MATS UTC

FINAL REPORT

Performance Measures for Freight & General Traffic: Investigating Similarities and Differences Using Alternate Data Sources

Date: June 5th 2015

Filmon Habtemichael, Old Dominion University
Rajesh Paleti, Old Dominion University
Mecit Cetin, Old Dominion University

Prepared by:
Transportation Research Institute
Old Dominion University
135 Kaufman Hall, Norfolk, VA 23529

Publication Number:

Prepared for:
Virginia Center for Transportation Innovation and Research
530 Edgemont Road
Charlottesville, VA 22903

TECHNICAL DOCUMENTATION

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.

4. Title and Subtitle Performance Measures for Freight & General Traffic: Investigating Similarities and Differences Using Alternate Data Sources	**5. Report Date** 6/5/2015
	6. Performing Organization Code
7. Author(s) Filmon Habtemichael, Rajesh Paleti, and Mecit Cetin	**8. Performing Organization Report No.**
9. Performing Organization Name and Address Old Dominion University 135 Kaufman Hall Norfolk, VA, 23529	**10. Work Unit No. (TRAIS**
	11. Contract or Grant No. DTRT13-G-UTC33
12. Sponsoring Agency Name and Address US Department of Transportation Office of the Secretary-Research UTC Program, RDT-30 1200 New Jersey Ave., SE Washington, DC 20590	**13. Type of Report and Period Covered** Final 2/6/15 – 5/24/15
	14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

Recent advances in probe vehicle data collection systems have enabled monitoring traffic conditions at finer temporal and spatial resolution. The primary objective of the current study is to leverage these probe data sources to understand if there are any significant differences between freight and general traffic characteristics. To this end, different types of performance measures (PMs) including speed, travel time, and reliability were estimated from three different data sources- INRIX, HERE, and ATRI. The analyses conducted for the Hampton Roads network suggest that while in most cases general traffic PMs serve as a good proxy for characterizing freight traffic, in some cases, there can be significant differences either in average speeds/travel times or reliability (*i.e.*, variation about mean). So, in the absence of freight-specific data, researchers planning to use general traffic data as a surrogate must exercise caution.

17. Key Words Performance measures; Probe vehicle data; Freight modeling; Reliability; Cluster Analysis	**18. Distribution Statement** No restrictions. This document is available from the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	48	

ACKNOWLEDGMENTS

The authors would like to thank Jungwook Jun at the Transportation and Mobility Planning Division (TMPD) of Virginia Department of Transportation (VDOT) for his help in procuring and understanding the HERE and INRIX datasets used in this study. Also, the research team thanks Mena Lockwood in the Traffic Engineering Division and Catherine McGhee at VCTIR for their feedback and suggestions over the course of this research project.

DISCLAIMER

The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the information presented herein. This document is disseminated under the sponsorship of the U.S. Department of Transportation's University Transportation Centers Program, in the interest of information exchange. The U.S. Government assumes no liability for the contents or use thereof.

TABLE OF CONTENTS

List of Figures

List of Tables

PROBLEM

Efficient and reliable movement of freight is an essential and critical requirement for a vibrant economy. A significant percentage of freight (both in terms of volume and value) is moved on trucks along the national highway system. So, it is important to monitor the transportation network conditions to identify bottlenecks and take adequate measures to alleviate traffic congestion along the major freight corridors. Furthermore, it would be useful to examine if there are any significant differences between the spatial and temporal profiles of bottlenecks of the freight and general passenger traffic. When possible, truckers try to avoid congested time periods and corridors to reduce operating costs incurred due to travel delays. In addition, banning truck travel in rush hours or other travel-hour restriction policies force truckers to operate at different "peak hours" than the rest of the traffic. Therefore, it is important to compare and contrast the network conditions experienced by trucks and the general traffic so that accurate and relevant performance measures are obtained. Such comparisons are also important for understanding whether a congestion relief program for a given bottleneck or corridor would benefit both freight and general-purpose traffic.

APPROACH

Recent advances in the probe-vehicle based data collection systems have enabled collection of key performance measures (PMs) including speed and travel time at a very high spatial and temporal resolution. In this proposed project, three probe-data sources: INRIX, HERE, and American Transportation Research Institute's (ATRI's) GPS data for the month of September, 2014 were used to compare the spatial and temporal profiles of bottlenecks of freight and general traffic for the Hampton Roads network. This was done by analyzing the variation in speeds and travel times as the main performance measures. The HERE data constitutes the National Performance Management Research Data Set (NPMRDS) acquired by the Federal Highway Administration (FHWA) to support freight management and operations. While passenger traffic speed estimates in the NPMRDS were obtained using HERE probe data sources, freight speed estimates were obtained using the ATRI data. For this research, the research team acquired the raw ATRI data that was used to derive freight speed estimates in HERE data to examine the potential use of disaggregate GPS data to support additional freight modeling work such as route choice modeling and OD matrix estimation. The research team also had access to INRIX freight data for July 2012. However, this data was available for a typical weekday (one typical Monday for the entire month) at 15 minutes resolution. Given that there can be significant differences in PMs across different days of the month, this aggregate 2012 INRIX data was not used in the analysis. Also, the research team obtained HERE data for July 2012 but this data was deemed unusable because, for any given day in the month, data was available for only a small subset 5 minute intervals out of 24 hours. Given the limited timeline for this project, the study focused on the Hampton Roads region (see Figure 1) that has several freight-significant transportation corridors (for example, I-64, US 13, and US17). Moreover, Hampton Roads is also a major freight hub with 80% of the entire freight tonnage being either inbound or outbound (instead of pass-through) due to three major port terminals in the region and nearly 60% of this freight tonnage (and 97% of freight value) is moved on trucks (Cambridge Systematics, 2010). HERE data is available mostly for major highways and preliminary data analysis indicated that among these major roadways only data along I-64 and I-664 corridors can support the refined spatial and temporal comparison analysis needed in this study. So, all the analyses were undertaken for these two roadways within the Hampton Roads region. Within the study region, I-64 extends about 53 miles in both WB and EB directions where as I-664 runs nearly 20 miles in the NB and SB directions. The research approach adopted to accomplish the study objective may be summarized as:

1. Generate freight-specific and general-traffic PMs for select freight corridors in Hampton Roads region based on the available datasets
2. Develop appropriate statistical methods to investigate the correlation between the freight and general traffic PMs based on probe-vehicle data
3. Compare and identify similarities and differences between the characteristics of general and freight traffic
4. Explore the potential of ATRI's truck GPS data to support advanced freight modeling (beyond PM) such as route choice analysis.

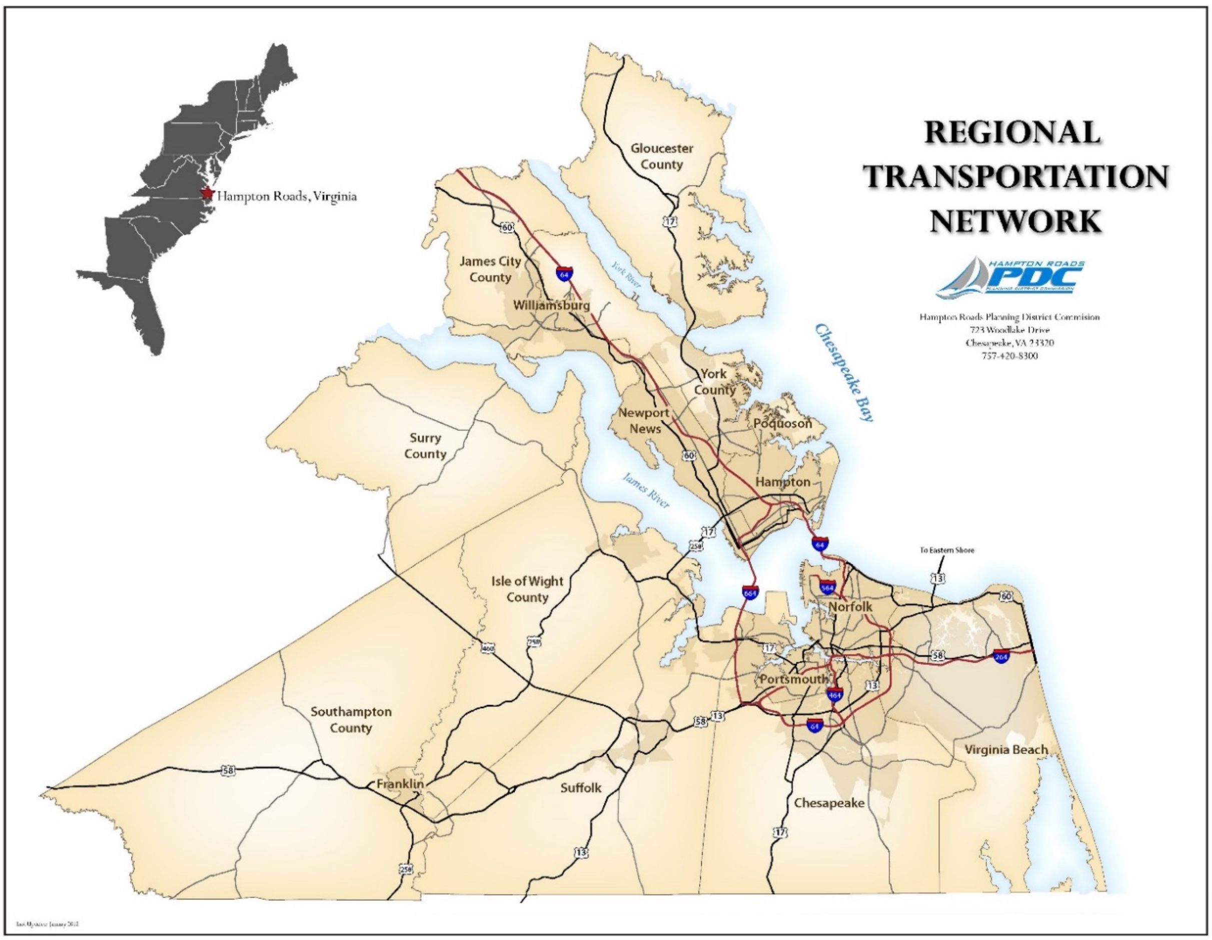


Figure 1 Hampton Roads Study Region[1]

[1] Obtained from the Hampton Roads Planning District Commission's "Maps and GIS" webpage: http://www.hrpdcva.gov/page/maps-and-gis/

METHODOLOGY

Reconciliation of Differences across Data Sources

One of the primary challenges associated with working on multiple data sources is reconciliation of differences across datasets to allow direct comparison. While HERE and INRIX data were reported at uniform time intervals and at a spatial aggregation of "Traffic Message Channels" (TMCs) that are non-uniform 1-3 mile road segments, the ATRI GPS data is a continuous stream of truck coordinates that were recorded at non-uniform time intervals. Also, not only are the TMC definitions in the HERE and INRIX data sources different but also INRIX data was recorded every 15 minutes while HERE data was reported at a five-minute resolution. To facilitate comparison across different data sources, the I-64 and I-664 roadways were divided into a uniform grid of one mile and all the speed and travel time estimates were obtained at this one mile resolution from all data sources. The truck GPS points from the ATRI data were snapped to this uniform grid using a distance buffer of 300 ft. Figure 1 and Figure 2 show the mile post locations along the uniform grid for I-64 and I-664, respectively.

Data Visualization Using Heat Maps

Speed heat maps were produced for the different corridors along I-64 and I-664 to visually investigate if there are any systematic differences in the PM estimates obtained from the HERE and INRIX data sources. The location of major bottlenecks along these roadways and congestion time periods can be easily identified using these heat maps. These trends observed in each dataset may be compared with general expectations and experience of travelers in the region. Also, the heat maps serve as a validation of the data processing and grid transition work that was undertaken to reconcile the differences in the TMC definitions in the INRIX and HERE data sources. More importantly, these heat maps are useful in identifying differences between freight and general traffic PMs because it is easy to visually compare the congestion zones (both spatial and temporal) for freight and general traffic.

Speed Comparison Using Box Plots

While heat maps are useful to visually understand the quality of data and identify bottleneck locations, it is difficult to summarize the variation in PMs across different days in the same map. To this end, box plots that indicate the difference between freight and passenger speeds were produced. For any given location and time period, general traffic and freight speed estimates were obtained from different data sources across all days. For instance, in the month of September 2014, there were five Mondays. So, for the evening peak period between 3 and 6 PM, 60 speed estimates (= 5 Mondays x 3 hours x 4 15-minute intervals in each hour) were obtained for each unit mile segment along each corridor from each data source. These speed estimates were used to produce box plots that show the distribution of speed for freight and general traffic over the entire length of I-64 and I-664 in both directions.

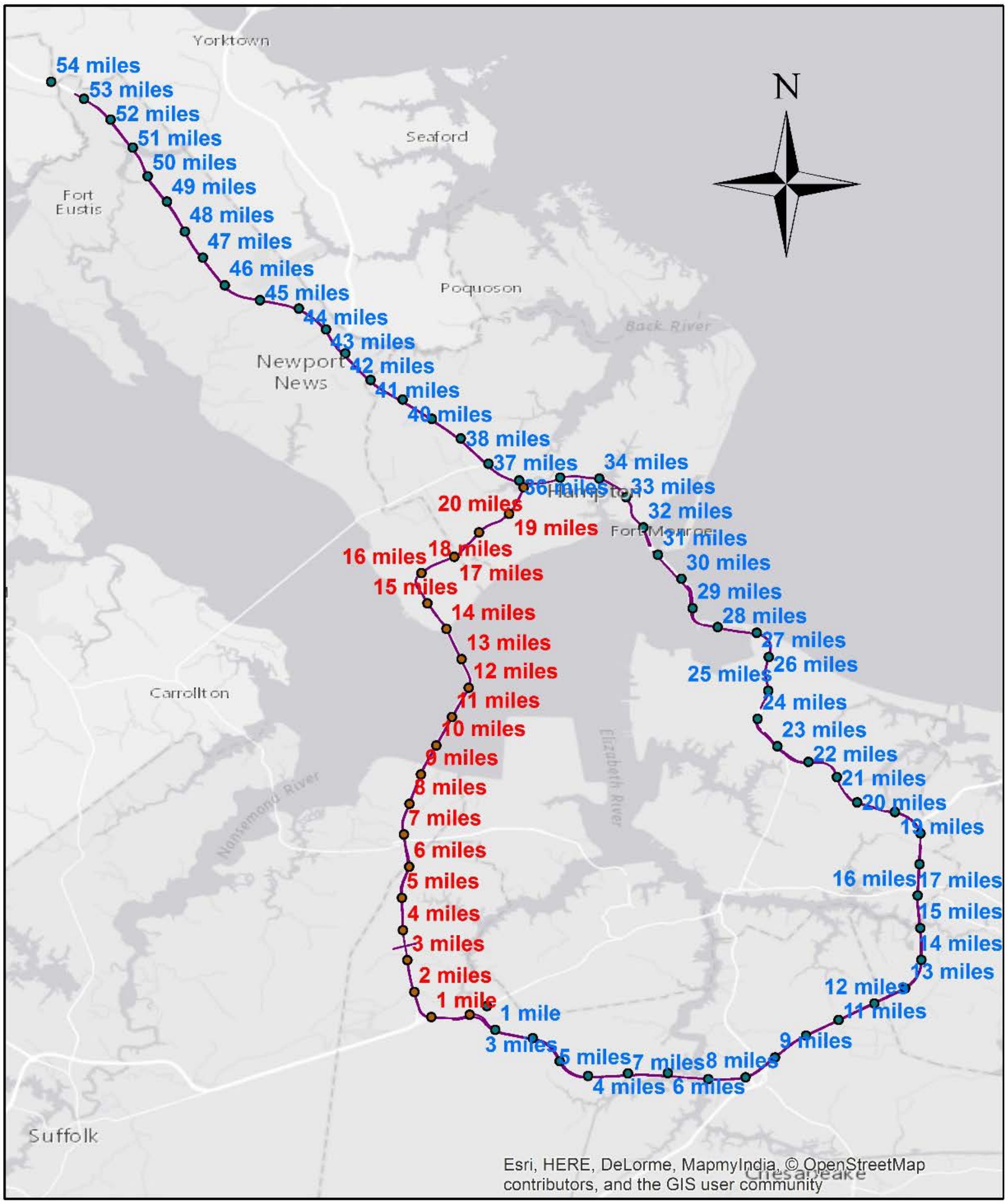


Figure 2 Mile Posts for I-64 WB and I-664 NB

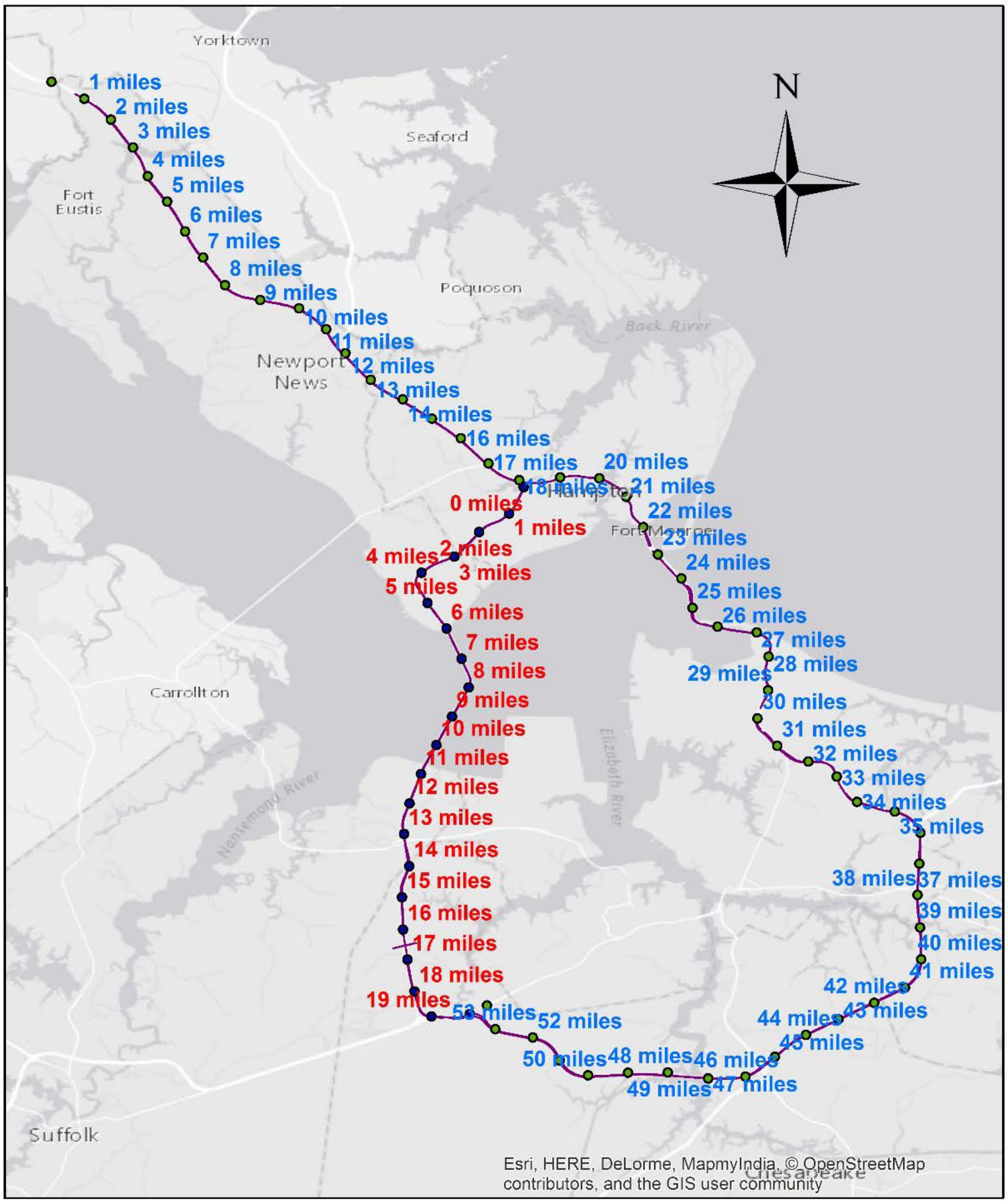


Figure 3 Mile posts for I-64 EB and I-664 SB

Reliability Comparison of Freight and General Traffic Travel Times

While box plots are useful to understand the extent of bias between freight and general traffic speeds, another metric of considerable importance would be the degree of variation in speeds at any given location and time. It is possible that the average freight and general traffic speeds are very close but there is greater variation in speeds for a certain type of traffic. To quantify this variation about the mean, travel time reliability metrics were computed for freight and general traffic. According to FHWA, travel time reliability is defined as the "*the consistency or dependability in travel times, as measured from day-to-day and/or across different times of the day*" (FHWA, 2006). The traffic speed estimates along the unit-mile grid were converted into travel times by dividing one mile distance by corresponding speeds and a reliability metric known as the "Buffer Time Index (BTI)" was computed for each unit mile segment and time period as follows:

$$Buffer\ Time\ Index\ (BTI)\ (\%) = \frac{95^{th} Percentile\ Travel\ Time - Average\ Travel\ Time}{Average\ Travel\ Time} \times 100$$

BTI is one of the metrics recommended by FHWA for quantifying reliability and is defined as the additional time (beyond the average travel time) that travelers must budget to ensure on-time arrival at a given level of confidence (commonly 95%). For example, if BTI is 50% and average travel time is 30 minutes, travelers must plan for 45 minutes (= 30+0.5x30) of travel time to ensure on-time arrival with 95% confidence. To ensure that extreme outliers in the data do not skew the average travel time measure used in the BTI calculation, a Winsorisation technique in which the most extreme outlier was recoded to the next highest value was used.

Clustering Analysis

For each unit mile stretch along the I-64 and I-664 corridors, for any given time period, three metrics can be computed– (1) average difference between freight and general traffic travel times, (2) BTI of freight traffic travel times, and (3) BTI of general traffic travel times. One of the questions of interest is "Which segments have significant differences between freight and general traffic characteristics?". Clustering techniques are particularly suited to answer this question because clustering will group data into different segments such that objects in different groups are very different from each other and objects in the same group are very similar to each other. There are two commonly used clustering methods in the data mining literature – K-means clustering and Hierarchical clustering. Within hierarchical clustering, there are two types of clustering algorithms – agglomerative and divisive. In the agglomerative clustering, every data record starts in its own cluster and iteratively clusters are merged based on dissimilarity metric. In the divisive clustering, all records are grouped together in a single cluster initially and are iteratively divided. For very large datasets, K-means clustering is preferred because hierarchical clustering can be very slow owing to its higher computational complexity. Also, unlike in hierarchical clustering, the number of clusters must be pre-specified in K-means clustering. In hierarchical clustering, the analyst can choose the number of clusters after the analysis using visual output referred to as "dendrogram" in which data records are arranged in a tree structure from the closest to the farthest (Everitt *et al.*, 2011, Zaki and Meira Jr, 2014). In this study, the Ward's agglomerative hierarchical clustering method was adopted using Euclidean distance as the dissimilarity metric because the number of data records is low (for instance, 54 records corresponding to 54 unit mile segments along I-64 WB) (Murtagh and Legendre, 2014). For each unit mile stretch, two attributes were chosen to undertake the clustering analysis - (1) the average difference between freight and general traffic speeds, and (2) the ratio of freight and general traffic BTIs.

Explore the Potential Use of ATRI GPS Data

Given that ATRI data tracks each truck as it moves along major freight corridors, this data can be used to support more advanced modeling efforts such as route choice analysis in addition to monitoring performance measures. For instance, in the current context, it would be useful to understand the route choice preferences of trucks that are leaving the Hampton Roads region. These trucks have two options: take I-64 WB or I-664 NB. These two alternate routes between the origin and destination points have about the same distance (~ 28 miles). I-64 WB usually experiences bottlenecks primarily due to congestion at the HRBT. Using the GPS data, the number of trucks that were traveling along each corridor and eventually exited the study region was obtained. From the raw ATRI data, the average travel time for this OD pair along the two routes was estimated. These travel time estimates along the two routes were then compared to check if they are consistent with observed truck flows along these two routes.

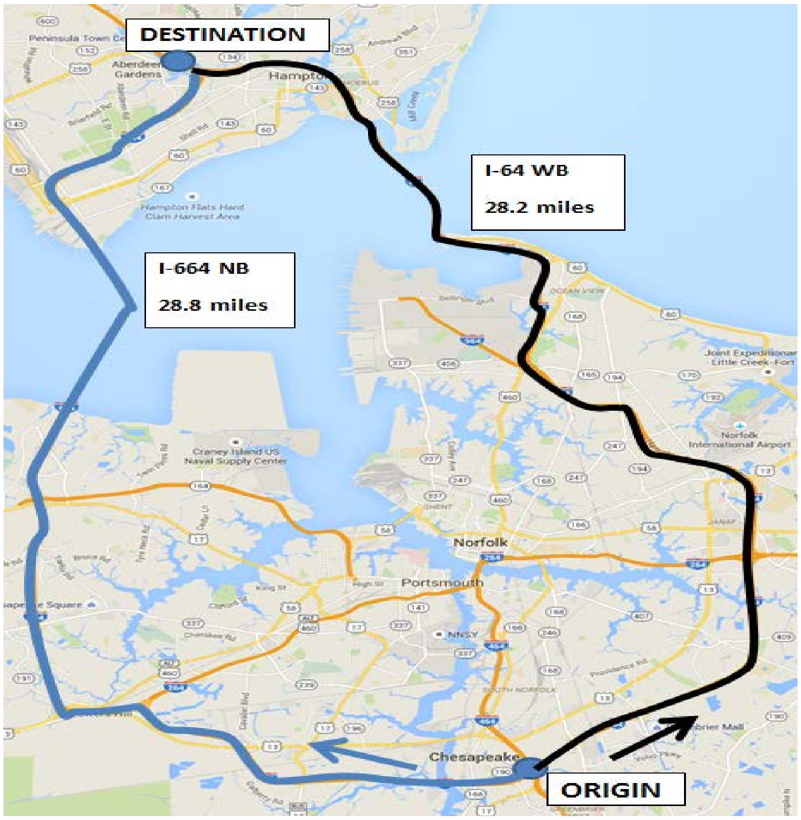


Figure 4 Two Alternate Routes for Exiting Hampton Roads Region

FINDINGS

The major findings from the study are summarized as follows:

1. Both HERE and INRIX data capture the major trends in PMs both temporally and spatially. There are no systematic differences in the spatial and temporal locations of congestion zones in the two data sources along the two corridors- I-64 and I-664.
2. Typically, freight speeds are lower than general traffic speeds at all locations (by about 3 mph in HERE data). Also, the magnitude of difference between freight and general traffic speeds is higher in INRIX data compared to HERE data. Specifically, speed differences computed with respect to general traffic speeds from INRIX data were nearly twice the speed differences computed with respect to HERE data. This suggests that INRIX general traffic speed estimates are higher than the corresponding estimates from HERE data.
3. The average differences between freight and general traffic speeds were similar during weekdays and weekends. However, there is greater variation in traffic conditions during weekdays compared to weekends. For instance, the BTI during peak period (3 to 6 PM) for I-64 WB was about 92.1% during weekdays compared to 67.1% during weekends.
4. Although the differences between general traffic and freight speeds vary by traffic conditions, it is difficult to discern the effect of traffic conditions just from the box plots. A more thorough investigation is needed to separate the effect of traffic conditions from other factors such as roadway geometry.
5. <u>I-64</u>: For I-64 WB, the differences were more pronounced slightly downstream the I-264 interchange, slightly downstream of the tunnel exit location, and the road segment past mile post 46. Truck speed reduction at this location can be linked to a number of factors particularly the geometric alignment or layout of the road segments, e.g., absence of shoulders and relatively narrower lanes inside the tunnel, high road gradient downstream of the tunnel exit, and two lanes being dropped past mile post 46 in the I-64 WB direction. Similar observations were made for I-64 EB at road segments slightly downstream I-664 interchange, Terminal Boulevard, I-264 and I-464 interchanges.
6. <u>I-664</u>: The travel times showed much lower variation along I-664 NB compared to I-664 SB. For instance, the average BTI during peak periods for general traffic along I-664 NB was about 23% compared to 54% along I-664 SB.
7. In general, the results indicate that there is greater variation in freight travel times as indicated by higher BTI values for freight traffic compared to general traffic. For instance, along I-64 WB during peak hours, the average BTI for freight traffic was 102% compared to 62% for general traffic.
8. <u>Cluster Analysis</u>: The hierarchical clustering analysis identified three types of segments along I-64 WB. The roadway segments in the cluster with highest membership (nearly 38 segments out of 53 belonged to this cluster) have similar average travel times as well as reliability for freight and general traffic. The second cluster constitutes 10 roadway segments with significant differences in reliability and slightly bigger differences in average travel times between freight and general traffic. The third cluster composed of 5

roadway segments with insignificant average travel time and reliability differences between freight and general traffic. The clustering analysis along I-64 EB also identified three clusters. The cluster with highest membership (42 segments out of 53) and the second cluster with high reliability differences (five segments) were very similar to the first two clusters identified along I-64 WB. However, there were slight differences in the third cluster along I-64 WB and EB directions. The third cluster in I-64 EB composed of seven segments with almost same average travel times for freight and general traffic but higher variability in freight travel times. Overall, these results suggest that while for most roadway segments general traffic PMs serve as a good proxy for freight PMs (both in terms of average travel times and reliability), the same does not hold true for several other roadway segments either because the average travel times are different or because the degrees of variation about mean are different between freight and general traffic.

9. Use of ATRI Data: The route choice analysis undertaken using ATRI data suggests that there can be nearly five minutes of travel time savings for each truck during peak hours if I-664 NB is chosen over I-64 WB. The travel time differences, however, are not significant during off-peak hours. Moreover, I-664 NB is more reliable (less variability in travel times) than I-64 WB both for general and freight traffic. Also, the truck flow numbers confirm that truck drivers are aware of the shorter travel times if they take I-664 NB rather than I-64 WB but there is still some potential for travel time savings. These findings underscore the utility of ATRI data to support freight modeling efforts including OD matrix estimation and route choice modeling. However, ATRI data provides good coverage primarily for interstates and may not be adequate to support similar analysis along other non-interstate roads with significant freight traffic (for instance, Hampton Boulevard and Terminal Boulevard in the current study region).

CONCLUSIONS

This research developed a general framework for comparing freight and general traffic performance measures (PMs) using multiple data sources that includes visual data analytics, statistical comparisons using box plots, reliability calculations, and cluster analysis. Irrespective of the data source, the results suggest that there are differences in freight and general traffic characteristics with freight travel times being slightly higher as well as having lower reliability (greater variation) compared to general traffic travel times. However, the differences between freight and general traffic are insignificant on less congested corridors (e.g., I-664) and more pronounced at bottleneck regions such as interchanges, lane reduction areas, and tunnel exit locations. Also, while the average speed and travel time reliability comparisons of freight and general traffic have similar trend for most segments, there are also a small cluster of segments that have bigger differences in average speeds but similar reliability and vice versa. Although the findings in this study may not be generalizable as the data used is specific to two major roadways in the Hampton Roads region, the methods and the framework developed in this study may be applied to any region to quantify differences between freight and general traffic PMs. In all data sources considered in this study, speed information was not available on a continuous time scale which hampered analysis at a finer temporal resolution (all the analysis was undertaken at 15-minute temporal resolution). Also, it is important to note that the analysis provided in this study is based on one month data for September, 2014. A similar study on yearly data is needed to capture the seasonal variation of traffic characteristics.

RECOMMENDATIONS

1. The analyses suggest that there can be considerable differences either between average travel times or variation about average travel times of freight and general traffic. However, these differences are more pronounced at congested locations along highways where there are major interchanges, lane drops, tunnel entrance and exit points, and upstream/downstream stretches of tunnels depending on the grade of on/off-ramps. So, VDOT can first identify all congested corridor segments in the state and purchase freight data only for these locations. For other locations, general traffic data would serve as a reasonable proxy for freight traffic.
2. Travel times during peak periods along I-664 are lower and have lower variation over different days compared to I-64. From a traveler's perspective, reliability is more important than average travel times. So, overall traveler experience can improve significantly if VDOT can effectively communicate higher reliability of I-664 compared to I-64 to the trucking community because choosing I-664 not only offers travel time savings to truck drivers but also potentially lessens congestion along I-64 and the HRBT corridor because of the reduced truck traffic.
3. As demonstrated in this study, ATRI data can be used to identify alternate routes and their corresponding travel times and reliability. Also, although not undertaken in this study, the ability to track each truck as it traverses along the highway network enables identification of major truck OD patterns in the region. Such analysis can not only be used to devise effective traffic management and operation strategies but also to support freight planning models. So, VDOT has considerable utility in purchasing the ATRI data at least for regions with major freight activity.

COMPLETE DOCUMENTATION

Data Gathered

Datasets used in this study were obtained from three different sources, namely: INRIX, HERE and ATRI. All of these datasets provided speed (travel time to cover a specific Traffic Message Channel or TMC) of general traffic, passenger traffic or freight traffic. Brief description of the datasets follows.

HERE Dataset: HERE data includes the travel time of general traffic, passenger vehicles and freight trucks for each TMC. The data also includes timestamp information of day-of-the-month and five minute epoch during the day. So, HERE dataset provided speed information at a resolution of 5 minutes aggregation intervals. Table 1 shows a sample of the HERE dataset. The temporal coverage of the HERE dataset was from January to December of 2014. However, since other datasets that were used in this study were available only for the month of September 2014, only HERE data corresponding to September 2014 were used for the analysis presented in this study. The spatial coverage of the HERE dataset includes all major highways in the nation. The total number of TMCs in HERE data is 5,937. However, since this study mainly focused on comparing the PMs of different vehicle types/classes in the Hampton Roads region of Virginia, only data records corresponding to the study area were extracted from the complete dataset. In total, there are 211 TMCs covering the Hampton Roads in HERE dataset. Together these TMCs cover I-64, I-664, I-264, I-564, I-464, US-17, VA-337 and US-60. The spatial coverage of HERE dataset in the Hampton Roads region is shown in Figure 5. In Figure 5, the dots represent the start and end of HERE TMCs.

Table 1 Sample of HERE Dataset

TMC	DATE (DDMYYYY)	EPOCH	TT_ALL_VEHICLES (in sec.)	TT_PASSENGER_VEHICLES (in sec.)	TT_FREIGHT_TRUCKS (in sec.)
110N04123	1022014	67	54	54	NA
110N04123	1022014	127	55	55	NA
110N04123	1022014	157	58	58	NA
110N04123	1022014	187	66	66	NA
110N04123	1022014	217	226	226	NA
110N04123	1022014	247	86	86	NA

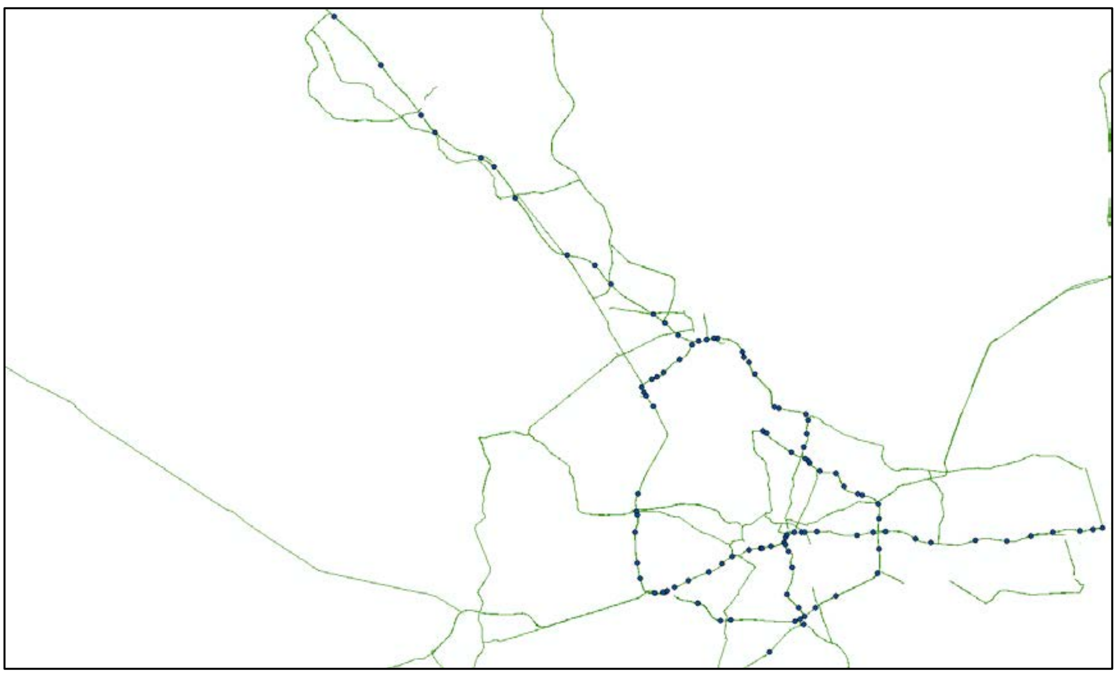


Figure 5 HERE TMCs in Hampton Roads Region

Preliminary analysis on the HERE dataset for the Hampton Roads region and for the month of September 2014 found that speed information was missing for a significant proportion of the records. Table 1 shows the proportion of missing entries corresponding to general traffic, passenger traffic and freight traffic in HERE dataset for the corridors I-64 and I-664. To further improve workability of the data, the aggregation interval of the datasets was increased from 5 minutes to 15 minutes. This way, the magnitude of the proportion of missing entries was on average reduced by nearly 50%.

Table 2 Proportion of Missing Records in HERE Data

Aggregation Interval	Proportion of missing data		
	TT_ALL_VEHICLES	TT_PASSENGER_VEHICLES	TT_FREIGHT_TRUCKS
5 minutes	29%	36%	66%
15 minutes	11%	16%	41%

Figure 6 shows the distribution of the missing entries for general traffic, passenger traffic and freight trucks by hours of the day aggregated across the entire month of September 2014. The majority of the missing entries correspond to off-peak hours, particularly during early morning hours. So, for the purposes of this study, peak period data is the most suitable dataset.

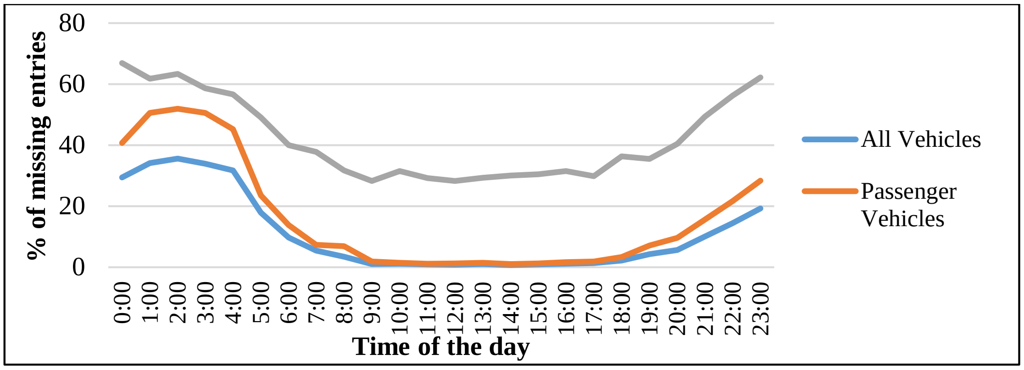


Figure 6 Proportion of Missing Records in HERE data by Traffic Type

INRIX Dataset: The second dataset used in this study was obtained from INRIX- freight data for the month of July, 2012 and both freight and general traffic data for entire 2014. The July 2012 freight dataset includes 537 TMCs that together span I-64, I-664, I-264, I-564, I-464, US-460, US-58, US-60, VA-164, VA-168, VA-199 and VA-337 (see Figure 7). Also, INRIX freight dataset provided speed estimates for a typical day (one average Monday, one average Tuesday *etc*) in July 2012 at a temporal resolution of 15 minutes. The dataset provides speed and travel time estimates along each TMC for freight traffic along with the count of the probe-vehicles and the standard deviation of the information obtained from the sample vehicles. Table 3 shows a sample of the INRIX freight data. The limitation of this dataset was that speed and travel time estimates were available for average days in a week in July 2012 rather than for each individual day in the month. For example, the freight speed information for all the Thursdays in July 2012 were averaged and provided as typical freight speed information for an average Thursday in July 2012. Such aggregated representation of traffic fails to capture the variability of traffic conditions across days of the weeks. Given that traffic is dynamic and its pattern does not remain the same for the same day across different weeks in a month, it was decided that such aggregation would create bias in the analysis. Figure 8 shows the traffic patterns during all Thursdays in July 2012 in terms of heat-maps of the speed of traffic for the HRBT corridor along I-64 EB. It is evident from this figure that the spatial and temporal extent of congestion (*i.e.*, the darker region in the heat-maps) is not the same in all the heat-maps. Therefore, INRIX freight data was deemed to be unusable for the purposes of this study and was excluded from further analysis. INRIX general traffic data for 2014, however, was available for all days in each month. So, this dataset was used for comparing freight and general traffic PMs in this study. For this purpose, INRIX speed and travel time dataset along the I-64 and I-664 was extracted for September 2014. Also, the INRIX general traffic data was reported at temporal resolution of 15 minutes.

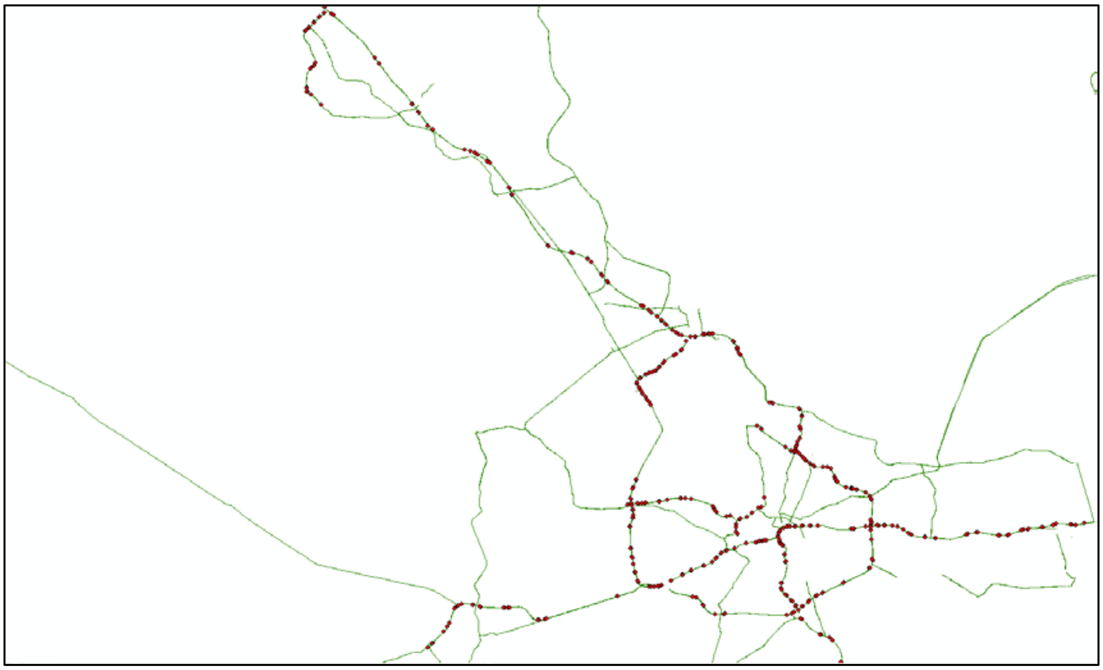


Figure 7 Spatial Coverage of INRIX Freight Data for July 2012

Table 3 Sample of INRIX Freight Data for July 2012

TMC	Day-Of-Week (1 to 7)	Minute	Speed (MPH)	Travel Time (Seconds)	Stddev	Point Count
110+04121	1	15	60	80	NA	NA
110+04121	1	30	60	80	NA	NA
110+04121	1	45	60	80	NA	NA
110+04121	1	60	60	80	NA	NA
110+04121	1	75	55	87	2.0817	3
110+04121	1	90	54	89	1.5275	3

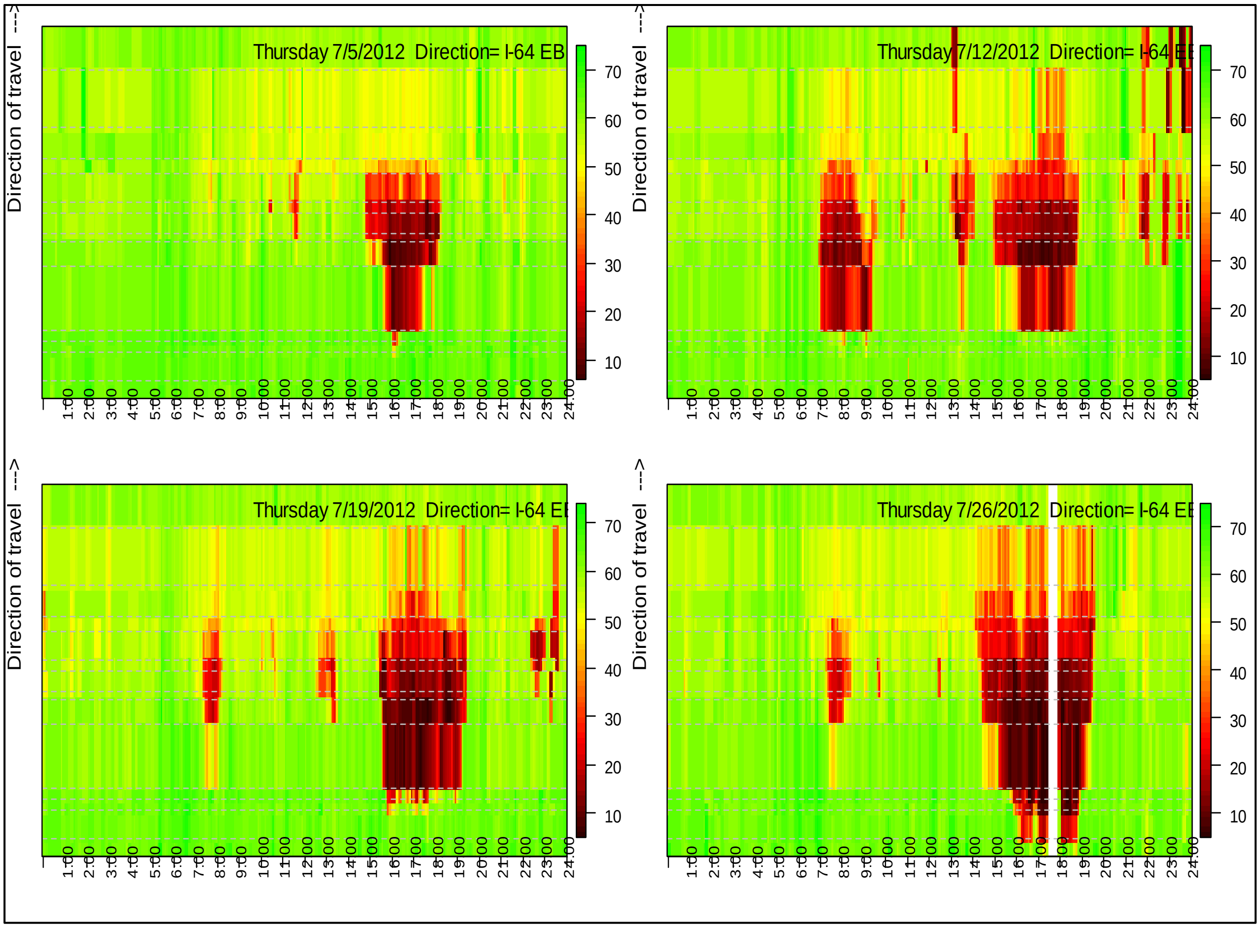


Figure 8 Variation in Traffic Patterns across Different Thursdays in July 2012

ATRI Dataset: The third data source used in this study was obtained from the American Transportation Research Institute (ATRI) for July, 2012 and September, 2014. ATRI data was collected by tracking GPS instrumented freight vehicles. The data consists of unique identification number of the trucks, location (latitude/longitude) of the trucks, their spot speed and direction of travel along with a timestamp indicating when the measurement was taken. In the recent times, ATRI data has been successfully used for freight modeling and planning (Pinjari *et al.*, 2014). The dataset for July 2012 consists of 761,595 entries while the dataset for September 2014 consists of 7,484,089 entries. Between 2012 and 2014, the coverage and quality of ATRI has improved significantly. Even though the number of data records may seem to be a lot, they represent point measurements of a sample of freight vehicles along the entire Hampton Roads for an entire month. Further analysis of the datasets revealed that majority of the measurements in the ATRI datasets were taken when the freight vehicles are parked or stopped for deliveries. Figure 10 shows the speed distribution of the vehicles in the 2014 ATRI dataset. The time interval between two successive measurements for a specific truck was also analyzed and was found to be non-uniform, *i.e.*, the measurements were not taken at a regular time intervals (see Figure 11). The analysis revealed that the average time interval between two successive ATRI data points is one minute but there are several records with time beyond successive recordings beyond 5 minutes.

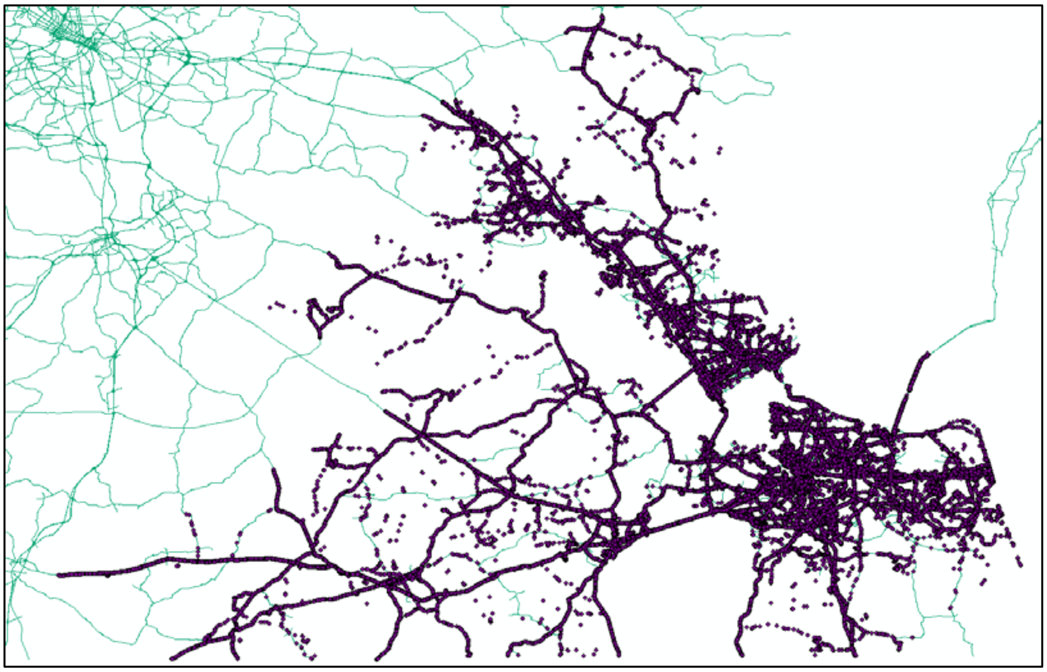


Figure 9 ATRI Truck GPS Points in the Hampton Roads region

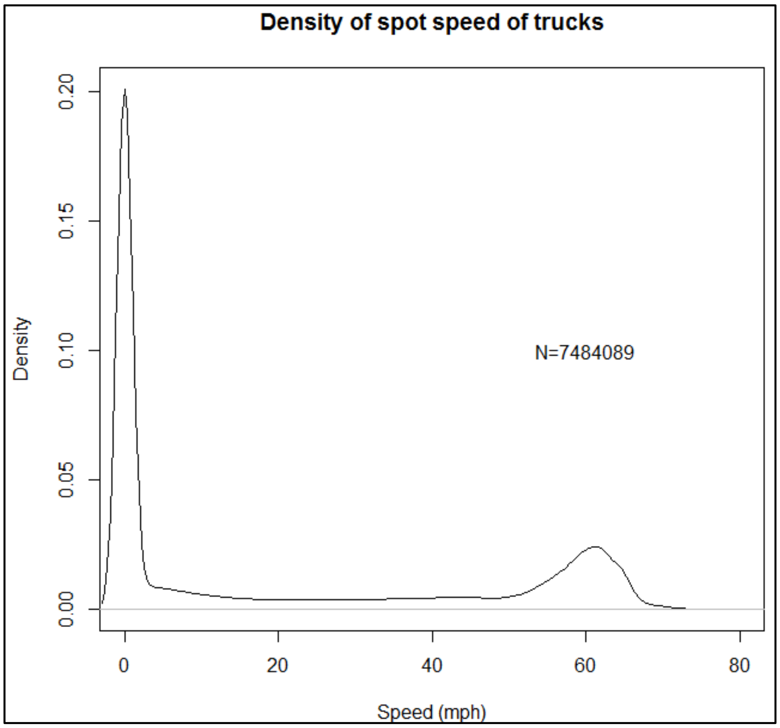


Figure 10 Density Plot of Spot Speed of Trucks in ATRI Dataset

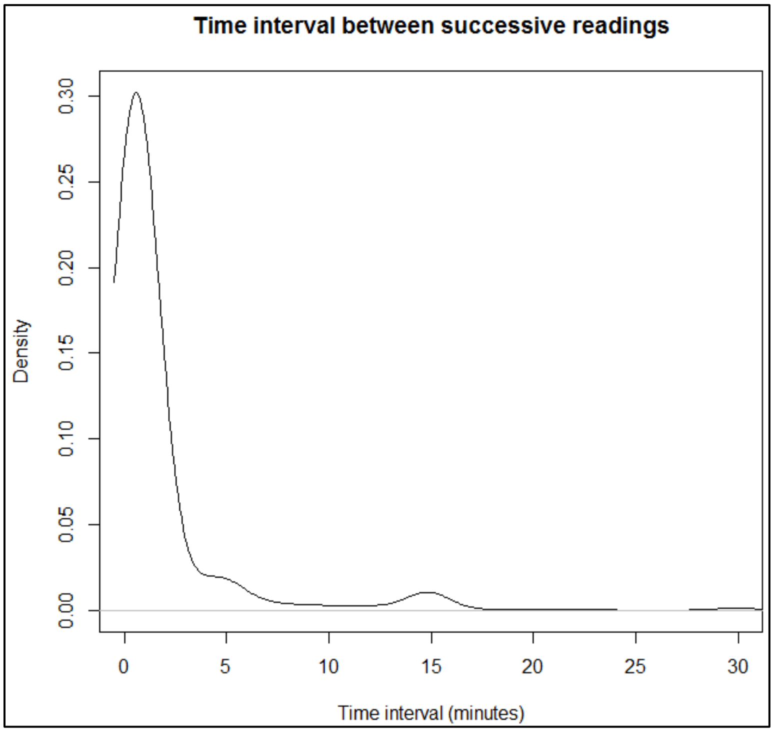


Figure 11 Density Plot of the Time Interval between Successive GPS Points

Proximity tools in ArcMap were employed to snap the GPS points in the ATRI dataset to the nearest INRIX TMC. INRIX TMCs along the major roads are shown in Figure 9 as green polylines (note that not all roads have TMC identifications, this is particularly true for local residential roads). Matching the ATRI data points to the nearest INRIX TMC was done so that speed information in the ATRI datasets can be appended to TMCs. This would make the comparison analysis of PMs across different data sources easier.

Figure 12 (a) shows the distribution of distances of ATRI data points from the nearest INRIX TMCs. It can be seen from the figure that most of ATRI data points are located very close to INRIX TMCs (more than 50% of the ATRI data points where located less than 100ft from the nearest TMC). As shown in Figure 12(b), a major proportion of the ATRI data points located farther than 300ft from the nearest TMC have very low speeds. This suggests that these GPS recordings were mostly taken when the trucks were parked. So, ATRI data points located farther than 300ft from the nearest TMC were excluded from the subsequent analysis.

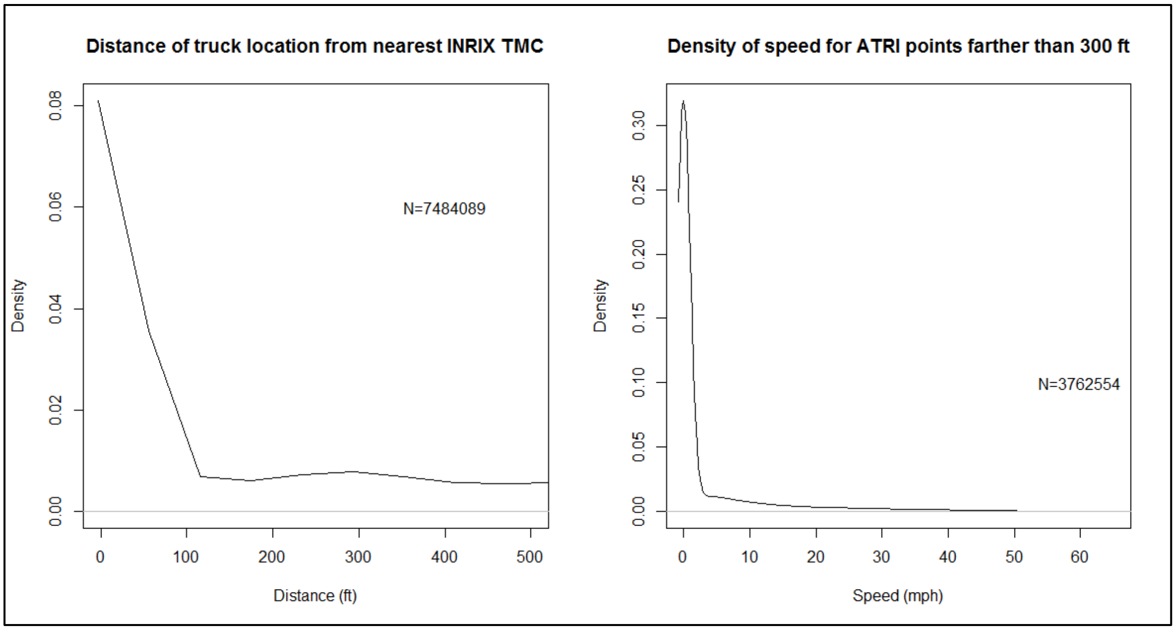


Figure 12 Density Plot of Truck Location & Speed Relative to INRIX TMCs

Figure 13 shows the distribution of speed of the ATRI data points after the data points located farther than 300ft from nearest TMC were removed. Even after removing these points, the speed density plot shows there are significant amount of ATRI points with low speed entries.

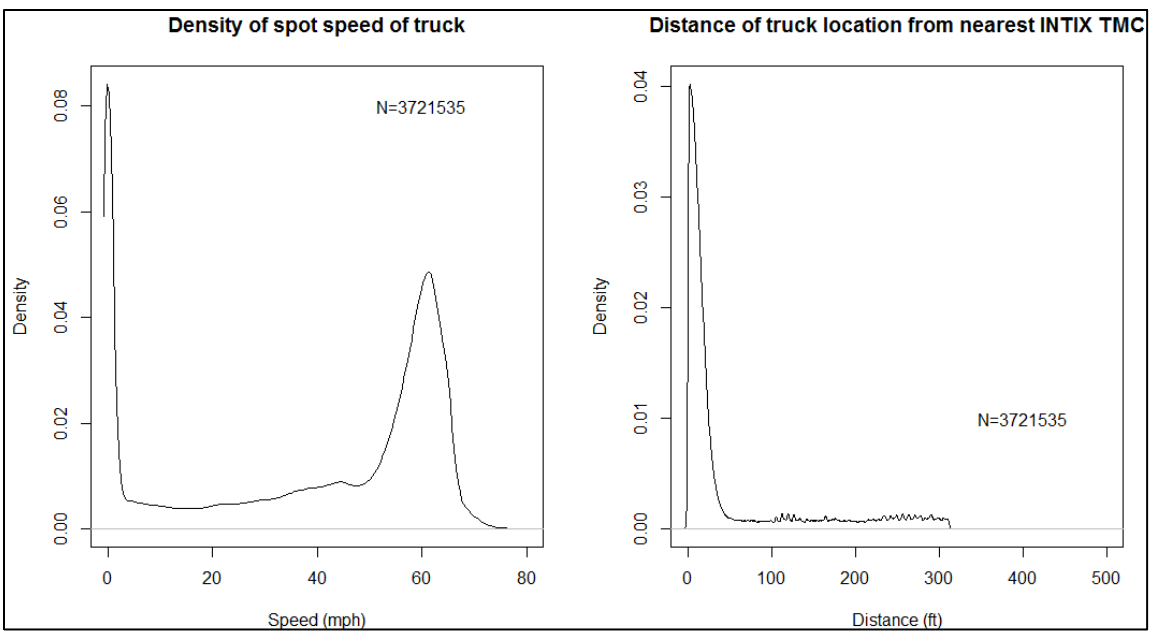


Figure 13 Density Plot of Truck Location & Speed Relative to INRIX TMCs (within 300 ft)

The distribution of spot speeds along I-64 and I-664 and their distance from the nearest TMCs are shown in Figure 14. The speed distribution shows that the truck speeds were mostly around the free flow speed of 60mph while lower speeds were also observed representing congestion.

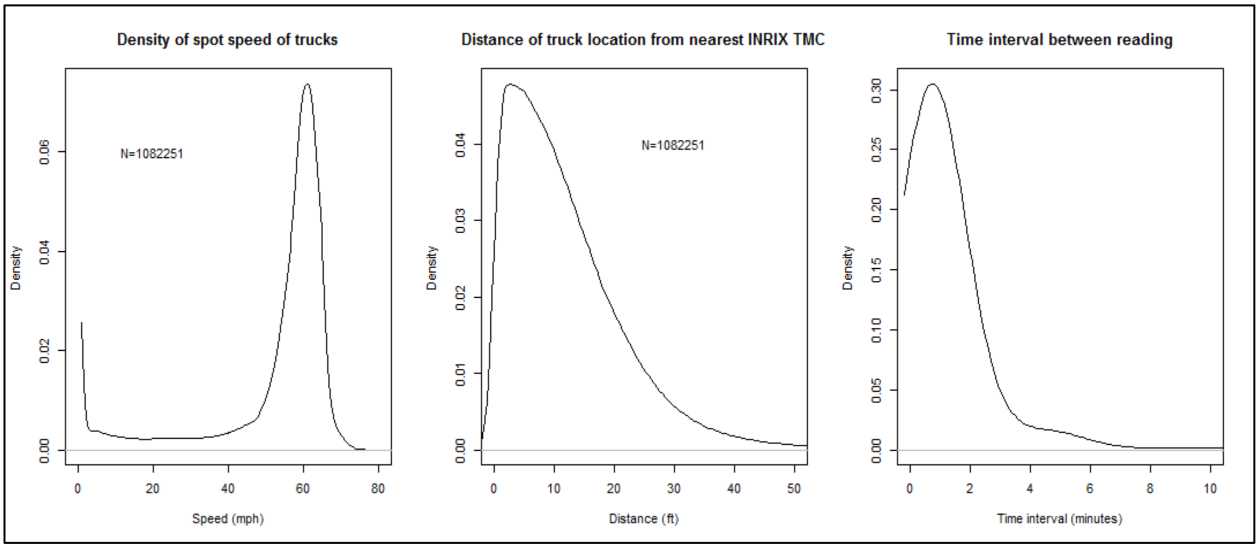


Figure 14 Distribution of Spot Speed, Distance from Nearest TMC and Time Interval between Successive GPS Recordings along I-64 and I-664

In some cases, even though successive GPS recordings for the same truck indicate that the truck has moved, spot speed was recorded as zero. In such cases, speed was estimated by dividing distance traversed by travel time. Travel time can obtained directly as the difference between the timestamps of the two ATRI data points. However, estimating the distance traveled is slightly challenging. Summation of the Euclidean distances between all consecutive ATRI data points between point A and point B may not provide accurate distance estimate because the freight vehicle might not have travelled along a straight path (due to roadway geometry). Instead, average speed was estimated by summing the length of the TMCs between point A and point B as shown in Figure 15. The distance travelled composes of three parts, namely:

- D1: Distance between the ATRI point A and the start of the first complete TMC
- D2: Summation of the length of complete TMCs between ATRI points A and B
- D3: Distance between the ATRI point B and the end of the last complete TMC

D1 and D3 were estimated using Haversine's great circle distance as (Hijmans *et al.*, 2015):

$$d = 2r \arcsin\left(\sqrt{sin^2\left(\frac{\phi_2 - \phi_1}{2}\right) + \cos(\phi_1)\cos(\phi_2)\sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right)$$

where: *d* is the distance along the great circle between two points
r is the radius of the earth (i.e., 6378137 meters)
$Ø_1$ and $Ø_2$ are the latitude of the points between which distance is being calculated
λ_1 and λ_2 are the longitude of the points between which distance is being calculated

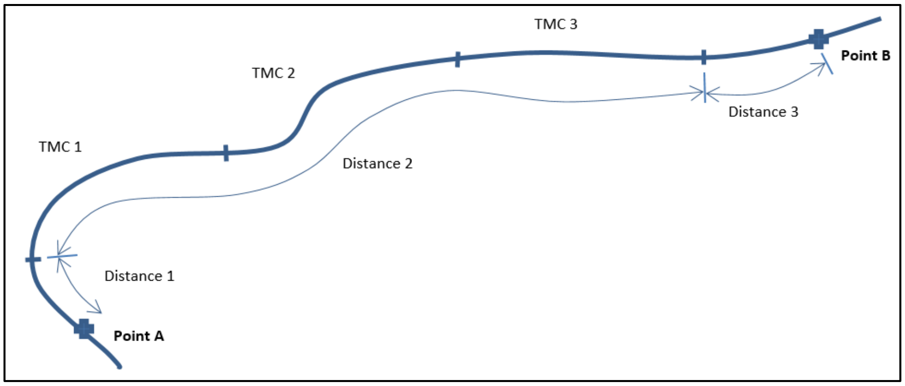


Figure 15 Estimating Distance between ATRI Data Points A and B

Using the speed estimates from time stamps and spot speed information from the ATRI data points, a new dataset of freight speed information aggregated at 15 minutes temporal resolution was developed. Figure 16 shows the availability of freight speed information for the major roads in the Hampton Roads region for different hours of the day across the entire month of September 2014. As can be seen in Figure 16, majority of the data for I-264, I-564, I-464, US-17, VA-337 and US-60 was missing. So, ATRI freight data for these roads was deemed unusable for the purposes of this study. Even along I-64 and I-664, ATRI data had good coverage peak hours compared to off-peak hours. It is worth mentioning that when assigning ATRI data points to the nearest TMC, the bearing of the nearest TMC and bearing of the ATRI data point should match. If the bearings of the ATRI data point and the first nearest TMC didn't match, the ATRI data point is assigned to the next nearest TMC whose bearing is matching. ATRI data points whose bearing did not match with the bearing of the first and second nearest TMCs were discarded from the data and were not used for any analysis (such ATRI data points were only 0.4% of the entire data set and most of them had low speed measurements).

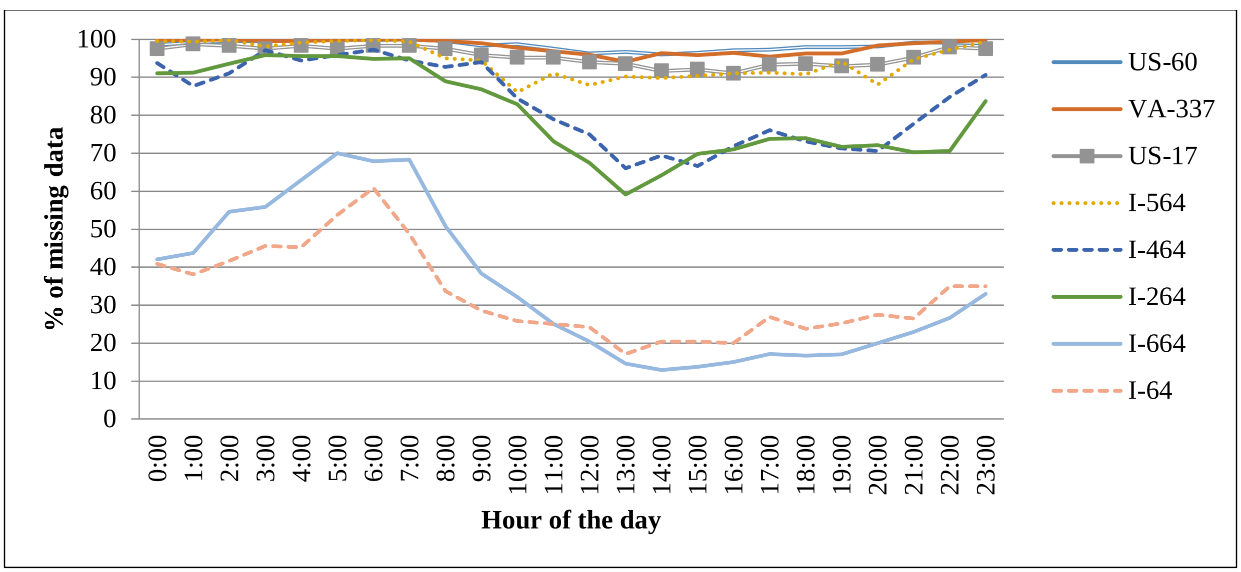


Figure 16 Proportion of Missing Speed Information in ATRI Data along Major Roadways

Figure 17 depicts the coverage of ATRI data at 4 pm along major roadways in the Hampton Roads region. It can be seen that for all roadways other than I-64 and I-664 the number of truck trips at 4 pm is mostly less than 50. Overall, the analysis suggests that ATRI data is primarily suited for analyzing truck traffic along interstates but not local roads.

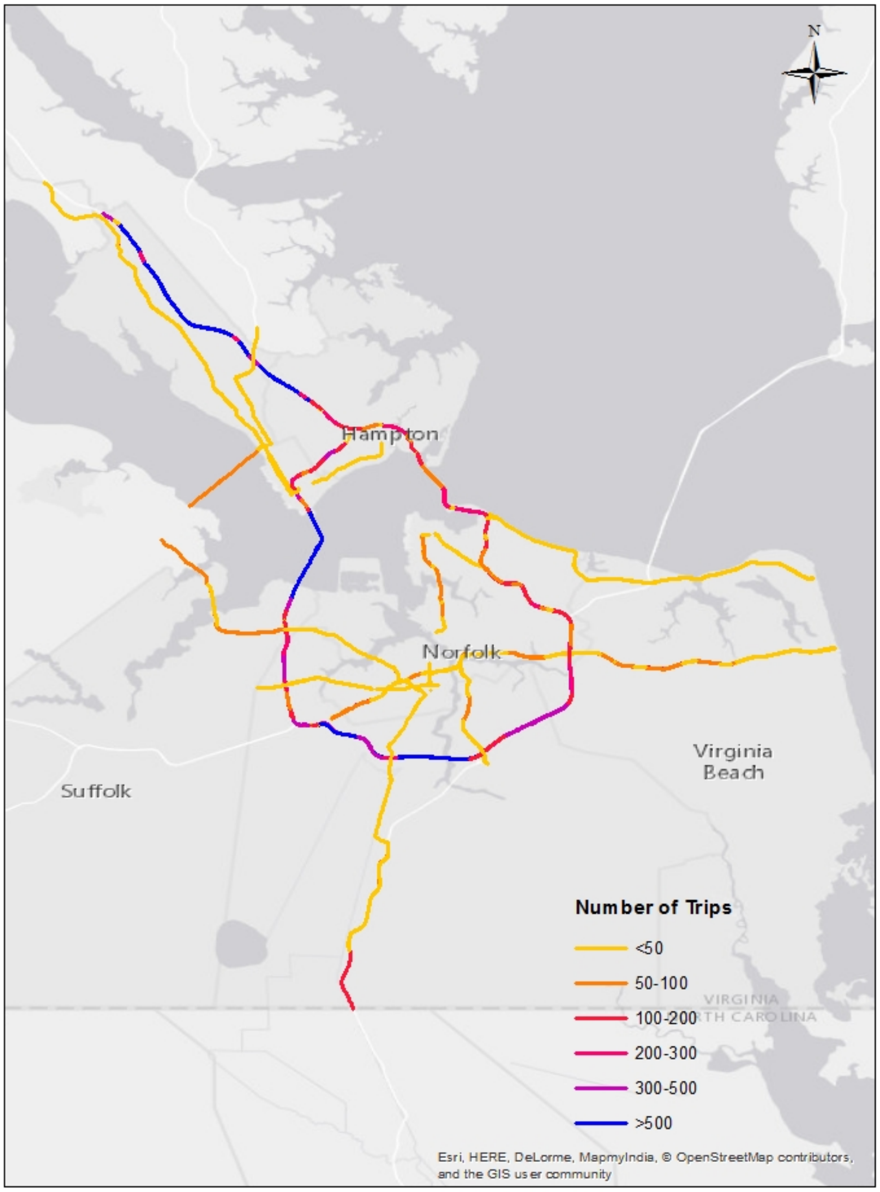


Figure 17 Coverage of September ATRI Data at 4 PM

\

Analyses Performed & Results Achieved

Data Visualization Using Heat Maps

Figure 18 depicts sample speed heat maps for and INRIX general traffic, HERE general traffic, and HERE freight speeds for the I-64 WB corridor between the 21st and 30th mile posts on September 8th (Monday), respectively. The white patches in the speed heat maps represent instances where speed information was missing in the corresponding datasets. In each of these maps, the horizontal axis indicates the hour of the day whereas the vertical axis indicates distance along the unit mile grid. It can be seen from these maps that both HERE and INRIX data capture the major trends in PMs both temporally and spatially. There do not seem to be any systematic differences in the spatial and temporal locations of congestion zones in the two data sources. These heat maps serve as a validation of the data processing and grid transition work that was undertaken to reconcile the differences in the TMC definitions in the INRIX and HERE data sources.

Speed Comparison Using Box Plots

Figure 19 to Figure 25 are box plots showing differences between freight and general traffic speeds for the evening peak period (between 3 and 6 PM) along I-64 WB for each average day (one typical Monday, one typical Tuesday, and so on). The horizontal axis indicates the mile post location along I-64 WB whereas the vertical axis indicates the difference between general traffic and freight speeds. At the top of each boxplot, there are two rows of numbers. The first (top) row of numbers indicates the average speed of general traffic at the corresponding mile post location where as the second (lower) row indicates the number of data points used to construct the boxplot at that mile post location. At any given mile post location along the horizontal axis, the number of data points used to construct the box plot is approximately 48 (=4x12) because each day appears about 4 times in a month and there are twelve 15-minute intervals in the 3 hour peak period (3 to 6 PM). In some cases, there can be fewer data points because of missing data or more data points because a typical day may appear five times in a month. The box plots also show the location of major interchanges, tunnel entrance and exit points, location of major drop in lanes, and other key locations along I-64 WB. Several important observations may be made from these box plots.

- First, typically freight speeds are lower than general traffic speeds at all locations.
- Second, the magnitude of difference between freight and general traffic speeds is higher in INRIX data compared to HERE data. Specifically, speed differences computed with respect to general traffic speeds from INRIX data are higher than speed differences computed with respect to HERE general traffic speeds. This indicates that INRIX general traffic speed estimates are higher than the corresponding estimates from HERE data.
- The average differences between freight and general traffic speeds were similar during weekdays and weekends. However, there is greater variation in traffic conditions during weekdays compared to weekends.
- Fourth, although the differences between general traffic and freight speeds vary by traffic conditions, it is difficult to discern the effect of traffic conditions just from the box plots. A more thorough investigation is needed to separate the effect of traffic conditions from other factors such as roadway geometry.

- Lastly, the differences are more pronounced slightly downstream the I-264 interchange, slightly downstream of the tunnel exit location, and the road segment past mile post 46. Truck speed reduction at this location can be linked to a number of factors particularly the geometric alignment or layout of the road segments, e.g., absence of shoulders and relatively narrower lanes inside the tunnel, high road gradient downstream of the tunnel exit, and two lanes being dropped past mile post 46 in the I-64 WB direction.

Similar boxplots were produced for other corridors- I-64 EB, I-664 NB, and I-664 SB and are included in the Appendix. It may be observed from boxplots for I-664 (Figure 33 and Figure 34 in Appendix) that the speed differences between freight and general traffic are very low and have much lower variability compared to I-64.

Travel Time Reliability Comparison

Figure 26 shows the profile of BTI for freight and general traffic travel times for I-64WB during a peak hour (4 PM) for Mondays, (Tuesdays, Wednesdays, and Thursdays), Fridays, and weekends. The numbers at the top of the plots in these figures indicate the number of samples underlying the BTI calculation. The BTI plots for other corridors are included in the Appendix. Four important observations may be made from these BTI profiles:

- There is greater variation in freight travel times compared to general traffic travel times.
- INRIX data has lower variation compared to HERE data as indicated by lower BTI values for general traffic in INRIX compared to HERE.
- There is greater variation in travel times on weekdays compared to weekends.
- The locations with greater variation in travel times are consistent with the locations identified using boxplots. For example, on Mondays, BTI values are higher slightly downstream the I-264 interchange, between terminal boulevard and 4^{th} view, slightly downstream the tunnel exit point.

Cluster Analysis

Figure 27 and Figure 28 provide the dedrograms obtained from the hierarchical clustering analysis of unit mile roadway segments along I-64 WB and EB corridors, respectively. Also, Table 4 and Table 5 provide the mean attribute values and membership of the clusters identified for these two corridors, respectively. It can be seen that majority of roadway segments belonged to the cluster in which the ratio of freight and general traffic BTIs was close to 1 and average speed difference was close to zero. This suggests that general traffic PMs can characterize freight traffic reasonably well in most cases. However, there are nearly 15 segments for which this assumption does not hold true either because there is greater variability in freight traffic conditions across different days or because the average difference in freight and general traffic speeds is non-zero.

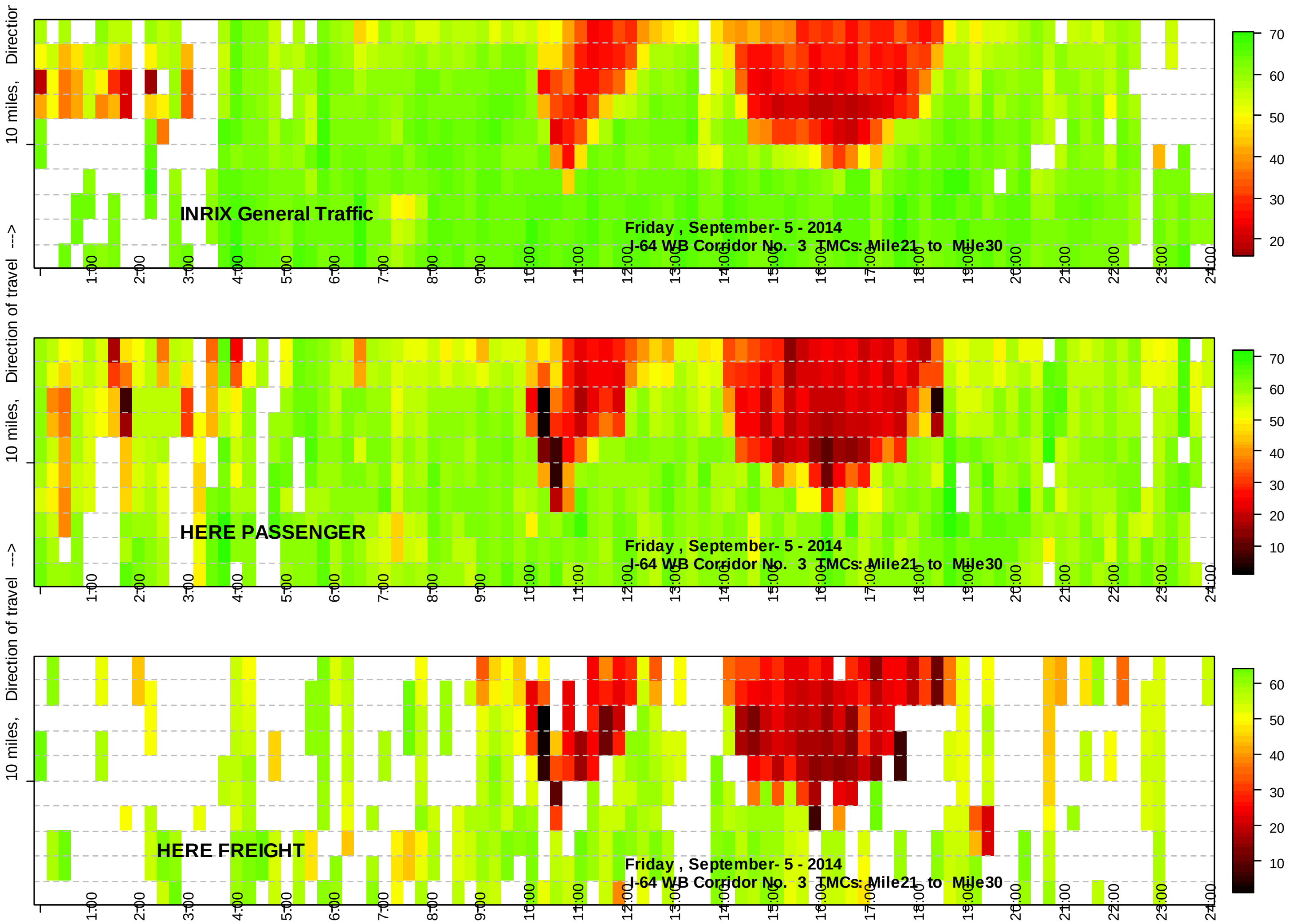


Figure 18 Sample Speed Heat Maps for I-64 WB

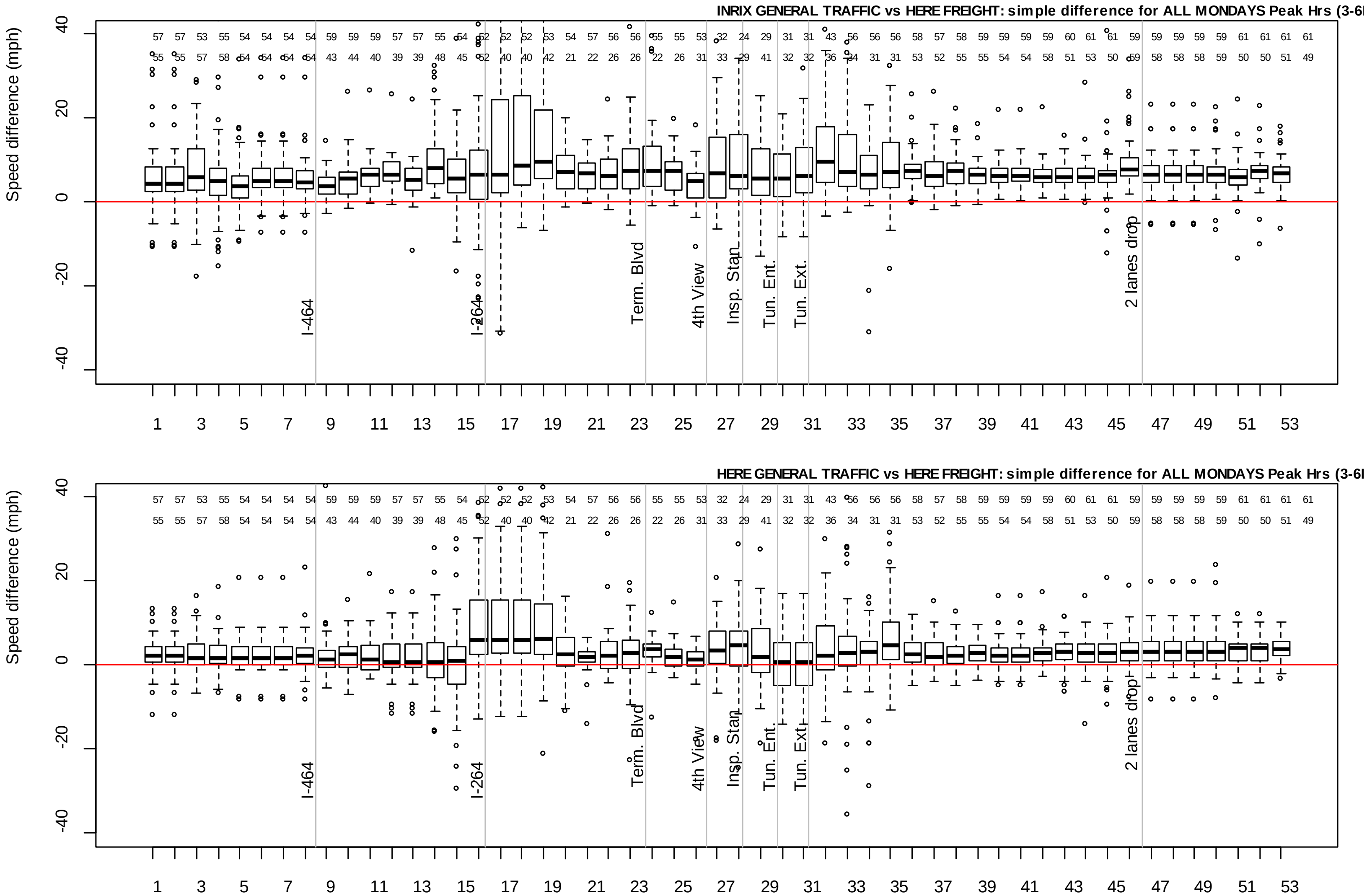


Figure 19 General Traffic & Freight Speed Comparison along I-64 WB for Mondays

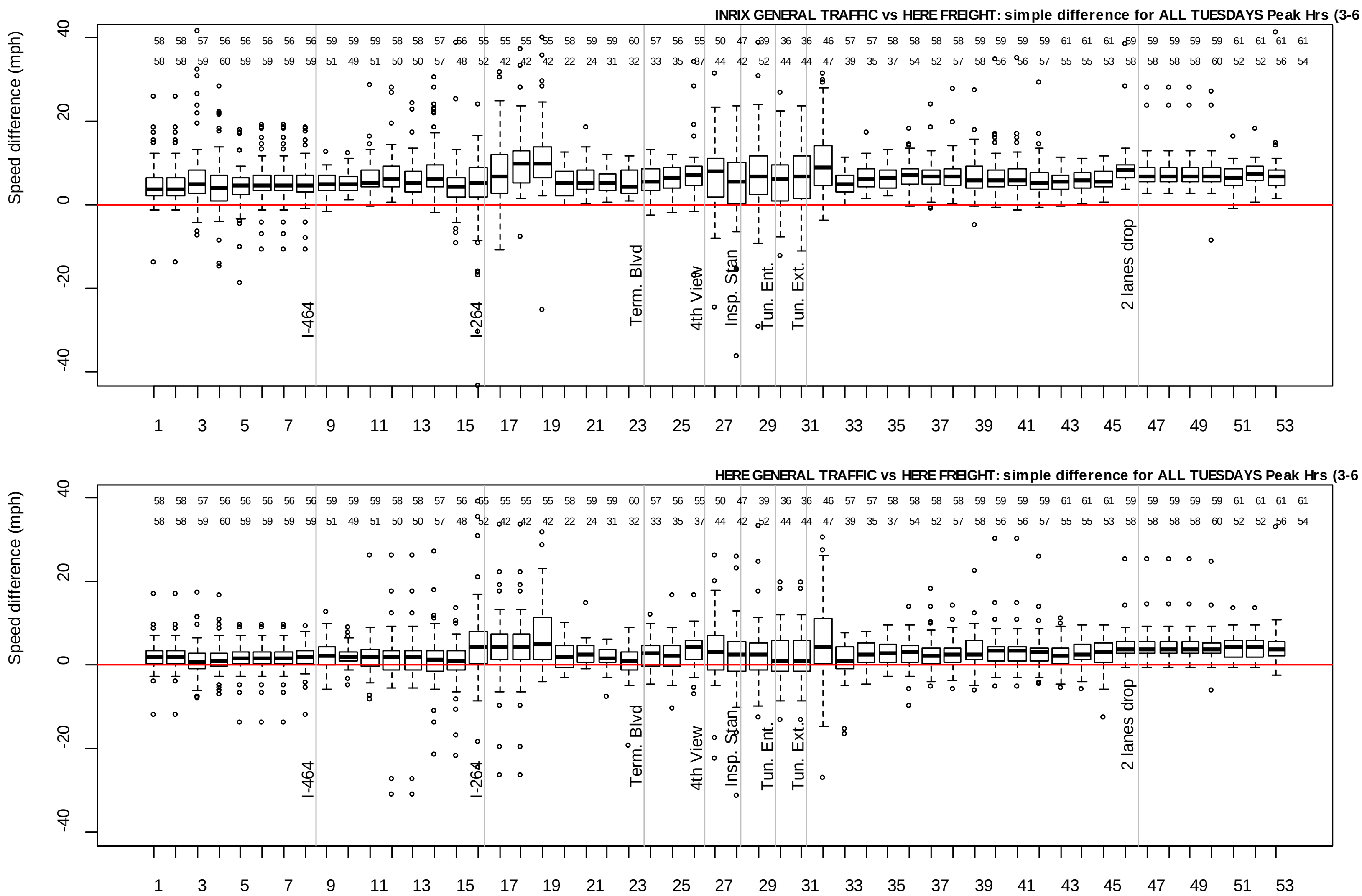
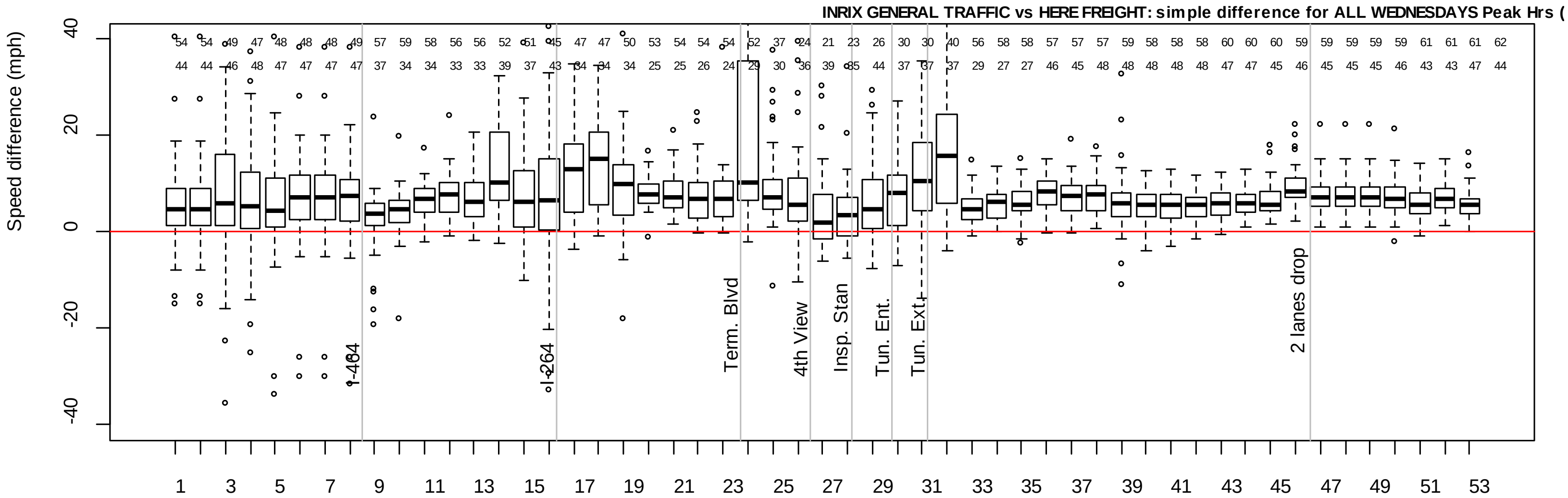


Figure 20 General Traffic & Freight Speed Comparison along I-64 WB for Tuesdays



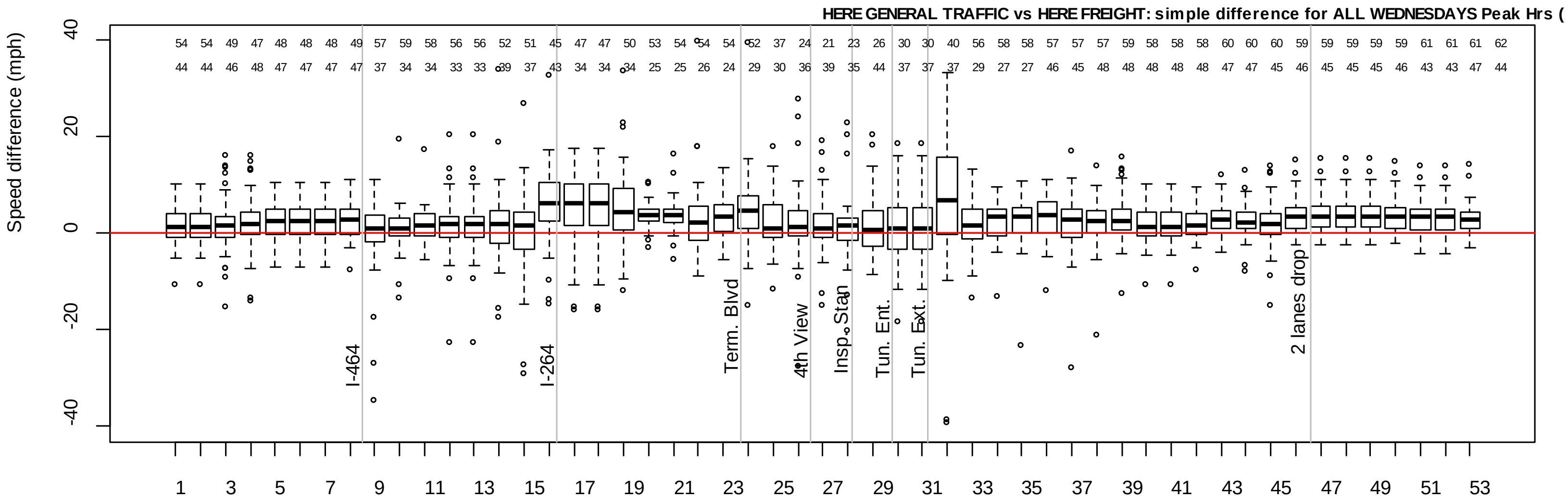
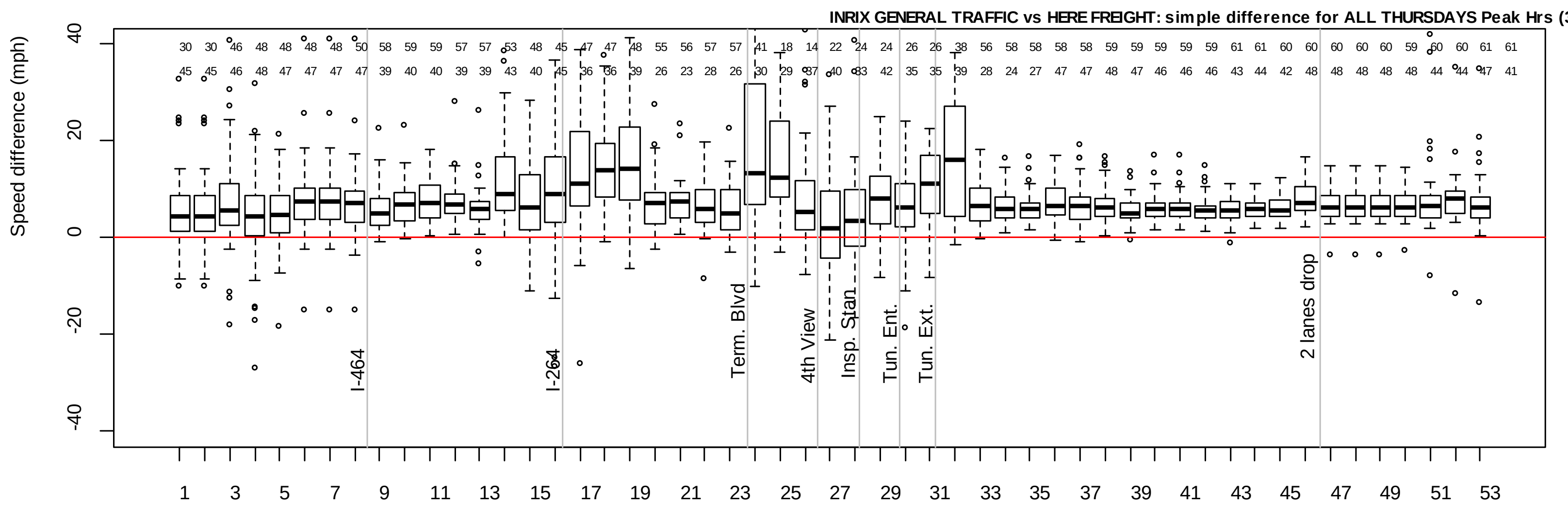


Figure 21 General Traffic & Freight Speed Comparison along I-64 WB for Wednesdays



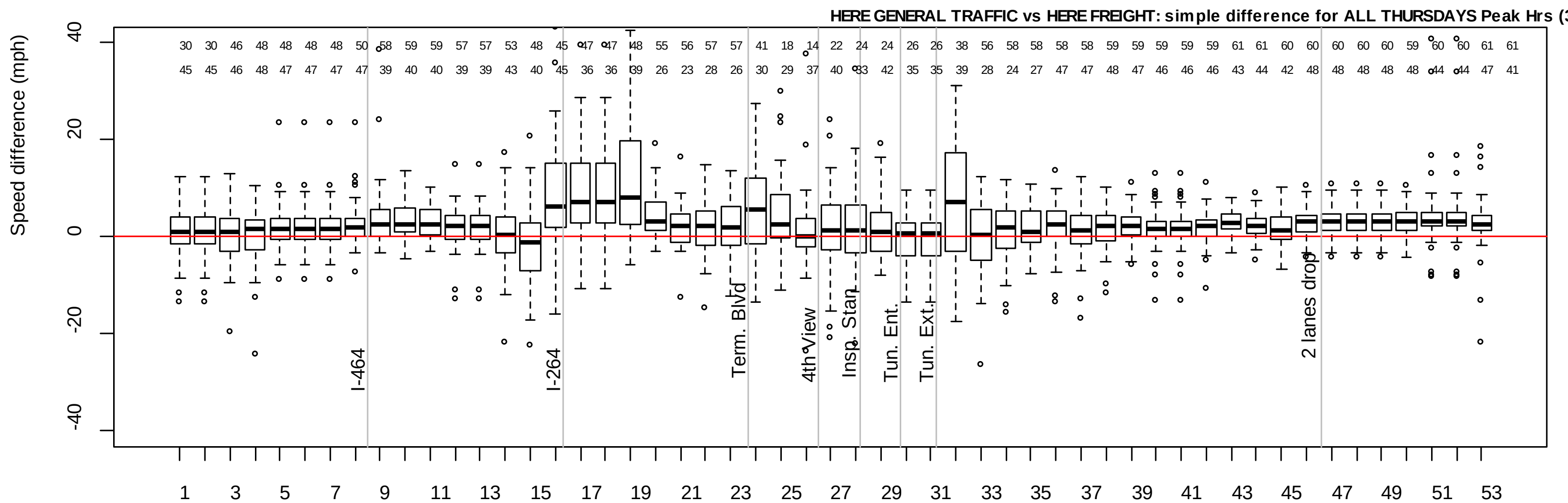
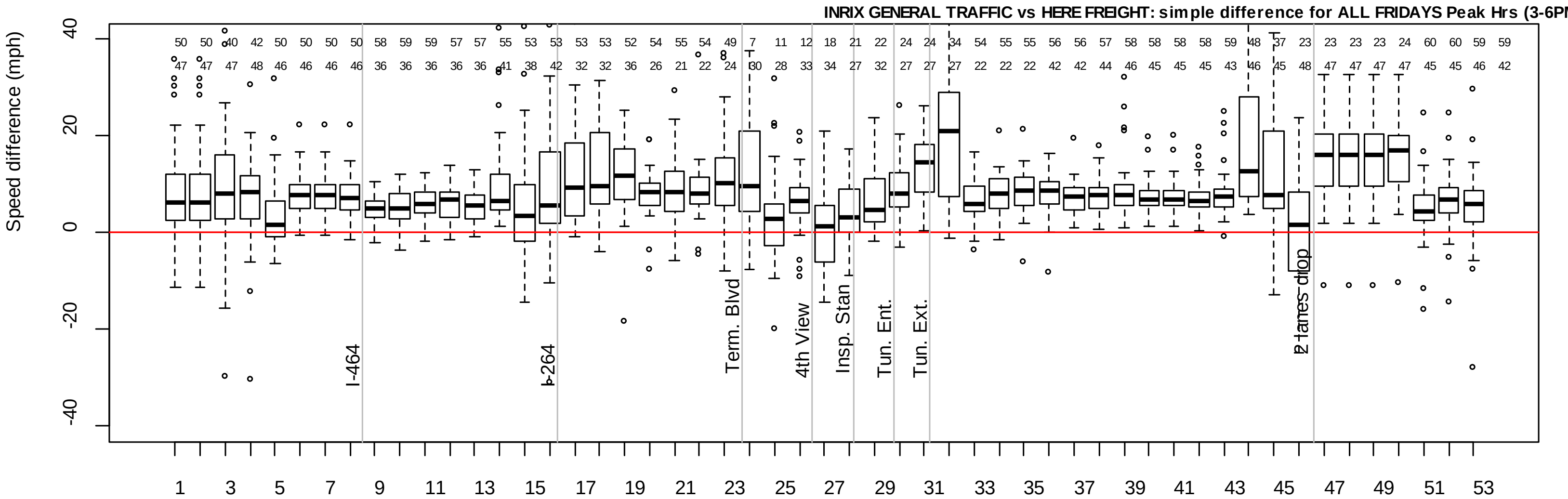


Figure 22 General Traffic & Freight Speed Comparison along I-64 WB for Thursdays



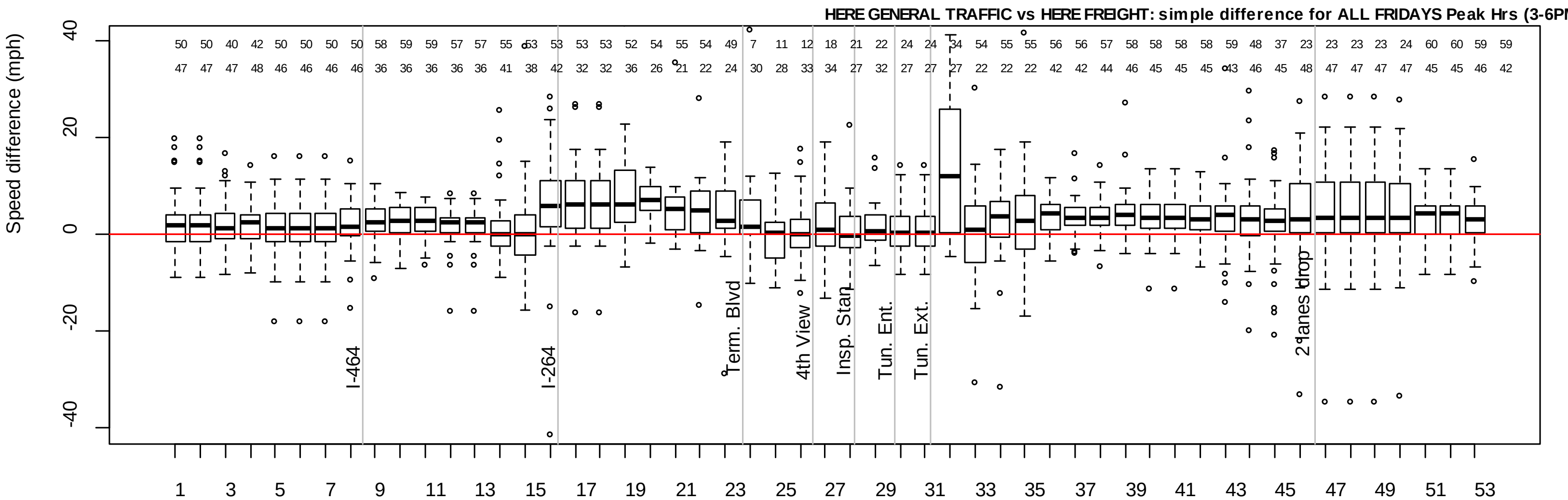
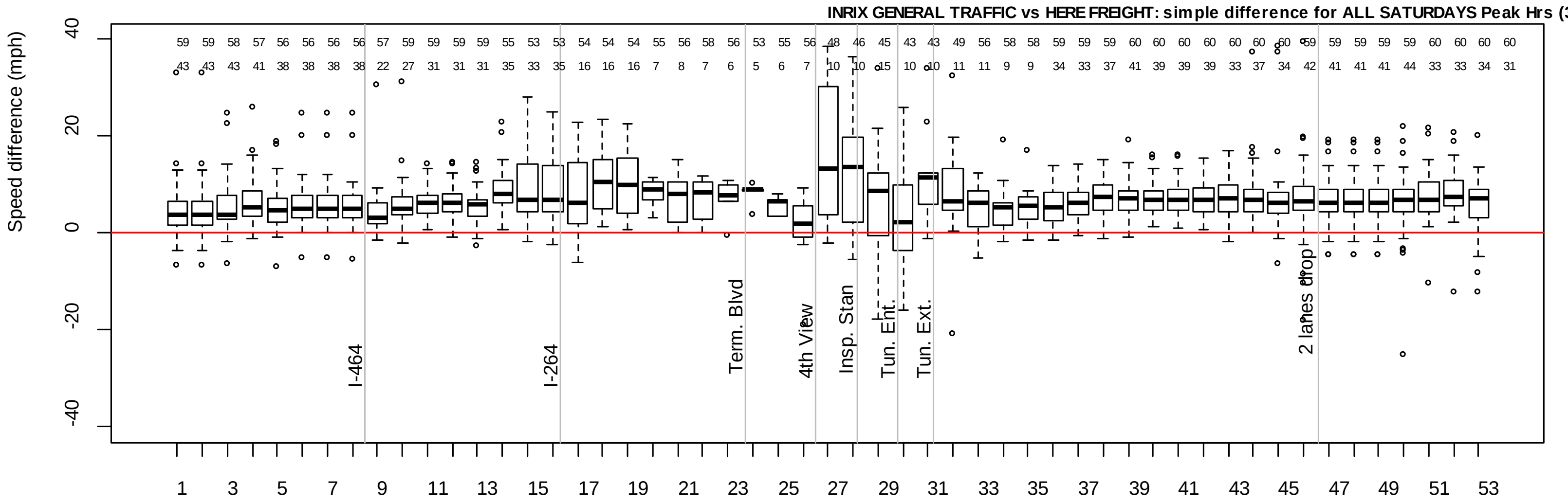


Figure 23 General Traffic & Freight Speed Comparison along I-64 WB for Fridays



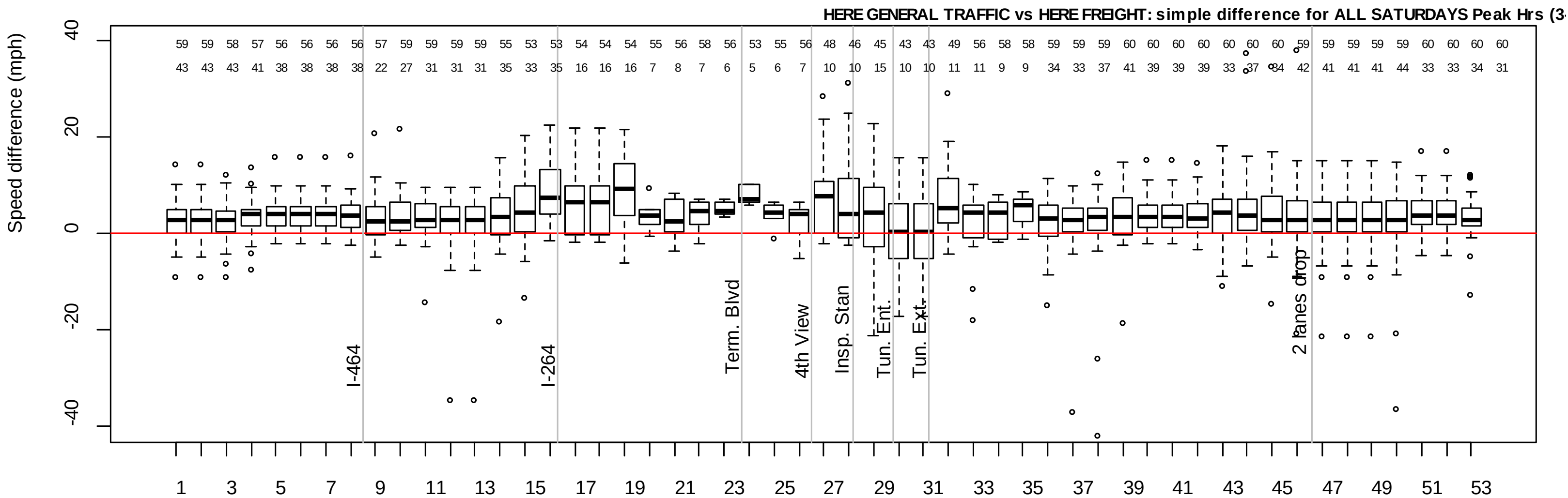


Figure 24 General Traffic & Freight Speed Comparison along I-64 WB for Saturdays

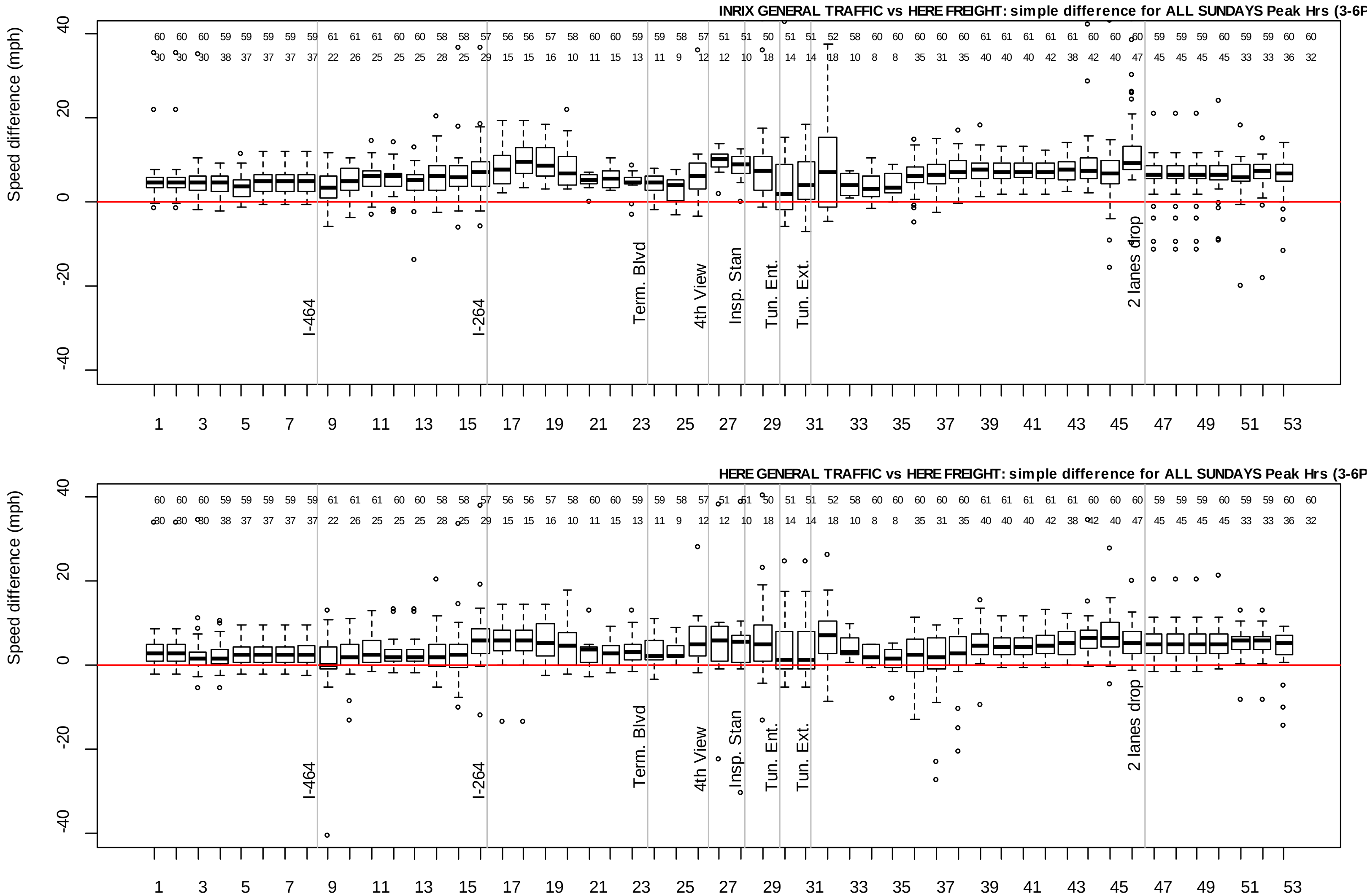


Figure 25 General Traffic & Freight Speed Comparison along I-64 WB for Sundays

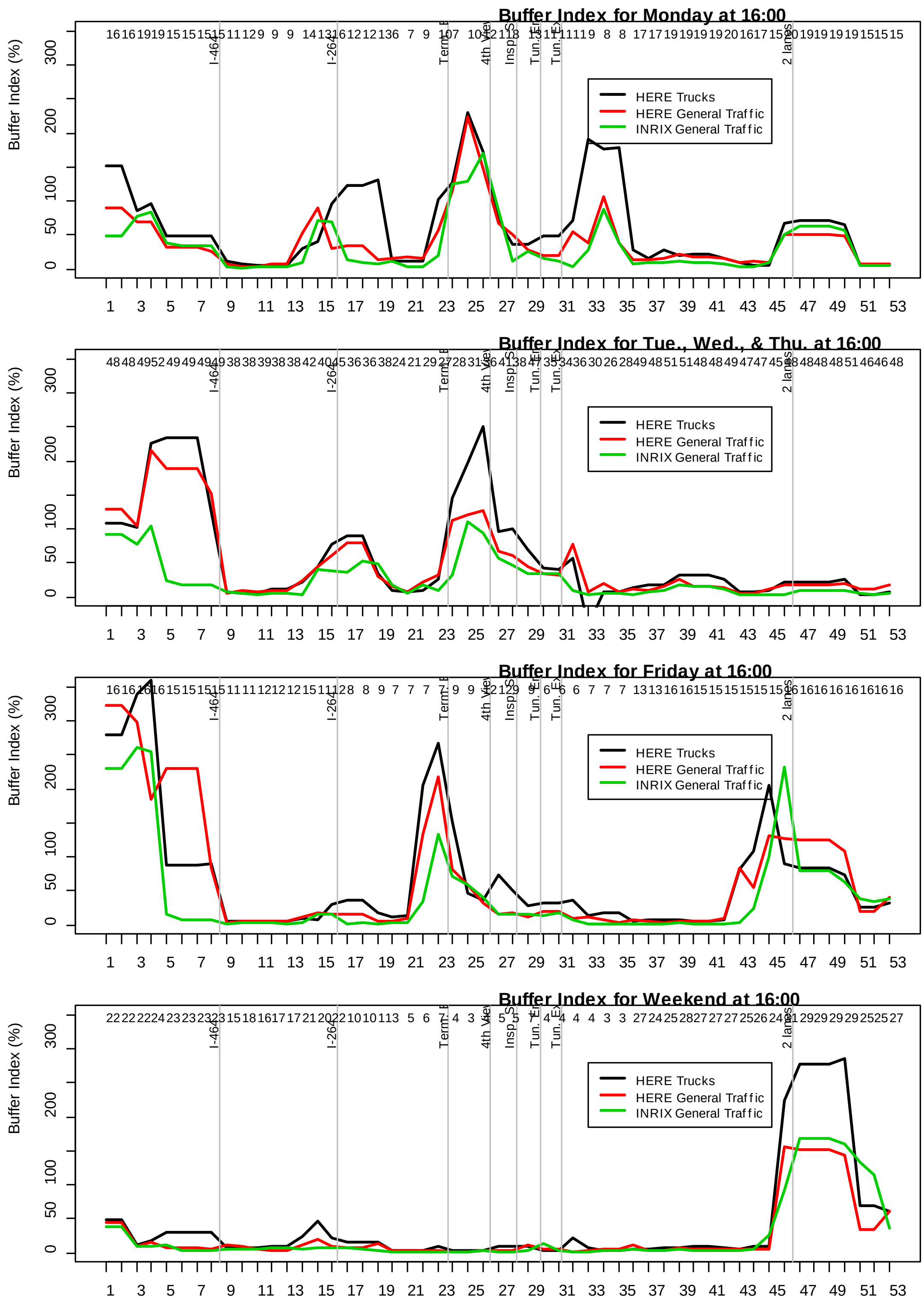


Figure 26 Reliability of Freight and General Traffic Travel Times at 4 PM along I-64 WB

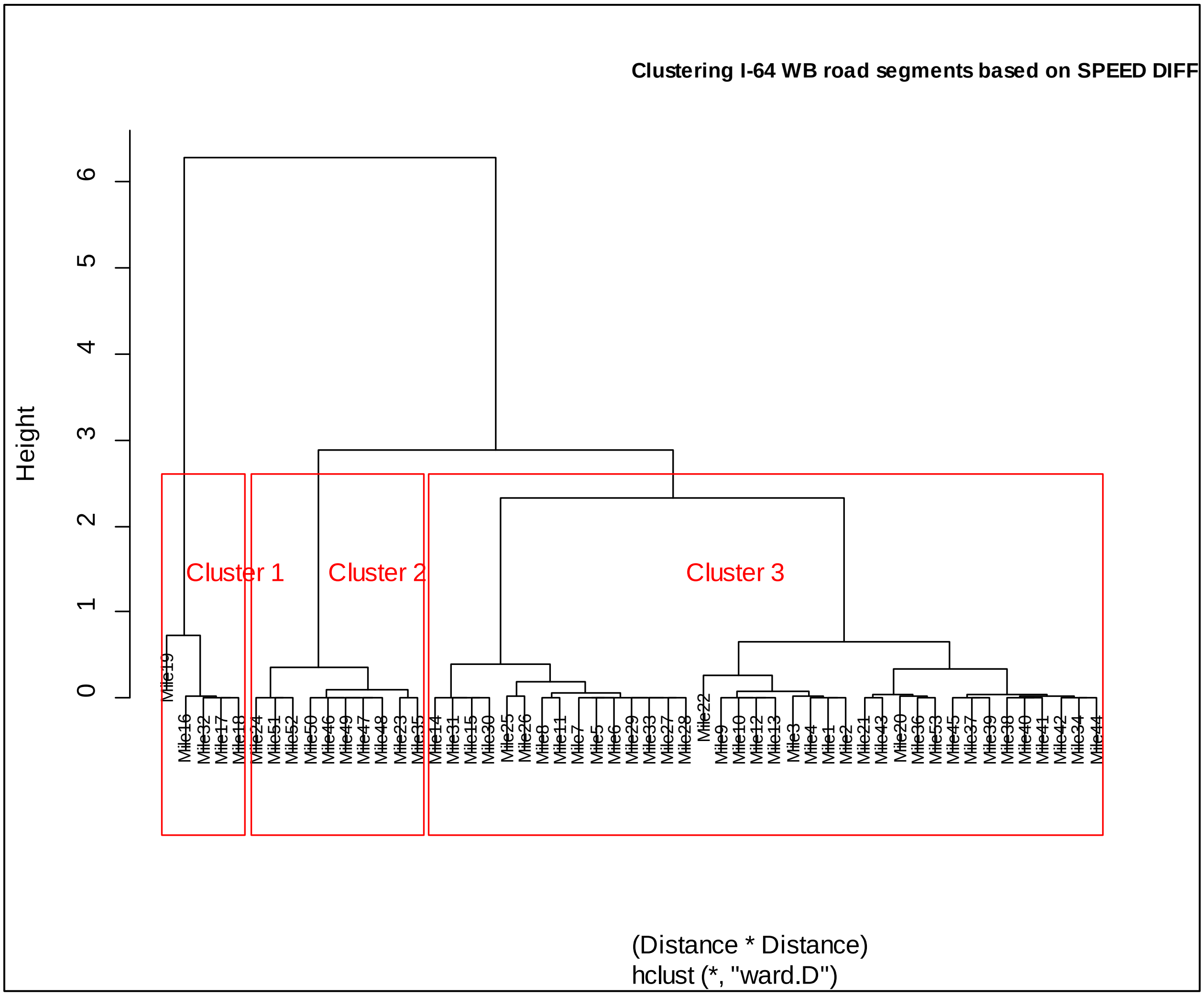


Figure 27 Dendrogram of I-64 WB Cluster Analysis

Table 4 Cluster Centers & Membership I-64 WB

	Average Buffer Index ratio between general traffic and freight	Mean Speed Difference between general traffic and freight	Count of members
Cluster 1	1.4	5.8	5
Cluster 2	1.7	3.4	10
Cluster 3	1.2	2.1	38

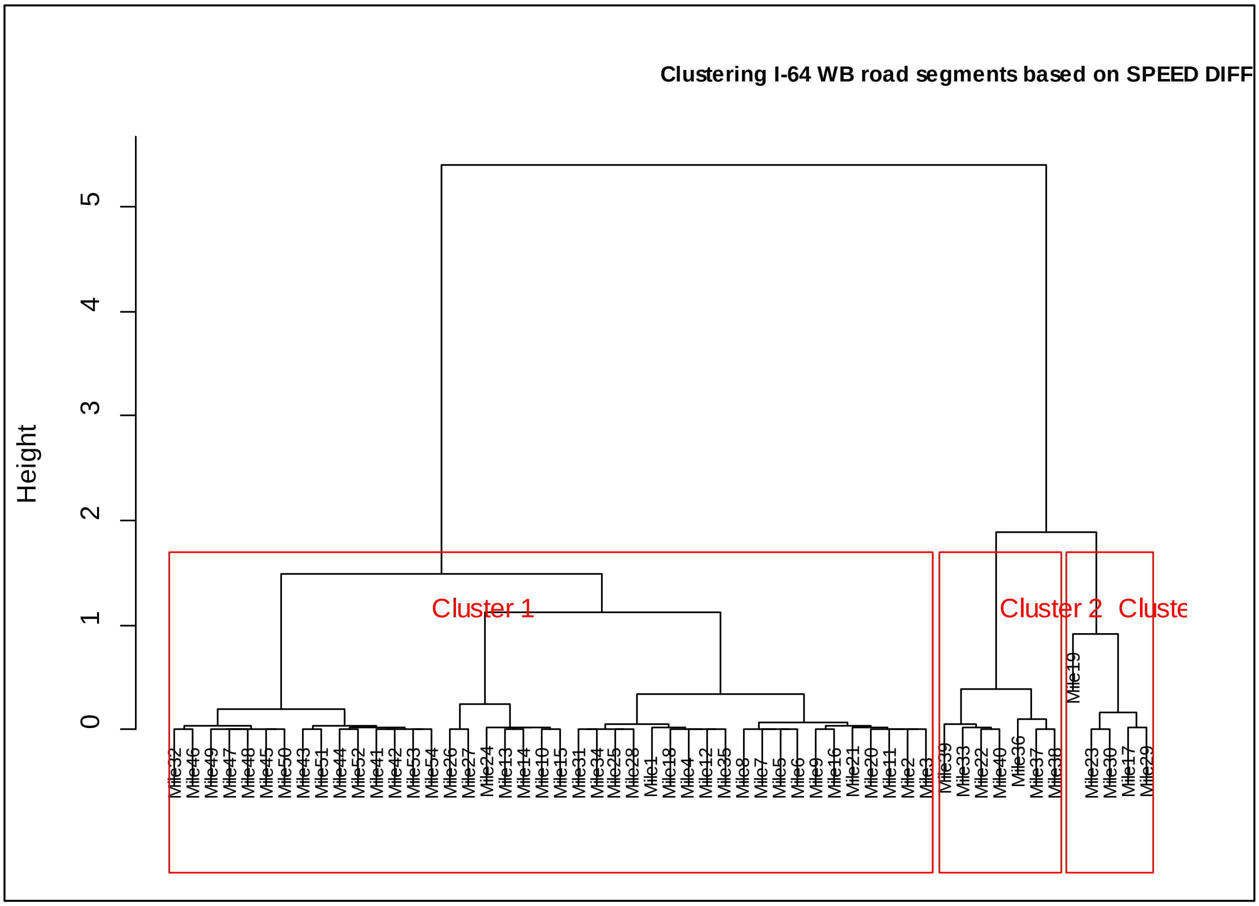


Figure 28 Dendrogram of I-64 EB Cluster Analysis

Table 5 Cluster Centers & Membership I-64 EB

	Average Buffer Index ratio between general traffic and freight	Mean Speed Difference between general traffic and freight	Count of members
Cluster 1	1.1	2.4	42
Cluster 2	2.1	4.0	5
Cluster 3	1.7	0.3	7

ATRI Data Analysis

Given that ATRI data tracks each truck as it moves along major freight corridors, this data can be used to support more advanced modeling efforts such as route choice analysis in addition to monitoring performance measures. For instance, in the current context, it would be useful to understand the route choice preferences of trucks that are leaving the Hampton Roads region. These trucks have two options: take I-64 WB or I-664 NB. I-64 WB usually experiences bottlenecks primarily due to congestion at the HRBT. Using the GPS data, the number of trucks that are traveling along each corridor and eventually exiting the study region was obtained. Figure 29 shows the count of these trucks that are leaving the study region along the two routes. For example, along the corridor upstream of HRBT, 1,163 trucks that are leaving the Hampton Roads region are traveling along I-64 WB whereas only 1 truck is using I-664 NB. Similarly, at the corridor downstream of the intersection of I-464 and I-64, 271 trucks are using I-664 NB whereas 301 trucks are using I-64 WB. It can be seen that as the distance from the tunnel entrance increases, there is a shift in the shares of trucks using the two routes.

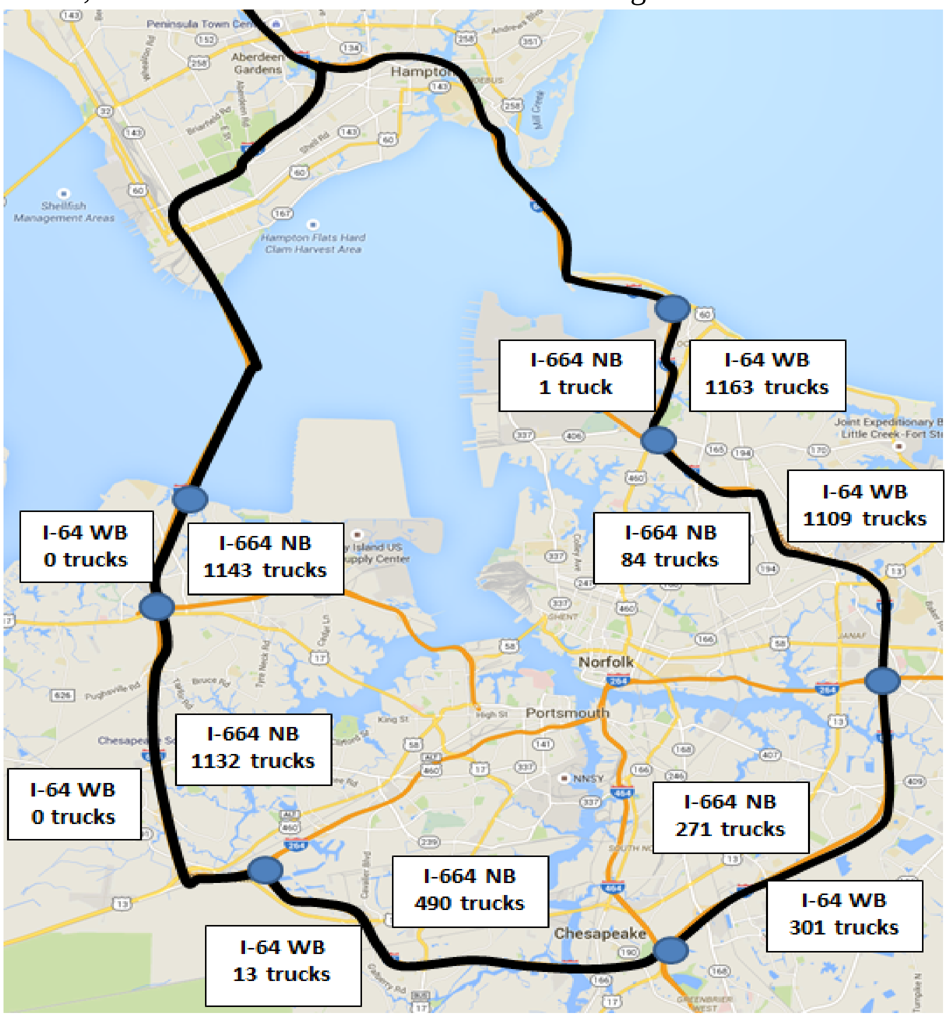


Figure 29 Trucks Counts along Alternate Routes (I-64 WB vs I-664 NB)

From the raw ATRI data, all trucks that traveled between this origin-destination pair were extracted and the average travel time and average speed for this OD pair was estimated. To be specific, average travel time for each time period across all days was computed for the two routes and summarized in Figure 30. It can be seen from these that travel time along I-664 NB is lower than travel time along I-64 WB (only week days were used for this analysis). There can be nearly five minutes of travel time savings for each truck during peak hours if I-664 NB is chosen over I-64 WB. The travel time differences, however, are not significant during off-peak hours. Nevertheless, the travel time profile along I-664 NB is more reliable (less variability) than that along I-64 WB. The speed profile along these two routes were also consistent with the travel time findings reported above (see Figure 31). The fact that more trucks chose I-664 NB confirms that truck drivers were aware of the shorter travel times if they take I-664 NB rather than I-64 WB (490 trucks on I-664 NB vs 301 trucks on I-64 WB).

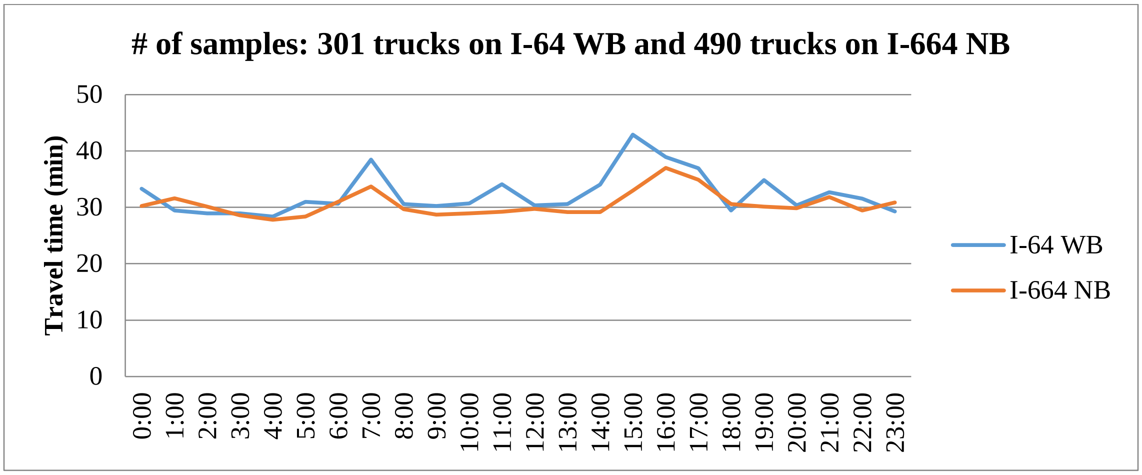


Figure 30 Average Travel Time Profile along Two Alternate Routes

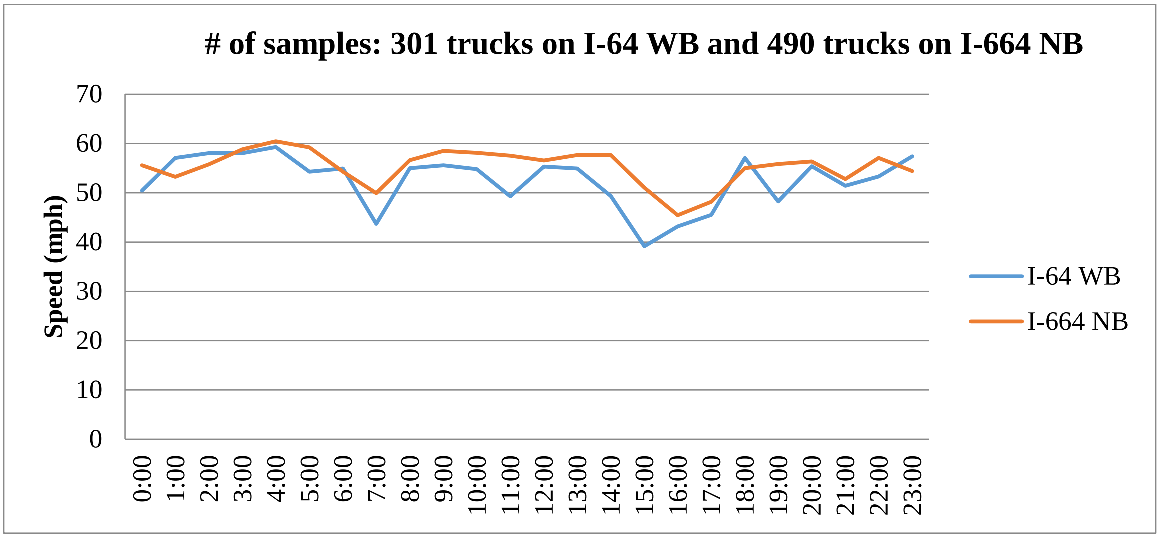


Figure 31 Average Speed Profile along Two Alternate Routes

REFERENCES

Cambridge Systematics, I. (2010). Virginia Statewide Multimodal Freight Study, Final Report, Prepared for Virginia Department of Transporation.

Everitt, B., S. Landau, M. Leese and D. Stahl (2011). Cluster Analysis (Wiley Series in Probability and Statistics), ,Chichester, UK, Wiley.

FHWA (2006). Reliability, Travel Time Making It There on Time, All the Time, Federal Highway Administration. Report FHWA-HOP-06-070.

Hijmans, R., E. Williams and C. Vennes (2015). Geosphere: spherical trigonometry. R package version 1.3-13, Available online at: http://CRAN. R-project. org/package= geosphere.

Murtagh, F. and P. Legendre (2014). "Ward’s Hierarchical Agglomerative Clustering Method: Which Algorithms Implement Ward’s Criterion?" Journal of Classification 31(3): 274-295.

Pinjari, A. R., A. B. Zanjani, A. Thakur, A. N. Irmania, M. Kamali, J. Short, D. Pierce and L. Park (2014). Using Truck Fleet Data in Combination with Other Data Sources for Freight Modeling and Planning. BDK84-977-20, Florida Department of Transportation.

Zaki, M. J. and W. Meira Jr (2014). Data Mining and Analysis: Fundamental Concepts and Algorithms, Cambridge University Press.

APPENDIX

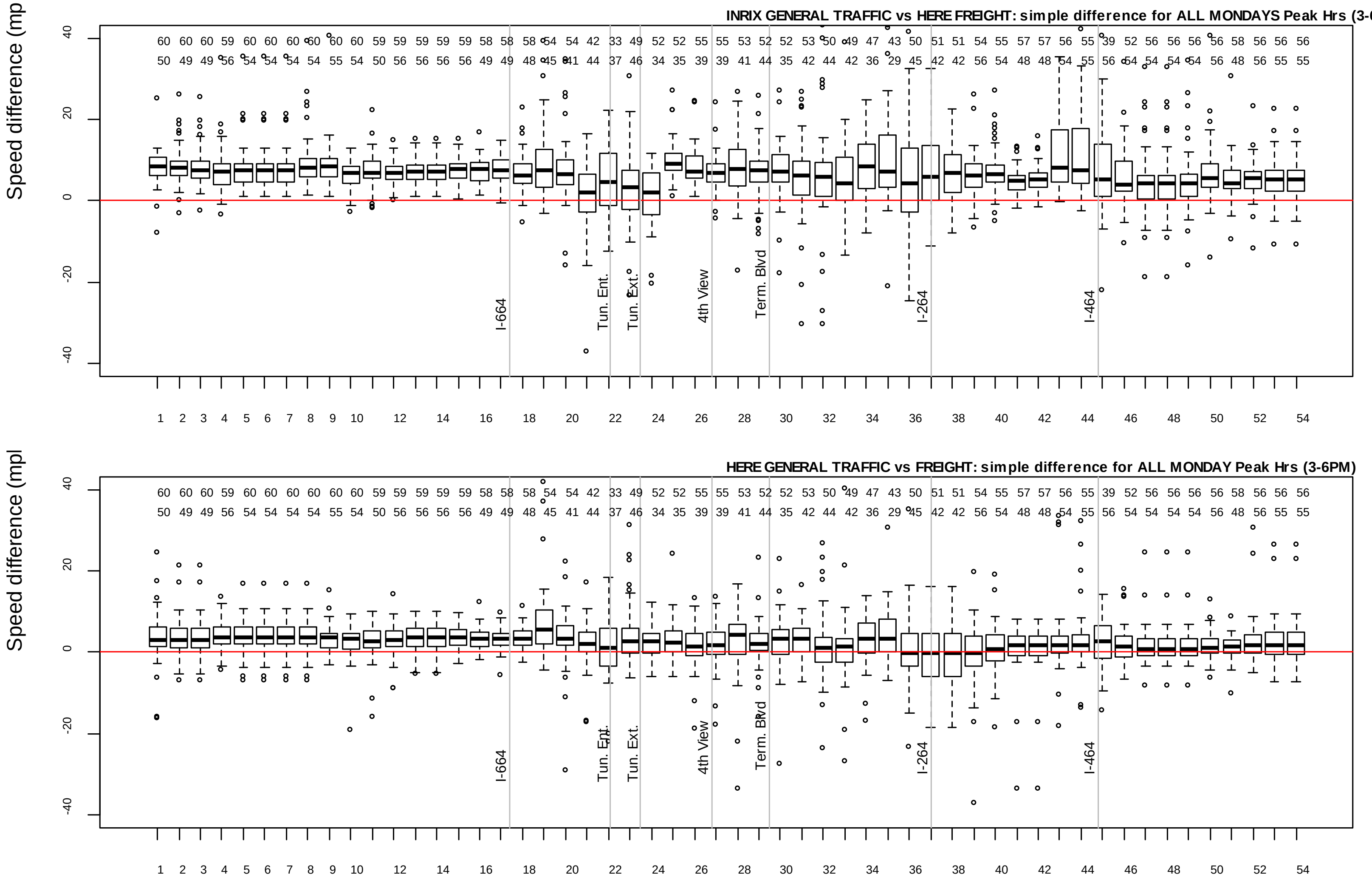


Figure 32 General Traffic & Freight Speed Comparison along I-64 EB for Mondays

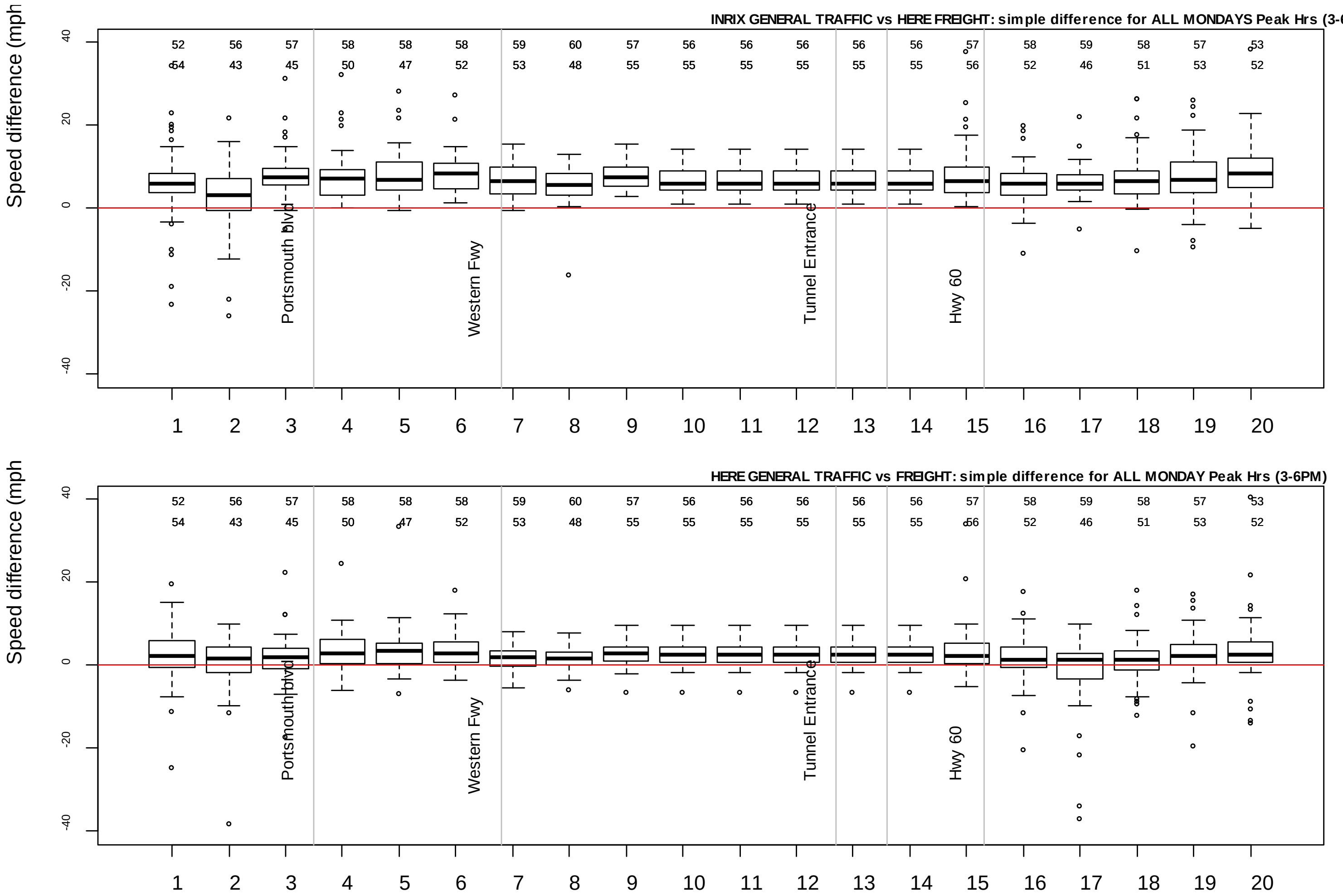
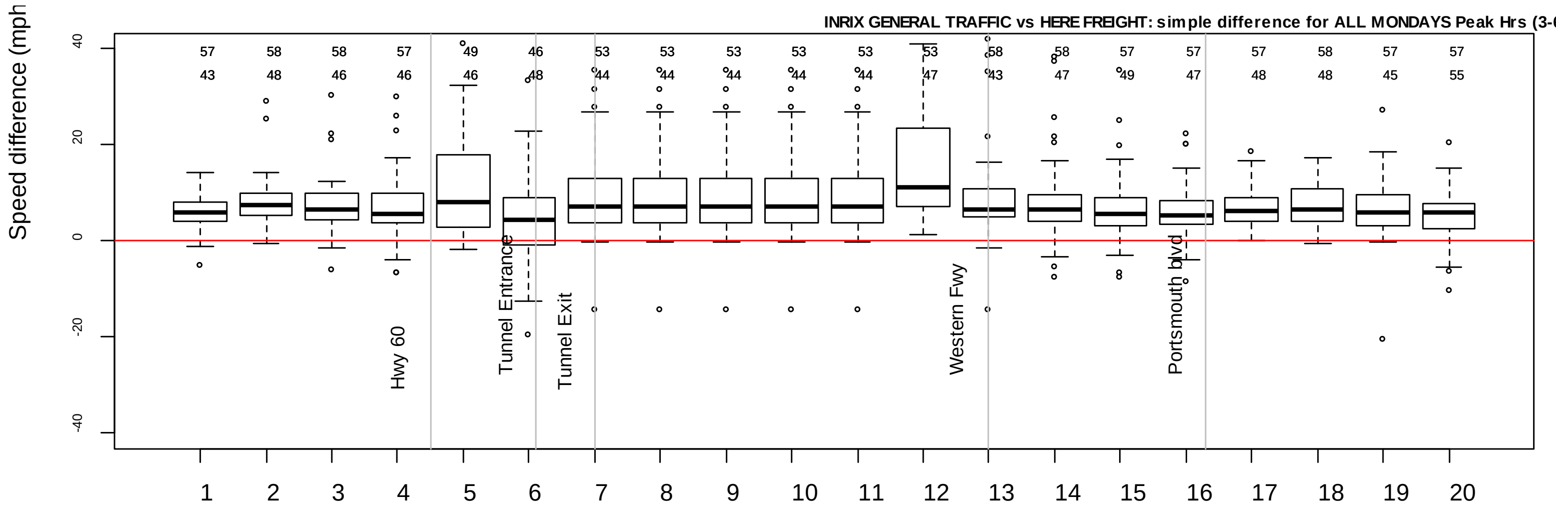


Figure 33 General Traffic & Freight Speed Comparison along I-664 NB for Mondays



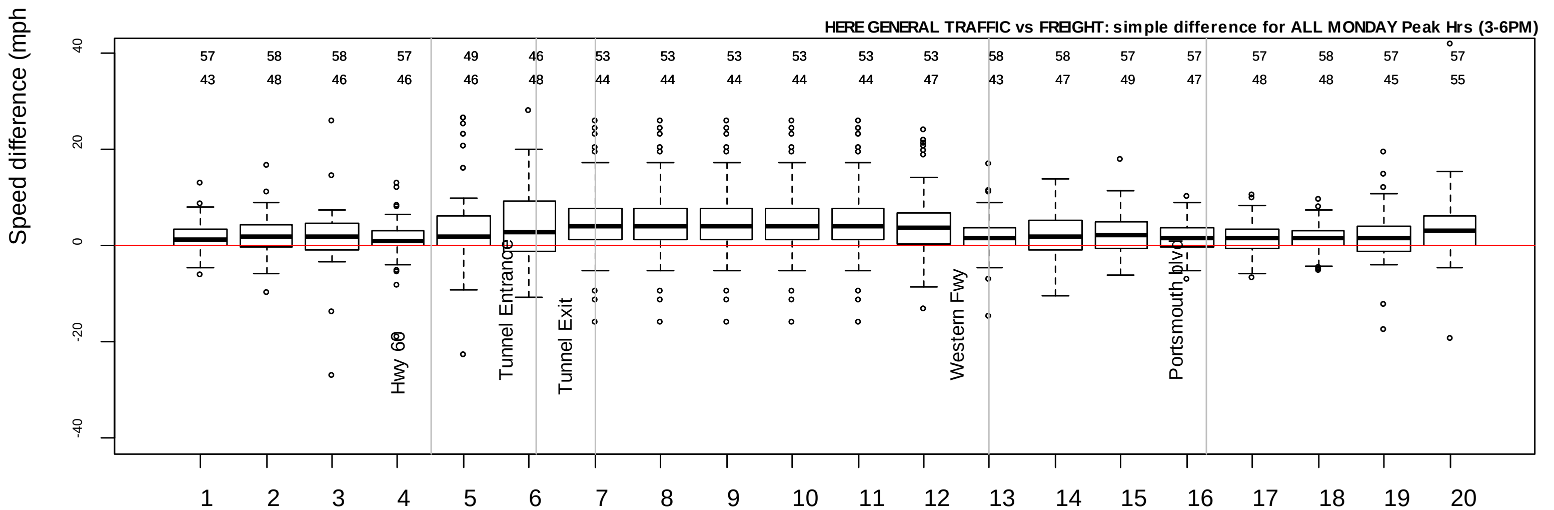


Figure 34 General Traffic & Freight Speed Comparison along I-664 SB for Mondays

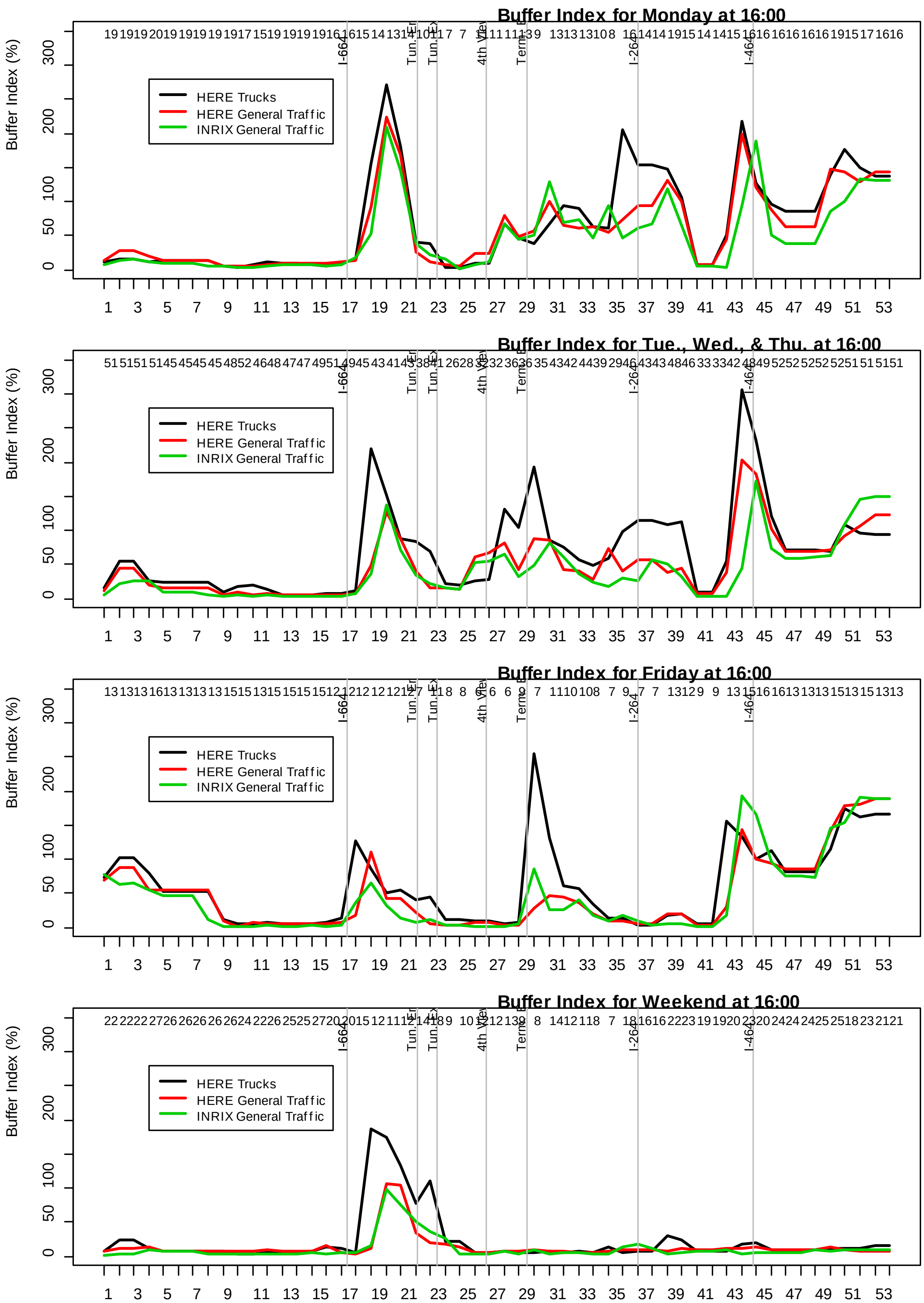


Figure 35 Reliability of Freight and General Traffic Travel Times at 4 PM along I-64 EB

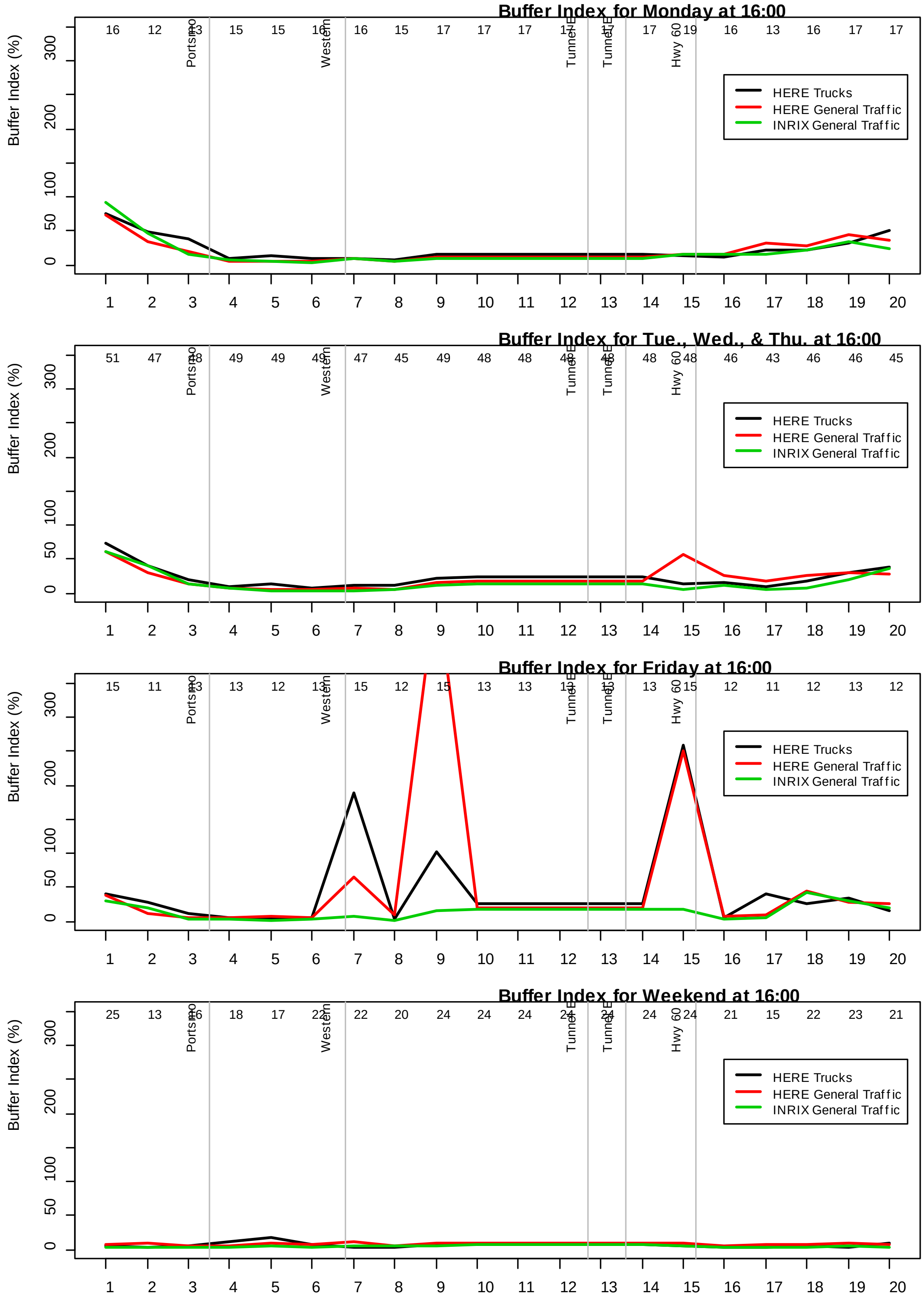


Figure 36 Reliability of Freight and General Traffic Travel Times at 4 PM along I-664 NB

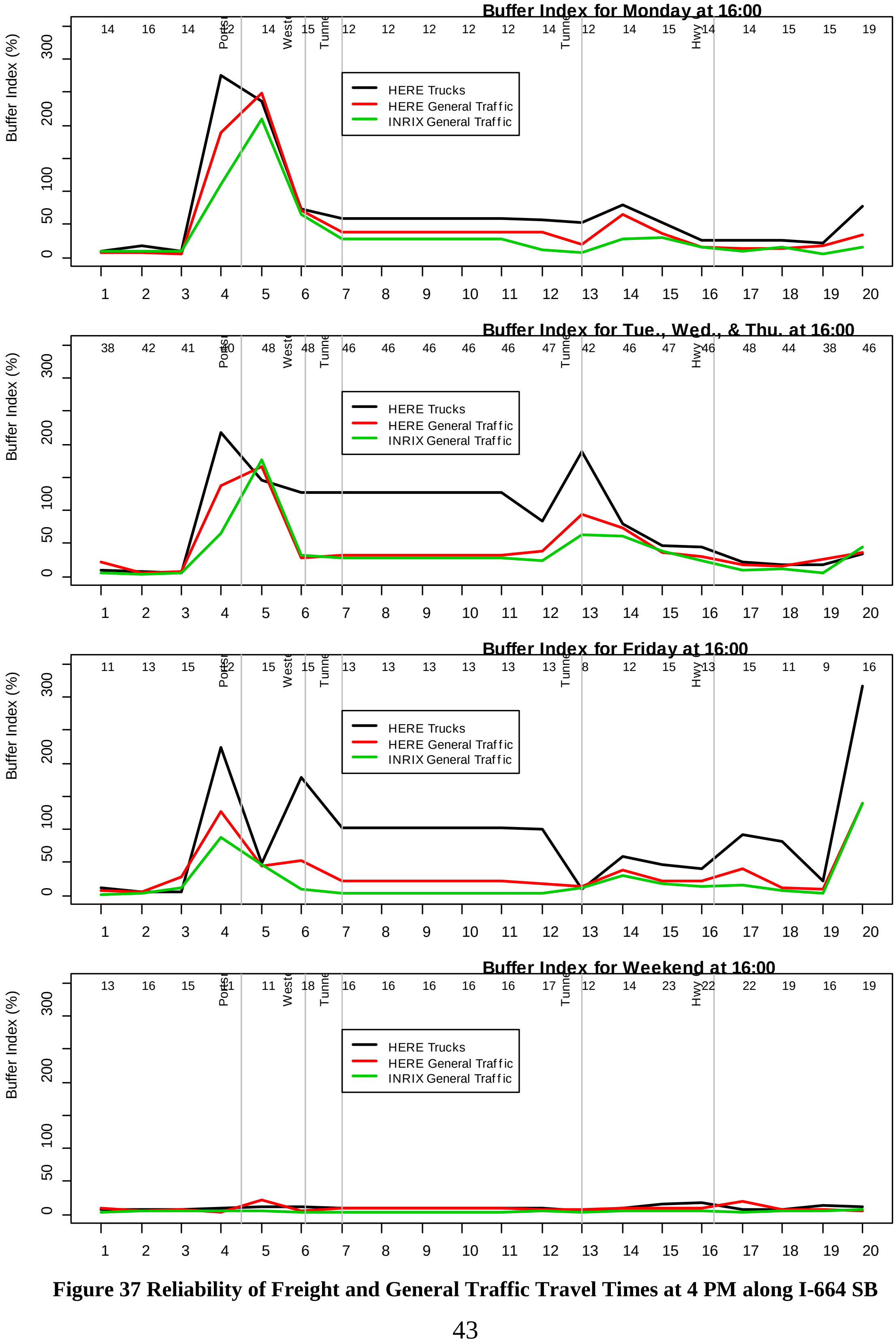


Figure 37 Reliability of Freight and General Traffic Travel Times at 4 PM along I-664 SB