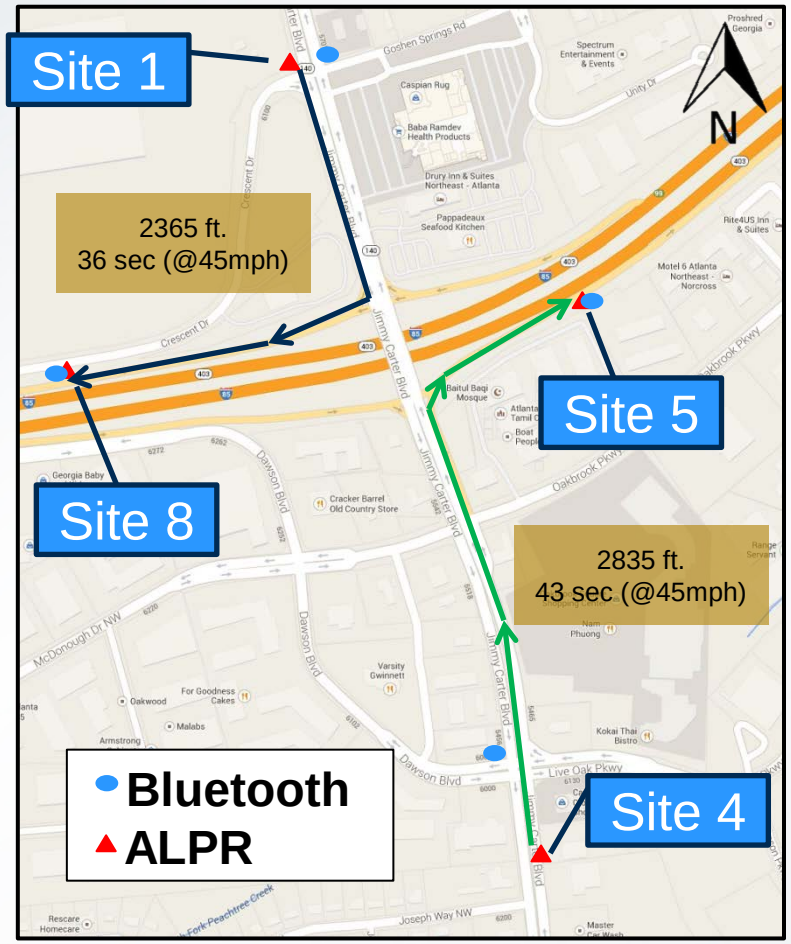
Image Source: Google Maps

Routes

4 to 2 and 1 to 3



4 to 5 and 1 to 8



4 to 8 and 1 to 5

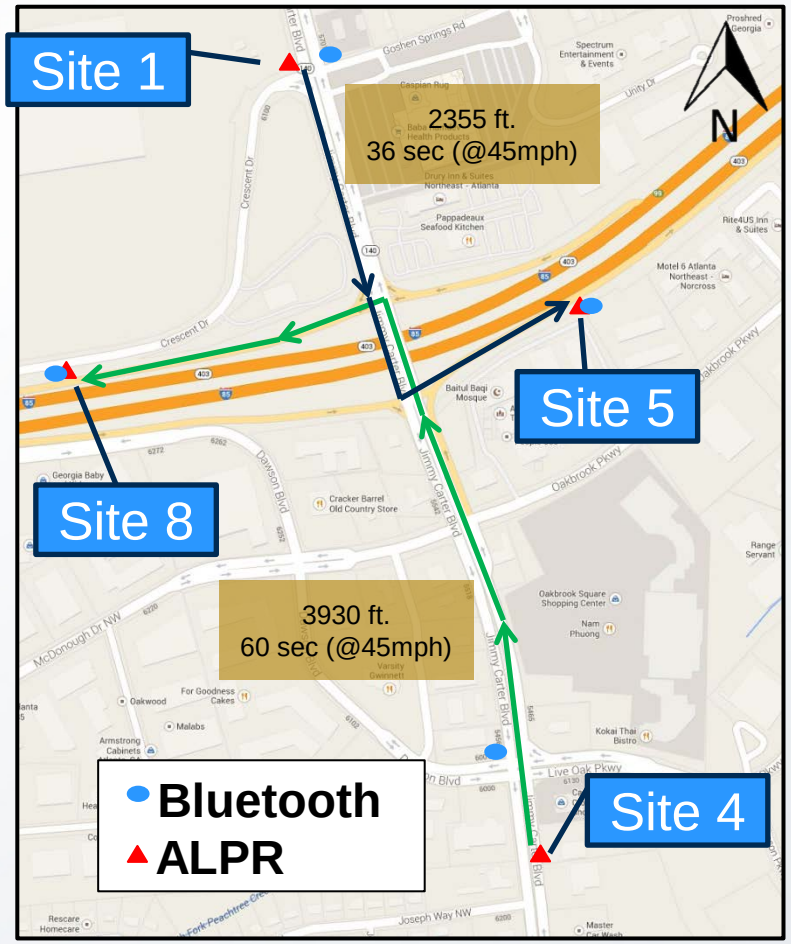
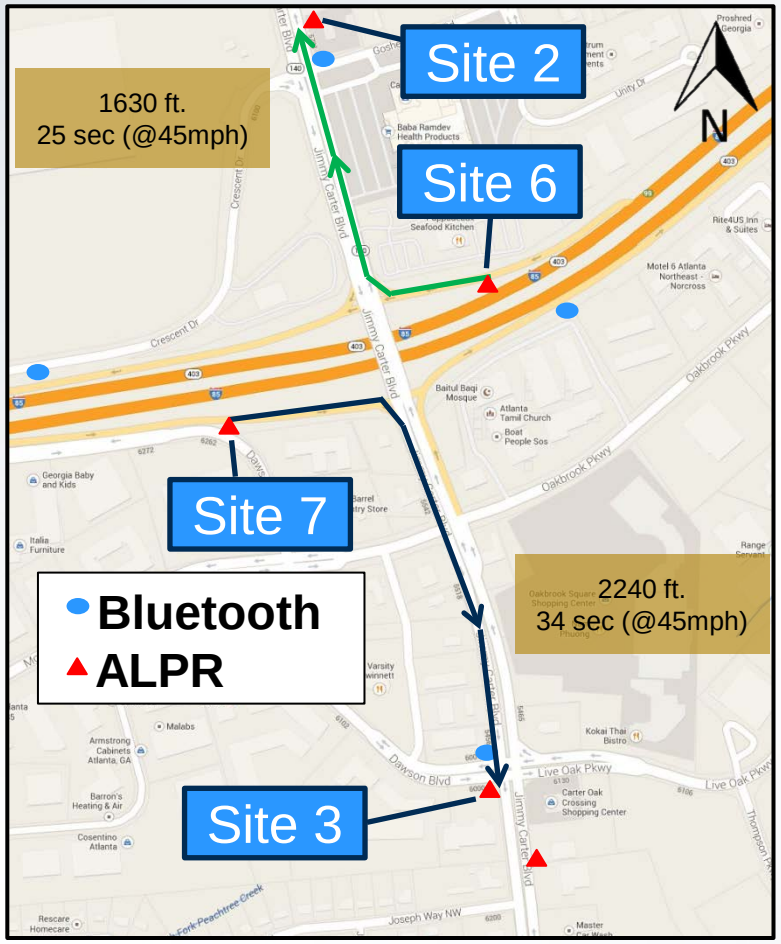


Image Source: Google Maps

Routes (Cont.)

6 to 2 and 7 to 3



6 to 3 and 7 to 2

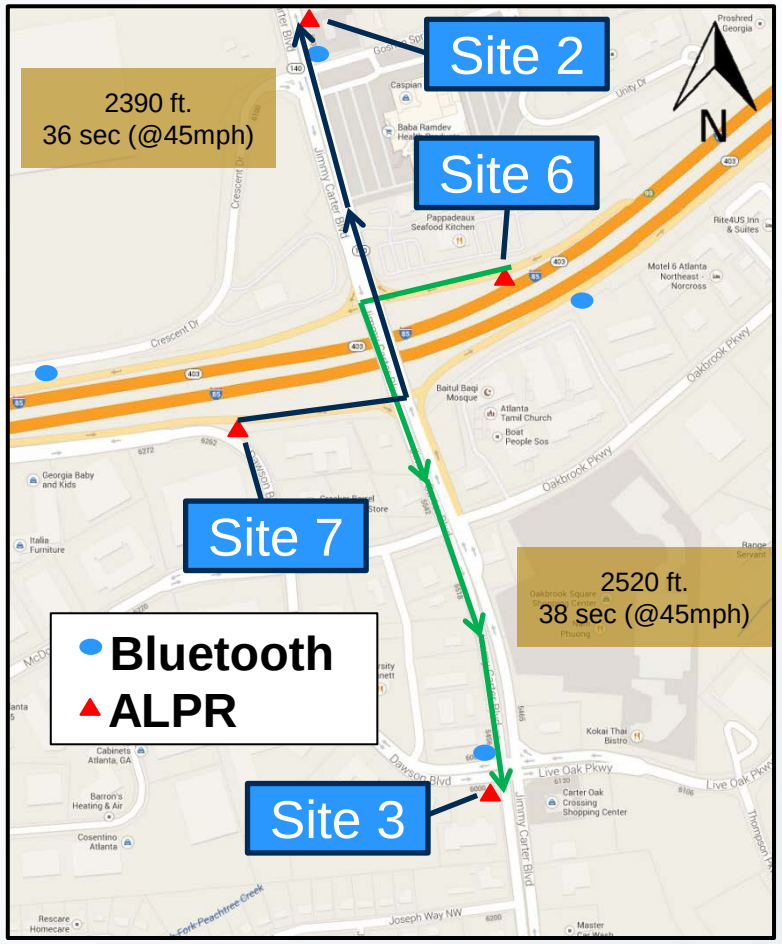
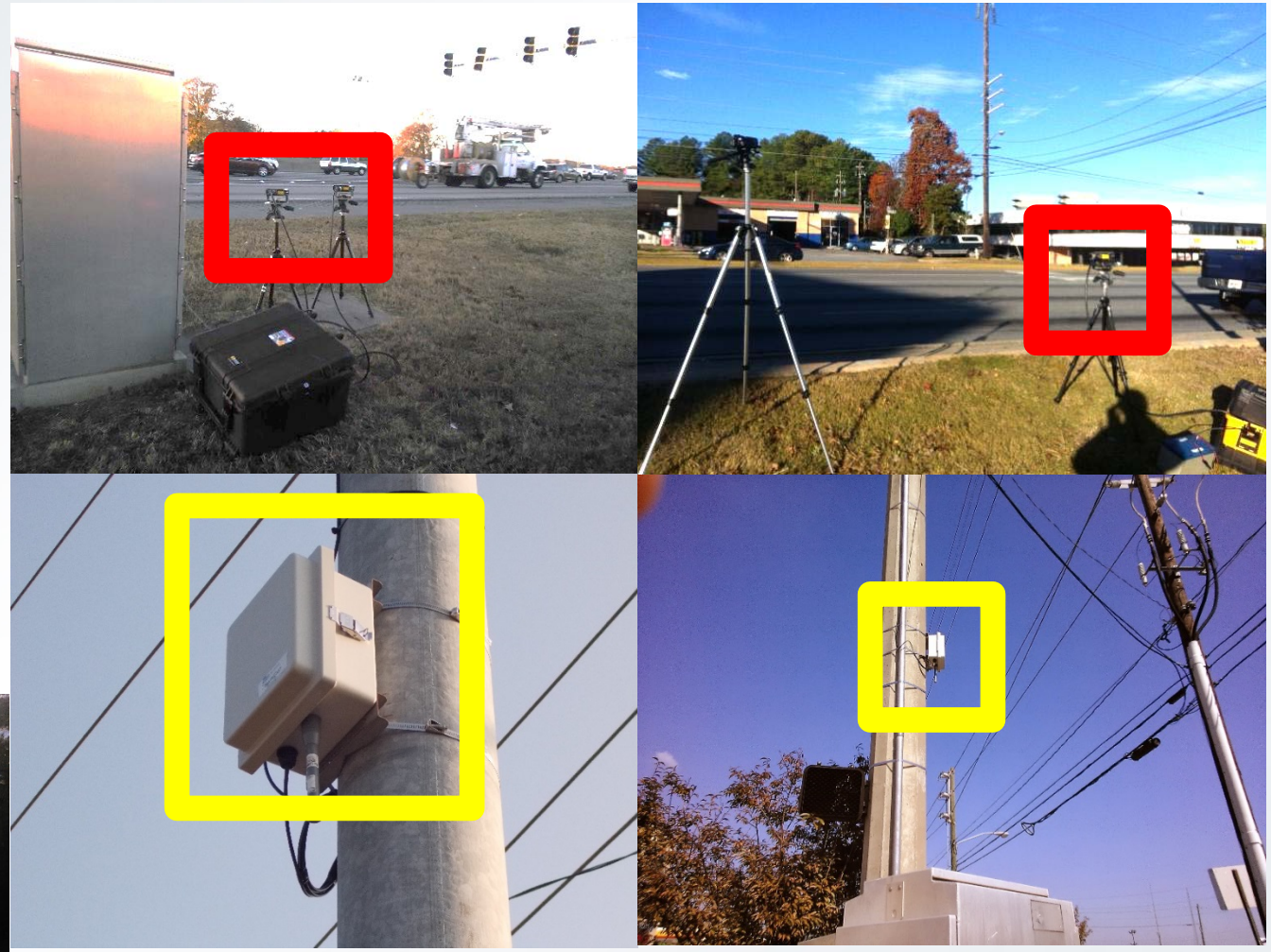


Image Source: Google Maps

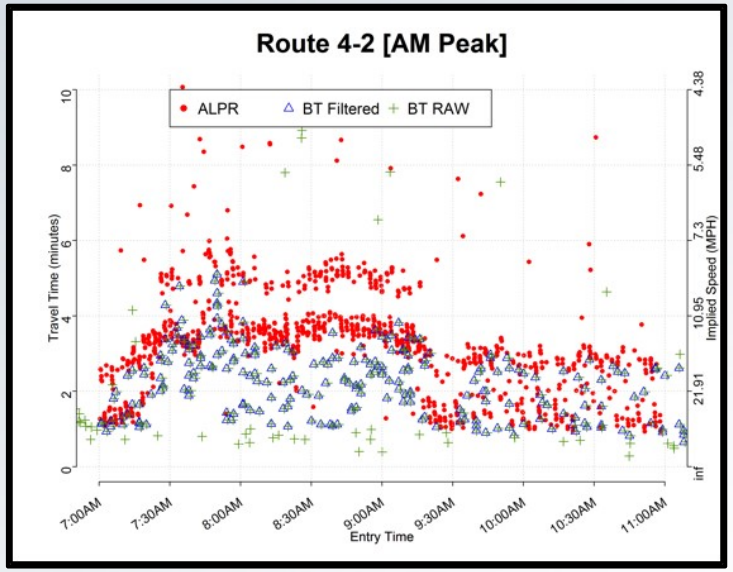
Supplementary Data Collection (ALPR)

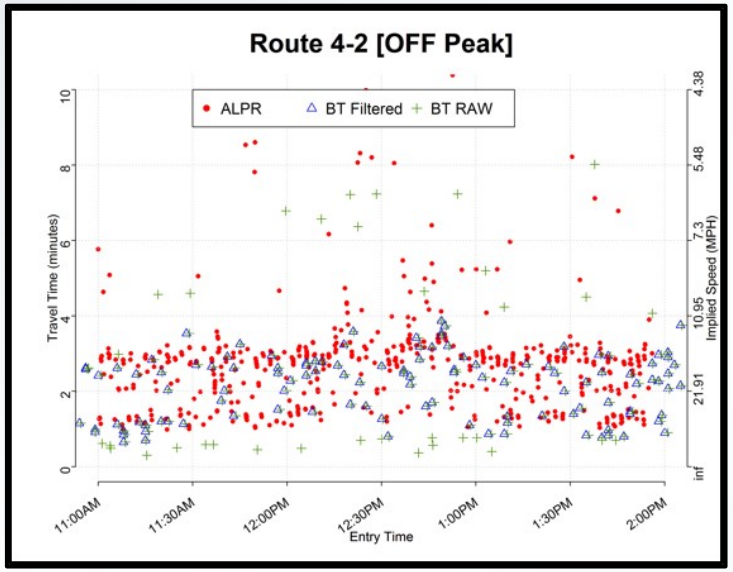
- Automated License Plate Recognition (ALPR) Cameras & HD Video Cameras
- Deployed at 8 locations
- Collected data from 7AM to 6PM
- November 21, 2014 (Thursday)

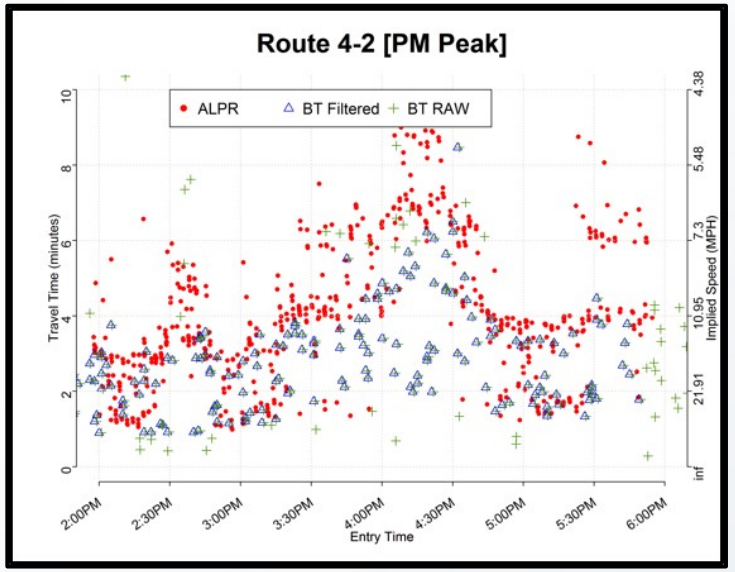


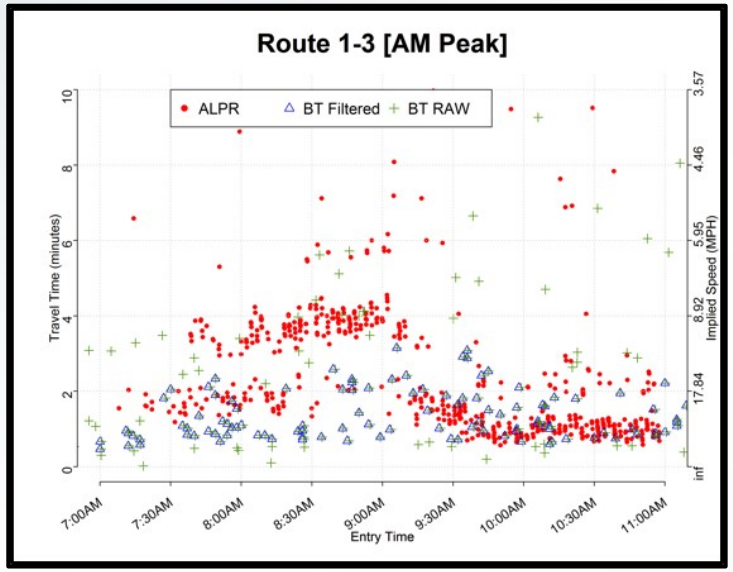
Images Courtesy: James Anderson and Edward Hightower

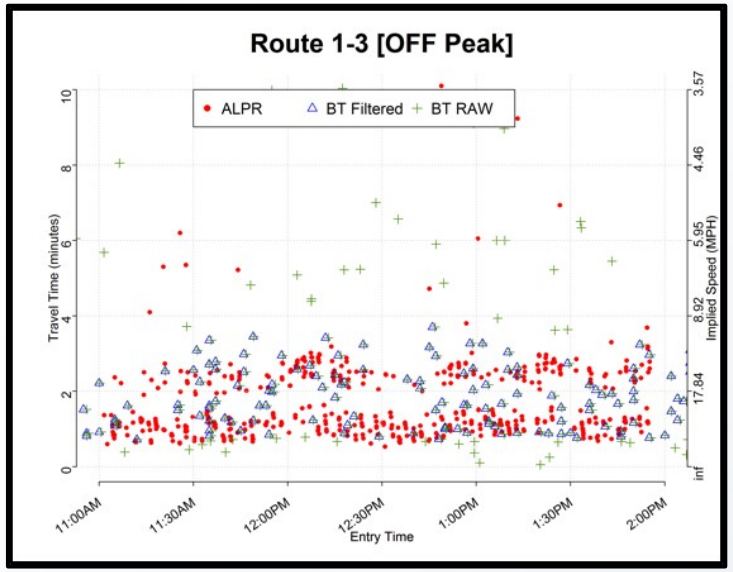
Travel-times on Through Movement Routes

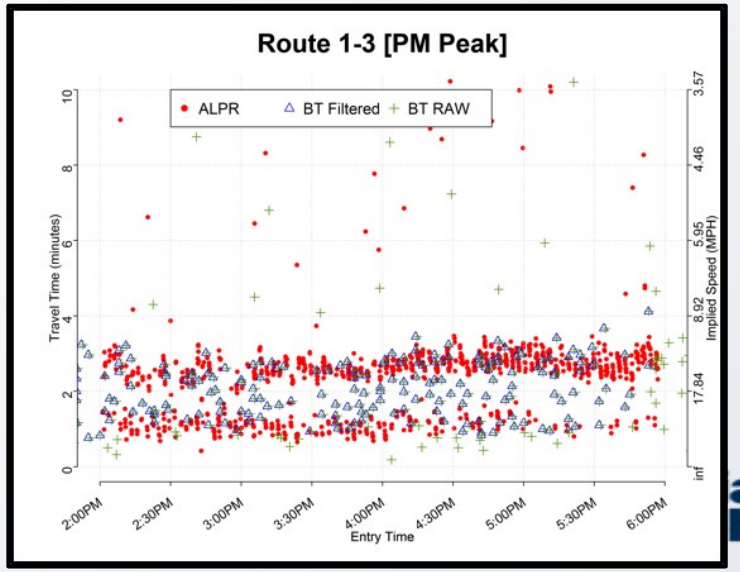




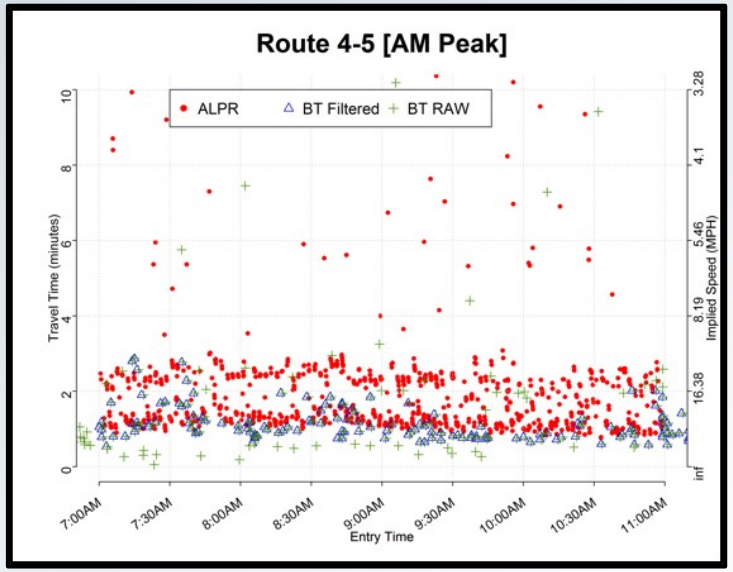


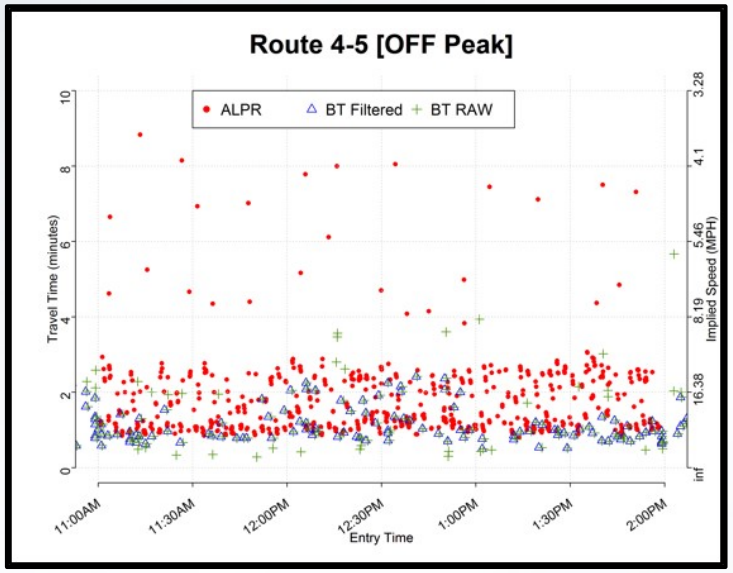


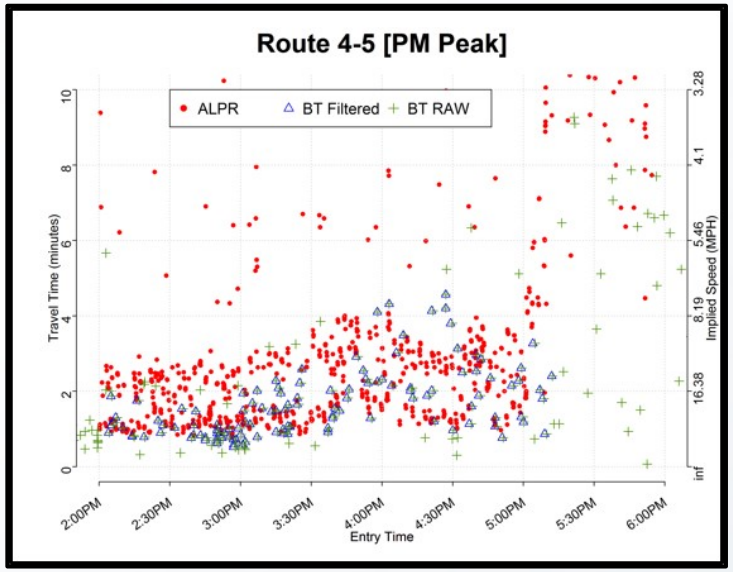


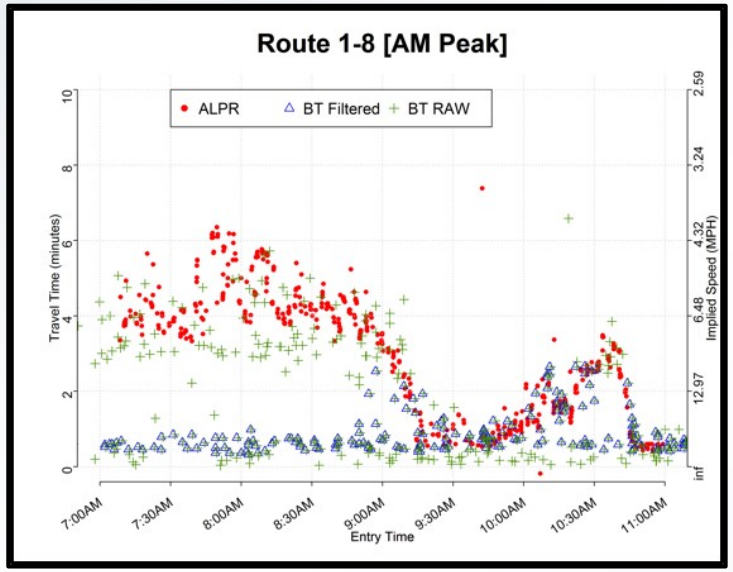


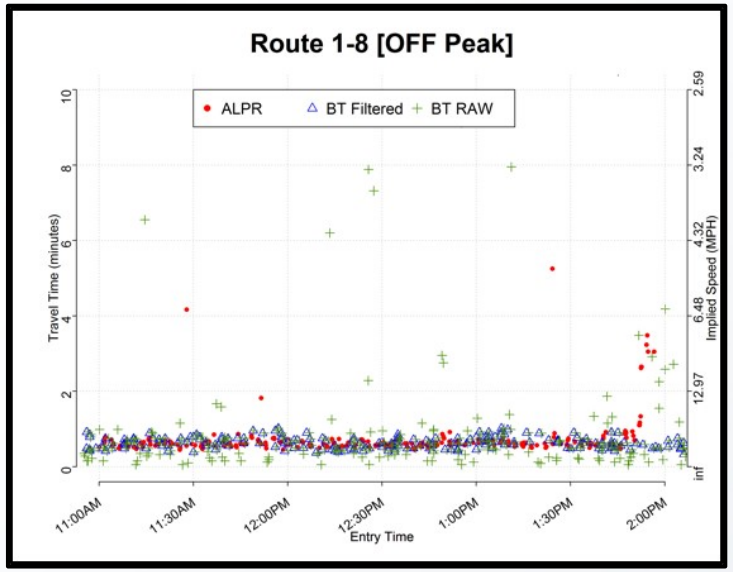
Travel-times on Right Turn Movement Routes

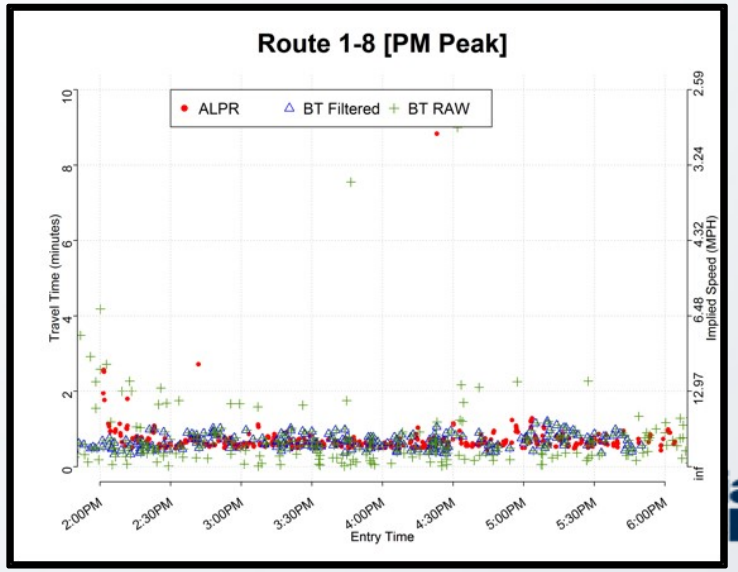




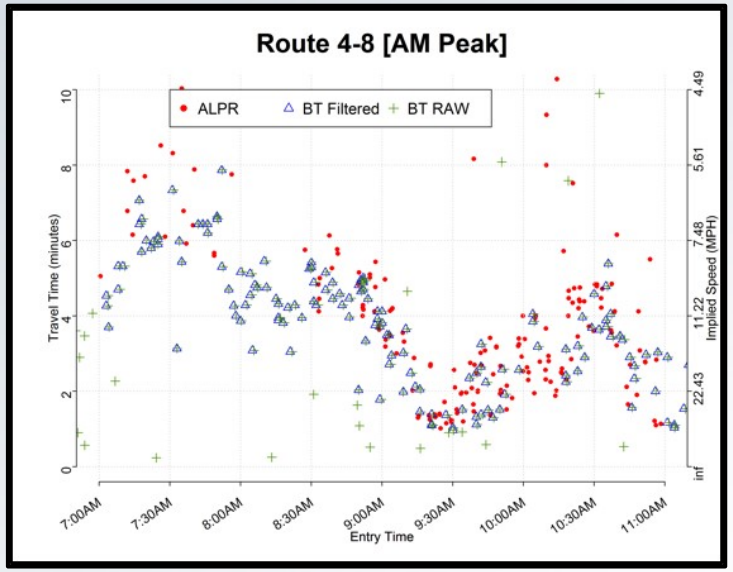


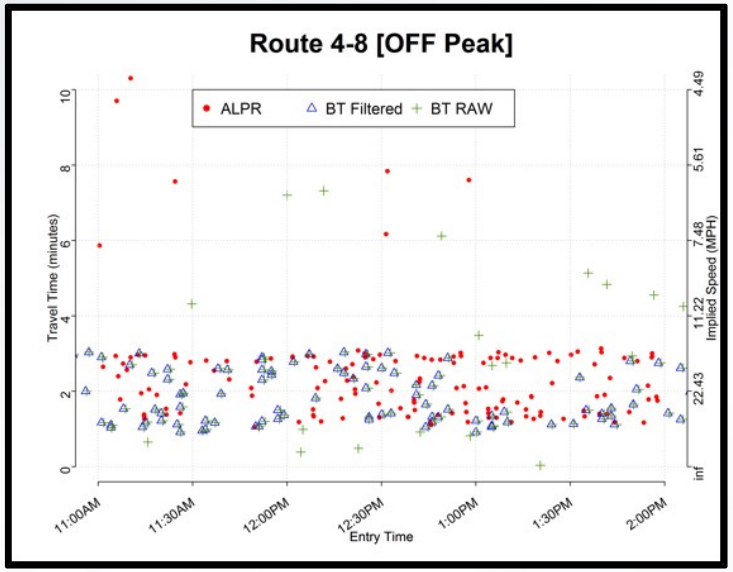


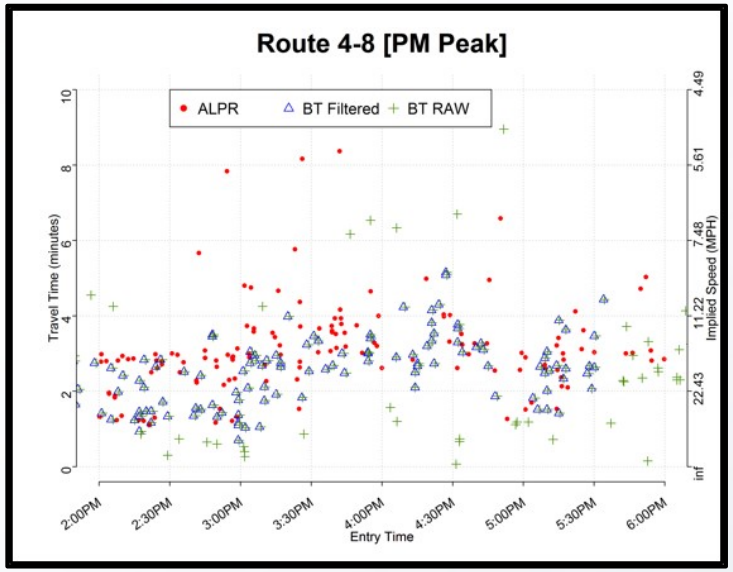


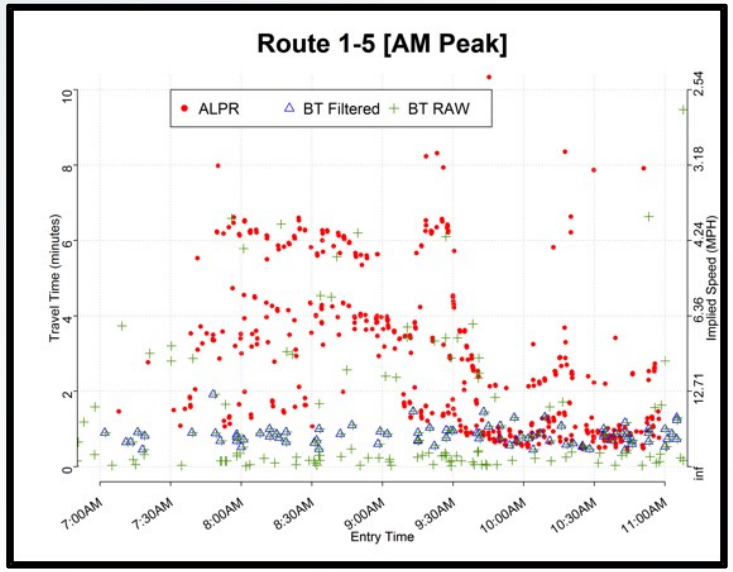


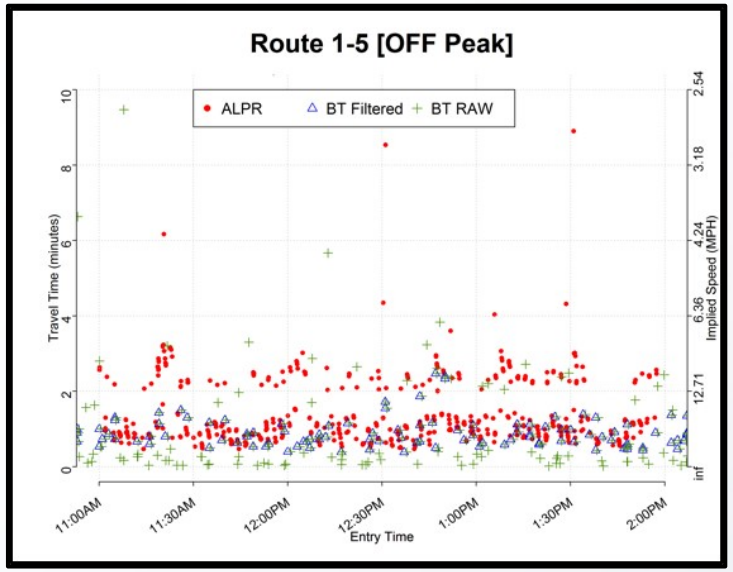
Travel-times on Left Turn Movement (to On Ramp) Routes

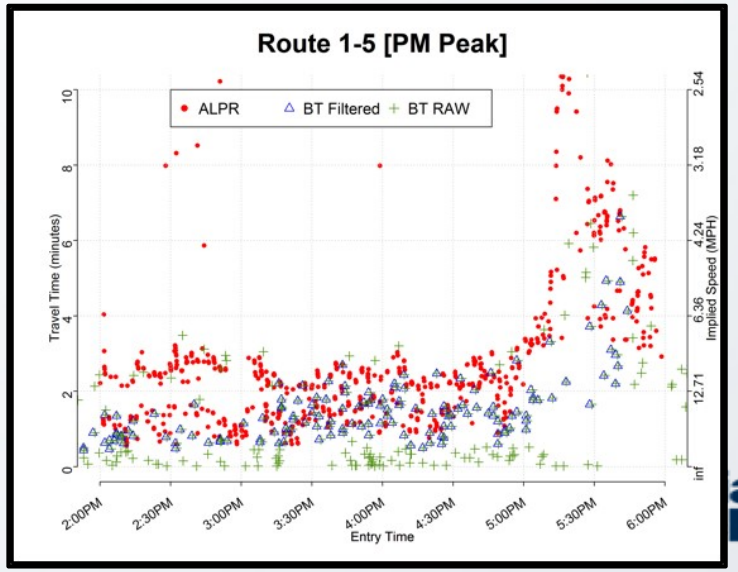




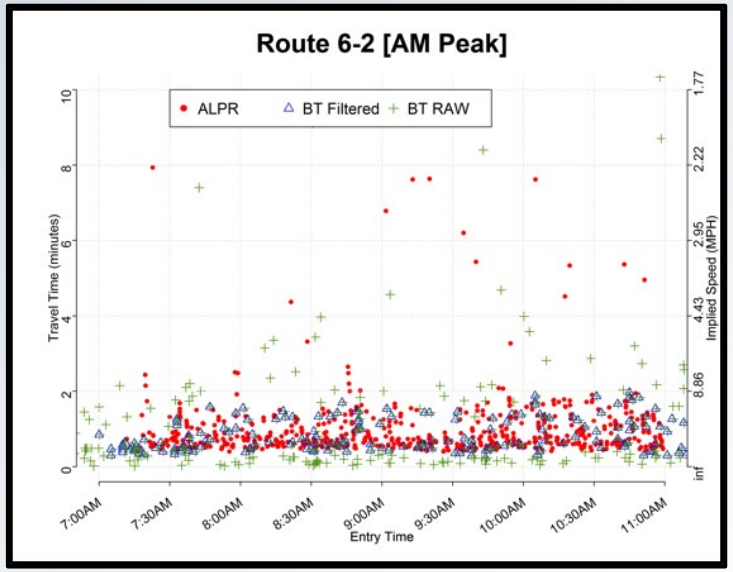


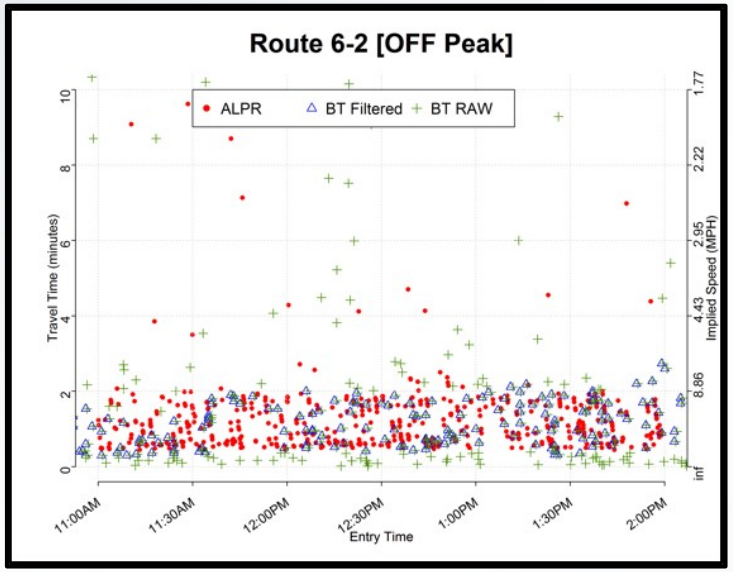


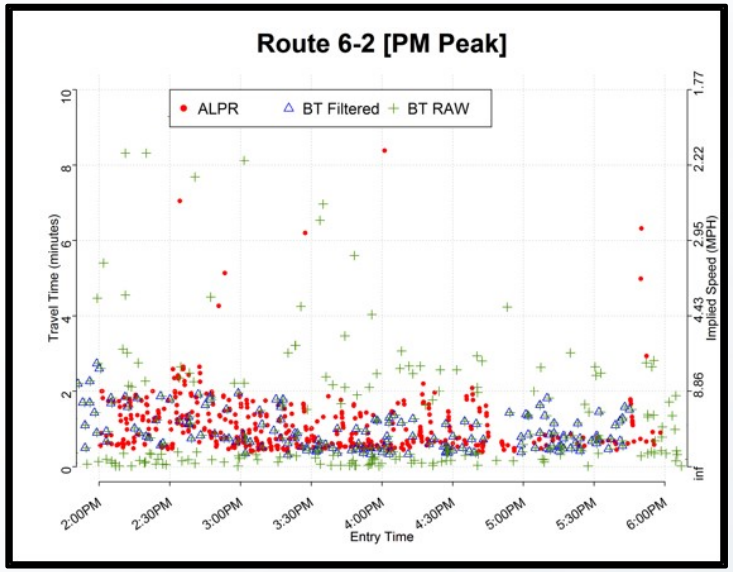


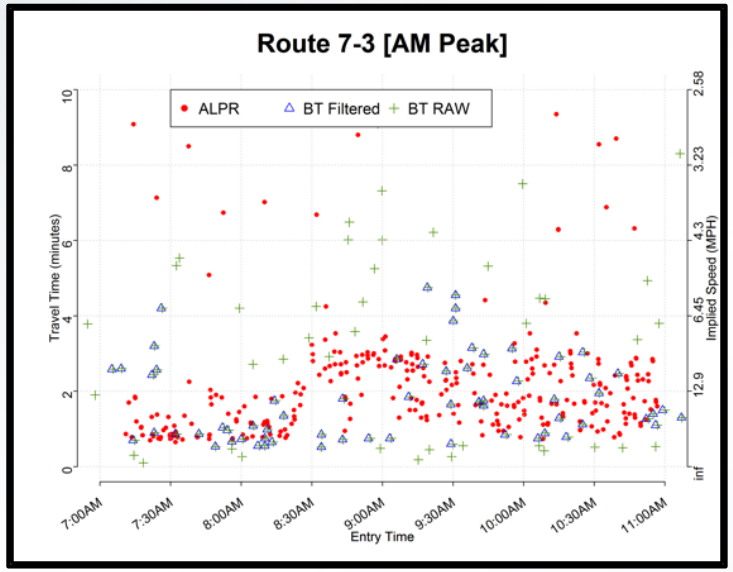


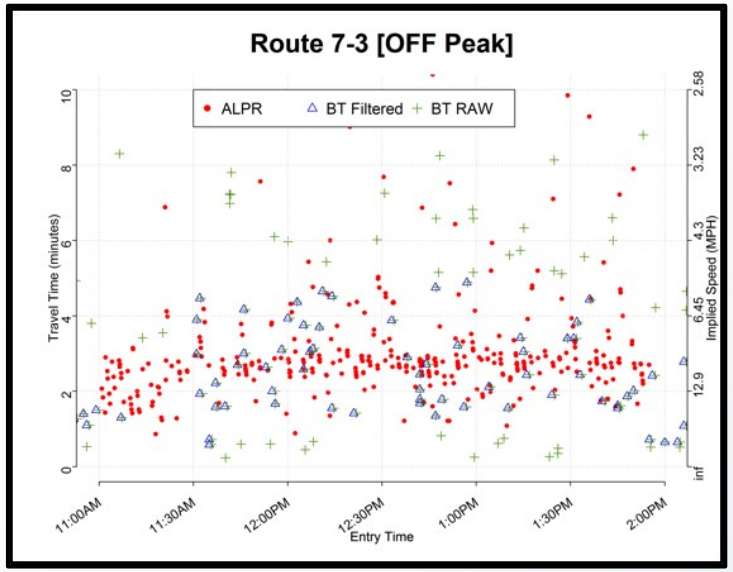
Travel-times on Right Turn Movement (Off-Ramp) Routes

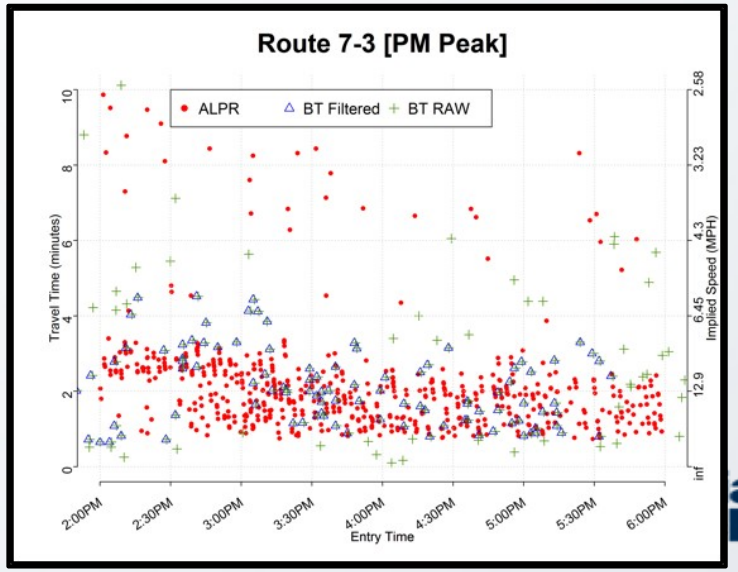




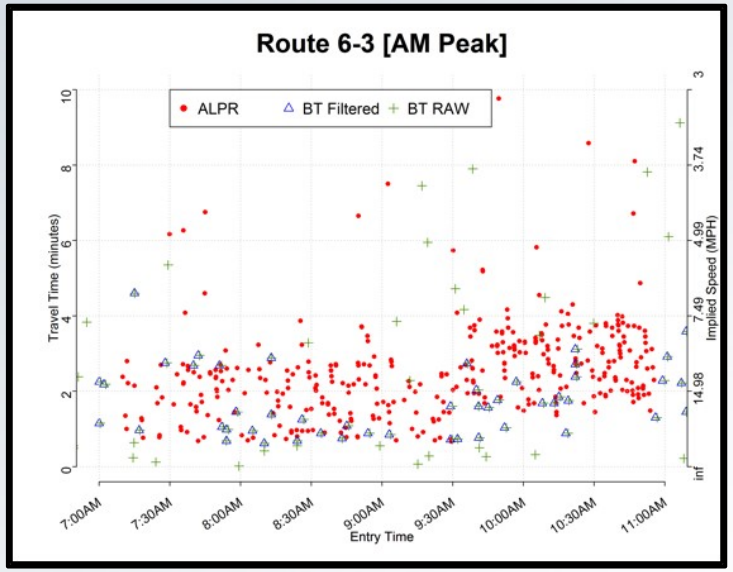


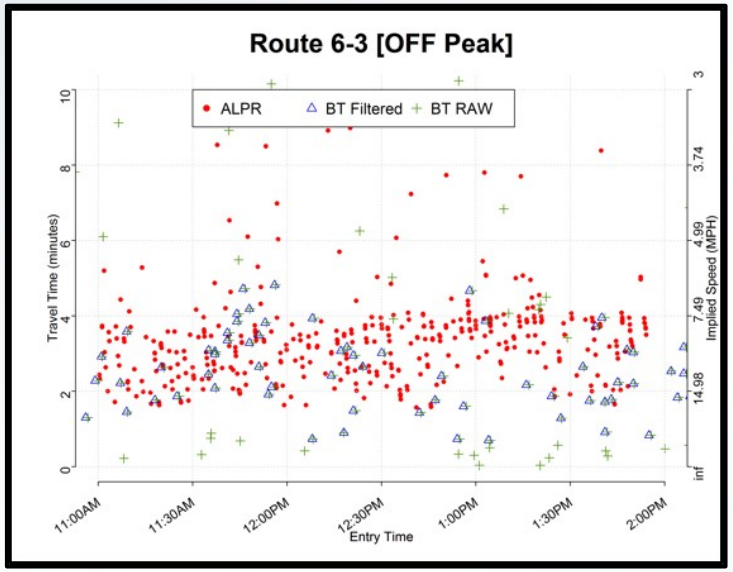


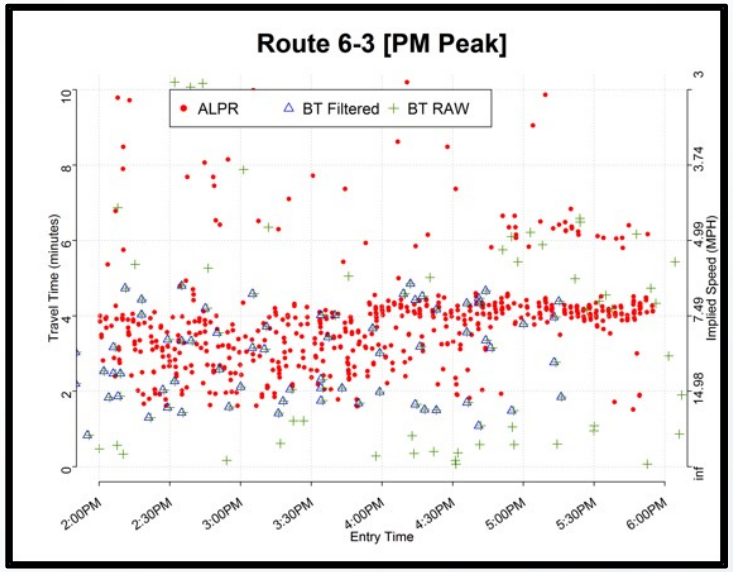


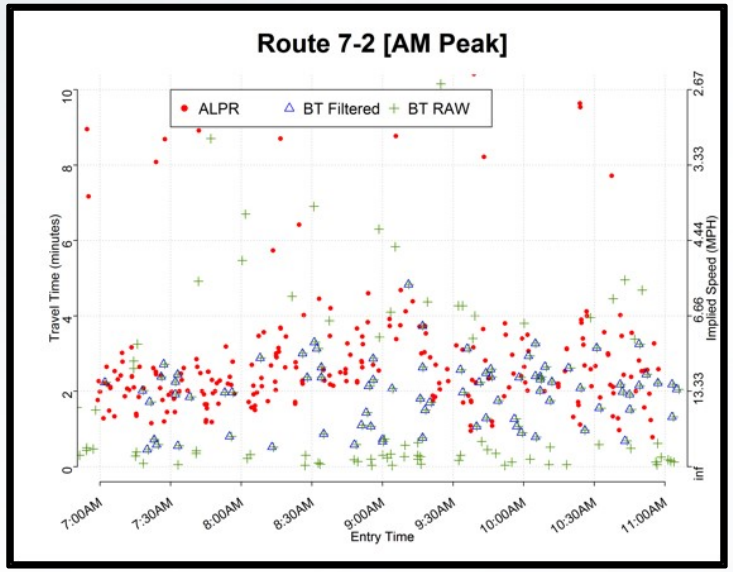


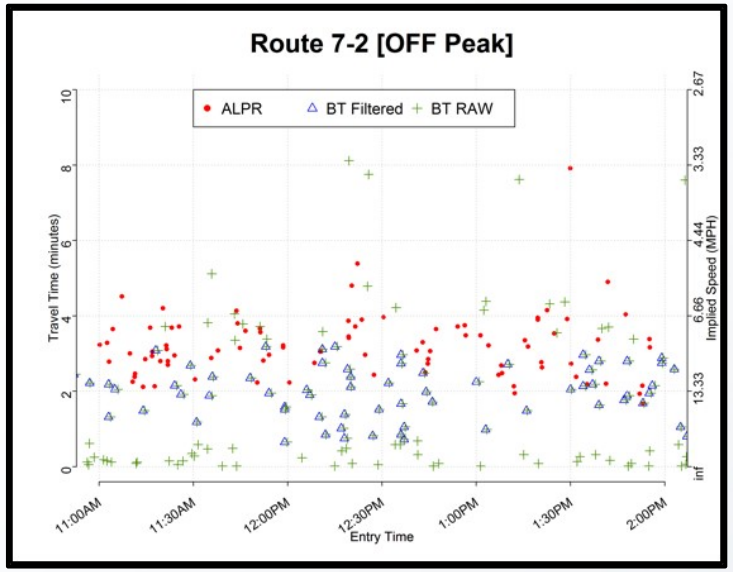
Travel-times on Left Turn Movement Routes (Off-Ramp)

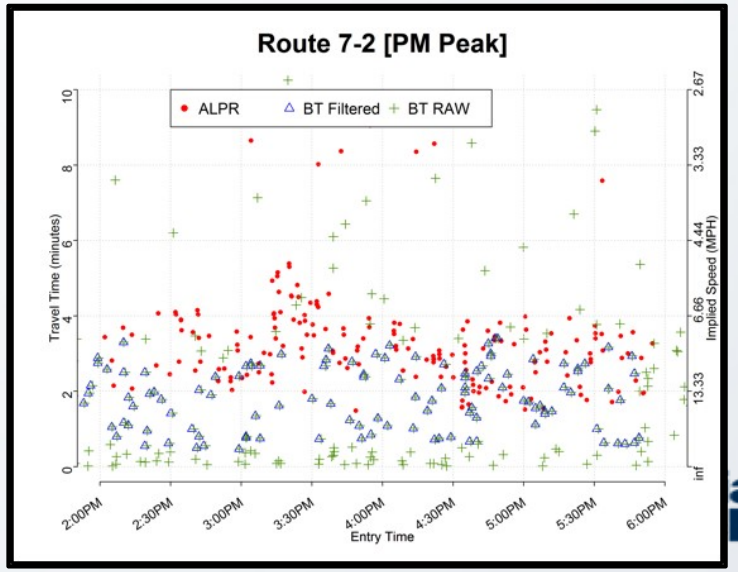






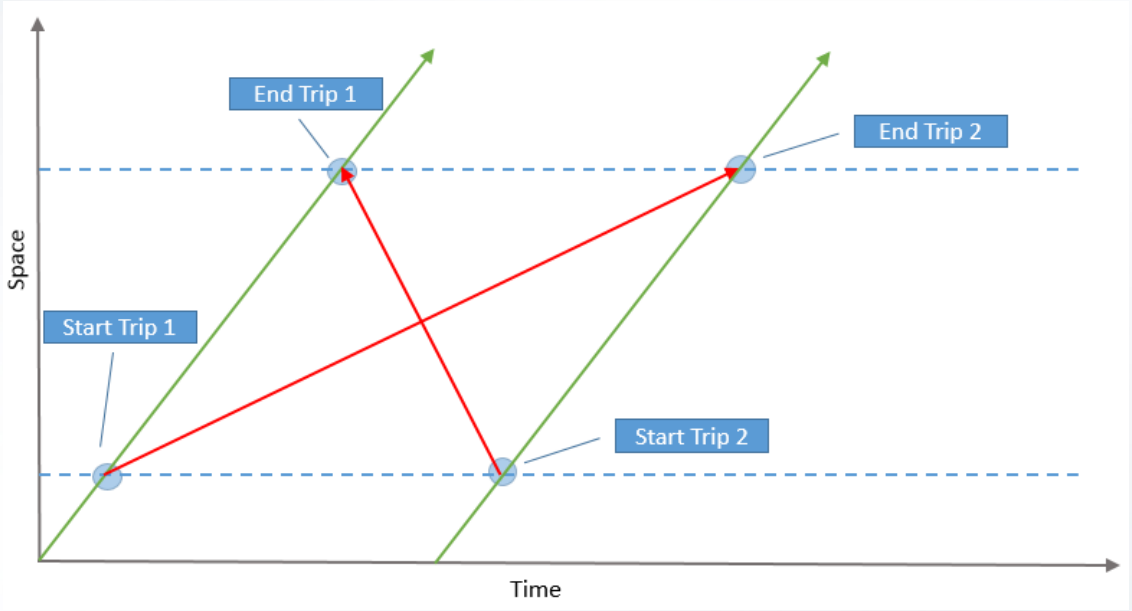




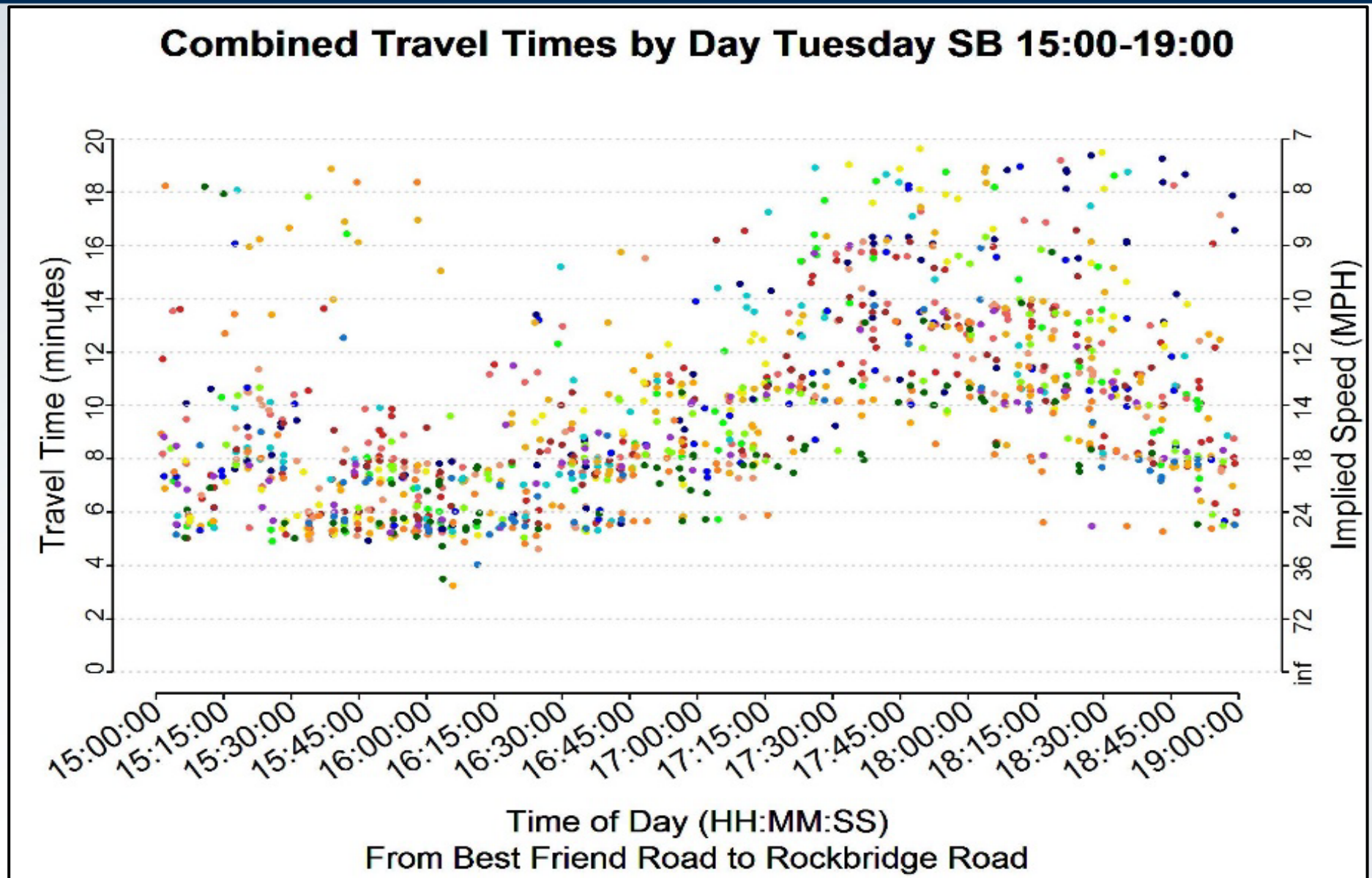


Variability Controls

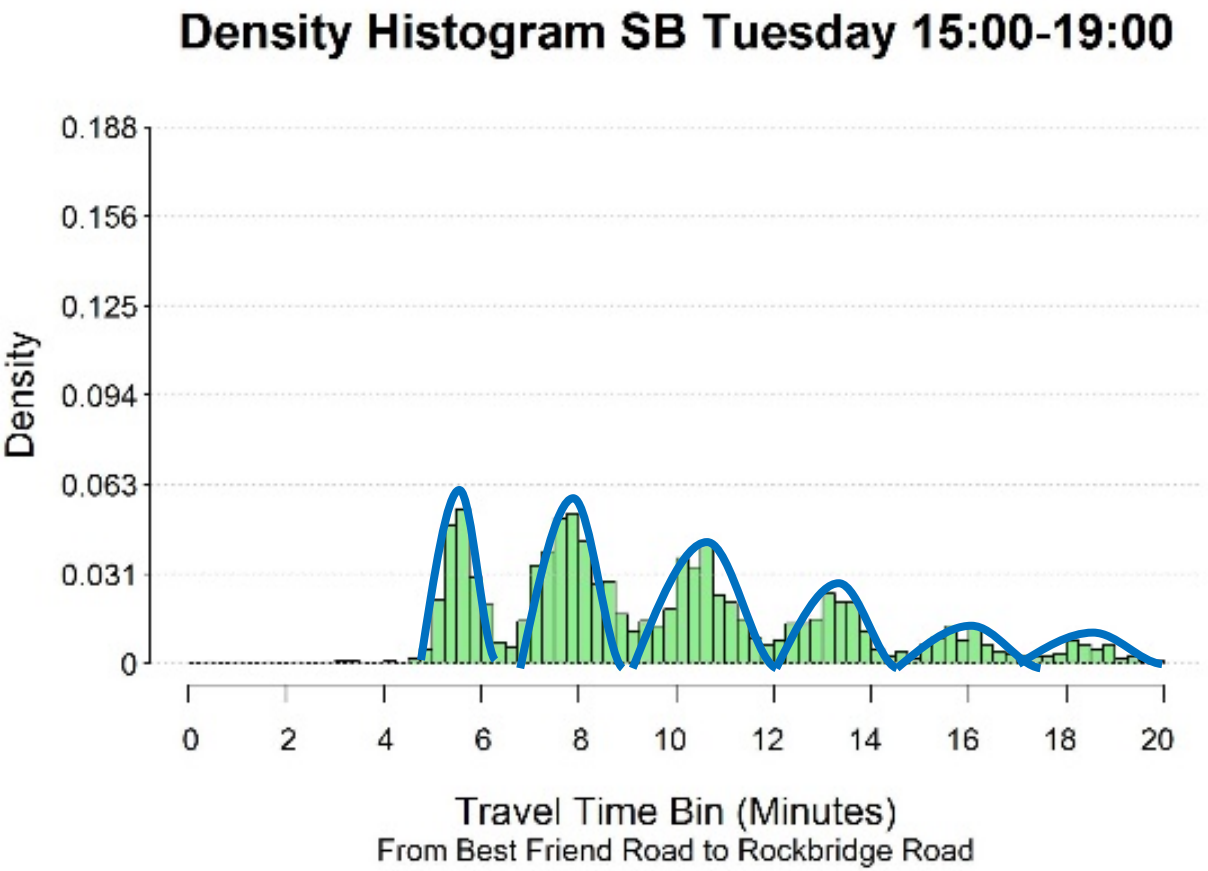
- Split dataset by signal plan
- Split dataset by day of week
- Outlier filter (upper and lower bounds)

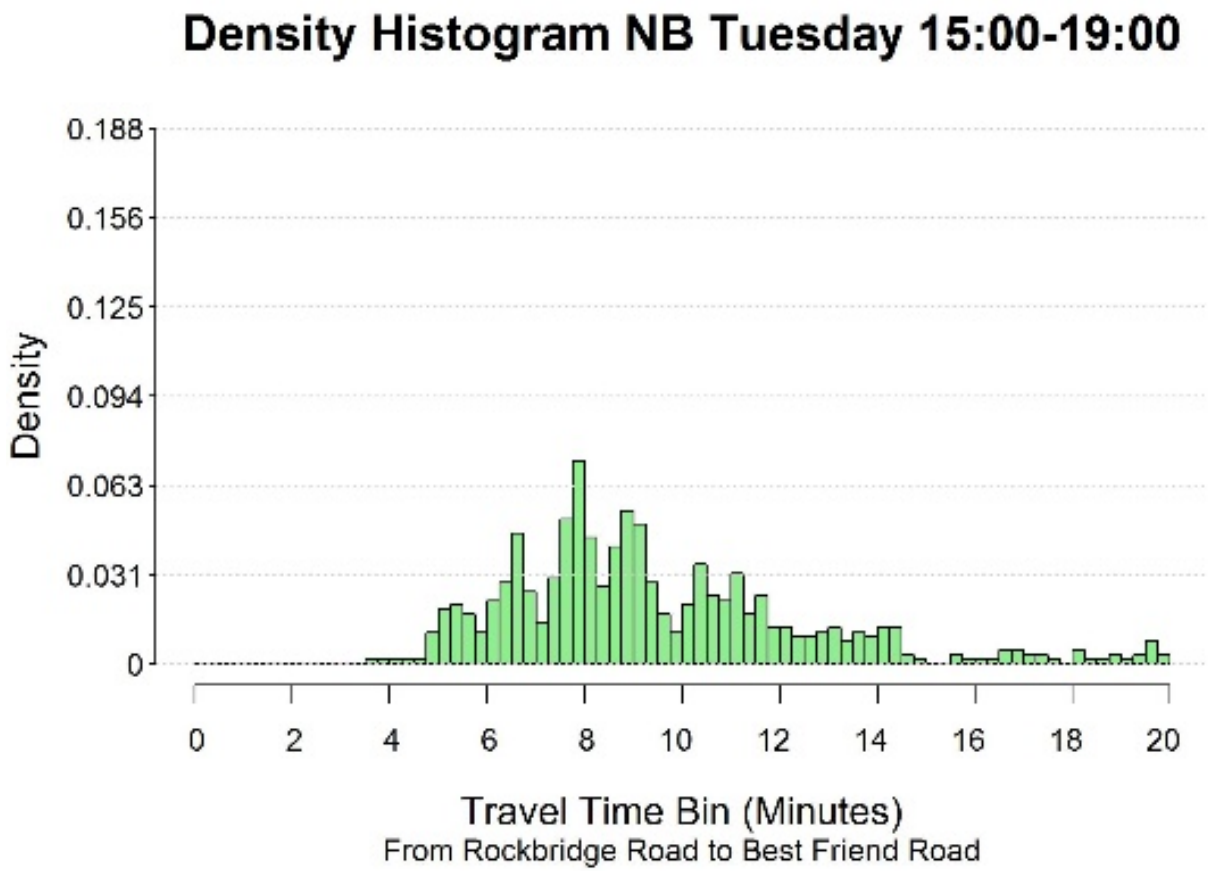


Identifying Patterns Over Multiple Days



Density Histogram

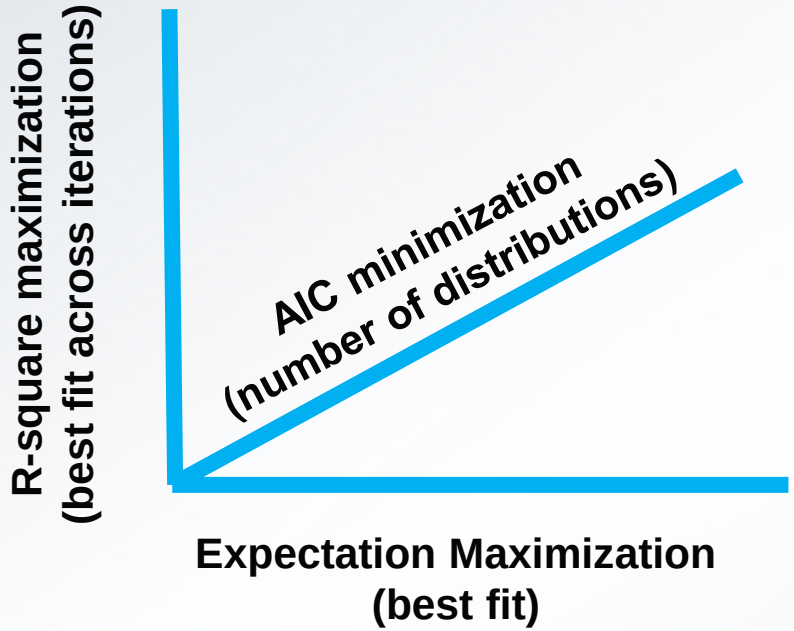




Methodology (1 of 2)

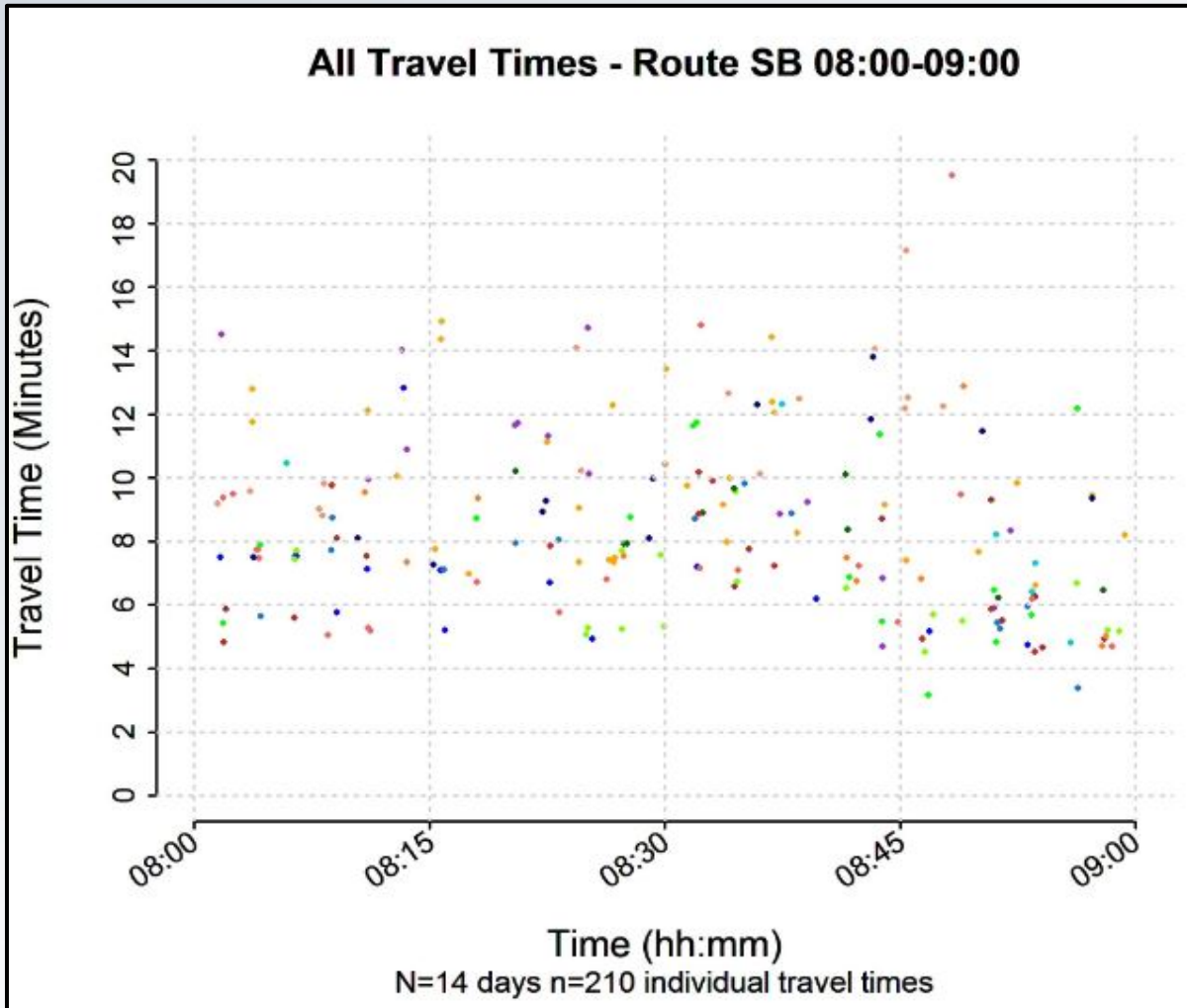
- Curve fitting
  - Choice of distribution: Gamma
  - Expectation Maximization algorithm for multiple curve fitting
  - Tools : R statisitcal software's mixtools package
- The EM process enhanced with Monte Carlo style method
  - EM process was run 100 times with different random starting values
  - Best fit according to R-squared value
- Number of underlying distributions was determined by fitting 1 to 5 distributions and using the Akaike Information Criterion ($AIC = -2 * \log(L) + 2 * k$) to determine which number of fits maximized the information content of the fitted function

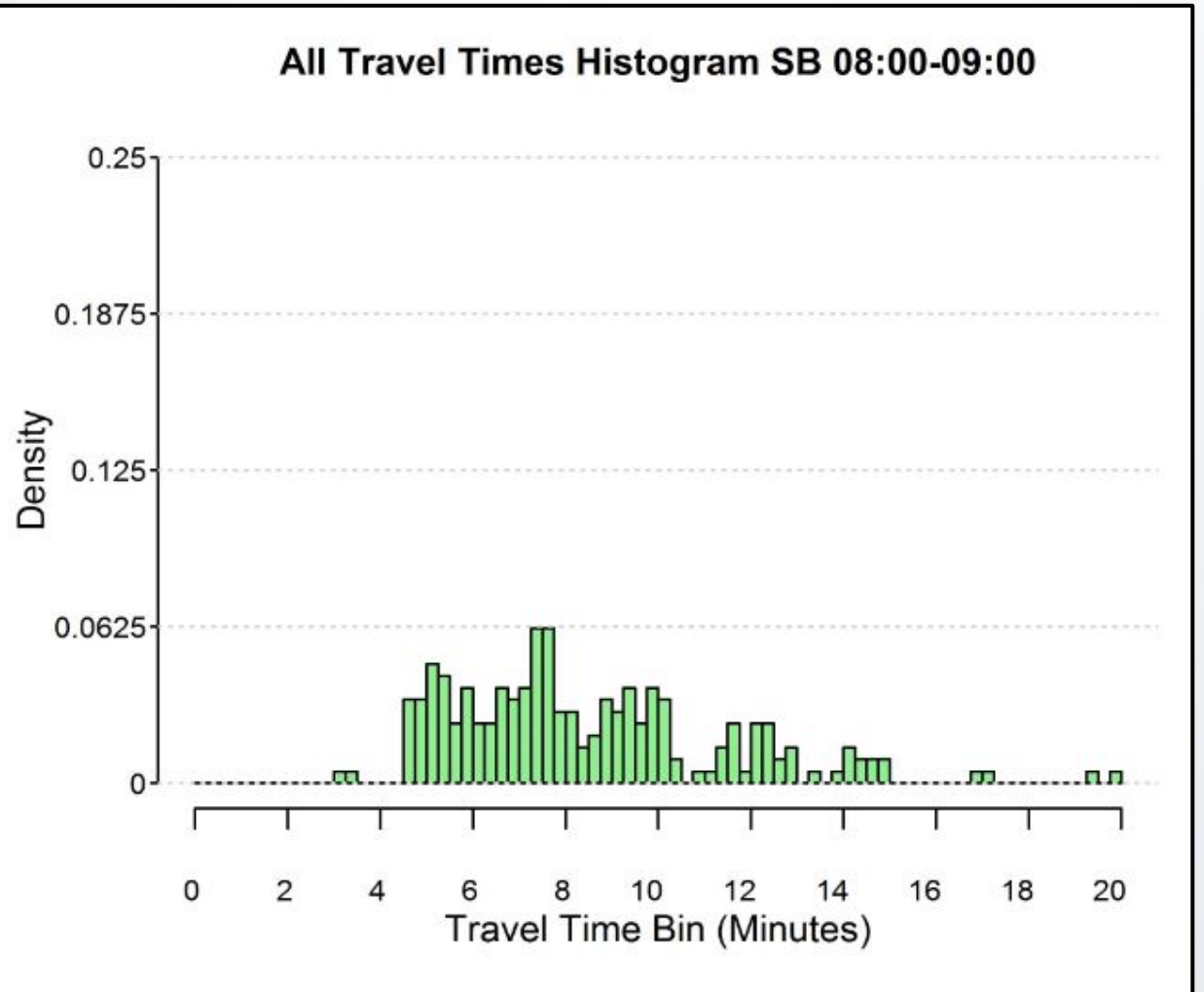
Methodology (2 of 2)



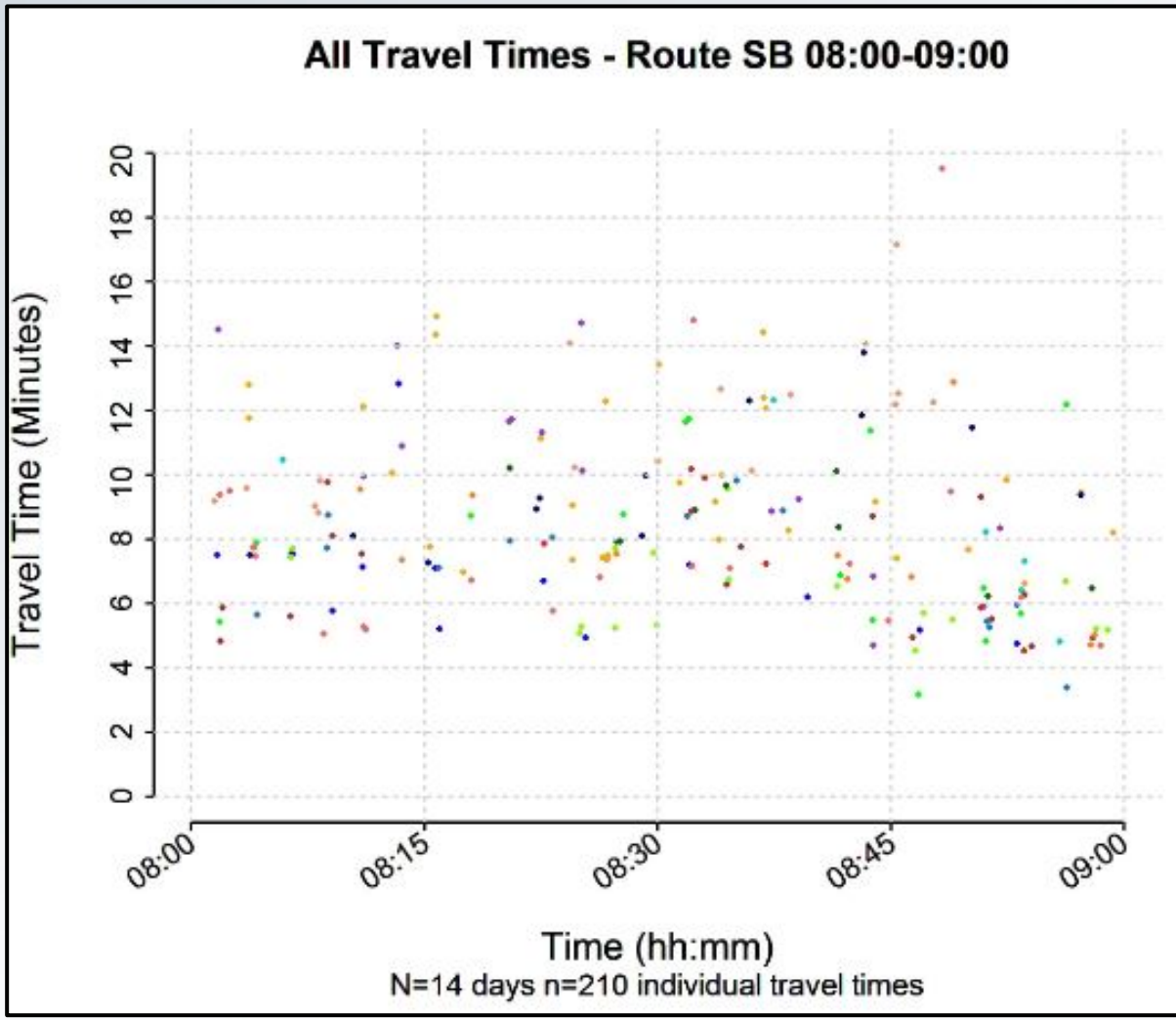
- Visualization of fit
  - Data points assigned to each distribution based on a posteriori probabilities and random uniform numbers
  - Random uniform number compared to a data point's a posteriori probability to determine its assignment to a distribution

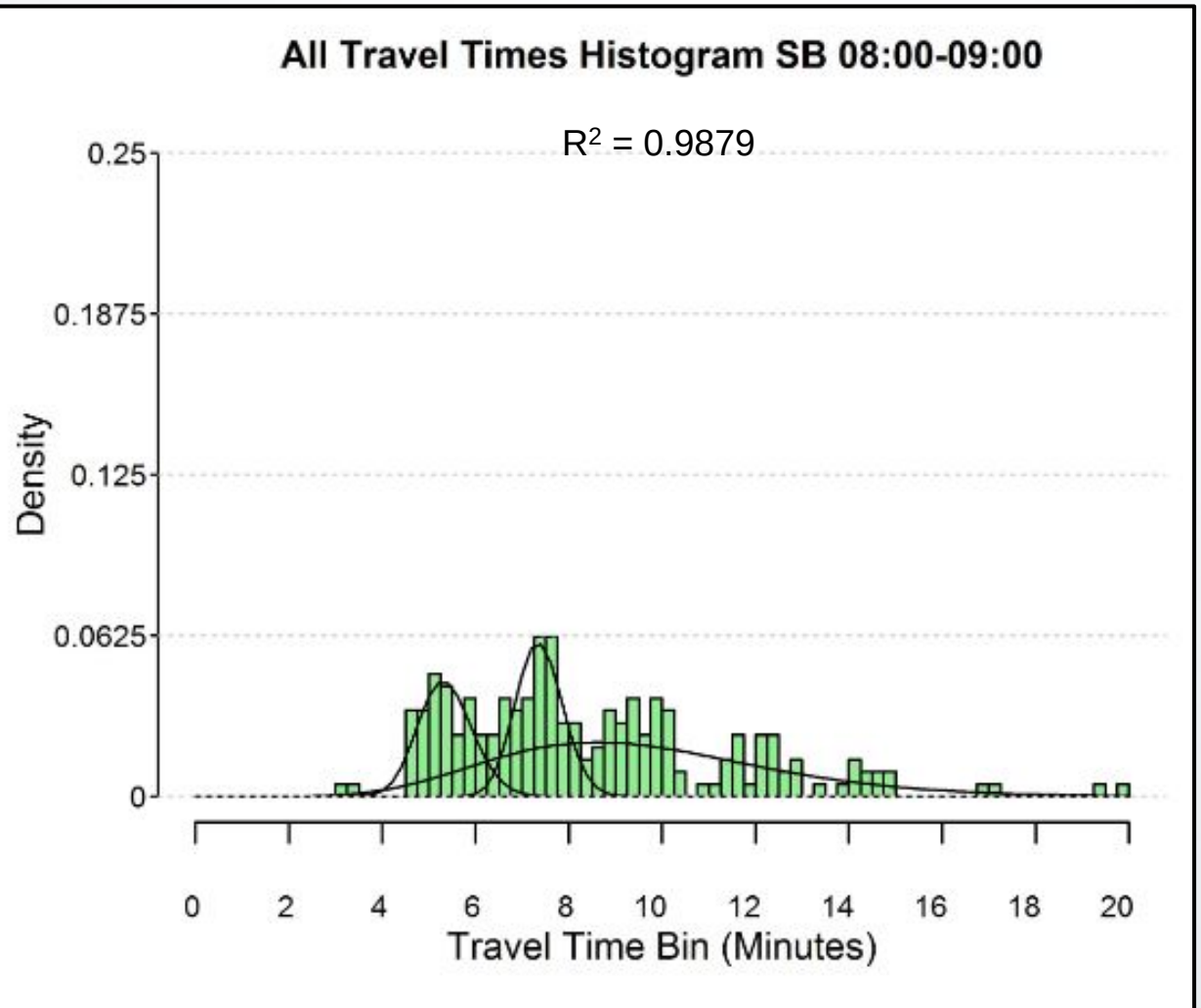
Composite Travel Time Separation (All Data, Tuesday AM)



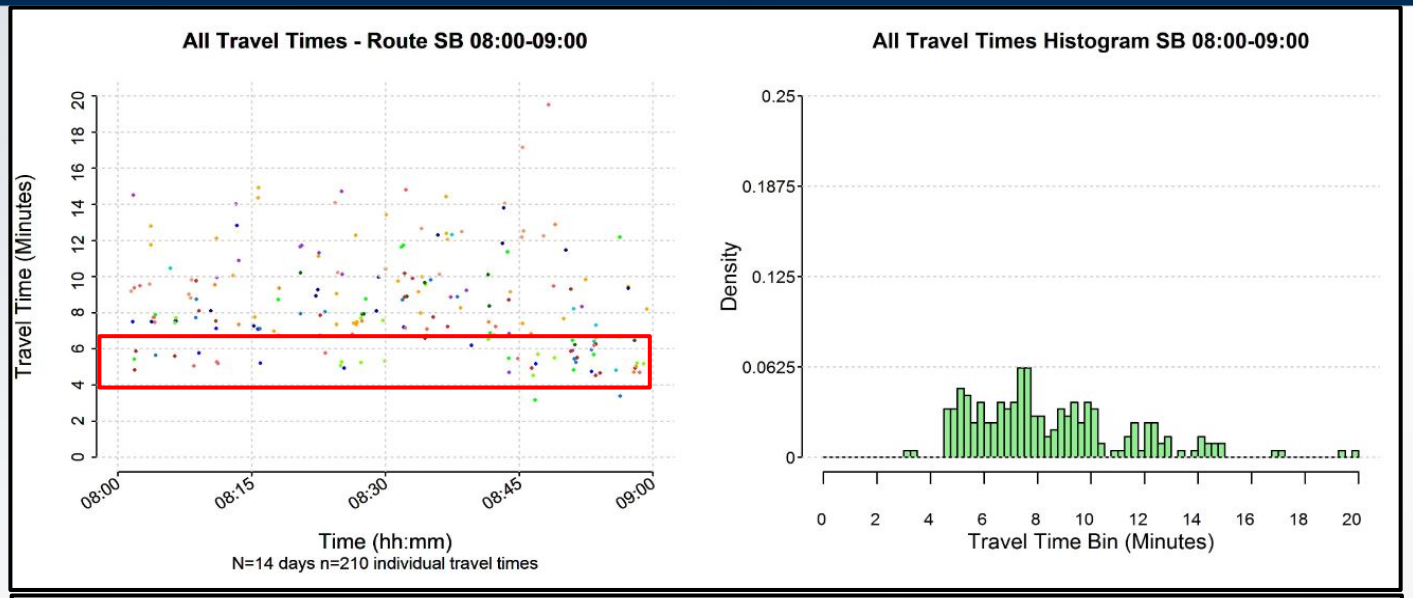


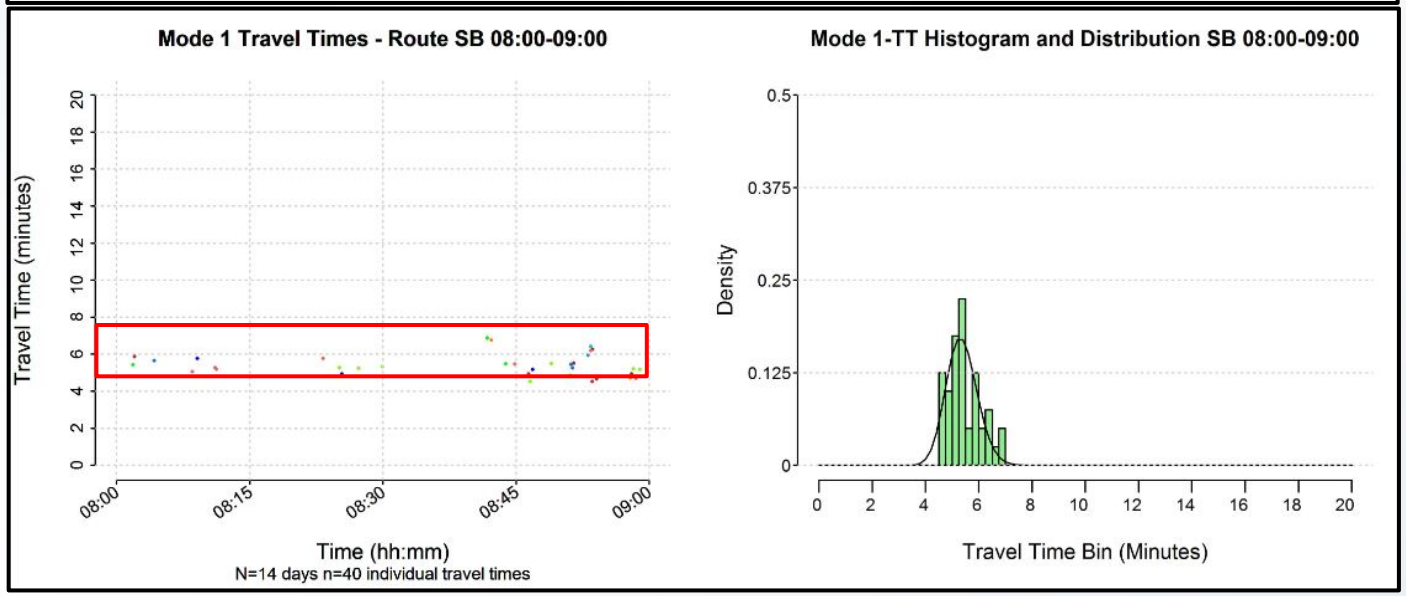
Composite Travel Time Separation (All Data, Tuesday)



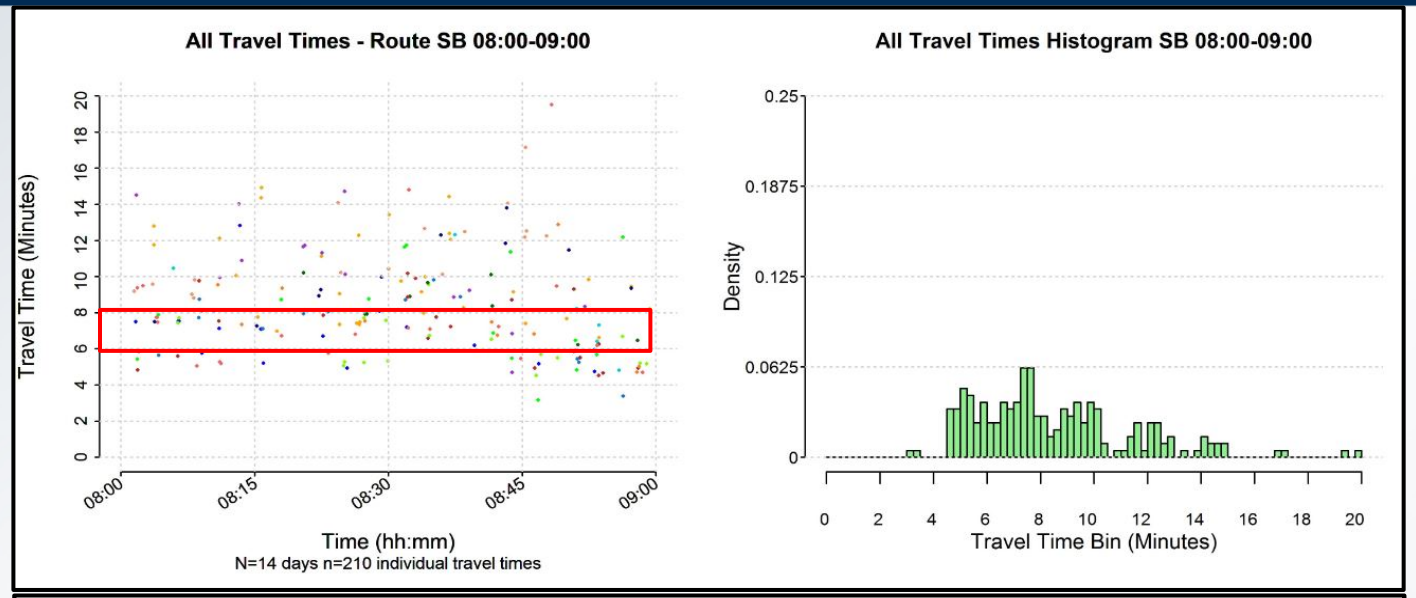


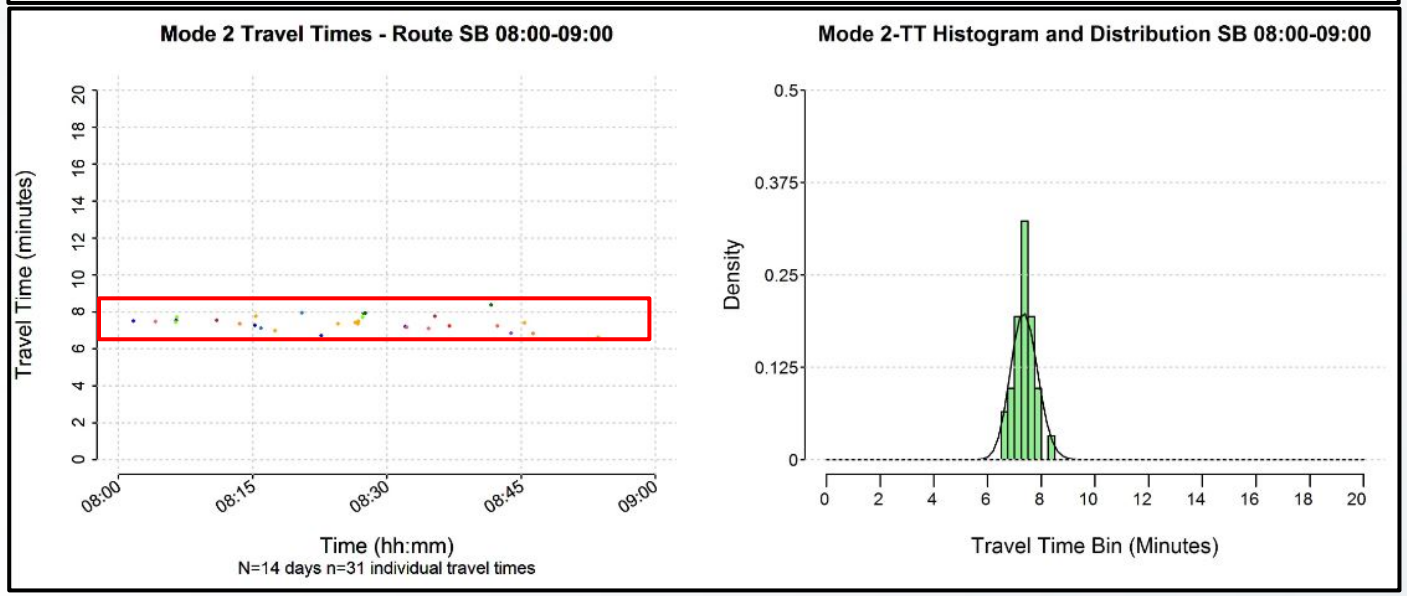
Composite Travel Time Separation (Mode 1, Tuesday AM)



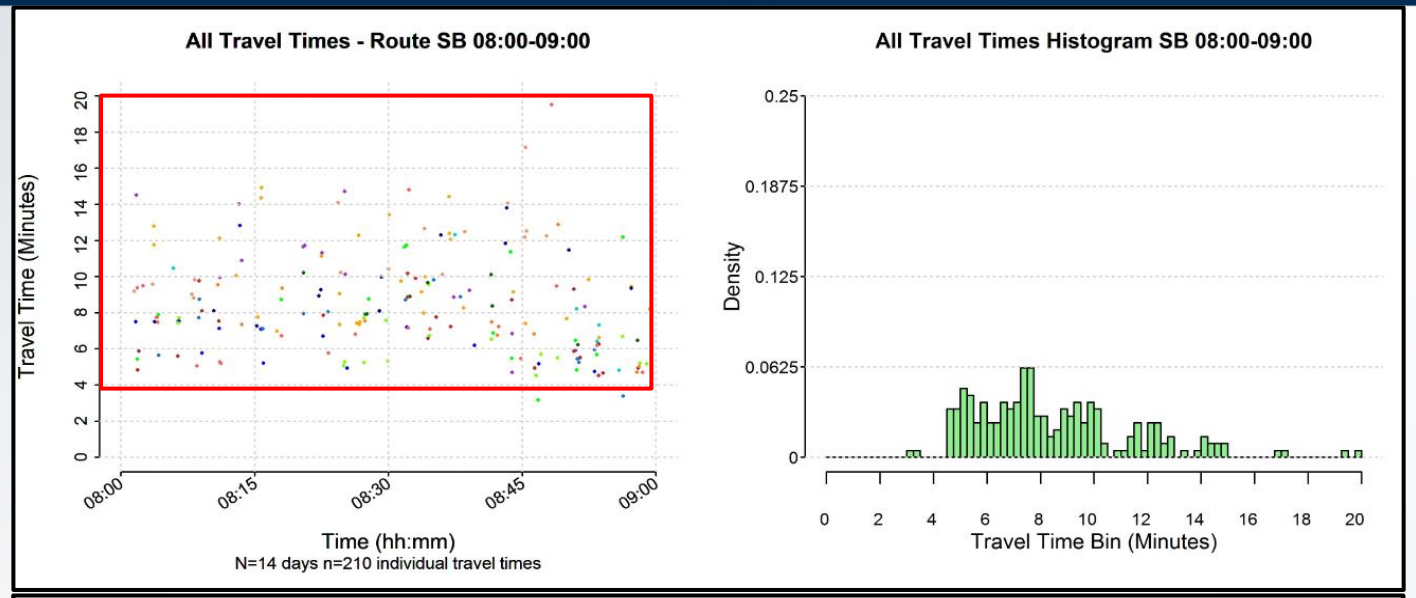


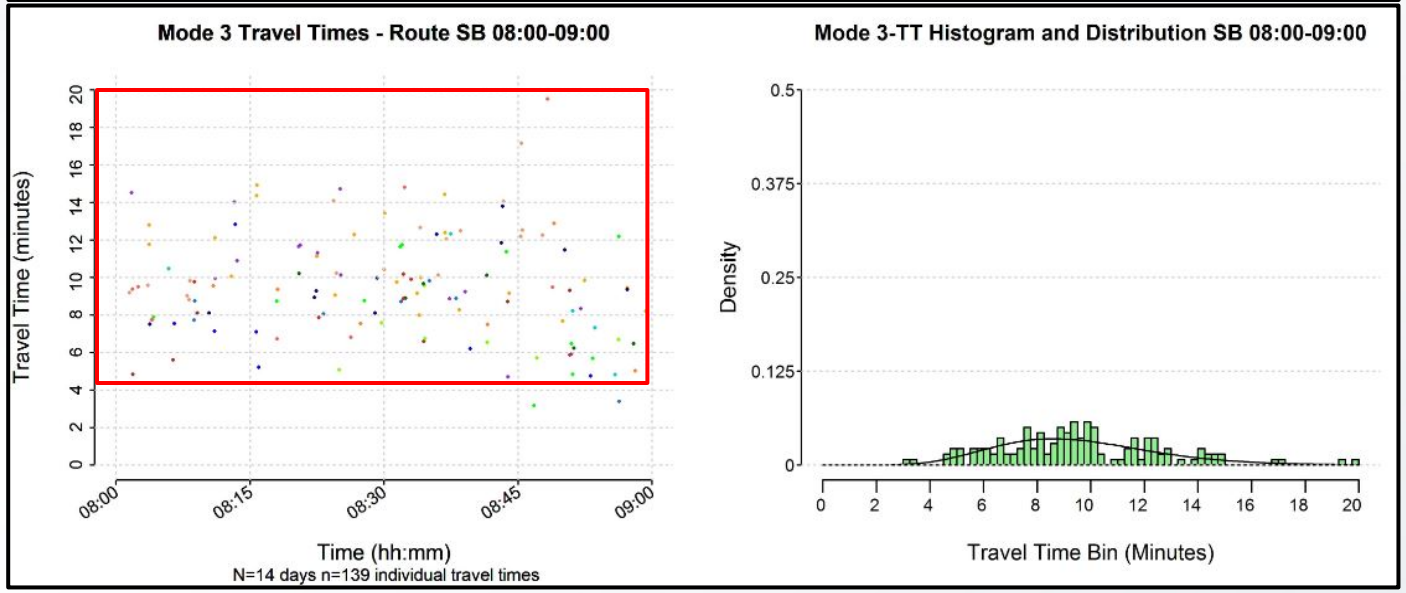
Composite Travel Time Separation (Mode 2, Tuesday AM)





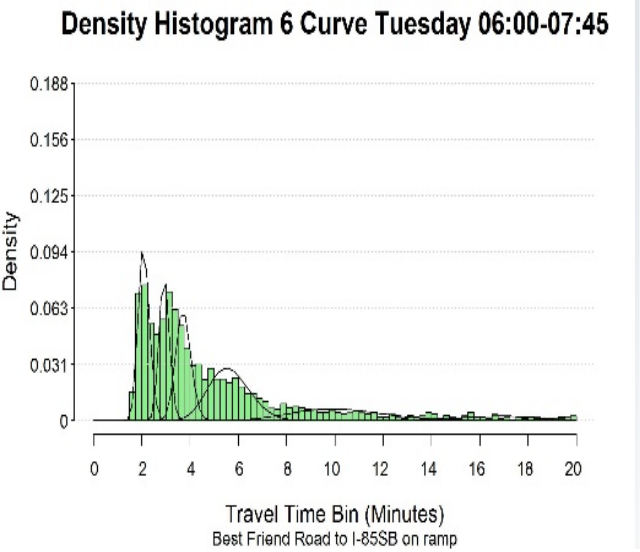
Composite Travel Time Separation (Mode 3, Tuesday AM)

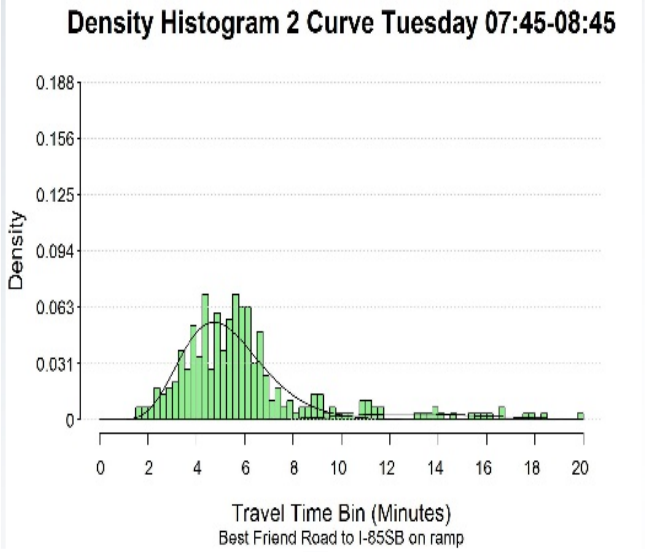


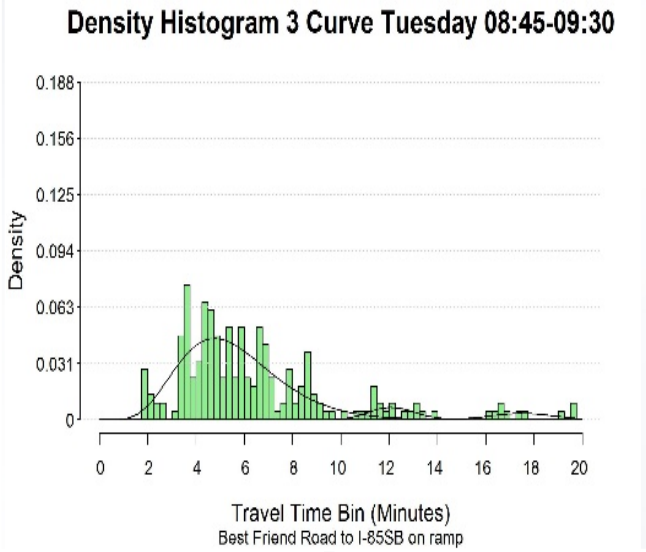


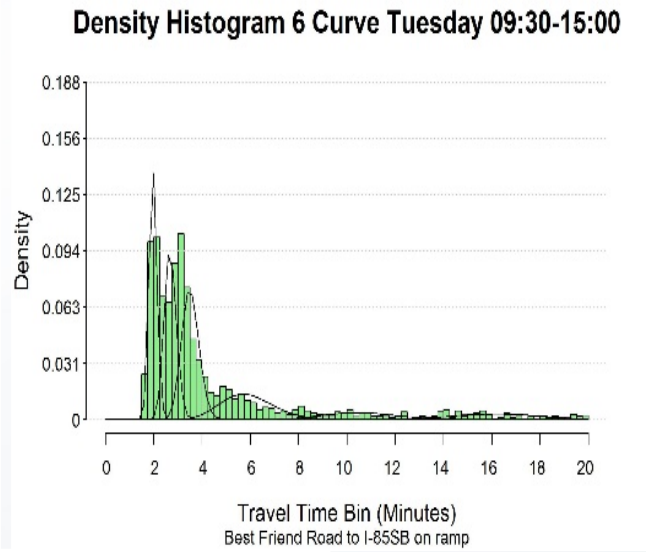
Composite Separation for Different Time of Day Periods

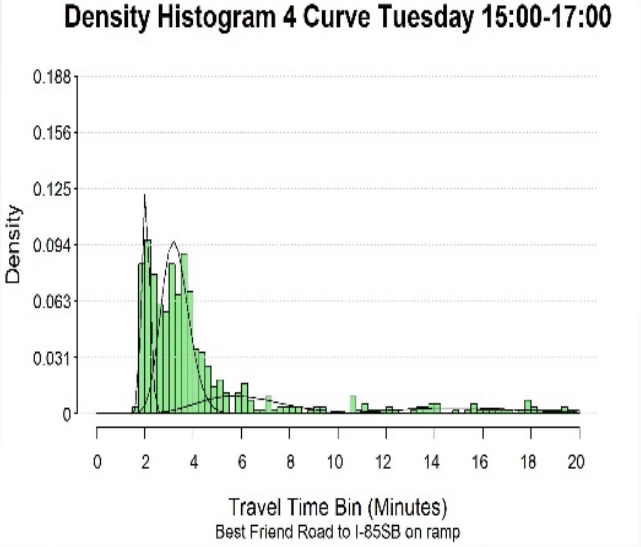
(Best Friend Road to I-85NB on ramp)

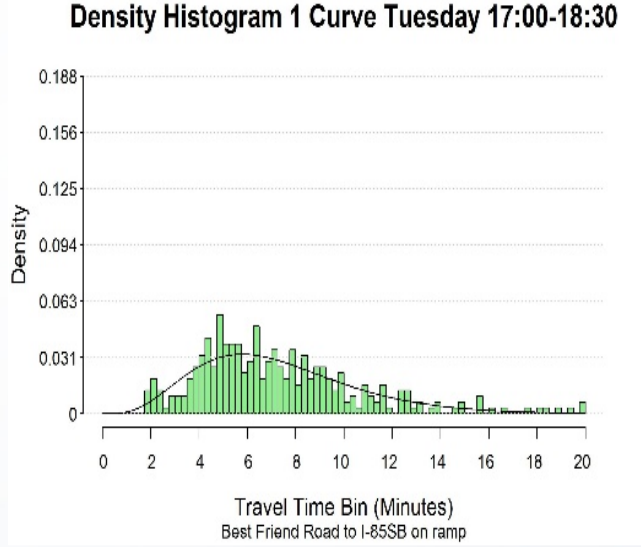


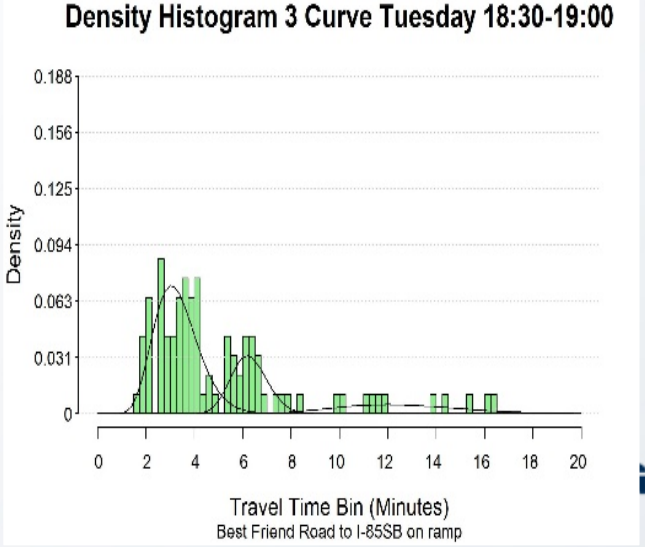












R^2 and AIC values for maximized R^2 output

(Best Friend Rd to I-85NB on ramp)

Day	Distribution Number	Hour	R-square	AIC
Tuesday	1	6:00AM - 7:45AM	0.000	16552.960
Tuesday	2	6:00AM - 7:45AM	0.433	15614.032
Tuesday	3	6:00AM - 7:45AM	0.759	15438.391
Tuesday	4	6:00AM - 7:45AM	0.813	15460.899
Tuesday	5	6:00AM - 7:45AM	0.906	15413.701
Tuesday	**6**	**6:00AM - 7:45AM**	**0.928**	**15372.831**
Tuesday	1	7:45AM - 8:45AM	0.000	1360.425
Tuesday	**2**	**7:45AM - 8:45AM**	**0.659**	**1314.288**
Tuesday	3	7:45AM - 8:45AM	0.703	1319.822
Tuesday	4	7:45AM - 8:45AM	0.860	1315.657
Tuesday	5	7:45AM - 8:45AM	0.907	1354.268
Tuesday	6	7:45AM - 8:45AM	0.923	1325.615
Tuesday	1	8:45AM - 9:30AM	0.000	1076.436
Tuesday	2	8:45AM - 9:30AM	0.632	1061.144
Tuesday	**3**	**8:45AM - 9:30AM**	**0.745**	**1058.056**
Tuesday	4	8:45AM - 9:30AM	0.799	1059.836
Tuesday	5	8:45AM - 9:30AM	0.897	1104.814
Tuesday	6	8:45AM - 9:30AM	0.926	1065.787

Day	Distribution Number	Hour	R-square	AIC
Tuesday	1	9:30AM - 3:00PM	0.000	7329.777
Tuesday	2	9:30AM - 3:00PM	0.542	6339.120
Tuesday	3	9:30AM - 3:00PM	0.835	6277.088
Tuesday	4	9:30AM - 3:00PM	0.903	6267.500
Tuesday	5	9:30AM - 3:00PM	0.910	6262.446
Tuesday	**6**	**9:30AM - 3:00PM**	**0.952**	**6235.012**
Tuesday	1	3:00PM - 5:00PM	0.000	2466.909
Tuesday	2	3:00PM - 5:00PM	0.737	2129.767
Tuesday	3	3:00PM - 5:00PM	0.853	2119.773
Tuesday	**4**	**3:00PM - 5:00PM**	**0.896**	**2098.818**
Tuesday	5	3:00PM - 5:00PM	0.919	2126.566
Tuesday	6	3:00PM - 5:00PM	0.941	2121.627
Tuesday	**1**	**5:00PM - 6:30PM**	**0.000**	**1581.146**
Tuesday	2	5:00PM - 6:30PM	0.358	1583.649
Tuesday	3	5:00PM - 6:30PM	0.609	1618.694
Tuesday	4	5:00PM - 6:30PM	0.806	1627.521
Tuesday	5	5:00PM - 6:30PM	0.882	1653.734
Tuesday	6	5:00PM - 6:30PM	0.918	1625.616
Tuesday	1	6:30PM - 7:00PM	0.000	447.467
Tuesday	2	6:30PM - 7:00PM	0.731	434.064
Tuesday	**3**	**6:30PM - 7:00PM**	**0.896**	**430.496**
Tuesday	4	6:30PM - 7:00PM	0.939	433.985
Tuesday	5	6:30PM - 7:00PM	0.952	438.389
Tuesday	6	6:30PM - 7:00PM	0.952	447.574

Application: Multiple LOS Analysis

- Mean speeds of each distribution calculated
- Assigned an LOS based on the HCM percentile speed of base free flow speed method

Travel Speed as a Percentage of Base Free Flow Speed (%)	LOS by Critical Volume-to-Capacity Ratio	
	≤ 1.0	≥ 1.0
>85	A	F
>67-85	B	F
>50-67	C	F
>40-50	D	F
>30-40	E	F
≤30	F	F

Multiple LOS Results Table (AM Peak Hour)

Direction	Day	Time	Mean Speed 1	Mean Speed 2	Mean Speed 3	LOS 1	LOS 2	LOS 3	Average Speed	Average LOS
Southbound	Monday	8AM-9AM	22.3	18.5	12.7	C	D	F	19.4	D
Southbound	Tuesday	8AM-9AM	26.1	18.9	14.8	C	D	E	17.4	D
Southbound	Wednesday	8AM-9AM	22.7	18.2	12.2	C	D	F	17.5	D
Southbound	Thursday	7AM-8AM	19.5	14.3	10.4	D	E	F	17.3	D
Southbound	Friday	8AM-9AM	23.6	15.1	12.3	C	E	F	17.3	D
Northbound	Monday	7AM-8AM	31.7	24.2	16.1	B	C	E	20.1	D
Northbound	Tuesday	7AM-8AM	36.0	22.9	14.1	B	C	E	18.0	D
Northbound	Wednesday	7AM-8AM	25.1	17.2	13.8	C	D	E	17.0	E
Northbound	Thursday	7AM-8AM	22.8	16.6	15.1	C	E	E	17.5	D
Northbound	Friday	7AM-8AM	35.5	23.6	15.4	B	C	E	19.5	D

Limitations

- Uniformity of sampling rates
  - Different probe technologies have well documented inherent limitations on sampling bias
    - Vehicle Speed
    - Distance from detector
  - Proportions of vehicles experiencing different levels of service are unknown
  - Use separated data sets as individual distributions and run robust statistical tests to determine if changes in the distributions are statistically significant

Questions?

Contact: guin@gatech.edu

Georgia Institute of Technology