Congestion Effects on Arterials as a Result of Incidents on Nearby Freeway
When should you get off the highway?

by

Kyle C. Weaver

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirements for the Degree
Master of Science
(Geographic Information Science and Technology)

May 2019

To my sweet Cakalakins

Table of Contents

List of Figures

List of Tables

Acknowledgements

I am grateful to my mentor, Professor Xavier, for walking with me along the path and the other faculty who gave me the confidence to embrace diversity.

Thank you also to COL [R] Steven D. Fleming, PhD, as my thesis advisor and the incredible team at USC Spatial Sciences Institute for providing an environment to explore my passions.

List of Abbreviations

ARIMA	AutoRegressive Integrated Moving Average
ATSAC	Automated Traffic Surveillance and Control
BP	Back Propagation
CHP	California Highway Patrol
CNN	Convolutional Neural Network
DBN	Deep Belief Network
GIS	Geographic information system
GJR GARCH	Glosten Jagannathan and Runkle Generalized Auto Regressive Conditional Heteroskedasticity
GMM	Gaussian Mixture Model
HMM	Hidden Markov Model
ITS	Intelligent Traffic System
LADOT	Los Angeles Department of Transportation
LSTM	Long Short-Term Memory
Mph	Miles per hour
MTL	Multi-Task regression Layer
RBF	Radial Bias Function
RNN	Recurrent Neural Network
USC	University of Southern California
XML	Extensible Markup Language

Abstract

Disruptions like accidents or closures on metropolitan freeways have the potential to increase traffic congestion on surface streets. Through spatiotemporal analysis, this project evaluates associations between traffic congestion spikes on arterial streets with freeway incidents. The unexpected increase of traffic on city streets from freeway overflow was expected to not only create severe gridlock negating the expected benefit for the motorist avoiding freeway delays, but also cause undue stress for local traffic normally on those streets. This thesis takes the initial steps in spatiotemporal analysis to assess how strong the associations are between incidents on the freeway and increased arterial traffic. Data preparation models from Alteryx are used in ESRI's ArcGIS Pro to provide a contextually rich multi-dimensional representation of sensor location, time and traffic speeds near freeway incident locations. This enables an intuitive way to recognize potential associations between speed data collection points. The use cases analyzed by this study were predicated on a long-duration traffic accident for which medical services were required. The results show that almost no clear association can be made for incidents of this magnitude. Using data about these effects and more severe use cases like complete freeway closures in concert with the visualization techniques presented, additional studies can be built to support determination of whether or not more significant disruptions may have clear associations. From that point, mitigation options can be designed to reroute traffic through techniques like optimizing traffic lights and active traffic rerouting.

Chapter 1 Introduction

Analysis of historical and current traffic information can predict the patterns that develop over time under normal conditions. This thesis explores whether the association between freeway incidents and increased arterial congestion can be measured using spatiotemporal analysis and presents visual representations to show relationships. Through spatial modeling of incidents, like accidents or road closures, and the subsequent traffic patterns that typically develop, road network modeling of discrete freeway events can identify areas where related increases in congestion on arterials develops.

1.1. What is Currently Used

The natural use for knowing the relationship between freeways and arterials would be to minimize travel times for travelers. Active manipulation of traffic flow through automated traffic systems used by large cities like Los Angeles already provides some relief. The arterial street traffic lights are controlled by the Los Angeles Department of Transportation Automated Traffic Surveillance and Control (ATSAC) system. ATSAC controls a network of over 4600 traffic lights, hundreds of video cameras, and 40,000 traffic loop sensors across over 460 square miles in the Los Angeles region (LADOT, n.d.). The ATSAC system depends on years of prior traffic data for its algorithms which are used to schedule traffic light timing to provide fluid traffic flow throughout the different times of day. Additionally, traffic management operators have the ability to override the programmed schedule to manually respond to backups if needed. In lieu of scheduled or manually controlled algorithms, the incorporation of expected effects on arterials following a freeway incident would theoretically allow the traffic light control system to actively minimize congestion on local streets through signage and traffic light manipulation. Using

predictions about how traffic patterns develop allows proactive actions to be taken and may have a positive effect on the increased congestion in the affected areas.

1.2. Other Traffic Analysis Approaches

Google provides a future travel time estimation in its maps service, but it is unclear as to the method used to calculate the times (Bell 2016). While Google and Waze algorithms are proprietary, research indicates that they model their predictions on historical averages on road segments and appear to provide time-phased results provided the road segments have enough data points in enough time intervals (Waze 2017). Both Google Maps and Waze services allow for input of desired departure or arrival time and displays a range of times that it could take to travel along the desired route. Often this travel time prediction results in a quite wide time uncertainty. This large uncertainty is likely based on the confidence of the prediction algorithm and its inability to consider events outside of normal traffic flow conditions to narrow the distribution of the predictions.

1.3. Benefits of the Results

Often, an unexpected increase of traffic from a freeway onto city streets can cause severe gridlock. This effect negates the expected benefit for the motorist who chooses to leave the freeway for a faster route, but also causes undue delays for traffic already on those local streets. With predictive data from an expansion of this analysis, it is possible that traffic lights can be optimized to account for the rerouted traffic. An assumption of direction of traffic flow can be made since the expected or intended route is predictable based on the direction of the freeway traffic being disrupted. This optimization has the potential to minimize surface street congestion providing benefit for local traffic as well as traffic that has deviated from the freeway.

1.4. Research Goals

With the ability to quantify the effects of traffic from a freeway using arterials to avoid delays, an intelligent traffic system (ITS) would be able to proactively adjust timing and directional control at key intersections that feed freeway on- and off-ramps as well as the connected streets along possible routes. As a first step towards that goal, this project shows the net effects which a disruption to the normal freeway flow has on the local streets. Various unplanned disruptions like an accident or temporary closure of the freeway may cause ripple effects on the roads with little warning. Normal daily traffic patterns due to commuting, or planned events like construction or special events may be predictable to a certain extent. Based on historical data of traffic speeds before and after discrete freeway disruptions, this paper explores visualizations of how traffic responds to incidents and identifies the magnitude and breadth of impacts that traffic leaving the freeway has on arterials in an area of Los Angeles.

The study proposes a process which takes raw speed data for a period of time and processes that data to provide a rendering of sensor locations with a time-speed representation of the underlying data. This process provides an effective means of visualizing speeds over time, geospatial relationships between various traffic data collection points and a means to assess how traffic responds to various incidents along the freeway. A three-dimensional map which includes contextual cues like elevation variations or labeling for distant points of interest as well as the coordinate driven data collection points is overlaid with a vertical representation of the analyzed data to allow for a simplified visual reference for how the data is potentially related. Subsequent studies can manipulate the process and visualization methods to adapt to a wide range of applications for spatiotemporal visualization.

Chapter 2 Related Work

Visualization is a necessary component of how we interpret traffic data. It is also a means of interacting with it. Whether it is traffic data on a map for commuting, routing for emergency services, or planning logistic movements of commercial goods, traffic data nearly always has a time component which is important addition to the dimension of space. Interactivity solves many of the challenges of the added time dimension. Using time sliders or animated maps allows the data to be shown sequentially, but they do not easily portray the changes and interpret the data across wider periods of time. Furthermore, a means to provide interaction with the data visualization is often absent, requiring a method for static representation. This study provides a method for achieving that.

Predictions are often an important part of traffic data interpretation. Existing literature for traffic prediction is divided into three primary areas: traditional geospatial analysis, machine learning, and hybrid methods. Until the last decade or so, the primary method for traffic modeling was based on traditional statistical and regression calculations for providing predictions. As the computing power for highly parallel processing became feasible, another genre of studies has emerged which uses machine learning to obtain high confidence results for traffic flow predictions. A third approach combines both traditional and machine learning techniques for traffic analysis to provide the best of both worlds: current geospatial analysis techniques plus modern insight to traffic analysis through machine learning. This thesis focuses on the association between freeway and arterial congestion through traditional methods, but the machine learning processes are germane since they provide context to how more advanced techniques can be employed and combined for extended analysis and a simplified method to visualize that analysis.

2.1. Visualization Methods for Traffic

Much of the existing work of multidimensional visualization in recent years is linked to the concepts present by Leland Wilkinson in *The Grammar of Graphics* (2005). Wilkinson proposes that proper visualizations contain specific components which interact to produce representations that serve the purpose of the study or analysis. These components are data, transformation, elements, scale, guide and coordinates. By specific choices in the selection of each of the components, complicated graphics can be achieved which represent multi-dimensional data that answers specific questions or reveals underlying trends. While the structure of the preparation of this data can vary, if the rules and best practices presented are followed, sensible, informative representations are possible.

Wickham (2010) builds on Wilkinson's work through the introduction of scalable dimensions. By using methods to achieve multidimensional data graphics through the use of coordinates, size, shapes, colors, facets and groupings, he provides a methodology to focus the analyst or reader on the important aspects of the data for their consumption.

Many statistical software packages like Matlab, SPSS, Tableau and others provide mechanisms to render data in three-dimensional space, but they often lack the ability to provide the context that this method provides with a synchronized map background. Pastizzo et. al (2002) created their own data visualization application called Multi-Dimensional Data Viewer (MDDV) which enhances the user interface from existing applications for manipulating the visualization, but still falls short of what a dedicated geospatial analysis can provide using a geographic information system program like ESRI's ArcGIS.

The methods and layered effects of these works help to provide a roadmap to achieving effective visualization. The data preparation required was automated and modeled to provide a streamlined process to allow for changing parameters for this study's product.

2.2. Traffic Spatial Predictive Analysis

The challenge of traffic prediction has been studied in transportation and urban planning since the 1930's (Albright 1991). The benefits of an effective process for managing and controlling urban traffic extend from economic benefits to livability and desirability for residents who live or work in major metropolitan areas. The typical traffic patterns that result from daily commuting, seasonal weather, or tourism can be disrupted in a number of ways. Accidents, construction, special events, or significant natural phenomena can all cause a change to normal flow. The nature of traffic flow presents spatiotemporal relationships that are often tough to correlate using traditional mathematical processes. Many studies focus on a particular segment of roads or on homogenous types of roads (e.g. interstates only, or arterial streets only) (Chow 2013; Yang 2010) and do not spatially or temporally consider the effect of these roads on other nearby roads in a different category. Even in studies which claim spatiotemporal predictions, these predictions often only refer to upstream and downstream effects (roadways before and after the study point) on the same road segment.

2.3. Traditional Methodology

Spatial algorithms like ordinary least squares, k-nearest neighbors, and a random forest are common in many geospatial analysis projects to determine spatial relationships (Ma et al. 2017). However, they fail to incorporate the essential multidimensional (including temporal) aspects needed for a correlation of traffic flow. It is important to note that these basic spatial techniques as well as the mathematical models in this section are critical foundational tools that can be combined with other techniques for effective predictions.

The autoregressive integrated moving average (ARIMA) model is one of the most used processes for traffic prediction (Williams, Durvasula, and Brown 1998). The ARIMA model is a parametric-based linear-regressive model that is effective in classifying time-series data. The

ARIMA process takes pre-selected time-series information, makes time-difference calculations recursively and extends existing data with a averaged regressive prediction (Williams and Hoel 2003). It is effective for traffic flow predictions on a specific road segment based on past flow conditions but fails to consider nearby road networks and their combined effects.

A Bayesian or hidden Markov model (HMM) can provide predictable results from a pseudorandom process like traffic flow. A Bayesian model provides insight into models with a number of unknown parameters through the use of probability assessments as a causal relationship. Similarly, a Markov model is a type of Bayesian model that takes the current behavior, or state, from a set of known parameters and uses that to predict how the model will react to new data assuming the new states are not necessarily causal (Ghahramani 2001). Using state transitions from an author-defined set of traffic states, Qi and Ishak (2014) provide estimates of future traffic conditions based on mean and contrast statistics over a given highway segment. The mean provides an assessment of the propensity of normal traffic observations to speed and volume. With this baseline, the HMM can be used to provide more insight through contrast as a second-order effect. Contrast is the measurement of the variability of traffic disturbances. Higher contrast indicators show more volatility and disturbances in the expected traffic flow. This can be caused by drivers switching lanes often or more frequent changes in flow above the baseline observation.

Zhang, Zhang, and Haghani (2014) combine several traditional analysis methods for short-term predictions. Spectral analysis, ARIMA, and a statistical volatility model provide pattern recognition from six radar detectors along a highway (US 290) in Houston, Texas. This multi-level approach achieves commendable results in classifying traffic in three categories: periodic trends, deterministic flow, and traffic volatility. In future studies, the measure of

volatility could provide useful information to the freeway/surface street boundary interaction and the likelihood and extent that overflow traffic affects arterial congestion.

In the Zhang et al. (2014) study, volatility of traffic is modeled to allow for a quantitative assessment of how variable the estimations of traffic flow can be, accounting for rate of change of the supplied traffic parameters. This estimation shows that in rapidly changing conditions, the estimates have less confidence and conclusions based on the estimate should be tempered with that knowledge. Through incorporation of a Glosten Jagannathan and Runkle (GJR) Generalized Auto Regressive Conditional Heteroskedasticity (GARCH) with the ARIMA process, the volatility of particular estimates can be measured acknowledging the asymmetric nature of traffic flow volatility. Whereas GARCH is optimized for single step predictions, GJR-GARCH allows for a threshold measurement for multi-step time predictions of the GARCH process (Monfared and Enke 2014).

One technique which is effective for smoothing variations in raw traffic data is Kalman filtering. In a general sense, Kalman filtering is commonly used to smooth various data inputs, eliminating anomalies in the data stream. Since a Kalman filter only relies on the previous state to make an update to its prediction for the next state, Kalman filters are widely used in real-time processes. A series of weights associated with inputs are updated with each state prediction thus providing a smoothing effect on changes from state to state and the ability to account for outlier data inputs. As such, an adaptive Kalman filter assumes non-gaussian distributions when determining what subsequent data will look like. With a recursive loop in the filter, real-time data can be implemented with less concern about significant changes in weighting of inputs due to very high or very low parameter values. This can be applied to traffic flow studies focused on prediction by effectively limiting the effects of outlier data on the overall average. In their study, Guo et al. (2014) utilize a modified ARIMA process smoothed with an adaptive Kalman filter to

create a traffic flow model. The use of an adaptive Kalman filter would provide a normalized traffic flow pattern that serves as a baseline for a prediction based on a localized disturbance.

2.4. Basic Neural Networks

The increase in computing power and technology associated with the widespread ability to produce neural network models that can be implemented has allowed expanded exploration in the area of machine learning for traffic prediction (Gebresilassie 2017). Initial traffic prediction machine learning models were simple, low-dimensional networks which only provided marginal improvements over traditional methods. As studies explored combinations of neural network architectures and new neural network implementations, the prediction confidence values increased and the ability to fine-tune solutions for particular use cases has become possible. The challenge of relating arterial and freeway networks has yet to be effectively shown in reviews of existing literature.

A recursive or a back-propagation (BP) neural network is essential to creating predictions based on time-phased inputs. As the complexity of these types of neural networks increases, so does the flexibility and confidence of predictions. As a BP network adjusts its weights to a particular system, the ability to generalize and tolerate errors can be optimized (Liu et al. 2010). Most of the successful neural network implementations below use some method of iteration in the solution.

As a feed-forward neural network, the Radial Bias Function (RBF) neural network provides some basic capabilities for continuous data with a predictable accuracy. The local details for a short-term prediction are demonstrated by in the conference proceedings from Zhu, Cao, and Fan (2013). An RBF's hidden layers allow for the incorporation of non-linear elements such as those from a heterogenous road network. The hidden layers are then weighted linearly to provide accuracy across the temporal horizon of the data. Zhu et al. also apply a swarm (or

flocking) transform to the output of the RBF which approximates the tendency of traffic to group itself.

One approach closely related to the problem explored in this study is the use of Convolutional Neural Network (CNN). A basic CNN is highly optimized for image recognition and works well on rasterized input. This method incorporates a representation of a wide area traffic pattern as an image and then combines several "images" together to form the input to the CNN (Ma et al. 2017). By creating a matrix of values with the x-axis as time and the y-axis as space (or the position on a road segment), each value (m) at m_{xy} represents the average speed. Transforming this into a raster image allows a CNN to recognize spatiotemporal correlations of each element in the matrix. A series of images was used to train the CNN and weight the relationships between matrix elements. In the Ma study, only one channel (speed as the relevant variable) was used, though more complex images with more variables could be used in the image generation at the cost of processing complexity in the CNN. This process leverages the image processing capabilities of a CNN and also provides spatiotemporal correlations across the study area. This proves to be especially useful when different types of roads are being studied.

2.5. Hybrid Modeling

The combination of data preparation and analysis using an amalgamation of traditional regression techniques and machine learning processes shows promise in several studies. The combination of basic building blocks in various ways builds on the proven techniques in traditional traffic flow analysis with the added advantage or deeper correlations only possible with machine learning. Combining methods can produce results with specific output characteristics useful in both long and short-term predictions.

One study uses a multi-task regression layer (MTL) in advance of a deep belief network (DBN) for its predictions (Huang et al. 2014). In this manner, the MTL produces results that

effectively provide “supervised learning” results without the need for manual interaction. Using these results, the DBN then operates in an unsupervised learning mode which it excels in doing. Internal hidden processing in a DBN assesses the contrast between time-phased inputs and essentially provides change detection which can then be interpreted into short-term predictions.

Dai’s DeepTrend method is a deep hierarchical network which combines a fully connected network for data extraction, and then feeds a long short-term memory (LSTM) network for prediction (Dai et al. 2017). The study focuses on the temporal relationships in traffic patterns, so the techniques would need to be adjusted to incorporate the spatial relationships needed in this thesis.

An LSTM is a very effective means to provide temporal awareness in a neural network and can be combined with a CNN in specific ways to bring both spatial and temporal predictions to the output sets. One drawback of a neural network without recursive memory is that change in state is presented to the next layer without regards to time. An LSTM manages the change in gradient through maintaining a memory of previous states and applying that change at an appropriate rate. With a cascading network of LSTMs and CNNs, time-phased spatially distinct inputs can initially be processed by the CNN for spatial relationships and then fed to two LSTM networks for time predictions (Yu et al. 2017). A process like this will help to connect the arterial congestion effects to the disruptions being observed on the freeway.

Zhao et al. provide a substantial use case of LSTMs for long-term (greater than 15 minutes) predictions. In contrast with Yu et al., this method uses a recurrent neural network (RNN) to achieve initial time-series weightings that the LSTM uses for correlation with an origin-destination matrix for the spatial element (Zhao et al. 2017). Since this thesis focused on traditional methods for short-term neighborhood effects, the application of Zhao’s LSTM method

is limited, but it addresses some challenges that a combined solution can help to mitigate in future studies.

A study of a Korean city's surface streets employed a combination of a gaussian mixture model (GMM) and an artificial neural network. This approach considers the effects of weather and road structure on congestion to provide an ITS with predictions based on current state changes (Oh, Kim, and Hong 2015). This approach is an effective one since it addresses the details of how various lane configurations can be affected by weather and construction projects but falls short of considering the effects of accidents. A variation on this approach is useful in designing a neural network solution for making predictions based on disruptions in nominal traffic flow. The GMM incorporates data which has several data types, like traffic, and combines the gaussian distributions for targeted variables into a cluster of predictions.

The ability to combine the concepts implemented in related work into a coherent strategy for traffic congestion visualization and future work in prediction is essential for success. The various combinations of techniques and tools used as well as the grammar framework discussed in paragraph 2.1 provide the building blocks for any analysis to meet its research objectives. In this case, the grammar and examples presented for multi-dimensional data visualization have been adapted to provide a geospatial representation of location, time and speed reports in a manner which provide insight into the association between freeway and arterial congestion.

Chapter 3 Methods

As seen in Chapter 2, there is a wide range of approaches to producing viable results for traffic predictions and associations. To limit the complexity of this project and provide a building block for future studies, variations on traditional spatial analysis were used. Machine learning will be reserved for future work. The key difference with this analysis compared to many existing studies is that the relationship between freeways and arterial streets is explored and any associations are visualized. The data required, and methods used are very similar for traditional and machine learning, but this analysis focuses on visualization techniques as a basis for showing how the speeds of the different road types change during a set time period and how connected roads affect each other.

3.1. Data Description

Since this thesis is focused on specific road segments and their attributes which are compared to changes over time and an average baseline, the data used was drawn from sources which provide flow and speed data for the roadways in the study area. Supporting data like the National Highway System data help to provide context which enables visual and mathematical analysis of the segment data. Table 1 describes the data used in this thesis.

Table 1. Data Descriptions

Source	**Dataset**	**Availability**	**Scale**	**Spatial Precision**	**Temporal Precision**	**Accuracy**
HERE.com	Traffic Flow	Via agreement	Filtered for Study Area	Road segment specific data specified by international standards (TMC), GPS accuracy is used in collection	Sampled at approximately 1 minute intervals	Traffic Message Channel (TMC) open standard using approximated WGS84 coordinates
CA Highway Patrol	Incident Data	Open Source	Statewide	CHP Incident data is approximated by standardized locations along a road segment with Lat/Long reference	Periodic reporting based on incident	Research indicates that most incidents are reported within 5 minutes of event.
CalTrans	Interstate Traffic Flow (Backup / Validation)	Open Source	District 7 (LA and Ventura Counties), Filtered for Study Area	CalTrans Flow data is collected by sensors at standardized locations along a road segment with Lat/Long reference	30-sec and 5-min sampling	GPS derived point sources for flow data
US Dept of Transportation	National Highway System	Open Source	National	Static	Not Applicable	Various accuracy based on input mechanisms

3.2. Research Design

As a spatiotemporal problem, traffic analysis needs to be particularly sensitive to the requirements to show how both time and distance affect traffic. Since roads are often bi-directional, or closely spaced with opposite directional flow, it is essential for a geospatial analysis to distinguish variations between road network direction and sensor data gathering results. The traffic speed data in this study clearly delineated direction, which allowed for specific routes to be analyzed properly. A directionally specific network dataset for the arterials and freeways in the study area was matched to the sensor data.

This thesis followed the methods for spatial analysis research as outlined in Montello and Sutton (2013). The selection of traffic flow data consisted of a single week chosen to create a baseline for normal traffic flow and several days of traffic in which an incident occurred. The data was collected at approximately one-minute intervals. The baseline data was analyzed and aggregated to provide a representation of nominal traffic flow for the morning rush hour (0600-1000). This sample set allows the assessment of the daily variation in traffic speeds for arterial city streets and highways. The incident days' data was compared to weekly average to determine if speed changes on freeway and arterials were associated. Selection of days with freeway incidents in the study area allowed for specific analysis of how a "non-standard" day is affected by an incident. Since workday traffic patterns are different compared to weekend traffic, this study only included weekday traffic in its analysis.

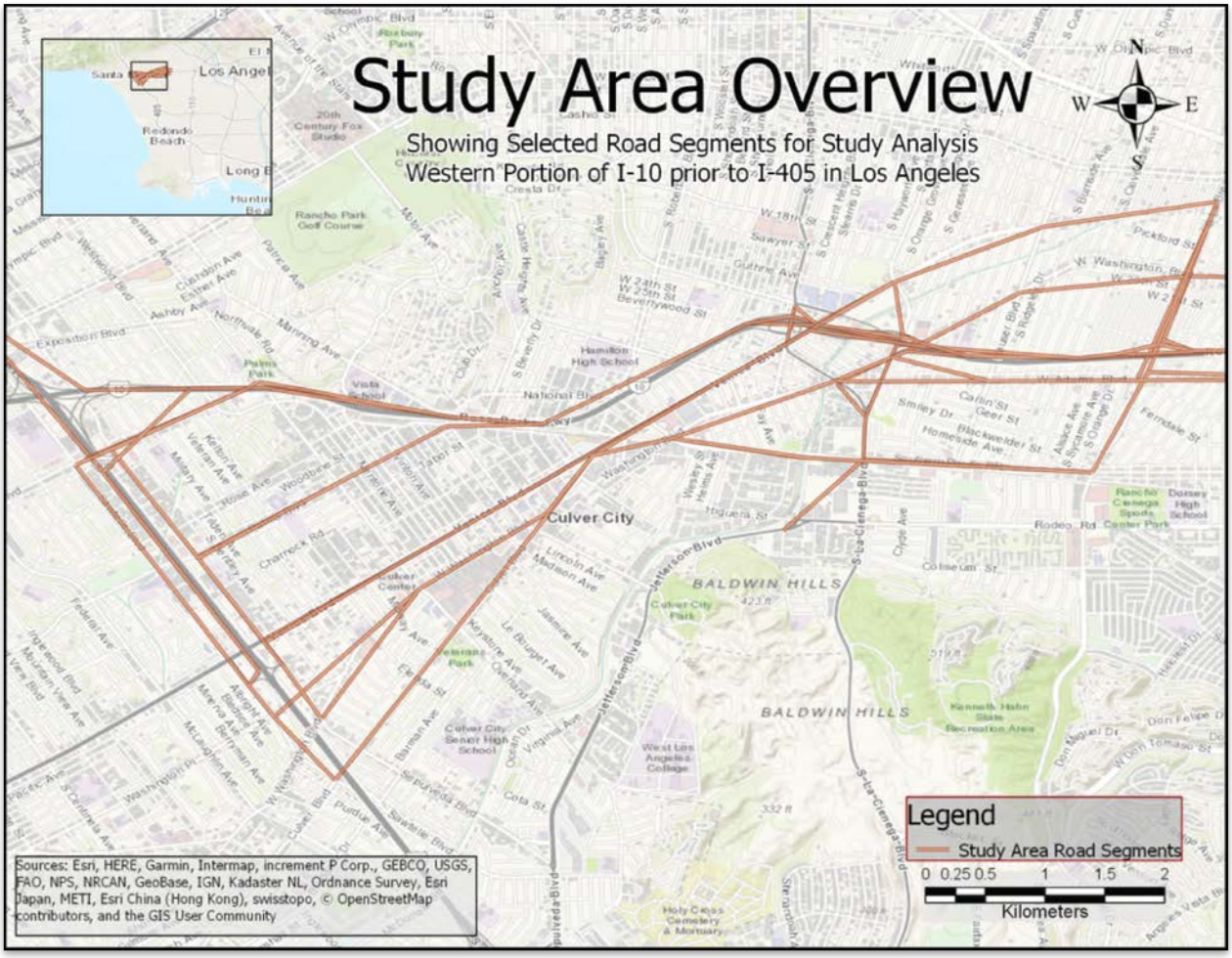

Figure 1: Study Area Overview

The traffic study area extends from mid-city Los Angeles to the Marina Del Rey area (Figure 1). This study area presented several options where the freeways (I-10W and I-405S) are often the fastest outside of rush hour times, but the arterial streets are viable options during peak traffic periods. Portions of Interstates 10 and 405 are included in the study. Within the study area, arterial roads were selected based on likelihood that they would be used as an alternative to freeway travel. Arterial roads included in the analysis were Venice Blvd., Washington Blvd., Adams Blvd., Jefferson Blvd., National Blvd., La Brea Ave., La Cienega Blvd., Fairfax Ave., Robertson Blvd., Sepulveda Blvd., Sawtelle Blvd., Palms Blvd., Culver Blvd., and Overland Ave. These are the primary arterials that can be used as alternate routes to freeway travel and

present logical alternatives with several options for travelers heading west. Other streets in the area were not represented in the modeling to simplify the variables associated with smaller side streets, even though personal routing systems like Waze or Apple Maps may use these smaller arterials as options. The sampled streets were used to provide the model with a directed graph for analysis. The traffic network nodes consist of attribute data for position and connections. Each edge in the spatially directed graph has length, free flow speed and directional attributes. In addition to the static spatially representative graph, a time-spatial graph was created containing the various states (current speed, variation from baseline, location and time series) for each network edge between the spatial node definitions.

The traffic flow samples for analysis were chosen to identify and show the effects of incidents like accidents or closures. Data from 15 minutes before and 45 minutes after an incident provides enough data to assess potential associations during the onset of each discrete incident and the aftermath of the incident. Using speed data from HERE.com and incident data from California Highway Patrol (CHP) reports, the freeways and arterials between mid-town LA and Marina Del Rey were examined to determine the relationship of interstate incidents to decreases in city traffic speeds. For each CHP incident report, the speed data was analyzed to ensure that the accuracy of the reported incident time corresponded with observed changes in traffic speed, particularly on the freeway.

An initial concern with the CHP incident data that delays in report timing compared to actual incident may affect the results by causing mis-correlations in timing has been mitigated with several sample analyses. Results show a correlation of reporting and resultant flow/speed changes within approximately five minutes of the reported time.

As suggested in various related works in Chapter 2, traffic speed, time, and differences from average speed were the primary variables considered. Traffic speeds allowed for empirical

determination of how much traffic the interstate can withstand even with a partial blockage of lanes before significant impacts on surface streets. As a further study, calculation of volatility would help to assess how quickly the situation on the freeway (and also on local streets) can change both in a positive as well as negative way. A measure of volatility will allow for future calculations of specific risk/benefit assessments for rerouted vehicles to help freeway travelers answer the question: “Is it worth taking a detour?”

3.3. Study Parameters

The study area encompasses a box with the Northeast corner in the Mid-city area of Los Angeles (34.0477415°N 118.3297119°W), and a Southwest Corner near Marina Del Rey, Los Angeles (34.0043824°N 118.4415229°W) as shown in Figure 1.

Two years of CHP incident data was examined for incidents during morning rush hour (0600-1000) along I-10 West. Ten incidents were initially reviewed which had potential to provide good case studies. After georeferencing this set of incident data and analyzing the time frame which traffic would be affected, the list was refined to five. These selected incidents fall within the rush hour timeframe (0600-1000) and range from 0651 to 0850 local time and are categorized by CHP with codes 1179 (Traffic Collision) and 1141 (Ambulance Enroute). The reported durations of the selected incidents range between 100 and 252 minutes indicating that the disruption was severe. See Figure 2for a map of the incidents notated with date and time and the sensor collection locations. Days of selected incidents were: April 17, 2017, July 17, 2017, July 21, 2017, November 16, 2017 and May 31, 2018.

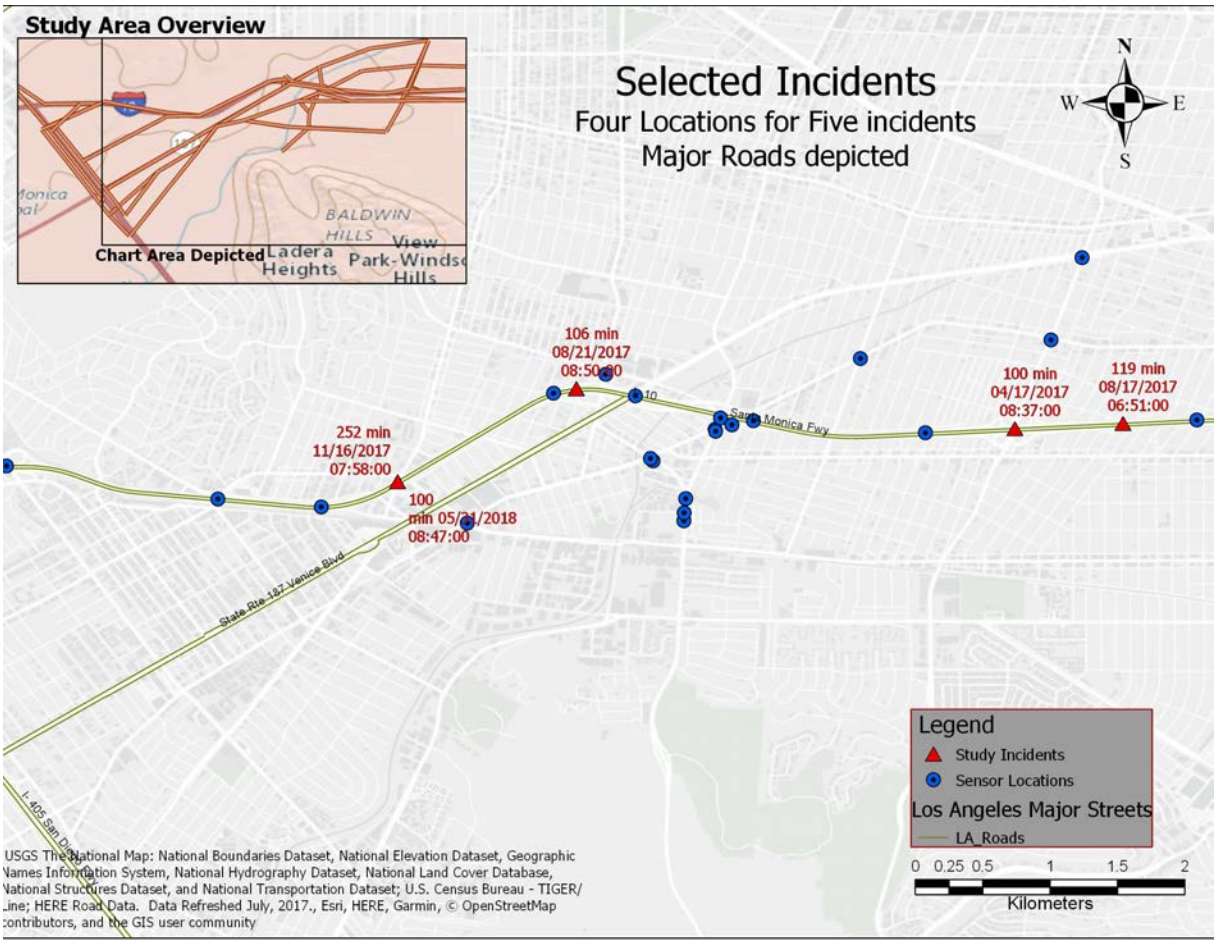

Figure 2: Selected Incidents

The speed data was collected for the 0600-1000 time range covering days with freeway incident reports that could affect arterial congestion. The data was presented as raw XML data from HERE.com and was filtered to the road segments in the study area. Baseline speed averages were calculated for each minute on each road segment from the week of April 16, 2018. This week was chosen because there were no significant incidents on the I-10 West or on the arterials in the study area which would impact a calculation of average speeds for each of the road segments during the morning rush hour.

Road network information was stored in a Neo4J database for directed graph analysis, Alteryx data stores for calculations and comparisons, and an ArcGIS geodatabase for spatiotemporal visualization. Graphical speed analysis in Alteryx and a series of ArcGIS Pro

spatiotemporal time-space cube visualizations provided the necessary tools for the determination of association between incidents and arterial road speeds. Since no predictions were made in this study, predictive tools like GJR-GARCH and ARIMA were not implemented; however, the data can provide inputs for these processes if future predictive studies are started.

Analysis of data for spatiotemporal correlation of traffic speeds resulting from freeway incidents through time-space visualization provides an effective method for the complicated interactions between physically separated freeways (due to limited access freeway exit structure) and arterial connectedness once travelers have left the freeway.

3.4. Data Preparation

Raw data for road segment speeds was presented to the Alteryx analysis program in a large set of XML data covering an area beyond the extent of the study area. Data files were sequentially parsed to include only the road segments in the study area. Data was then joined with static road segment data from the National Highway System database and Here.com road segment definitions. Figure 3 shows a portion of the model for the initial data file iterations and XML data extraction.

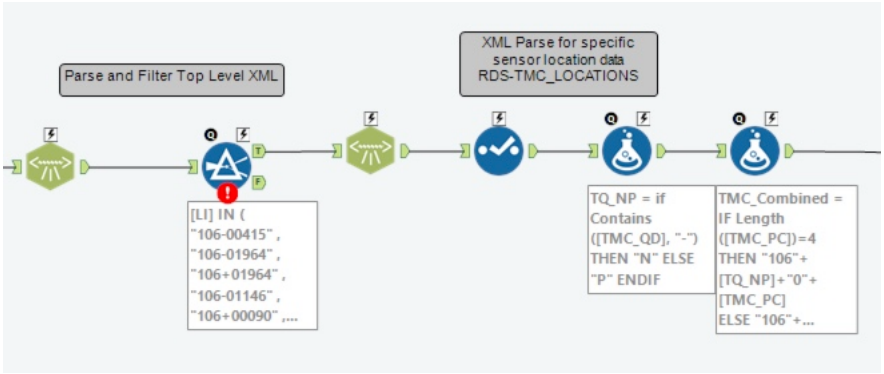

Figure 3: Alteryx model for iteratively parsing XML files with raw sensor data.

Segment averages were calculated from the raw data to provide a flat, single row per minute, reference for each sensor reading to ease manipulation. Final data fields are shown in Table 2. Times were rounded to the nearest minute to normalize the sensor data since raw data was sampled at different second intervals within each minute.

Table 2: Description of Flow Data After Preparation Used in Analysis

Traffic Flow Data		
	Actual_PBT	Actual Time of Sensor reading
Super Key	Rounded_PBT	Time of Sensor Reading to the nearest minute
Super Key	TMC_Ref	Road Segment Code
	TMC_DE	Road Segment Text Description
	CF_SU	Uncapped Speed reading from sensor
	CF_SU_Avg	Average Speed on road segment from baseline week
	CF_SU_Pct_Diff	Ratio of sensor reading to baseline average
	CF_FF	Free-Flow speed for road segment
	CF_Pct_FF	Ratio of sensor speed to free-flow speed
	LongX	Sensor Longitude (WGS-84)
	LatY	Sensor Latitude (WGS-84)

Line charts included in the analysis for each incident show the Speed Uncapped (labeled as measured speed in the figures) and the Baseline Average Speed for each road segment during the time period shown. The uncapped speed is the actual speed measured by the sensors. A separate data element from the raw data limited the reported speed capped to the speed limit of the road segment; this artificially limited speed was not used in the analysis. The average speed is the calculated value from the reference week as comparison. Figure 4 shows the model used for combining sensor data with the reference week and calculating the difference. Each record contains the actual data and the reference data to allow ArcGIS Pro to use in the time-space cube analysis tool.

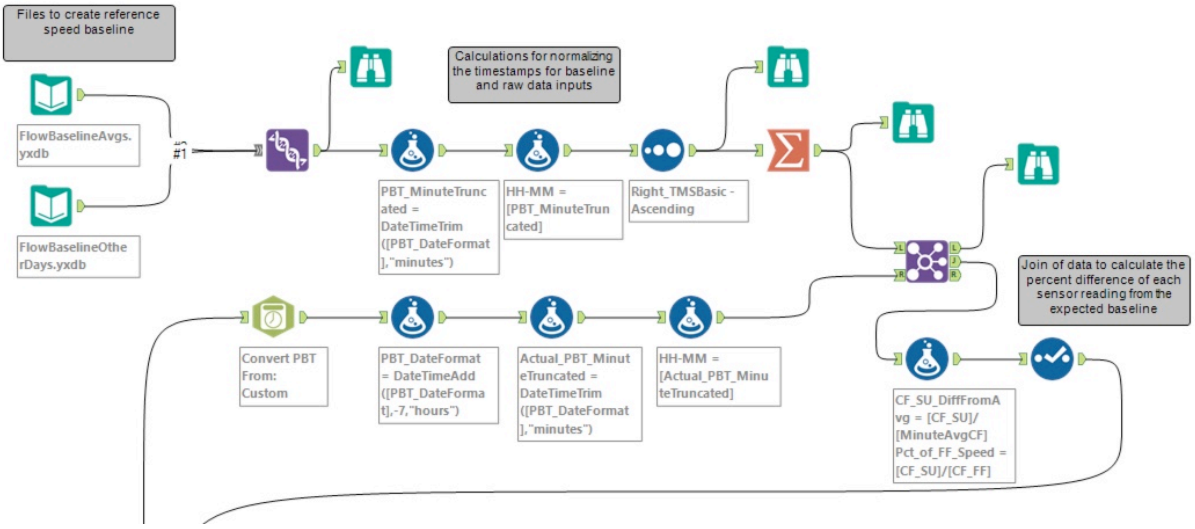

Figure 4: Alteryx model for joining reference data and calculating difference with raw sensor data.

The raw data provided free-flow speed for road segments as well. Free flow speed was used to compare how slow the general traffic flow was for that time of day. In ideal traffic conditions (e.g. late night or very low congestion), uncapped speeds were at, or above, the free-flow speeds. Extraction of speed data within timeframe before (15 minutes) and after (45 minutes) incident was done to show localized impacts to traffic speed.

Initially, ten incidents were considered for comparison from the CHP Incident database from the section of I-10 West in the study area. Following spatial and time analysis, the list was pared to five specific incidents which could potentially have had the greatest impact on arterial congestion. Incident data was georeferenced and imported to ArcGIS Pro for processing. Figure 2 shows the location of the selected incidents relative to the speed collection locations.

Selected road segments were also georeferenced, simplified to straight lines, and categorized to show which segments had sensor data, and which did not, as shown in Figure 5. Road segments with sensors are labeled in the figure. The relatively low number of sensor locations (27) of the selected road segments (93) may have contributed to results with fewer associations. This potential drawback was mitigated by the selection of incidents which took

place before or after a high density of arterial sensors near the intersection of Washington Blvd and I-10 in the center of the study area.

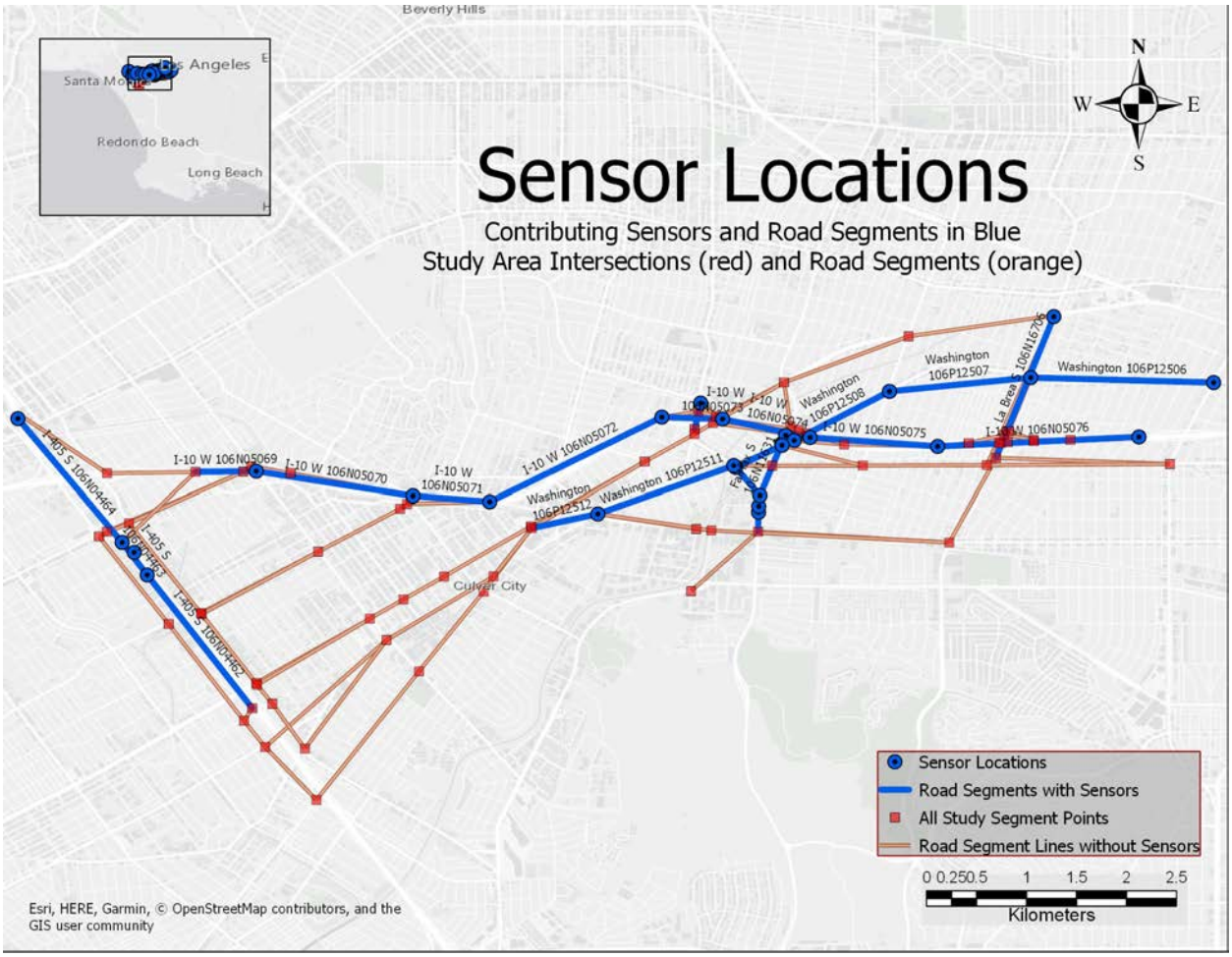

Figure 5: Selected Road Segments and Sensor Locations

The methodology and data preparation produced a streamlined model to ingest raw data and create a visualization of traffic speed variations across a designated time period. This model can be utilized for any set of data for any time period. In the final step, manipulation of the ArcGIS Pro time-space cube is a manual process providing the analyst with a means to best represent the particular dataset provided and create a customized visualization for effective understanding of any sensor location-traffic speed relationships over time.

Chapter 4 Results

The results of this study show that there is not a significant association between increased arterial traffic as a result of a major incident on nearby freeways for incidents lasting one and a half to three hours which did not result in a freeway closure. While one incident (April 17, 2017) shows a higher possible association than the others, the overall analysis does not support the hypothesis that an increase in arterial congestion can be observed as a result of an incident on the freeway. The following sections will describe the analysis for two incidents and data resulting in the unsupported hypothesis. The three remaining incidents are discussed in the appendices.

4.1. Time Series Speed Analysis Representation

The raw XML data was parsed and then prepared to include the various attributes and calculated data needed. Free-flow speed for each segment was recorded, the baseline average speed was calculated (averaged from the week of April 16, 2018), and the percent difference between these numbers and the measured speed was calculated to enable a time-series comparison of speed data for measured segments. This data is represented in the following sections. Only chart data showing pertinent trends and variations from averages are shown.

The first figure in each section is a time-space speed representation of the overall incident area. Figure 6 is a sample graphic created using the ArcGIS Pro time-space cube analysis tool. The column of traffic speed data was averaged into two-minute results to allow for simplified presentation and positioned over the sensor which collected the data. Column data points (circles in the stack) show percent difference from baseline average speed at each sensor location for a one-hour period. The legend shows the color divisions for the speed data. Size of the circle increases the further below average based on the reported speed. The semi-transparent gray circles represent differences from the average which are much higher than expected (greater than

133%) and are therefore considered outliers. The circles are stacked in time. The earliest time is at the bottom. The column starts 15 minutes prior to the incident for the date and time noted on the chart. The top of the column is 45 minutes after the incident. Each graphic was then oriented in three dimensional space to provide clear observation of the important time-speed columns for comparison.

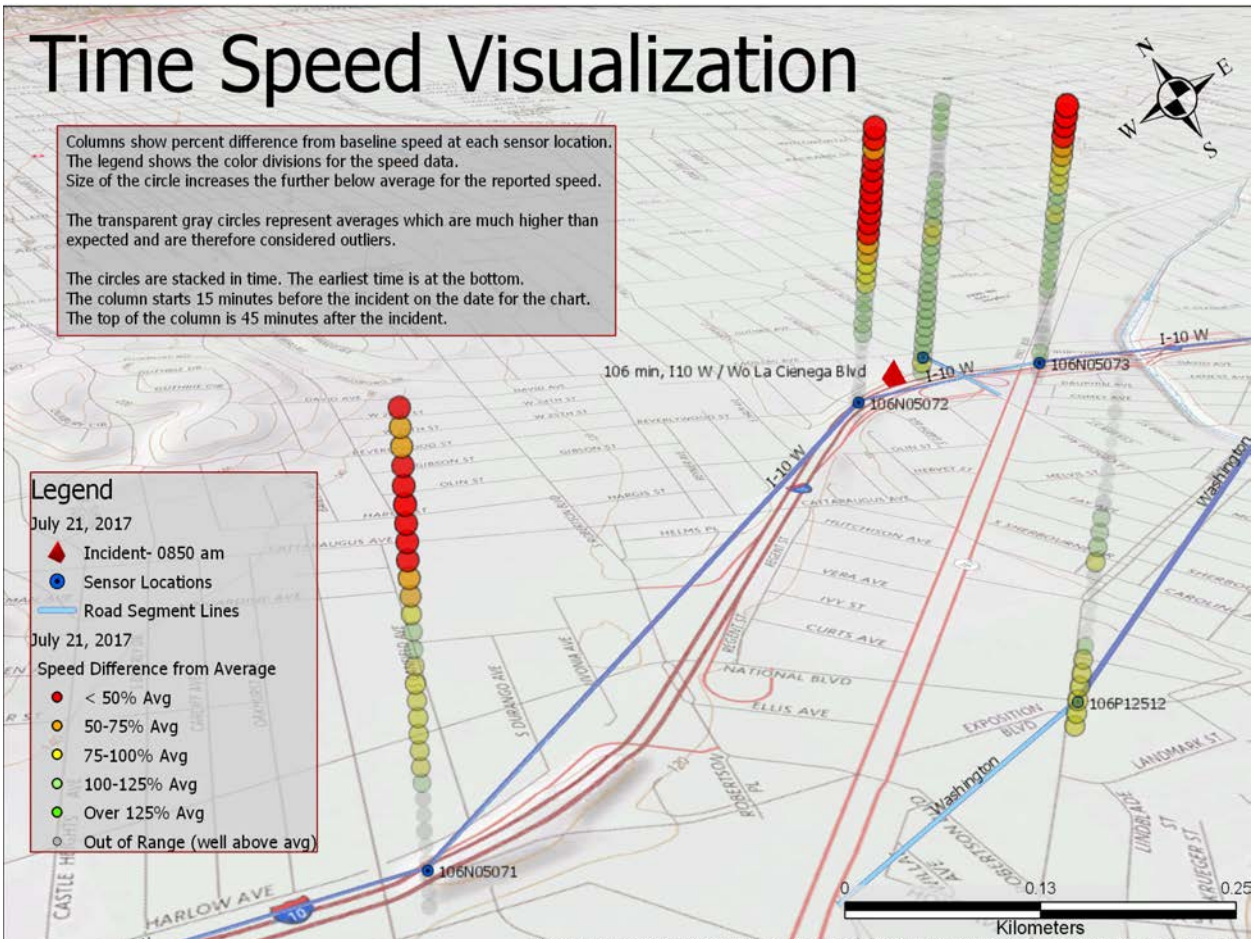

Figure 6: Sample ArcGIS Pro Time-space cube analysis tool output. 3-D orientation to enhance readability and highlight possible association between data stacks.

The second figure in each section is a satellite image showing the incident area and several of the sensor locations and analyzed road segments. This provides contextual information about the area and the street layout. The extent of each image was selected to provide the ability to analyze the way that vehicles exiting the freeway will enter the arterial network.

The final figures are line graphs for some of the sensors for the date of the incident. These graphs were output and annotated from Alteryx as part of the data preparation modeling. The particular graphs were selected to show data which was pertinent to the analysis discussion. Similar to the time-space speed column representation, the horizontal axis is the time axis and begins 15 minutes prior to the incident and extends to 45 minutes after the incident. The vertical axis is the speed in miles per hour. Two series of data are shown in the chart. The green lines are the actual sensor readings for that day (Speed Uncapped in the data description [Table 1]). The blue lines represent the average speeds expected on a normal workday as calculated from the reference week. The incident time is marked by a vertical red line for reference. The alphanumeric code (e.g. 106N0575) is the unique identifier for a particular sensor known as the Traffic Message Channel (TMC). A textual description of its location is annotated at the top of each graph.

Two analyses are presented in the main body of this study. The selected incidents show scenarios where the best indicators for associations can be shown. The first has weak associations, but not clear causality. The second shows an ambiguous data representation which could indicate inaccurate incident data but could also show early indications of arterial as a result of the freeway incident. The details of these scenarios are discussed in their respective sections. Three additional incidents are described in the appendices. The incidents in the appendices demonstrate either, or both, a lack of association as well as unexpected behavior of the traffic, geographically, before and after the incident location.

4.2. Incident on April 17, 2017

This incident took place near the La Brea exit on I-10W at 0837 and had a duration of 100 minutes as reported by California Highway Patrol. Figure 7 shows much higher than average speeds along I-10W at locations 1400m ahead (sensor 106N05076), 700m after (sensor 106N0575) and 2km after (sensor 106N05074) the incident location, before the time of the incident and for nearly 30 minutes following the reported incident time.

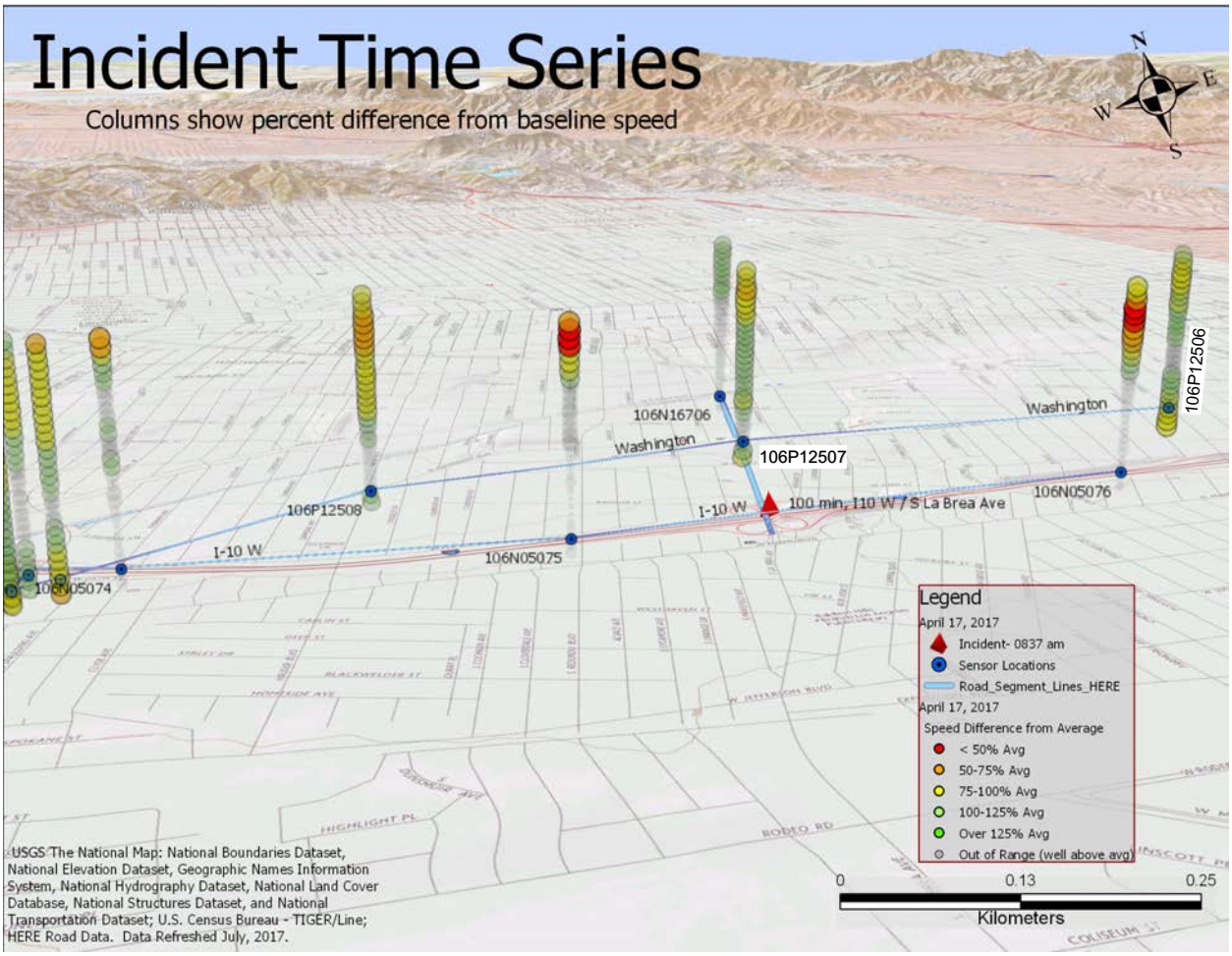

Figure 7: April 17, 2017 Incident Overview

The satellite imagery in Figure 8 shows a separated lane structure for vehicles exiting onto South La Brea Ave. and potentially passing the collection stations on Washington Blvd. Depending on the actual lane where the incident occurred, there could be different results for overall restriction of traffic movement in this area since the bypass or main freeway lanes could provide a way to avoid the accident without leaving the freeway.

Figure 8: April 17, 2017 Incident Area Imagery

At 30 minutes following the reported incident, both 106N5075 and 106N5076 show a drastic reduction in speed (Figure 9 and Figure 10), but the trend is slowing beginning at the incident time.

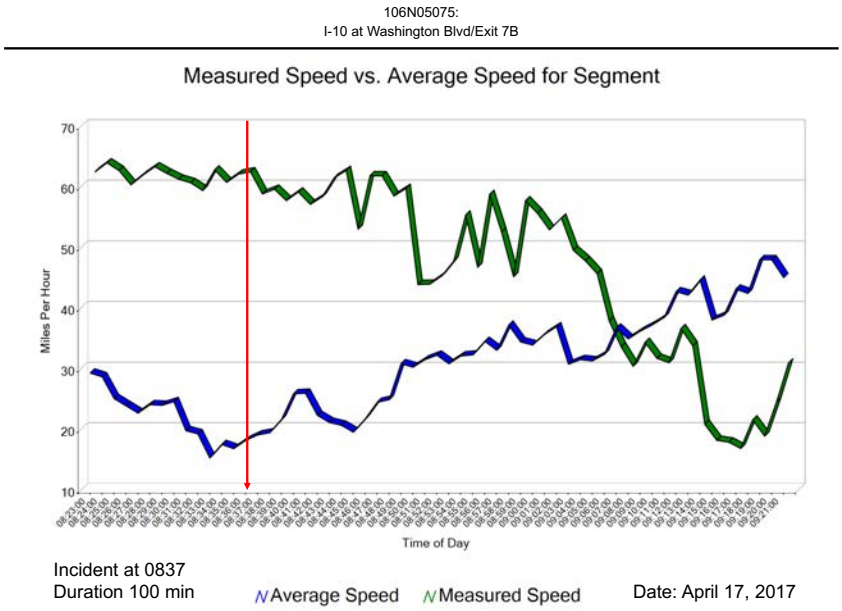

Figure 9: April 17, 2017 106N0575: Free-flow speed=63mph

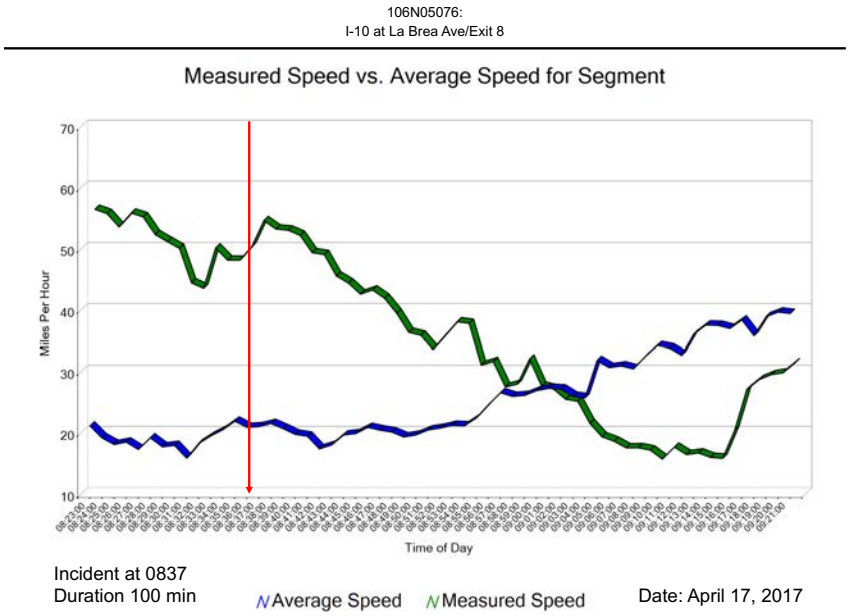

Figure 10: April 17, 2017 106N0576: Free-flow speed=63mph

While Washington Blvd (106P12507) and its westward continuation (106P12508) initially show above average speeds, they also show a slowdown well below average about 30 minutes after the incident (Figure 11 and Figure 12).

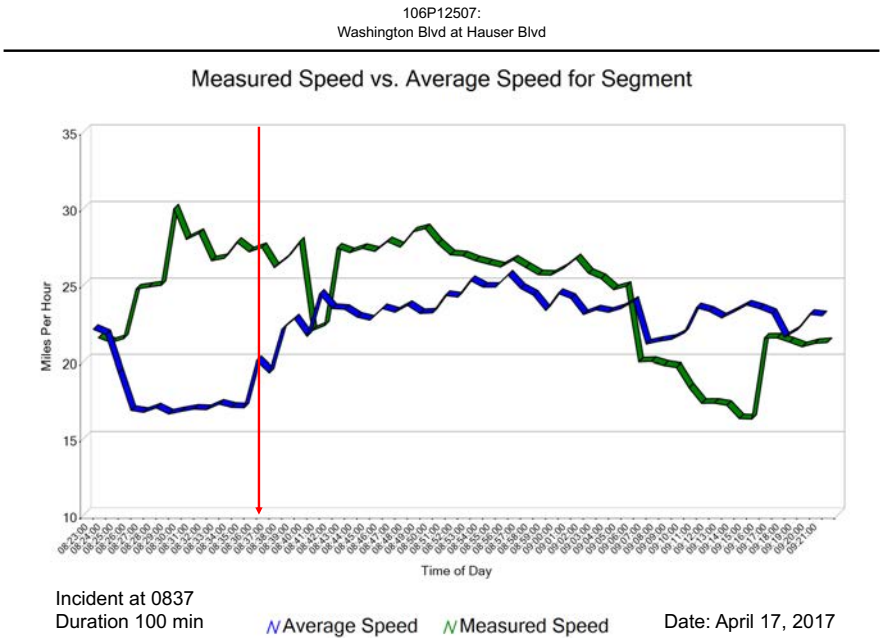

Figure 11: April 17, 2017 106P12507: Free-flow speed=31mph

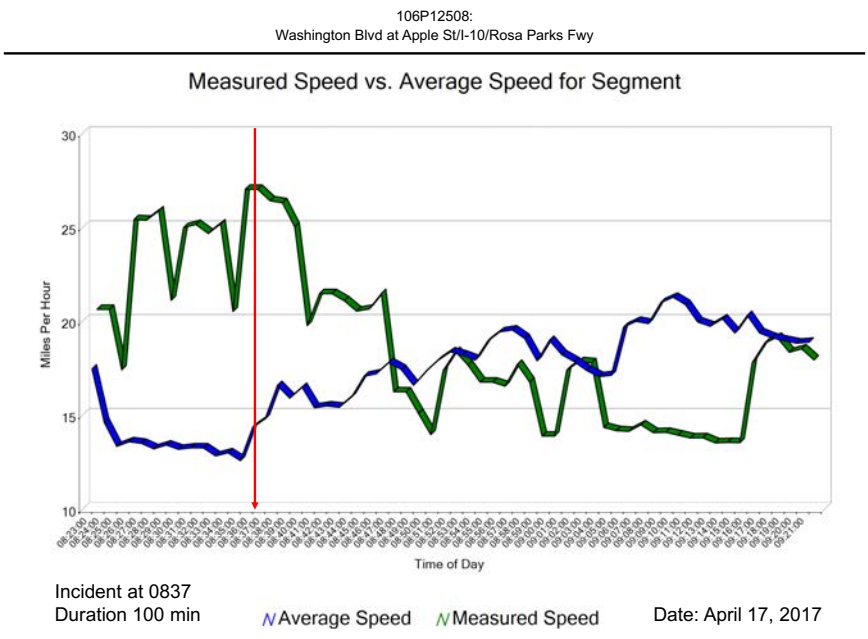

Figure 12: April 17, 2017 106P12508: Free-flow speed=24mph

If the assumption is made that incident did not fully develop into increased congestion until 15-20 minutes after the reported time, it is possible that a weak association could exist in the noted road segments. The timing of the slowdowns on Washington Blvd. are consistent with the incident causing slowdowns on the freeway and more than normal amounts of traffic potentially self-redirecting to Washington Blvd. westbound via La Brea northbound as an alternate route to the west. The data does not indicate a specific causal effect as a result of the slowdown, but the data is consistent with higher than average volume on Washington Blvd. which would result in lower than average speeds at a time consistent with how long it would take for the traffic to flow to the sensor locations on Washington. 106P12508 (the more western sensor on Washington Blvd.) shows below average speeds earlier than the sensor at the intersection of Washington and La Brea to the north of the incident location (106P12507). This discrepancy weakens the argument for causality since driving distance and other potential factors in the traffic flow on Washington could be affecting speeds for different reasons.

4.3. Incident on July 21, 2017

This incident is reported to have taken place just west of the La Cienega interchange on I-10W at 0850 and had a duration of 106 minutes as reported by California Highway Patrol. Figure 13 shows above-average speeds along I-10W before and for nearly 30 minutes following the reported incident time.

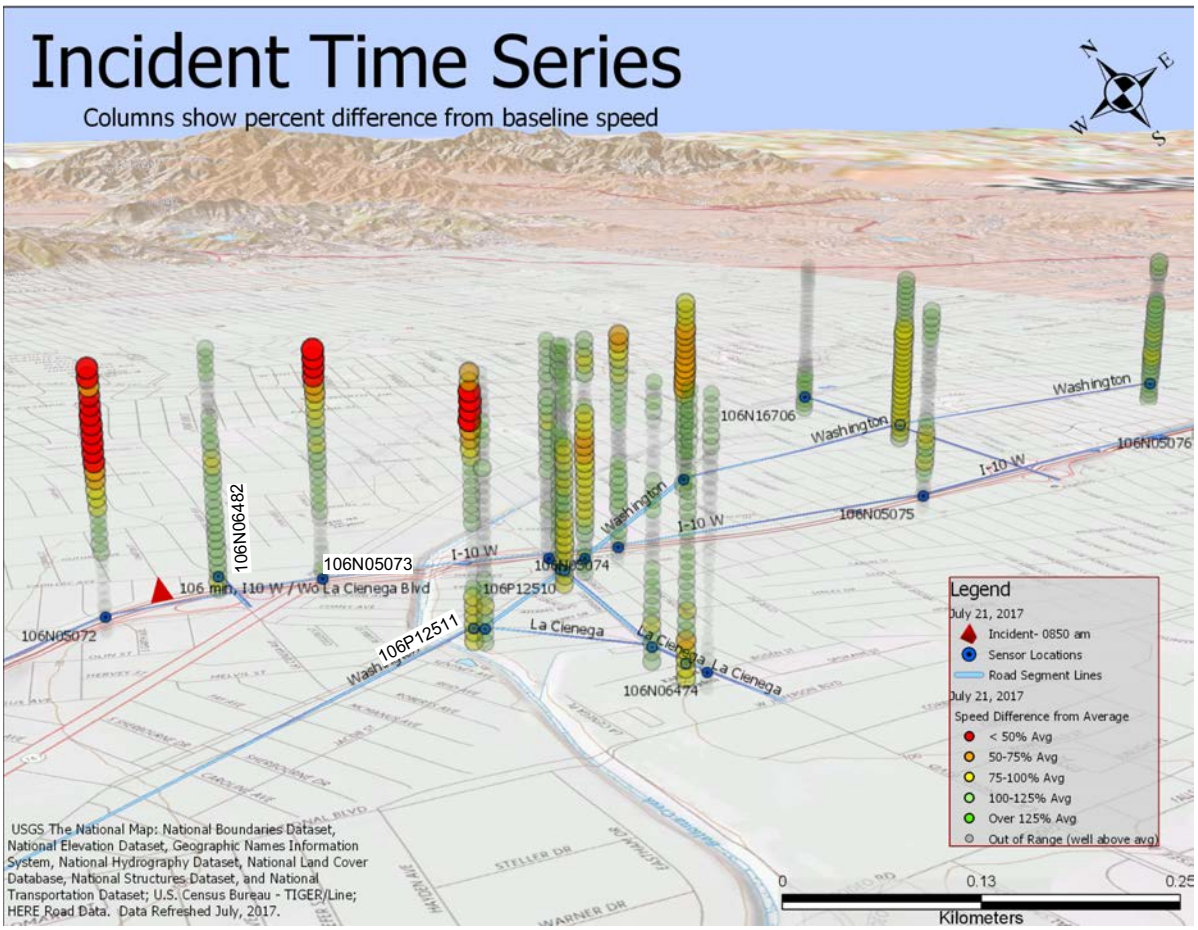

Figure 13: July 21, 2017 Incident Overview

Figure 14 shows a complicated picture of how vehicles can exit the freeway. Two nearby opportunities for leaving the freeway serve Washington Blvd and Fairfax Ave. about 1.4km prior to the incident location with a nearer exit for Venice Blvd. and La Cienega Blvd at 1km prior to the reported incident location. Due to the multiple arterials being fed and the subsequently larger number of routing possibilities for vehicles to take as an alternative to the freeway, this could have caused the ambiguous results noted for the July 21 incident.

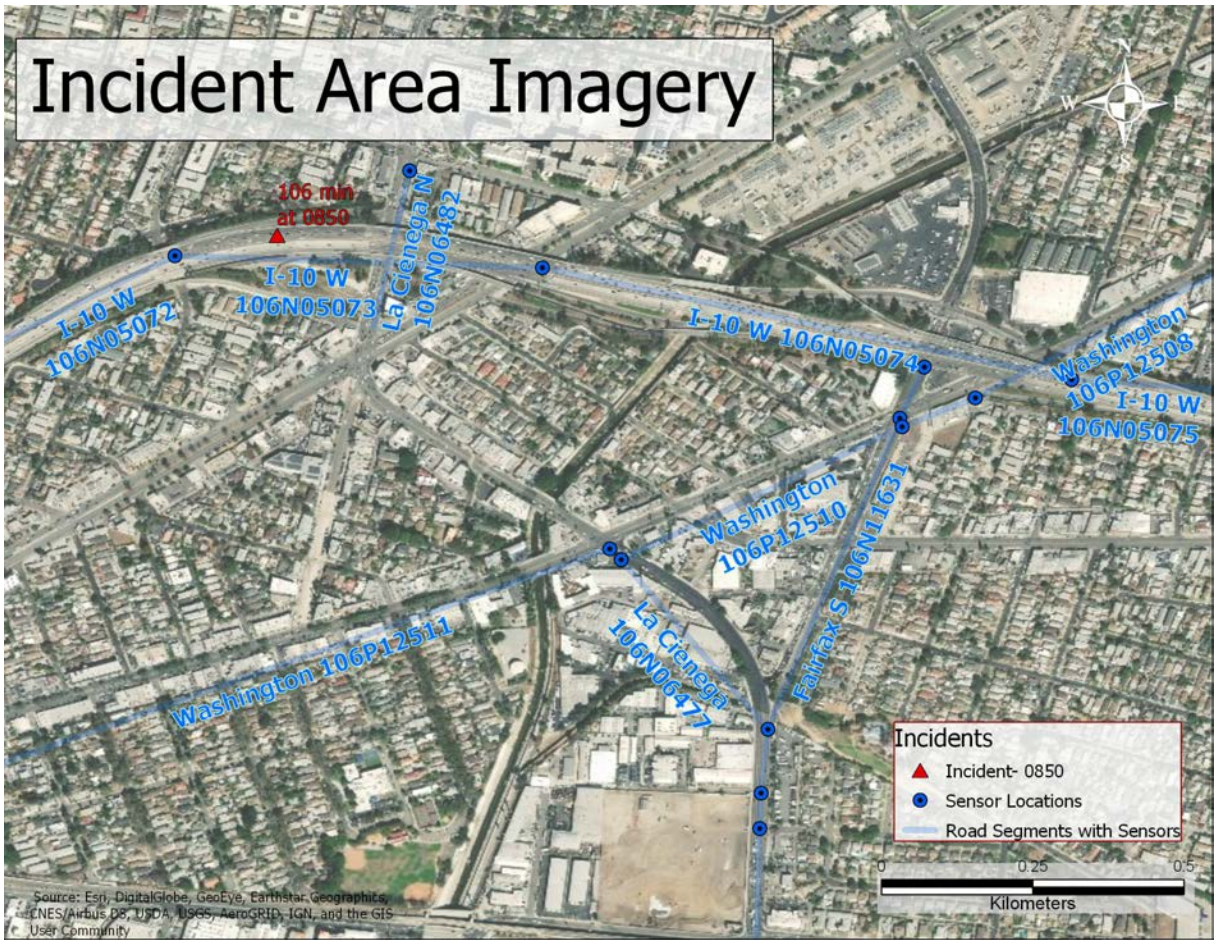

Figure 14: July 21, 2017 Incident Area Imagery

At 30 minutes following the reported incident time, both 106N05072 (leftmost column, with red on the top half, graphed in Figure 15) and 106N05073 (third from left column, with a red top, graphed in Figure 16) show a dramatic reduction in speed 15 and 30 minutes

respectively after the incident. Since the location of 106N05072 is further down the freeway from the reported incident location, it is suspected that the actual incident may have occurred at a location west of the reported position since a natural pattern would be for 106N05072 to positively diverge from average speeds due to a backup or slowdown prior to the sensor location and traffic flowing smoothly after the incident and 106N05073 to slow down earlier. The dotted green arrows show the expected behavior for these locations. In this graphic, a dramatic slowdown several minutes after the incident at 106N05072 is mirrored several minutes later at the previous sensor. This would be normal if the incident happened further west on the freeway from 106N05072.

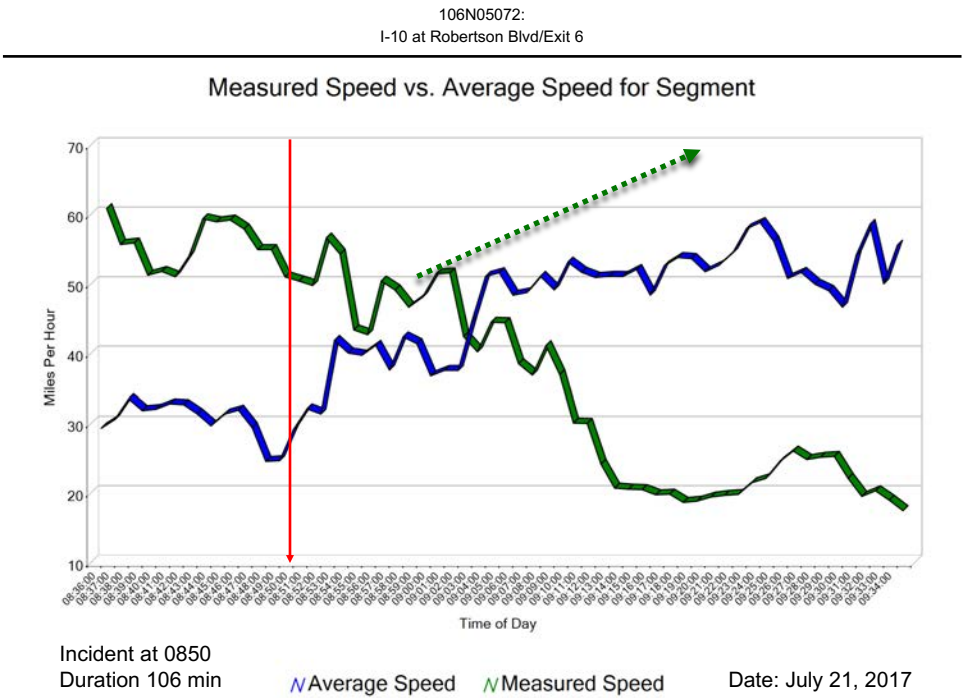

Figure 15: July 21, 2017 106N05072: Free-flow speed=65mph. Expected speed behavior is noted with the dotted green arrow.

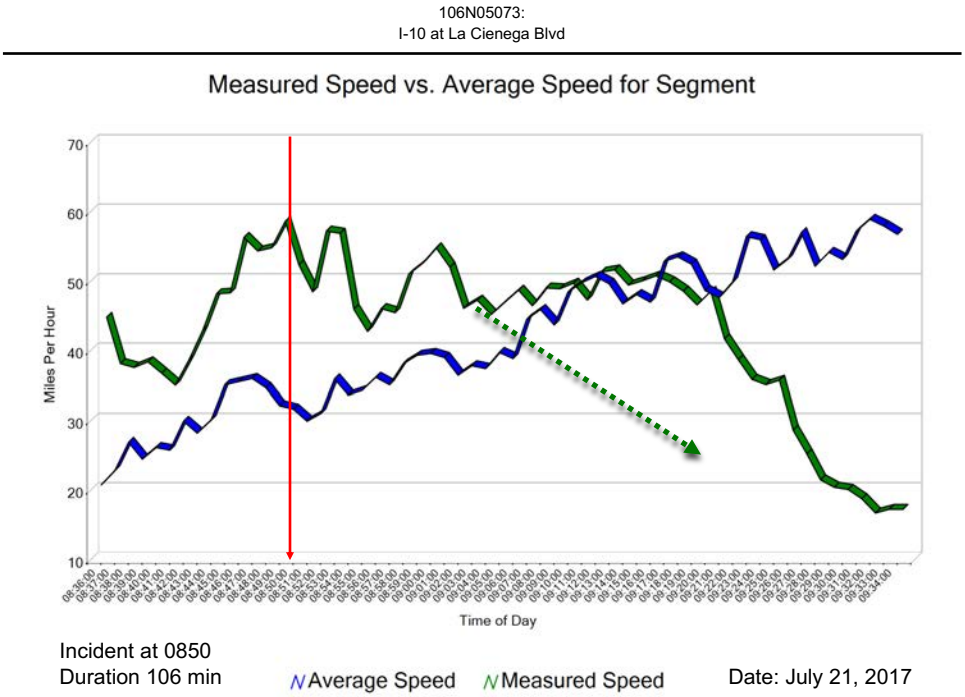

Figure 16: July 21, 2017 106N05073: Free-flow speed=65mph. Expected speed behavior is noted with the dotted green arrow.

In the overview graphic for this incident, the second column from the left is 106N06482 (La Cienega Blvd, Figure 17) and shows mostly higher than average speeds during the period following the incident. The exit from I-10W for that road segment is well before the I-10W sensor, 106N05073. This could be an indication of a delayed effect of less traffic exiting the freeway than normal resulting in higher than average speeds on that segment of La Cienega. The fourth column from the left (which has red at the top) is Washington Blvd. (106P12511, Figure 18). The data points in this column show a potential association, but the timing and lack of similar effects on nearby sensors make it an unlikely resultant effect of the incident.

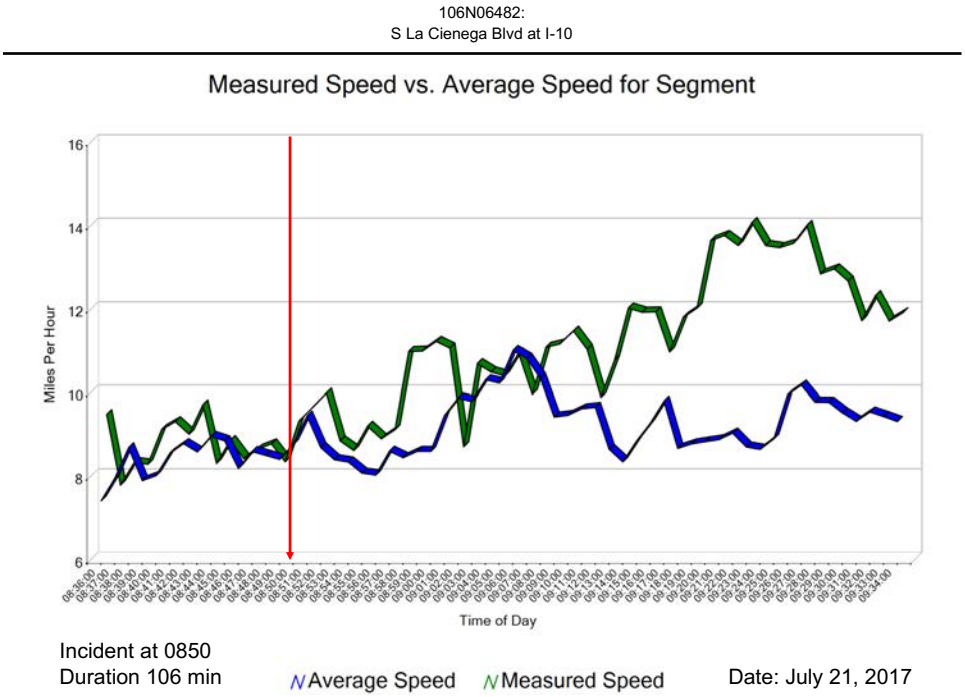

Figure 17: July 21, 2017 106N06482: Free-flow speed=26mph

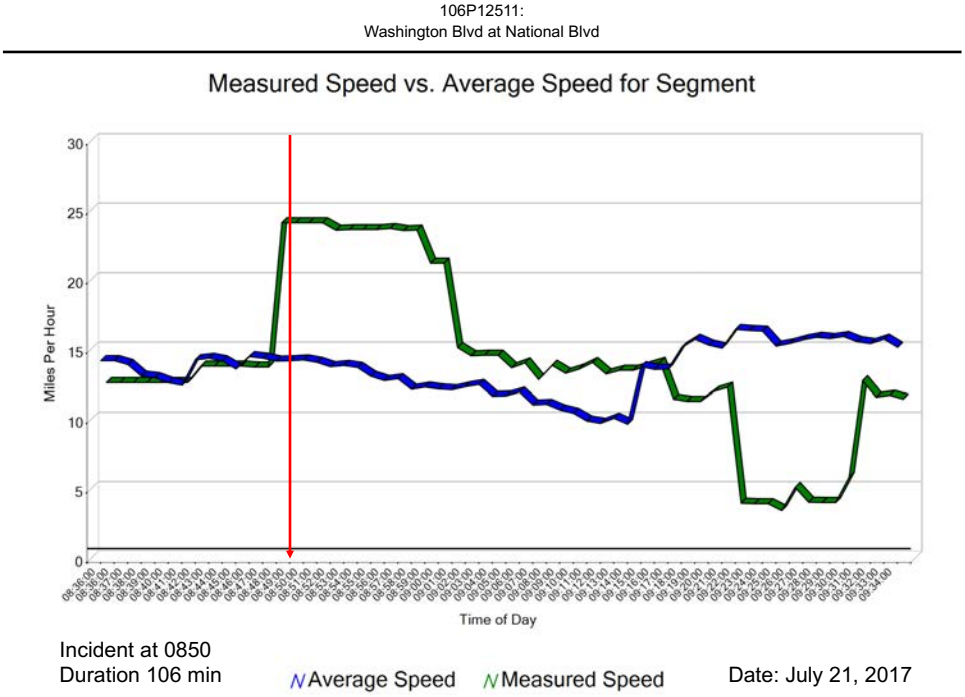

Figure 18: July 21, 2017 106P12511: Free-flow speed=25mph

Chapter 5 Conclusion

The hypothesis for this study was that arterial traffic would be distinctly affected by significant slowdowns on the freeway. As seen in Chapter 4, this could not be expressly demonstrated. Several possibilities exist in why this may be the case. The overall summary of results, explanation of the challenges experienced with the study, and future work possibilities will assist researchers to reference this work in developing more robust research plans to confirm or refute these findings.

5.1. Summary of Results

With the exception of a weak potential association in the April 17, 2017 incident (Section 4.2), the incidents analyzed did not show any particular relationships with the variations in traffic flow on the arterials from expected behaviors. Inertia could play a key role in the explanation of these results. The inertia, in this case, is commuters becoming accustomed to a particular route and the expected traffic conditions along that route and thereby resisting choosing an alternate route. In the case of the study area, free-flow speeds of 65mph or higher on the freeways provide incentive to remain on the freeway.

Even in rush hour periods like this study examined, the average freeway speeds were near free-flow speeds until around 0730. After 0730, the speed normally drops to the 20-40 mph range (Figure 19). Since commuters are typically travelling at the same periods every day, the patterns become inherent in their expectations. A difference of 10 mph below the expected speed may not be enough of an inconvenience to prompt the commuter to seek an alternative route. Short of a complete closure, it is possible that if the traffic on the freeway continues to move, however slowly, there will not be enough traffic choosing to leave the freeway to materially affect the arterials.

106N05075

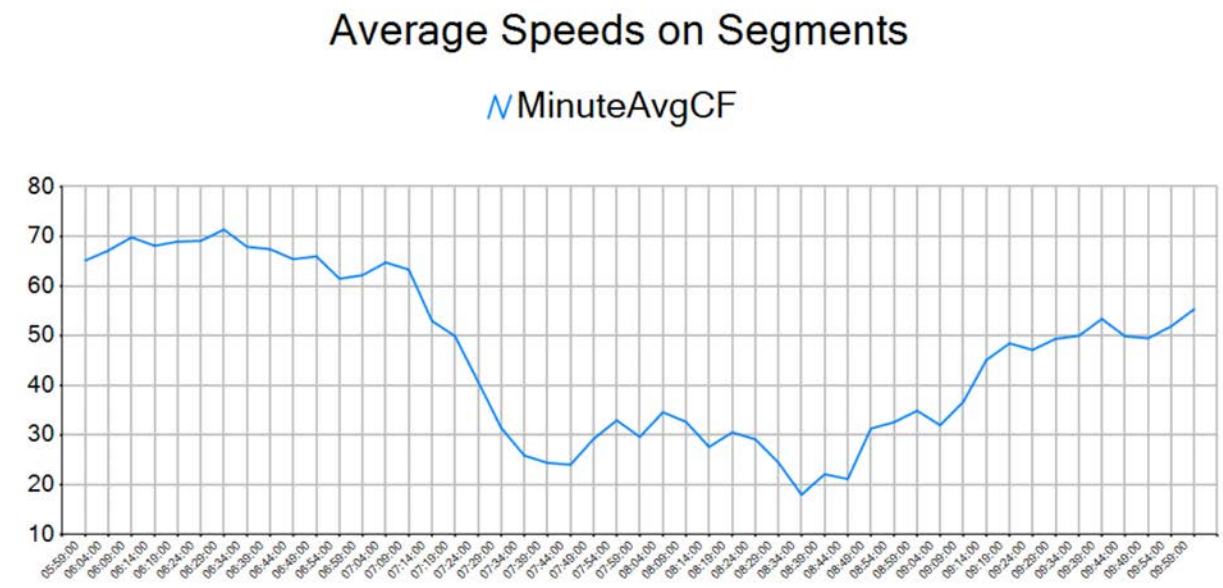

Figure 19: Typical Pattern for Freeways from 0600-1000. This graph is from I-10W at Washington Blvd. Free flow speed: 63 mph. MinuteAvgCF denotes average continuous flow (CF) for each minute.

Arterials showed a wider difference in daily patterns. Each arterial has a unique signature for average speeds during the study time period, but the general pattern for all arterials is a gradual slowdown until 0800 and slow recovery beginning around 0900 (Figure 20). Furthermore, the range of the changes in arterial speeds during the study period hover within 5-10 mph of the mean for the time period providing even less incentive for using those routes in lieu of the freeway.

106N06482

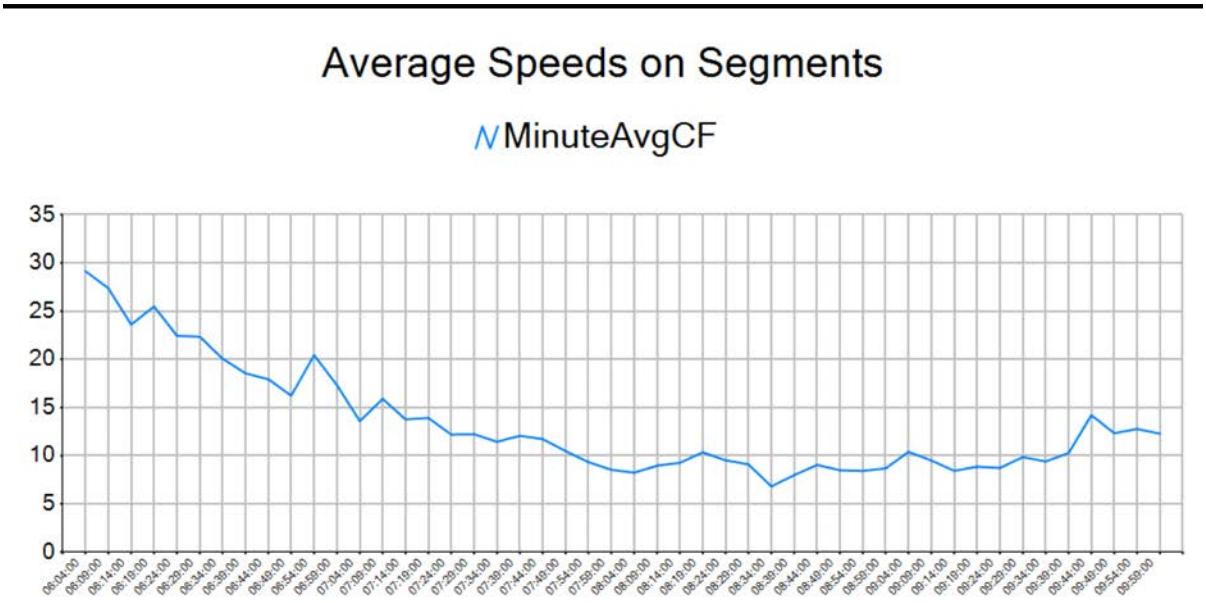

Figure 20: Typical Pattern for Arterials from 0600-1000. This graph is from La Cienega near the I-10 intersection. Free flow speed: 26 mph. MinuteAvgCF denotes average continuous flow (CF) for each minute.

Time and distance also play a role in the decision to leave the freeway. This study looked at arterials which provided a relatively direct route for commuters whose destination was south of I-10 along the I-405 corridor. If the perceived savings in time to use the arterials did not outweigh the cost of enduring the slower traffic along the freeway, then even commuters with enough situational awareness of alternate routes would not be tempted to leave the freeway. The proliferation of tools like Waze would assist commuters in making that decision and have more comprehensive information available for individual commuter routing needs.

5.2. Challenges Experienced

Of primary concern during the data exploration phase of this research was the development of best methods for parsing and filtering the raw data. Over 2000 individual XML files with varying timeframes and data elements were used to provide the study with the required data. Computer storage and processing power was insufficient for a complete concatenation and XML parsing of the data prior to filtering. Additionally, due to network accessibility restrictions and file size limitations, it was not feasible to import the bulk files directly into a relational database system. A workflow was developed to iterate through files and append filtered data by date and time-range to a composite file with a manageable size.

Sensor location data was a second concern. While the 27 sensors from the selected road network were enough to analyze many of the eastern intersections in the study area, there were some notable gaps in coverage which could have provided additional data points in the analysis. Additional sensors on Venice Blvd near the I-10W intersection would have helped to provide a direct comparison to the ones on Washington Blvd. Sensors along this roadway would have allowed redundant verification of the behaviors seen on two nearby parallel streets.

Visualization of temporal data is often difficult using a static method like a chart or graph. The ability to show an interactive animation or graphic provides easier understanding of the data presented. In this study, many spatiotemporal representations were evaluated, but a time-space cube visualization using ArcGIS Pro provided the most clarity by manipulating the tool parameters to show just the sensor locations and their data and orienting the perspective to best show geographic relationships as well as the temporal variations.

5.3. Improvements for Future Work

Even speeds significantly below average did not appear to provide enough impetus for commuters to leave the highway. Complete freeway closure may cause the expected effect but

was not studied in this analysis. During a research interview, it was learned that early results of a study at USC of broad neighborhood effects from a freeway closure using machine learning show similar outcomes (Yan 2018). That study uses road closures as the freeway incident input and compares a wide array of traffic light sensor loop data as flow information for training data in various neural networks. At the time of the interview, various machine learning methods had been tried, but predictive confidence values were still extremely low. From an analysis standpoint, this could confirm the weak association relationships like those experienced in this thesis.

Possible improvements to this study include selecting a time of day where average speeds were not already significantly below free-flow speeds. During the morning rush-hour, since speeds are already low, slowdowns would be easily overlooked. This could be further improved by selecting incidents which resulted in complete closure of the freeway or choosing more stringent criteria for incidents with clear indications of a severe impact on the freeway.

An alternative to selecting just severe incidents would be to provide the analysis model with a range of incidents from minor slowdowns to complete closures. With this wide range of incident types, it would be possible to determine if there is a "tipping point" when arterial traffic is clearly affected by freeway disruptions. As shown in this study, the criteria of a long duration traffic collision with medical services required did not seem to reach that tipping point.

Additional sensor data to create a more complete sensor network would help to more clearly show any correlations between sensors along the same roadway as well as along nearby streets. With an expanded network, the overall association of impacts could be better assessed since more intersections would be covered. Additionally, it would allow direct comparisons between nearby streets to support findings.

Study the past, if you would divine the future. - Confucius

References

Albright, David. 1991. “History of Estimating and Evaluating Annual Traffic Volume Statistics.” *Transportation Research Record: Journal of the Transportation Research Board* 1305: 103–7.

Bell, Karissa. 2016. “Waze Update Will Predict Traffic Conditions before You Leave.” Mashable. Mashable. March 2016. http://mashable.com/2016/03/17/waze-planned-drives/.

Chow, Joseph Y J. 2013. “On Observable Chaotic Maps for Queuing Analysis.” *Transportation Research Record: Journal of the Transportation Research Board*, no. 2390: 138–47. https://doi.org/10.3141/2390-15.

Dai, Xingyuan, Rui Fu, Yilun Lin, Li Li, and Fei-Yue Wang. 2017. “DeepTrend: A Deep Hierarchical Neural Network for Traffic Flow Prediction.” http://arxiv.org/abs/1707.03213.

Gebresilassie, Mesele Atsbeha. 2017. “Spatio-Temporal Traffic Flow Prediction.” *Kth Royal Institute of Technology School of Architecture and the Built Environment*.

Ghahramani, Zoubin. 2001. “An Introduction to Hidden Markov Models and Bayesian Networks.” *International Journal of Pattern Recognition and Artificial Intelligence* 15 (01): 9–42. https://doi.org/10.1142/S0218001401000836.

Guo, Jianhua, Wei Huang, and Billy M. Williams. 2014. “Adaptive Kalman Filter Approach for Stochastic Short-Term Traffic Flow Rate Prediction and Uncertainty Quantification.” *Transportation Research Part C: Emerging Technologies* 43: 50–64. https://doi.org/10.1016/j.trc.2014.02.006.

Huang, Wenhao, Guojie Song, Haikun Hong, and Kunqing Xie. 2014. “Deep Architecture for Traffic Flow Prediction: Deep Belief Networks with Multitask Learning.” *IEEE Transactions on Intelligent Transportation Systems* 15 (5). https://doi.org/10.1109/TITS.2014.2311123.

LADOT. n.d. “LADOT ATSAC.” ATSAC Website. http://trafficinfo.lacity.org/about-atsac.php.

Liu, Yuanlin, Wusheng Hu, Shujie Xin, and Li Li. 2010. “Prediction for Short-Term Traffic Flow Based on Modified PSO Optimized BP Neural Network.” *Research on Influence of Aggregate Gradation on the Performance of Porous Asphalt Pavement*, 3738–46.

Ma, Xiaolei, Zhuang Dai, Zhengbing He, Jihui Ma, Yong Wang, and Yunpeng Wang. 2017. “Learning Traffic as Images: A Deep Convolutional Neural Network for Large-Scale Transportation Network Speed Prediction.” *Sensors (Switzerland)* 17 (4): 1–16. https://doi.org/10.3390/s17040818.

Monfared, Soheil Almasi, and David Enke. 2014. “Volatility Forecasting Using a Hybrid GJR-GARCH Neural Network Model.” *Procedia Computer Science* 36 (C): 246–53.

https://doi.org/10.1016/j.procs.2014.09.087.

Montello, Daniel R., and Paul C. Sutton. 2013. *An Introduction to Scientific Research Methods in Geography and Environmental Studies*. 2nd ed. Thousand Oaks, CA: SAGE Publications. https://doi.org/10.1111/gean.12022.

Oh, Se-do, Young-jin Kim, and Ji-sun Hong. 2015. "Urban Traffic Flow Prediction System Using a Multifactor Pattern Recognition Model." *IEEE Transactions on Intelligent Transportation Systems*. https://doi.org/10.1109/TITS.2015.2419614.

Pastizzo, Matthew J., Robert F. Erbacher, and Laurie B. Feldman. 2002. "Multi-Dimensional Data Visualization." *Behavior Research Methods, Instruments, and Computers* 34 (2): 158–62. http://www.google.com/patents?hl=en&lr=&vid=USPATAPP11370606&id=PpijAAAAEBAJ&oi=fnd&dq=Multi-Dimensional+Data+Visualization&printsec=abstract.

Qi, Yan, and Sherif Ishak. 2014. "A Hidden Markov Model for Short Term Prediction of Traffic Conditions on Freeways." *Transportation Research Part C: Emerging Technologies* 43: 95–111. https://doi.org/10.1016/j.trc.2014.02.007.

Waze. 2017. "Routing Server - Waze." Wazeopedia. 2017. https://wazeopedia.waze.com/wiki/USA/Routing_server.

Wickham, Hadley. 2010. "A Layered Grammar of Graphics." *Journal of Computational and Graphical Statistics* 19 (1): 3–28. https://doi.org/10.1198/jcgs.2009.07098.

Wilkinson, Leland. 2005. *The Grammar of Graphics*. 2nd ed. New York, NY, USA: Springer Science & Business Media. https://doi.org/https://doi.org/10.1007/0-387-28695-0.

Williams, Billy M., Priya Durvasula, and Donald Brown. 1998. "Urban Freeway Traffic Flow Prediction: Application of Seasonal Autoregressive Integrated Moving Average and Exponential Smoothing Models." *Transportation Research Record: Journal of the Transportation Research Board*. https://doi.org/10.3141/1644-14.

Williams, Billy M., and Lester A. Hoel. 2003. "Modeling and Forecasting Vehicular Traffic Flow as a Seasonal ARIMA Process: Theoretical Basis and Empirical Results." *Journal of Transportation Engineering* 129 (6): 664–72. https://doi.org/10.1061/(ASCE)0733-947X(2003)129:6(664).

Yan, Weifeng. 2018. "Interview by Kyle Weaver." Los Angeles, CA.

Yang, Jun. 2010. "Highway Traffic Modeling." The Catholic University of America In.

Yu, Haiyang, Zhihai Wu, Shuqin Wang, Yunpeng Wang, and Xiaolei Ma. 2017. "Spatiotemporal Recurrent Convolutional Networks for Traffic Prediction in Transportation Networks." *Sensors (Switzerland)* 17 (7). https://doi.org/10.3390/s17071501.

Zhang, Yanru, Yunlong Zhang, and Ali Haghani. 2014. "A Hybrid Short-Term Traffic Flow

Forecasting Method Based on Spectral Analysis and Statistical Volatility Model." *Transportation Research Part C: Emerging Technologies* 43: 65–78. https://doi.org/10.1016/j.trc.2013.11.011.

Zhao, Zheng, Weihai Chen, Xingming Wu, Peter C. Y. Chen, and Jingmeng Liu. 2017. "LSTM Network: A Deep Learning Approach for Short-Term Traffic Forecast." *IET Intelligent Transport Systems* 11 (2): 68–75. https://doi.org/10.1049/iet-its.2016.0208.

Zhu, J., J. Cao, and Y. Fan. 2013. "Short-Term Traffic Flow Prediction Based on Flocking Theory and RBF Neural Network." *ICTIS 2013: Improving Multimodal Transportation Systems - Information, Safety, and Integration - Proceedings of the 2nd International Conference on Transportation Information and Safety*, 954–59.

Appendices

The following appendices show the analysis of the three incidents not covered in the main body of the thesis. The incidents have a distinct lack of association and show speed data which is counter to the expected results of an incident at a certain location. In these cases, it did not appear that the data was simply not supporting the expected hypothesis, but that the data presented as incident data was incorrect. As discussed in the analysis below, in normal incidents, there are clear correlations for the speeds of traffic on the freeway itself, and these incidents do not support that.

A1. Incident on July 17, 2017

This incident took place east of the La Brea exit on I-10W at 0651 and had a duration of 119 minutes as reported by California Highway Patrol. Figure 21 shows near average speeds along I-10W before and for nearly 30 minutes following the reported incident time. At that time, both 106N5075 and 106N5076 show a slight reduction in speed. During this time of the morning, the average speeds along I-10W are near the free-flow speeds of 65mph until about 0715 when overall speeds drop on an average day. 106N5076 shows a faster drop than normal to just above 50% of the average speed for that time of day. The arterial streets in the area do not show any significant timing correlation with the incident. This could be a result of most arterial sensors being positioned after the location of the incident. A sensor on Washington Blvd (106P12509, which contains a long string of orange circles in the center of the figure) does not provide any significant supporting information since it mirrors the behavior of the average speeds and nearby road segments do not show similar slowdowns.

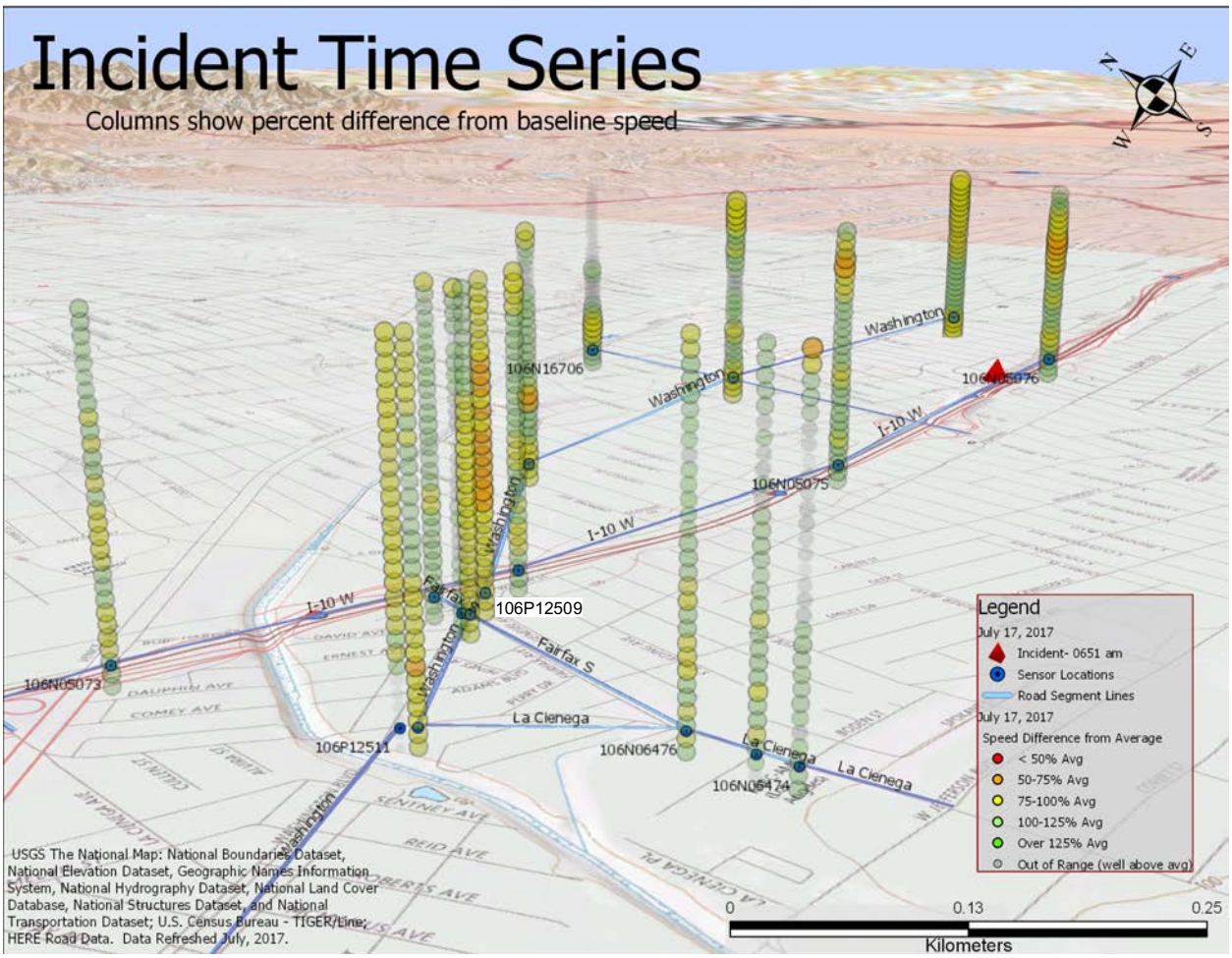

Figure 21: July 17, 2017 Incident Overview

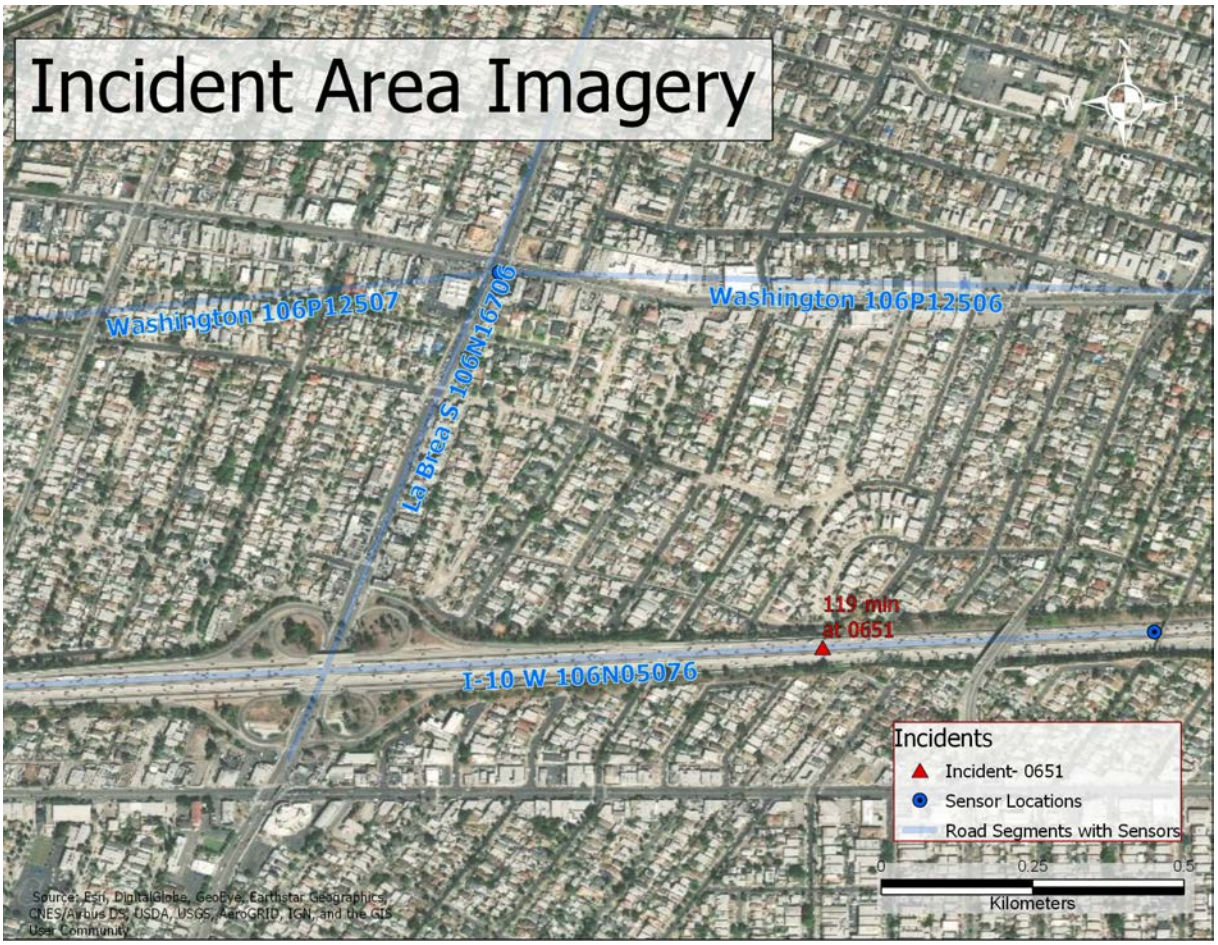

Figure 22: July 17, 2017 Incident Area Imagery

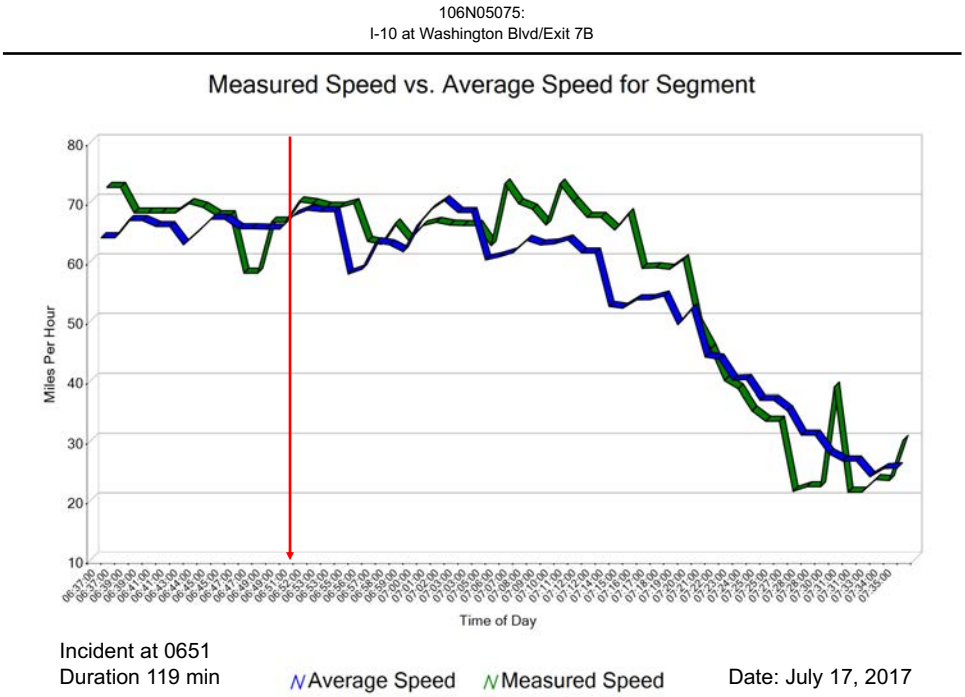

Figure 23: July 17, 2017 106N05075: Free-flow speed=63mph

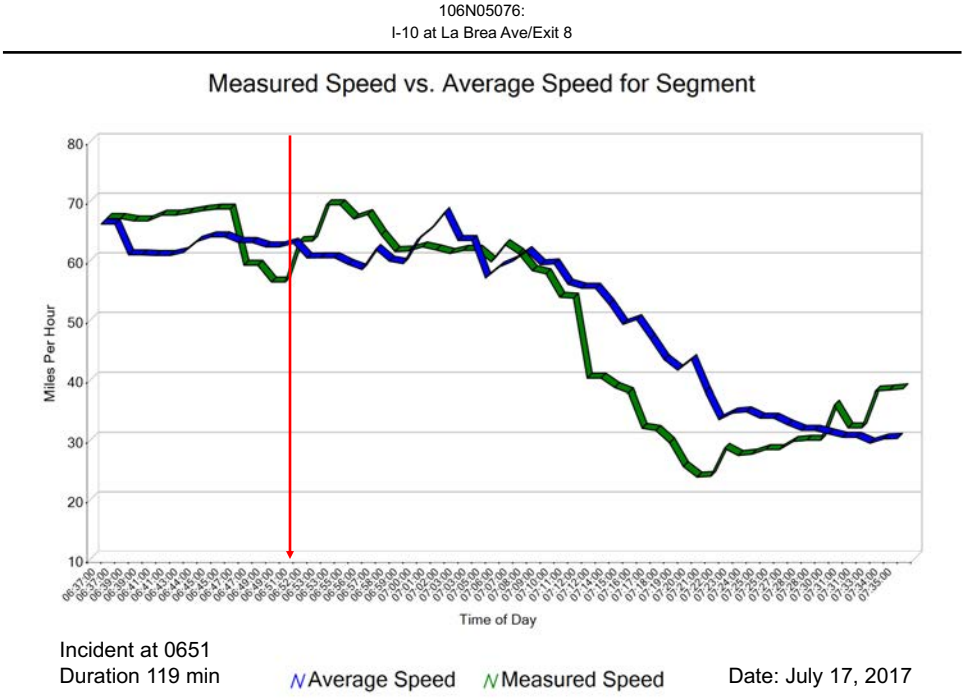

Figure 24: July 17, 2017 106N5076: Free-flow speed=63mph

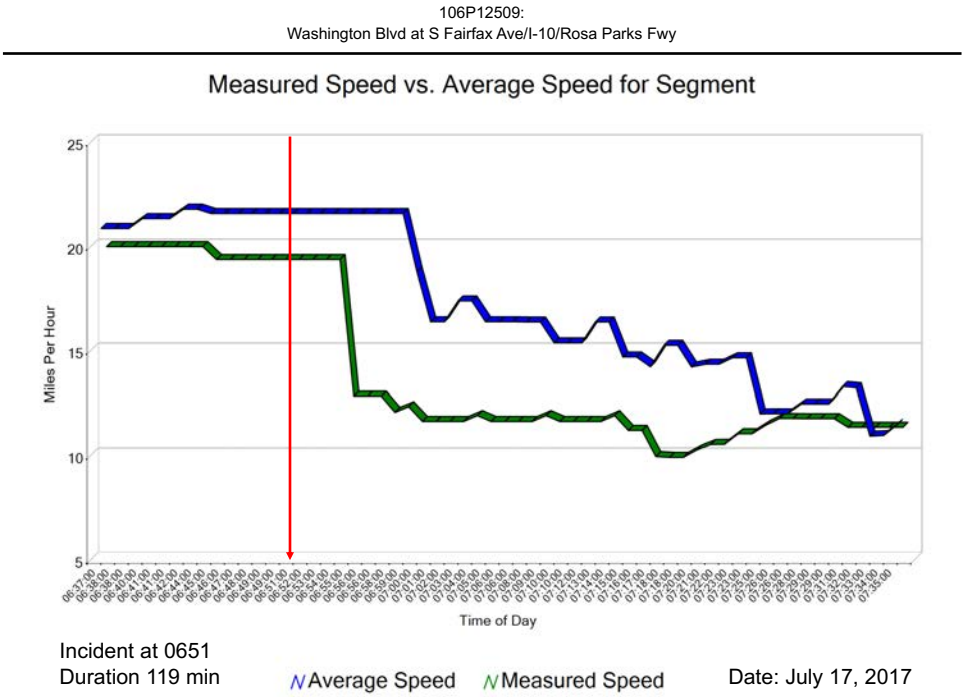

Figure 25: July 17, 2017 106P12509: Free-flow speed=22mph

A2. Incident on November 16, 2017

This incident is reported to take place near the overpass of Robertson Blvd. on I-10W at 0758 and has the longest reported duration of the study at 252 minutes per California Highway Patrol. Figure 26 shows the least potential association of any incident. The freeway sensors on either side of the reported incident site both show a slight dip below average around the incident time, but a high rate percentage increase in speeds above average following the incident. To support the hypothesis, it would be reasonable to assume that traffic on the freeway would slow following an incident especially one which has a long duration and an ambulance was called to the scene. This is not the case with the November 16 incident.

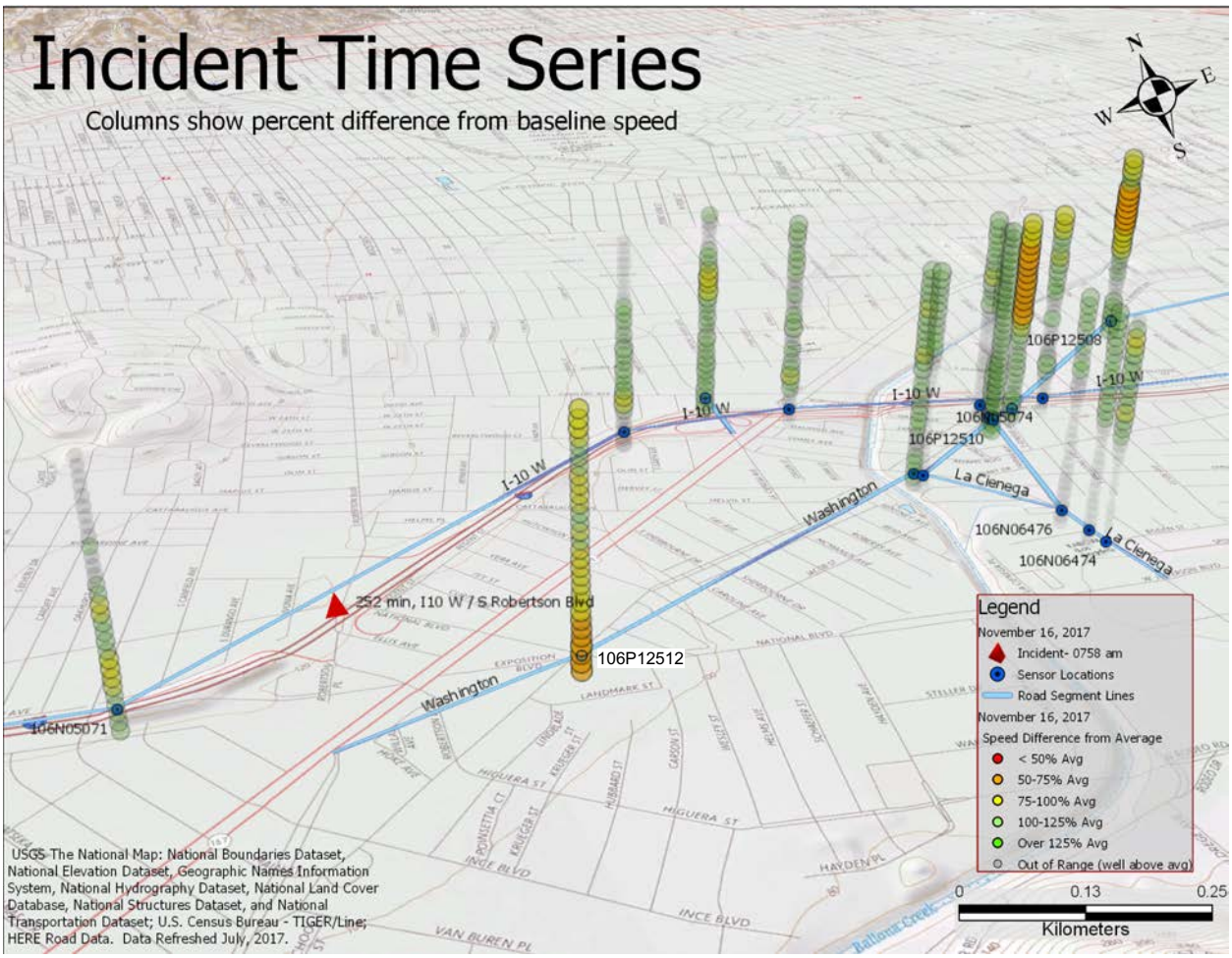

Figure 26: November 16, 2017 Incident Overview

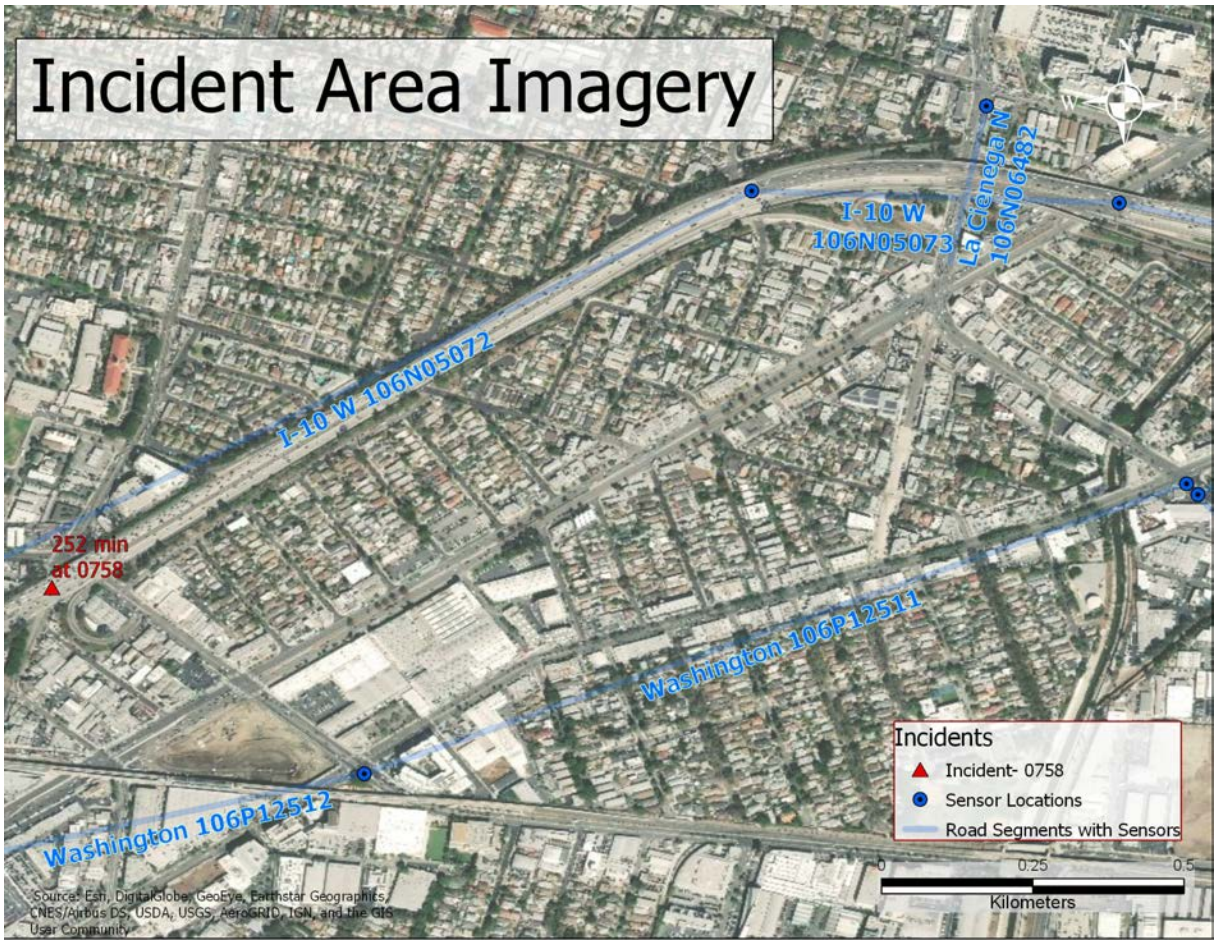

Figure 27: November 16, 2017 Incident Area Imagery

A3. Incident on May 31, 2018

This incident took place at the same reported location as the one on November 16, 2017 (A2. Incident on November 16, 2017) on I-10W at 0847 and had a duration of 100 minutes as reported by California Highway Patrol. In this case, the freeway traffic is reacting generally as expected: the speeds are converging from well above average to near (but still above) average (106N05071 and 106N05072 in Figure 28) following the incident time. Figure 30 and Figure 31 show this trend. Likewise, 106N06482 (a short segment of La Cienega Blvd which is fed directly from an exit from I-10W) begins a rapid drop from well-above average a few minutes before the incident. This does not appear to be caused by freeway traffic reacting to the incident since the

freeway speeds remain well above average during the drop. Despite this, none of the arterials accessible from the freeway eastward of the incident show any significant association with the drops in speed along the freeway.

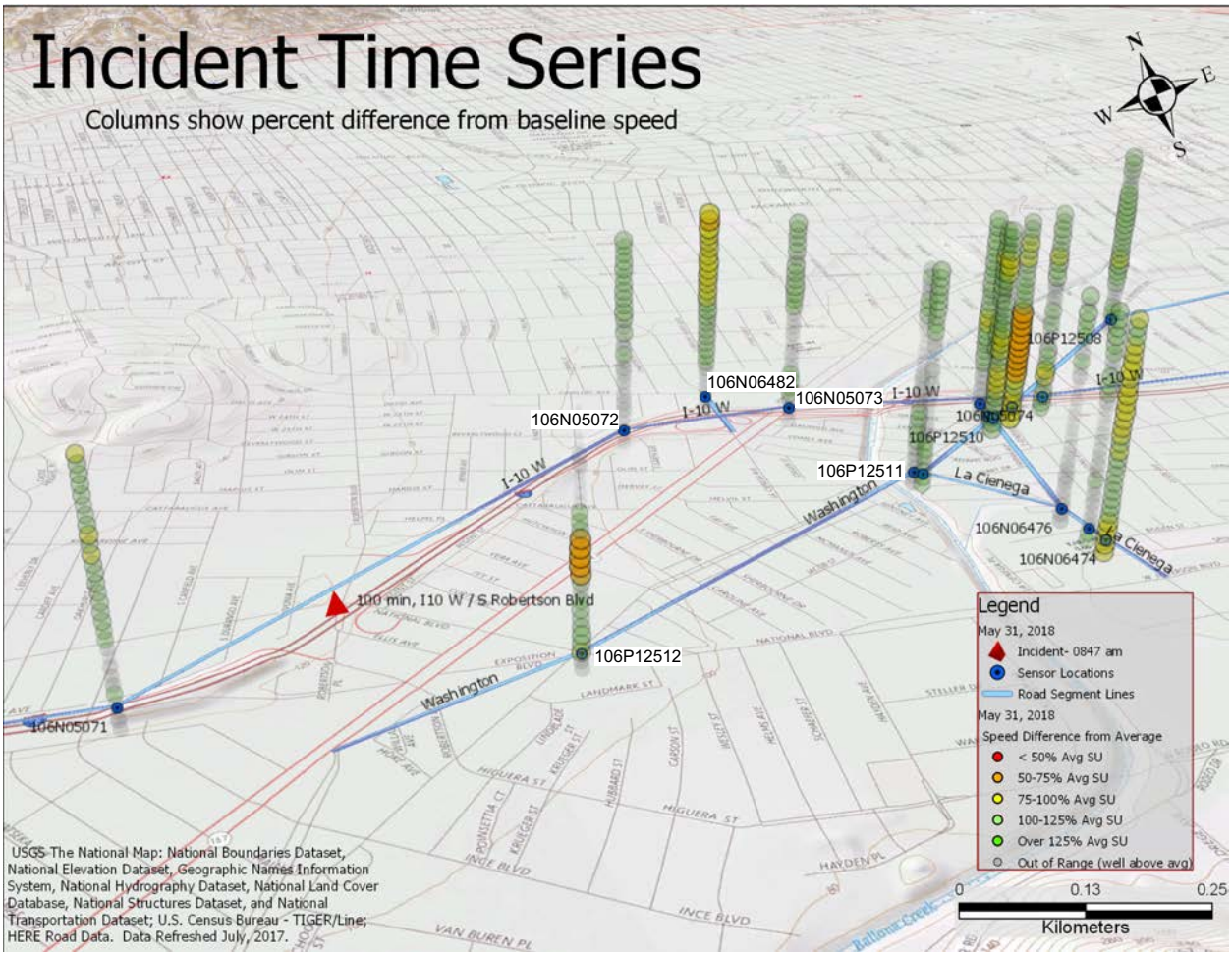

Figure 28: May 31, 2018 Incident Overview

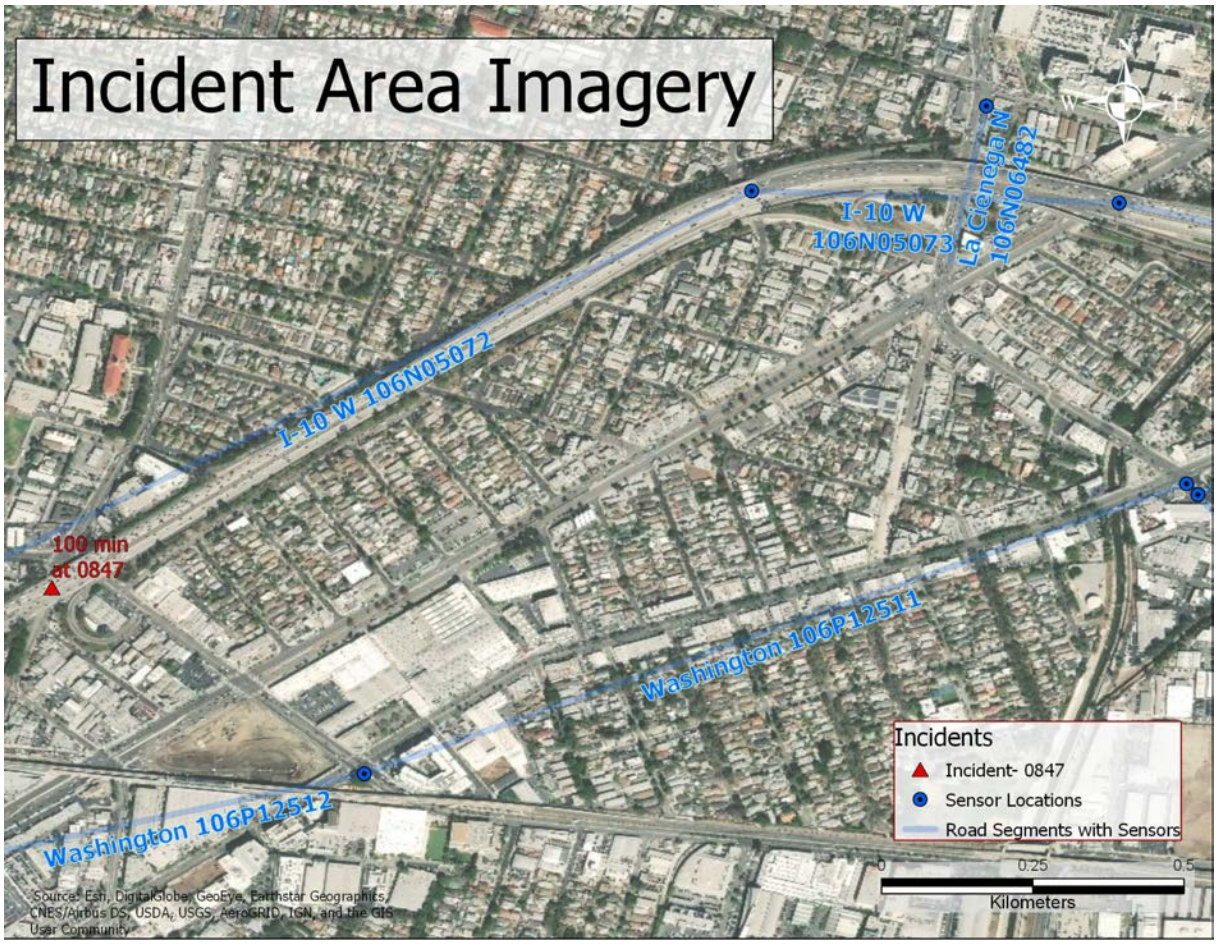

Figure 29: May 31, 2018 Incident Area Imagery

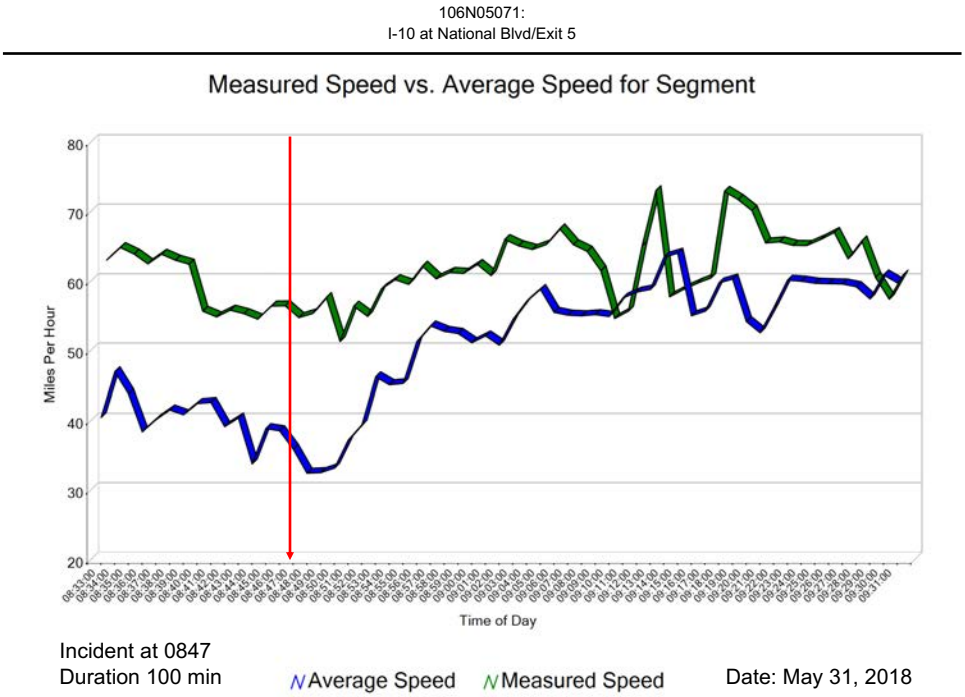

Figure 30: May 31, 2018 106N05071: Free-flow speed=65mph

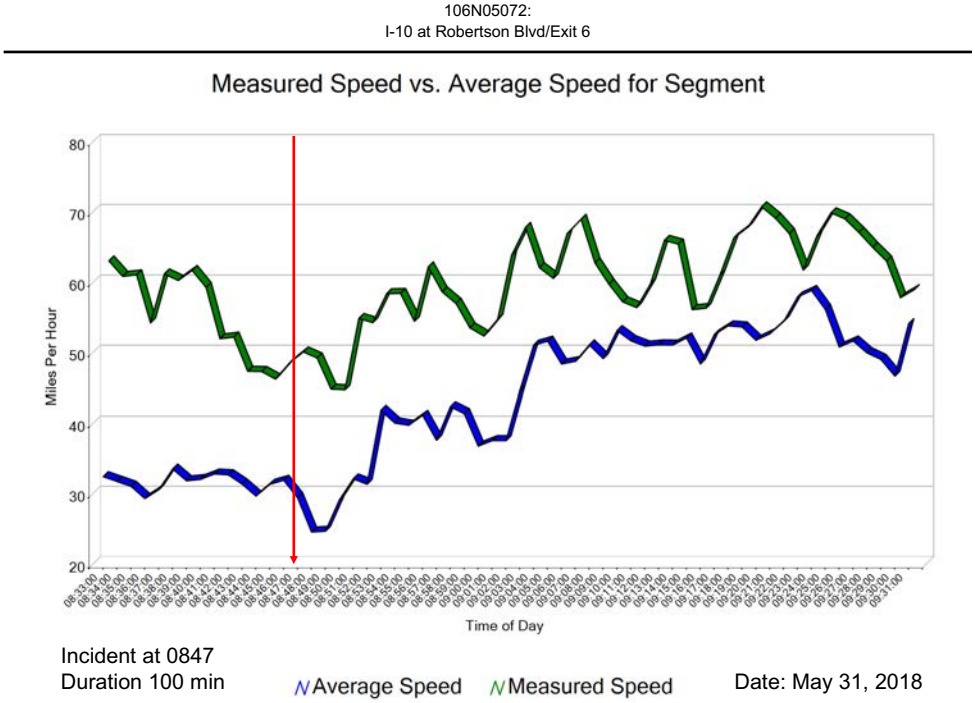

Figure 31: May 31, 2018 106N05072: Free-flow speed=65mph

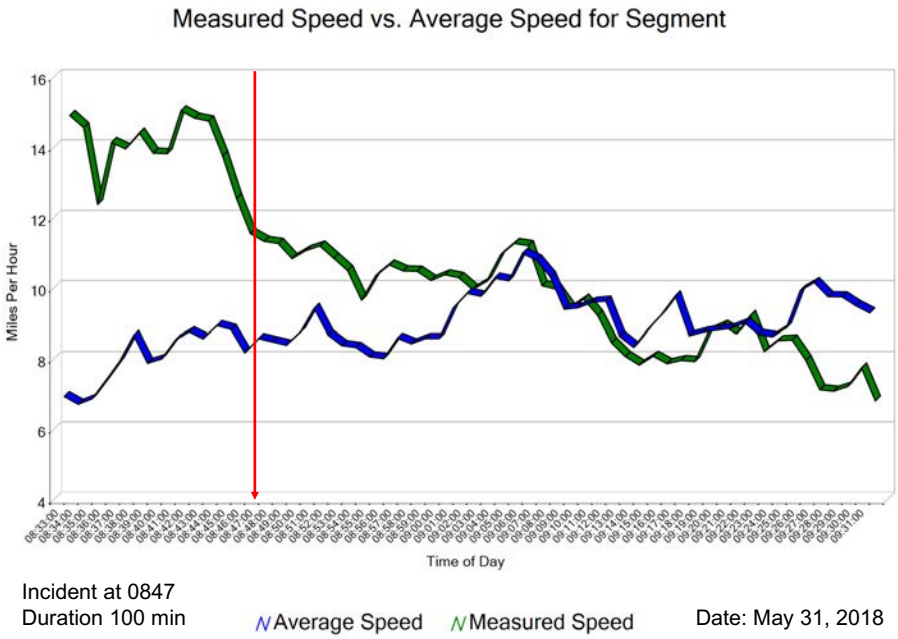

Figure 32: May 31, 2018 106N06482: Free-flow speed=26mph