2015

An investigation of freeway standstill distance, headway, and time gap data in heterogeneous traffic in Iowa

Andrew Jeremy Houchin
Iowa State University

An investigation of freeway standstill distance, headway, and time gap data in heterogeneous traffic in Iowa

by

Andrew Jeremy Houchin

A thesis submitted to the graduate faculty

in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Jing Dong, Major Professor
Anuj Sharma
Simon Laflamme

Iowa State University

Ames, Iowa

2015

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGMENTS

I would like to thank my major professor Dr. Jing Dong as well as the other members of my program of study committee Dr. Anuj Sharma and Dr. Simon Laflamme. They challenged me to do my best work while also being supportive and providing important insights. I also want to thank fellow graduate students Navid Shafieirad and Chaoru Lu for their assistance with the research project from which this thesis derived. I also thank the undergraduate student hired to work on the project, Mary Warhank, for her efforts processing a large amount of video for standstill distance measurements and radar detector validation. Finally, I want to thank all my family and friends for their great support and encouragement – whenever I doubted my ability to finish this thesis, you always kept me going. I offer my thanks to all of you, as I could not have done without you.

ABSTRACT

Microsimulation models have been growing in popularity in traffic engineering in recent years, and are often used as an important tool in the decision making process on large roadway design projects. In order to get valid results, it is necessary to calibrate such microsimulation models to local conditions. This is frequently achieved through a trial and error process of adjusting model parameters to get simulation results to match real world calibration data. Rarely is data collected on the model parameters themselves to provide a physical basis for the selection of their value. Two of the most important microsimulation model parameters for freeway models are standstill distance (the distance between stopped vehicles) and preferred time headway or time gap (the time between successive vehicles). Many simulation models treat these values as constants for all drivers and do not allow them to be set separately for different vehicle classes. This study presents a repeatable methodology for collecting standstill distance and headway/time gap values on freeways (mostly urban, with one rural location). It applies that methodology to locations throughout the state of Iowa. It continues by analyzing that data and comparing it for different locations and conditions. It finds that standstill distances vary by location and vehicle pair type. Headways/time gaps are found to be consistent within the same driver population and across different driver populations when the conditions are similar. An initial comparison between headways/time gaps at three urban areas to one rural location indicates a potential difference in driver behavior between those two conditions. Both standstill distance and headway/time gap are found to follow fairly disperse and skewed distributions. As a result of these findings, it is

recommended that microsimulation models are modified to include the option for standstill distance and headway/time gap to follow distributions as well as be set separately for different vehicle classes. Additionally, the standstill distances and headway/time gaps found in this study may be used as a starting point for future microsimulation calibration efforts on urban freeways in Iowa.

CHAPTER I: INTRODUCTION

Background

Traffic models have been used for decades by metropolitan planning organizations (MPOs) to predict traffic demands and create comprehensive plans for their cities using this traffic information (Wang 1996). The models allow them to prioritize which projects need to be completed first as well as experiment with different project alternatives to observe their effects on the network. Software utilizing these models are considered macroscopic simulation (or macrosimulation) software. Macrosimulation generally deals with large scale modelling in that it is designed to take on problems at a city-wide or region-wide level. This makes it ideal for MPOs and it is still used today by many engineers and agencies.

In the past decade or two, however, microscopic simulation software has gained increasing popularity as stakeholders require more and more detailed information about projects (Sbayti and Roden 2010). Microscopic simulation (or microsimulation) is defined by the FHWA as "the modeling of individual vehicle movements on a second or subsecond basis for the purpose of assessing the traffic performance of highway and street systems, transit, and pedestrians" (Dowling et. al. 2004). This outlines one of the main dividing lines many modelers point toward to differentiate between micro and macrosimulation: microsimulation models individual vehicles making separate decisions while macrosimulation models traffic as a continuous flow similar to a fluid.

Microsimulation software is often used to assist in the design process for large projects such as interchange justification reports (IJRs) (FHWA 2010). These IJRs are used in the recommendation of interchange designs for major freeways, whose costs are measured in the tens of millions of dollars. So naturally, it is important to have as much highly detailed data about the traffic operations of the different design alternatives as possible. Incorrect predictions for the traffic operations of the alternatives could lead to significant errors in decision making. These errors could be selecting an alternative which is not appropriate for the actual traffic operations or selecting an alternative which is much more expensive than another alternative which would be just as effective at handling the traffic.

In order to get the more detailed simulation results that microsimulation provides, it is necessary to provide a much larger number of parameters than with macrosimulation. Most microsimulation programs provide default values for these parameters, but these almost never provide accurate results. Therefore, it is necessary to calibrate the parameters for local conditions. There has been a lot of research on this topic in the past ten years or so, and all the approaches so far focus on adjusting parameters, either manually or automatically, to get simulation results which match real traffic data. This process is customarily started from scratch for every new simulation model. However, there are few methods which have been presented which actually collect data on any of these parameters directly in order to provide a physical basis for the model from local traffic conditions. If some of the more important modeling parameters can be found by directly from field measurement, and if those parameters to not vary substantially within similar driver populations and driving situations, then those values could be used as a starting point for such situations to reduce the level of effort needed for calibration.

Two of the most important parameters for freeway modelling in a microsimulation environment are the standstill distance and preferred time headway values. Both parameters have to do with the amount of space a following vehicle leaves between itself and the leading vehicle. Standstill distance is the distance between the back bumper of the leading vehicle and the front bumper of the following vehicle, and its average value can be shown to be closely related to the jam density of the road. Time headway is the amount of time that passes from the instant the front bumper of the leading vehicle passes one point on the road until the instant the front bumper of the following vehicle passes that point. Headway is the inverse of flow rate, so if every driver on the road is following at their preferred headway, the roadway is operating at its capacity. The accuracy of these parameters extremely important when it comes to actually designing roadways as well as accurately modeling the traffic operations. For example, if the projected traffic on a roadway does not quite warrant four lanes under the assumption that the default headway value is correct when the true headway value is actually substantially higher than the default, then four lanes becomes the necessary option. If such a road were built with 3 lanes, then it will likely have to be improved before it reaches its design life. While not every microsimulation program has two parameters labelled as “standstill distance” and “headway” specifically, almost all will have something similar or equivalent. For example, models will often be based on time gap rather than headway, but in that case the only difference is the leading vehicle’s length is excluded from the calculation (so it is the time from back bumper to front bumper).

The objective of this research is to develop a methodology for collecting standstill distances and headways on freeways and apply that methodology to compare average values and distributions in different parts of the state of Iowa. If these values are consistent across the state, they will be able to serve as a valuable starting point for future microsimulation calibration efforts in Iowa. In addition, the methodology could be repeated in other regions given the proper resources.

Overview of this Paper

This paper is organized into six chapters: Introduction (this chapter), Literature Review, Data Collection, Data Validation and Analysis Methodology, Results, and Discussion and Conclusion. Each of the middle four chapters have their own introduction and conclusion section as well as subsections to help break up the text and organize similar topics together. It is organized such that each chapter begins broadly, becomes more specific in the middle, and broadens out again at the end.

The Literature Review will investigate literature pertaining to microsimulation calibration practices, standstill distance and headway data collection efforts, and the distribution of standstill distances and headways. It will show a trend of microsimulation calibration not collecting traffic data directly on the parameters they are measuring. Some microsimulation software also does not allow for different preferred headway based on vehicle class. It will also show a lack of studies collecting standstill distance data, especially for freeways, as well as a lack of studies comparing headway values and distributions at different locations with different driver populations.

The Data Collection chapter will detail how the standstill distance and headway data were collected. The standstill distance data was processed from video of stop-and-go incidents on urban freeways in Iowa. The headway and time gap data was collected using a side-fired radar detector made by Wavetronix called SmartSensor HD on freeways in three urban areas and one rural location in Iowa. These urban areas were Des Moines, Council Bluffs, and the Quad Cities, which are widely separated geographically (see Figure 1).

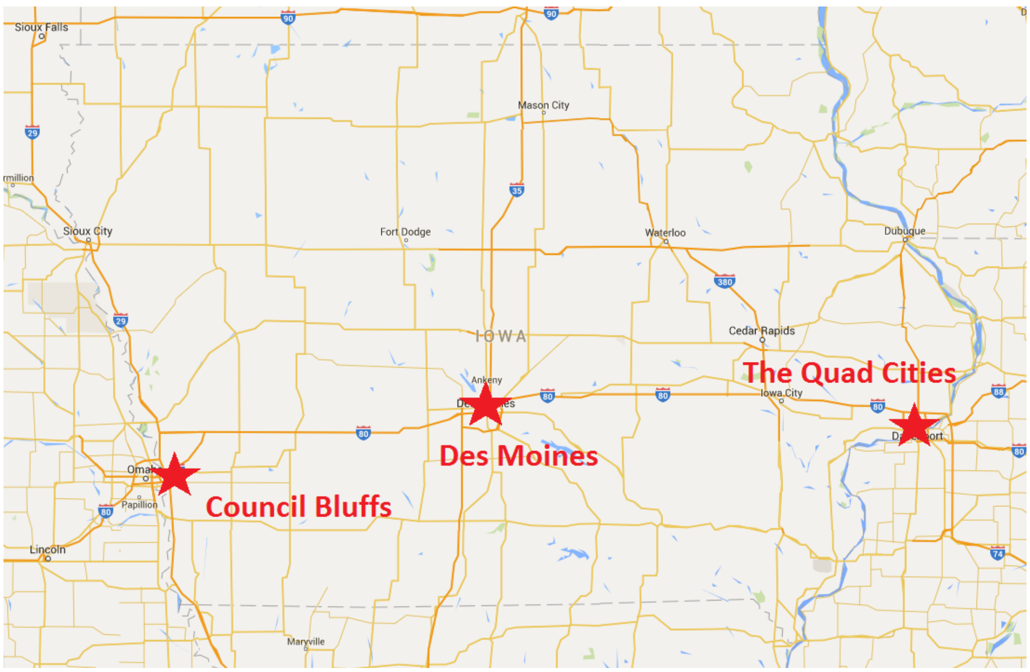

Figure 1. Urban headway data collection locations

In the Data Validation and Analysis Methodology chapter, the accuracy of the process measuring standstill distances and the accuracy of the Wavetronix detectors is established. Additionally, the methodology used to analyze both data sets is described, along with the filtering process for the headway and time gap data. The statistical software R was used to analyze both data sets. Standstill distance group means were compared using the p-values from t-tests. Due to extremely large sample sizes, mean headways were compared using a practical significance threshold of 0.1 seconds.

The Results chapter provides a detailed breakdown of the summary statistics for standstill distances broken down into different groups and for headways and time gaps for each detector used. For the headway and time gap data, the results are presented for each type of car following scenario when considering only cars and trucks as vehicle types (car-car, car-truck, truck-car, and truck-truck). Additionally, the distributions of standstill distances and headways/time gaps are examined.

Finally, in the Discussion and Conclusion chapter, the findings are summarized and the implications are discussed. For the first time, standstill distances on a freeway were collected, and it was found that they do vary at statistically significant level in different locations in Iowa, however this represents a difference of only four to five feet. They also vary for car-car following compared to when a truck is involved. Headways and time gaps were fairly consistent across the different urban areas, but were somewhat different in the rural location. It seems that headways tend to be influenced more by the following vehicle rather than the leading vehicle. It is also found that car-car combinations maintain somewhat similar time gaps as truck-truck combinations, but cars tend to follow trucks more closely and trucks tend to follow cars from further away than they do when they are following the same type of vehicle. Both standstill distances and headways/time gaps follow fairly disperse distributions. These results indicate that microsimulation models should allow for these parameters to be set separately for different vehicle classes and variance within each class should be included for both parameters. Finally, the consistency of the headway and time gap values within an urban freeway setting indicates that the values found in this research can be used as a starting point for model calibration efforts in such settings in Iowa, and the average standstill distances can be approximated to within four to five feet of their actual value.

CHAPTER II: LITERATURE REVIEW

Introduction

With the growing popularity of microsimulation models in transportation fields, it comes as no surprise that there has also been an increase in research efforts with respect to them. Additionally, two of the most important and most often calibrated parameters in microsimulation are the average distance left between stopped vehicles (standstill distance) and the average preferred time between a leading and following vehicle (headway/time gap). In this section, since it is the motivation of this thesis, the body of literature pertaining to microsimulation calibration will be reviewed. Additionally, literature related to the collection of standstill distance and headway data as well as their distributions is also reviewed.

Microsimulation Calibration

The first segment of literature examined pertained to methods of calibrating microsimulation models. There appears to be two predominant methods of calibrating microsimulation models. Both methods involve selecting one or more measures of effectiveness on which data is collected from the existing traffic conditions. These data serve as the baseline to which the model builder attempts to match microsimulation results. Matching the measures of effectiveness is achieved through adjustments to the model parameters, and this is where the two main calibration methods differ. In the first method, the parameters are adjusted manually in a trial and error process, while in the second method, the parameters are changed automatically through the use of an algorithm implemented in computer program. In both methods, once calibrated, the model is applied to a new time period, and compared to the existing traffic during that time to assess the model's predictive abilities – this is referred to as validation.

Procedure

Two studies that were among the first to propose a methodology for calibrating microsimulation models were published in 2003. A study titled *A Practical Procedure for Calibration Microscopic Traffic Simulation Models* (Hourdakis et al. 2003) proposed a general methodology with three calibration stages, with the final stage being optional. The first stage is volume-based calibration, the second stage is speed-based calibration, and the final (optional) stage is objective-based in which the model can be fine-tuned to project specific objectives. They then applied their methodology to a case study in Minnesota and found it to be quite effective in improving the model's performance with respect to the actual traffic patterns (Hourdakis et al. 2003). A similar study in 2003 laid out a more step-by-step calibration procedure. It involved determining the measures of effectiveness to be used, collecting the data, identifying the calibration parameters, implementing an experimental design (to reduce the number of parameter combinations), running the simulation multiple times for each parameter set, developing a function relating the measures of effectiveness to parameters, determining parameter sets, evaluating the parameter sets, and validating the model with new data. The authors also implemented their methodology with a case study, and again noted the benefits of calibrated results compared to uncalibrated results (Park and Schneeberger 2003). While these two methodologies may not appear extremely similar on the surface, they are actually structured quite similarly in essence: they match measures of effectiveness in the simulation results to the data in the field by altering simulation parameters.

In 2004, the Federal Highway Administration (FHWA) released its *Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software* (Dowling et al. 2004) which covered all aspects of microsimulation modelling, including a chapter on calibration. Their calibration procedure to a large extent mirrors that of the two studies mentioned above. There are some differences in the methods, though: for example, rather than calibrating based on demand as in Hourdakis et al. they recommend calibrating based on capacity. However, the main structure of their method was the same – alter the simulation parameters to match simulation results in the different measures of effectiveness to the observed traffic data. They also provided different calibration target values for a variety of measures of effectiveness (Dowling et al. 2004). The Oregon Department of Transportation later created their *Protocol for VISSIM Simulation*, which applied the FHWA's guidance to a specific modeling software and further refined the calibration process (Oregon DOT 2011).

Manual calibration

While manual calibration is not usually the recommended procedure due to the vast number of combinations of parameters in microsimulation software, there are some advantages to this method, and it is still frequently used, particularly in private consulting. A few of its advantages are that it is low on computational demand, relatively simple to implement, and compatible with qualitative measures of effectiveness such as bottleneck length, time, and location and general driver behavior because of the analysts ability to view the model animation and compare it with his or her experience. Some of its disadvantages are that the solution reached will likely be less optimal than one reached by an automated process,

One study that used manual calibration was *Congested Freeway Microsimulation Model Using VISSIM* (Gomes et al. 2004). In this study, the authors modeled a 15 mile stretch of I-210 West in Pasadena, California, which is a congested and complex segment. There were high occupancy vehicle (HOV) lanes, metered on-ramps, and three interacting bottlenecks. Due to the unique situation, the authors did not use typical measures of effectiveness such as volume, travel time, or delay. Instead, they attempted to match qualitative aspects of the freeway including the location of bottlenecks, start and end times of queues, and length of queues. Manual calibration was used in large part due to a lack of computing power (Gomes et al. 2004). Additionally, this study took place in 2004, early in the body of literature examined in this study, when automated methods may not have been as well developed or well researched.

Automated calibration

While some research studies use manual calibration for their microsimulation models, the vast majority use some form of automated calibration. This is likely because it can come the closest to providing an optimal solution, which is what researchers are often interested in finding. It has also become more and more feasible for researchers to use computationally intensive automated methods as computing power has increased.

One study that used this approach was *Microsimulation Calibration Using Speed-Flow Relationships*. In this study, the authors selected five VISSIM driver behavior parameters to use for the calibration and ran an evolutionary algorithm to select the optimized parameter set. The evolutionary algorithm starts with several initial parameter sets, selects the ones that perform the best, combines them, and repeats this until it converges to one set. The objective function that determined which parameter sets performed the best was based on pattern recognition of speed-flow graphs (Menneni et al. 2008).

Another study that used an automatic method of adjusting parameters is *Methodology for the Calibration of VISSIM in Mixed Traffic* (Manjunatha et al. 2013). Though it focuses on signalized intersections, the study calibrated driver behavior VISSIM parameters which are typically used for freeway sections. The authors calibrated all nine main driver behavior parameters, using a method similar to the evolutionary algorithm in the study by Menneni et al. (2008). The measure of effectiveness used to evaluate the parameter sets in this case was delay.

Though the vast majority of research into microsimulation calibration will select a few parameters which it will adjust for the calibration, one recent study adjusted all the parameters of a microsimulation model at one. In *Calibration of Micro-simulation Traffic-Flow Models Considering All Parameters Simultaneously*, Paz et al. (2014) used a simultaneous perturbation stochastic approximation algorithm to calibrate all the parameters in CORSIM at the same time based on several measures of effectiveness.

Rahman et al. (2014) delved deeper into calibration looking specifically at calibrating the car-following models themselves in *A Parameter Estimation and Calibration Method for Car-Following Models*. The authors used a large number of vehicle trajectories to improve the accuracy of car-following models by making them more closely replicate driver behavior. Monteil et al. (2014) also calibrated car-following models in *Calibration, Estimation and Sampling Issues of 2 Car-following Parameters*.

Clearly, there is a common thread in calibrating microscopic simulations: adjust simulation parameters until the simulation results match the data collected on the real roadway as closely as possible. This works well for the site which is calibrated, but may not translate well to other projects, study sites, or potentially even future traffic patterns at the same site if major characteristics of it change. This incongruity is possible because, as numerous studies have pointed out, there are multiple sets of parameters that may provide similar results with respect to a few measures of effectiveness. Because these parameter sets are not based on the actual behavior of the vehicles (that is, the parameters were adjusted essentially at random), it is possible that a selected parameter set would not produce similarly accurate results when applied to other sites. This paper attempts to take a different approach by collecting data on two of the most important parameters themselves, with the hope that such data could be used as the basis for a more stable and transferable parameter set.

Parameter Collection and Distributions

The capacity and queuing behavior of roadway sections, particularly freeway sections, are significantly influenced by the standstill distance and headway distribution of the population traversing the section. Because these parameters control the amount of roadway space available in a given lane (ignoring lane changing), they have a substantial impact on the facility's operations. This is true both in reality and in microsimulation models. This section will examine literature validating the importance of standstill distance and headway parameters in microsimulation. It will also examine past data collection efforts and literature regarding the forms of their distributions.

Importance of standstill distance and headway/time gap

Standstill distance and headway are important both in theory and in practice when it comes to microsimulation models. The standstill distance controls the maximum density (jam density) of vehicles on a roadway section, because if all the vehicles are at their standstill distance, they will be as close together as possible. Likewise, headway controls the capacity of the section. Once all vehicles are following at their headway, then any additional density will cause vehicles to begin braking and thus cause congestion. In fact, headway is the inverse of traffic flow, so if the average headway is known, the flow rate can be found, and vice versa. Time gap is closely related to headway, except it is defined as the time that elapses between the *back* bumper of the leading vehicle to the front bumper of the following vehicle, whereas headway is front bumper to front bumper. A number of studies have investigated the importance of standstill distance and headway/time gap to road operations. In addition to their theoretical importance, the importance of these parameters is further demonstrated in practice because most calibration studies include them in their parameter selection, and sensitivity analyses show them to have large impacts on microsimulation results.

Many textbooks and classic research studies have established that vehicle spacing is the inverse of density, and time headway is the inverse of volume (Elefteriadou 2014). This means that the smallest vehicle spacing will lead to the largest jam density, and the smallest headway will lead to the capacity of the facility. While standstill distance is not the exact same thing as spacing (because standstill distance ignores vehicle length), it is closely related to spacing and can be used to approximate the jam density of a facility. Likewise, the average headway value can be used to approximate the facility's capacity. The jam density and capacity are two of the most important macroscopic characteristics of a roadway from a traffic operations perspective, so clearly, their corresponding microscopic characteristics will have a significant impact on the facility's operations as well (Elefteriadou 2014).

In discussing the difference between macrosimulation software and microsimulation software, the Highway Capacity Manual 2010 (HCM 2010) uses the fact that headway and flow are inverses of each other to help compare HCM results to microsimulation results:

> Microscopic simulation tools, however, do not have an explicit capacity input. Most microscopic tools provide an input that affects the minimum separation for the generation of vehicles into the system. Therefore, specifying a value of 1.5 s for this input will result in a maximum vehicle entry rate of 2400 (3600/1.5) veh/h/ln.

This reaffirms the theoretical importance of the following headway value and directly establishes the relationship between the selected headway value and its impact on the maximum capacity in microsimulation.

One recent study (Wu and Liu 2013), investigated how the uncertainty of time gap selection affects traffic flow and the fundamental diagram, which displays macroscopic operation characteristics (speed, flow, and density). While this study focused on an arterial with signalized intersections, some of the same concepts apply to the urban freeways discussed in this research. The authors focused on congested flow conditions and found that drivers typically do not display as much variation in their time gap selection at constant speeds as they do when they are accelerating or decelerating. They also found that the variation in time gaps contributes to the scatter of the fundamental diagram, and that when traffic is accelerating or decelerating the shape of the diagram changes (Wu and Liu 2013).

In addition to the numerous studies pointing out the theoretical importance of standstill distance and headway to traffic behavior on uninterrupted flow facilities, the majority of microsimulation calibration efforts include these two parameters if they choose to calibrate a subset of all the changeable parameters. Sensitivity analyses have also shown these two variables to be among those having the largest effect on a number of measures of effectiveness in microsimulation models.

In a case study using the microsimulation software VISSIM included with Parker and Schneeberger's (2003) proposed calibration methodology, standstill distance and headway were two of six parameters the authors chose to calibrate (Parker and Schneeberger 2003). While they do not explain their rationale behind the selection of calibration parameters, it stands to reason that they selected those parameters that would have the largest impact on the model in their experience, and, according to the company that makes VISSIM, PTV Group, the headway parameter has the largest impact on capacity (PTV Group 2011). In 2004, in the FHWA's *Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software*, four examples of capacity-related parameters for freeways included "mean following headway" and "minimum separation under stop-and-go conditions" (Dowling et al. 2004). The FHWA also released guidelines for one specific microsimulation program, CORSIM, in which two car following parameters and a factor for the minimum distance between vehicles were included in the “candidate list of key parameters for calibrating freeway capacity” which comprised four parameters total (Holm et al. 2007).

In yet another microsimulation calibration project, standstill distance and headway were two of three VISSIM driver behavior parameters that were adjusted to calibrate a 15 mile long, complex stretch of highway in California (Gomes et al. 2004). The fact that the authors were able to successfully calibrate such a large model with so few driver behavior parameters illustrates the important role standstill distance and headway play. Another study in California, which calibrated the model using speed flow charts as measures of effectiveness, included the headway parameter among the five they calibrated, but not the standstill distance (Mennini 2008).

Some studies have undertaken the task of conducting sensitivity analyses on the various microsimulation programs, and the results tend to agree with those found in the case studies that standstill distance and headway are two of the most important parameters for capacity, particularly headway. One study in India found that standstill distance and headway were among five VISSIM headway parameters that had a significant effect on capacity (Manjunatha et. al 2013). A different sensitivity analysis only indicated headway as one of three VISSIM parameters with the greatest influence capacity (Woody 2006). A third study found that both standstill distance and headway could have a statistically significant impact on capacity when they are far enough away from their calibrated value (Lowens and Machemehl 2006). Despite the clear importance of standstill distance and headway for microsimulation models, and despite the obvious differences in behavior between cars and trucks, some models do not include an option for different preferred standstill distances and headways for different vehicle classes (e.g. VISSIM).

Separate efforts to model standstill distance and headway

Aside from the studies on calibrating microsimulation software, there have been a number of studies investigating the distributions of following and free headways. However, there have not been nearly as many attempts to collect data on standstill distances, particularly on freeways, and observe their distribution. As with any driver behavior parameter, not every driver will behave the same; some will be more conservative, while others are aggressive, etc. This variance is often not accounted for in microsimulation models, despite the abundance of research that indicates that standstill distance and headway are not constant parameters.

There have not been many efforts to collect standstill distances at all, let alone on freeways. Most of the efforts have been focused on signalized intersections where standstill distances are important for queue lengths and much easier to collect. Because traffic on each approach is guaranteed to stop every time there is a red light, one can simply create a scale on the pavement or next to the traffic that can be used to estimate the distance between vehicles.

One such study focused on calibrating a variety of VISSIM parameters to local conditions in Delaware (Delaware Valley Regional Planning Commission 2013). This was one of the only studies found that focused on calibrating microsimulation software by collecting data on the parameters themselves. Standstill distance was one of the parameters calibrated for urban and suburban settings at signalized intersections. The authors collected the data by marking 5 foot increments on the approach of a number of different intersections and approximated standstill distance to the nearest foot. They compared urban and suburban settings and compared through/right-turn lanes to left-turn lanes. The average standstill distance they found was about 9, feet with little variation across the different conditions. This value is greater than the default VISSIM parameter. They also noted a wide variation in the measurements even within the same queue (even after they excluded “drivers who were not paying attention or left an unreasonable large gap”). Finally, they found the standstill distances when a truck was involved to be “comparable” to those of car-car pairs (Delaware Valley Regional Planning Commission 2013).

Another study collected data on the spacing of queued vehicles at a traffic signal and compared average of these values to the default value in the microsimulation software CORSIM and to commonly used assumptions for queue length calculations used in roadway and signal design (Long 2002). In this report, the author lamented a lack of recent data on the spacing of vehicles queued at signals. The author collected data from four locations in Florida and two locations in Chicago, Illinois that spanned many different traffic conditions and driver populations. He found an average spacing of 12 feet with no significant differences between sites, which is significantly higher than the CORSIM default and commonly assumed values in roadway and signal design. The distribution of spacing was not directly discussed (Long, 2002).

One study, interestingly found in a physics journal, measured the standstill distances of vehicles at a traffic signal in Prague. The distances between vehicles were measured with laser technology. The study focused on modeling the distribution of the standstill distances as well as the inter-vehicle distances once the light turned green. It was found that the stopped traffic and its progression through the signal acted as a "thermodynamical [sic] gas of dimensionless particles exposed to a thermal bath" (Krbálek, 2008).

While a few studies have investigated standstill distances at signalized intersections, these are quite scarce, and no studies were found that did the same for a freeway. It would not make sense to extrapolate the data from signalized intersections to freeways, because the two facility types require entirely different driving behaviors. At most signalized intersections, the lane changing immediately upstream from the signal is minimal, while it is present on freeways. Another difference is that at signalized intersections there is defined period in which all vehicles at a signal must stop and then a period when all vehicles can go, whereas in stop-and-go conditions on a freeway, one lane may advance slowly while another is stopped or vehicles ahead may stop or go without warning. These differences in conditions make a separate study specifically on freeway standstill distances necessary.

As far as headway distributions are concerned, many more studies have been able to collect freeway headway data and observe the distribution of such data. One study created a car following model based on NGSIM data that used headway as the main determinant at high speeds and spacing at low speeds (Chen et al. 2014). This study found that the headway data are distributed approximately log-normally, or in the more general case, follow the gamma distribution. However, the authors did not describe to which distributions the log-normal distribution was compared. They found the mean headway value to be around 2 seconds (Chen et al. 2014).

The same study which looked at the impacts of the uncertainty of time gap selection also looked at how those time gaps were distributed (Wu and Liu 2013). The authors compared the log-normal, gamma, and Weibull curves using log likelihood values. The study found that the log-normal and gamma distributions were approximately the same in terms of goodness of fit and that both were better than the Weibull distribution. All three time gap distributions had the same means, 1.22 sec for all vehicles (Wu and Liu 2013).

Another study looked at the distribution of headways by examining loop detector data on an urban freeway in Seattle, Washington (Zhang et al. 2007). Of particular interest to this study was comparing the headway distribution of normal lanes to that of high occupancy vehicle (HOV) lanes. It compared a number of different single and mixed distributions based on well they fit the headway data. Single models are standard statistical distributions that treat all the headways the same and include normal, gamma, and lognormal distributions. Mixed models have separated components to distinguish between following vehicles and free vehicles and they include Cowan M3, Cowan M4, generalized queuing, and double-displaced negative exponential (DDNED) distributions. The study found that the lognormal distribution was “adequate in fitting headways on general purpose lanes under most circumstances” but the DDNED model better described both the general purpose lanes and especially the HOV lanes. In addition, the results showed average headways of about 2 seconds during the busiest portion of the day (Zhang et al. 2007).

A similar study was conducted on individual vehicle data from a freeway in France (Ha et al. 2012). The authors compared statistical distributions which had previously been used to model headway data to some new distributions which they proposed. The study found that a gamma-generalized queuing mixed model provided the best fit for the headway data, though the log-normal distribution performed best of the single distributions that were tested (Ha et al. 2012).

Conclusion

This section reviewed the research literature related to microsimulation calibration and the distribution of standstill distances and headway values. Studies related to calibration procedures were reviewed to provide a background on the current practices and illustrate the importance of standstill distance and headway for these models. Further support for the importance of these parameters was presented through calibration case studies that had selected standstill distance and headway to be among their calibration parameters, as well as several sensitivity analyses that tended to show that standstill distance and headway have significant impacts on calibration, particularly when calibrating by capacity. Finally, studies that collected data on and investigated the shapes of the distributions of standstill distances and headways were reviewed.

Overall, there is clearly a common thread in the approaches to calibrating microscopic simulations: adjust the simulation parameters until the simulation results match the data collected on the real roadway. This method works and has been shown to be effective many times over, but its results may not translate well to other projects, study sites, or potentially, if major changes occur, even the projection of future traffic patterns on the same roadway, because the parameter values are not selected with a physical basis. This paper attempts to take a different approach by collecting data on two of the most important themselves in the hope that such data or the data collection process could be used to streamline the calibration process by acting as a solid starting point with a basis in empirical data that would require only small tweaks.

Additionally, it was discovered that efforts to collect standstill distances have been scarce and have focused solely on signalized intersections; it is believed that the present study is the first to collect standstill distances in a freeway setting. Also, despite obviously different driving behaviors between car and trucks, some simulation programs do not provide the option for different vehicle classes to have different standstill distance or time headway preferences. Finally, while there were an abundance of papers modeling the distributions of individual vehicle headways, there were not many which compared these for different locations. One of the goals of this paper is to demonstrate that one or both of the parameters studied is consistent across different driver populations in the same region, which would be a useful stepping stone for future simulation calibration efforts.

CHAPTER III: DATA COLLECTION

Introduction

This research required the collection and analysis of time stamped individual vehicle data. By using such data, it is possible to measure the headways and time gaps of individual vehicles and observe their distributions. Additionally, this research required the collection of data pertaining to the distance between stopped vehicles, which was not found by the literature review to have been collected in any past freeway studies. This section will detail the methodology for the data collection process.

Data Collection

There were two main data collection efforts in this research. The first major data collection effort for this research was acquiring standstill distances on freeways. The literature review revealed no studies that directly collected standstill distance data on freeways. One challenge related to collecting these data on urban freeways is the lack of reoccurring stop-and-go traffic, particularly in Iowa. Unlike urban freeways in some other cities, there are no known locations where stop-and-go traffic can be observed on a regular basis. If such conditions were present, a scale could be set up next to the road at these locations, and the traffic could be recorded with video and processed relatively easily. This has been the strategy used by past studies at signalized intersections. Additionally, the methodology developed for this research proved to be fairly time consuming and required special access to the Iowa Department of Transportation’s (IDOT’s) network of cameras and dynamic message signs (DMS). It involved using a program to view dynamic message signs' message histories, downloading recorded video from IDOT cameras, and manually measuring the distance between stopped vehicles using Photoshop CC 2014. Without such access to the IDOT network, collecting standstill distances using this process would have been impossible.

The second data collection effort was collecting individual headway/time gap data. In order to collect individual headway data, a number of options were investigated, including manual collection, loop detectors, laser-based collection, video/image processing, and radar-based collection. These options were evaluated with a number of goals in mind for the data collection, including the desire to have time stamped individual vehicle data, especially speed, vehicle class, and lane assignments. Manual collection was deemed too resource intensive, loop detectors could not be moved to different locations, and no laser-based or video processing options were found to meet the goals of the data collection as well as the selected option. In the end, it was determined that Wavetronix's SmartSensor HD side-fired radar detectors best accomplished all of these goals.

Standstill distances

Unfortunately, the process for collecting standstill distance measurements was not nearly as automated as that for collecting the headway data. There was some discussion of trying to find video processing software that could automate some of the measuring process or of crowd sourcing some of the steps of the process. However, ultimately, it was decided that those processes were either not feasible or the return would not be worth the time and resource investment. However, an undergraduate student at Iowa State

University, Mary Warhank, was hired to do the majority of the video collection and measuring once the process was established.

The first step in the process was to identify locations where stop-and-go traffic would be likely to have occurred. In order to do this, a report was created using TranSuite TIS software, which listed every message posted on any DMS in Iowa during the time period specified at the creation of the report. Then, the message of any sign whose message was changed by an event manager was examined. The locations of signs that displayed messages indicating an accident, slow traffic, roadwork, or anything else that may cause congestion were recorded for later use. Almost all of these DMSs are located in urban areas, so, unfortunately, this process made it extremely difficult to collect standstill distances for rural locations. However, the fact that standstill distances were directly collected at all was a contribution to this research area.

The next step was to use the IDOT network video recorder (NVR) software to access recorded video from IDOT cameras to visually review each of the potential stop-and-go incidents. If the incident in question did cause stop-and-go traffic, video from the time period in which traffic was affected was downloaded. After all the relevant videos for that report were downloaded, each video was watched, and each time there were stopped vehicles in the frame a screenshot was taken and the vehicles that were moving were marked (so the moving vehicles would not be measured by mistake).

Finally, the vanishing point filter option in Photoshop CC 2014 was used to find the actual bumper-to-bumper standstill distance measurements. The vanishing point filter allows the user to create a flat plane on which measurements are to be made. If there is an object or a mark of known length in the established plane, that reference object can be used to create a baseline measurement that Photoshop uses to measure anything else in that plane. In the case of this research, the painted lane lines were used as the baseline measurement. The standard for painted lane lines on freeways is that they be 10 feet in length, and they are painted using an automated system. Google Earth was used to measure these lines and confirm they meet the 10 foot standard. An example of one fully processed image is shown in Figure 2. In addition to the standstill distance measurement, the following combination for each pair was also recorded. The conditions surrounding the incident were also noted.

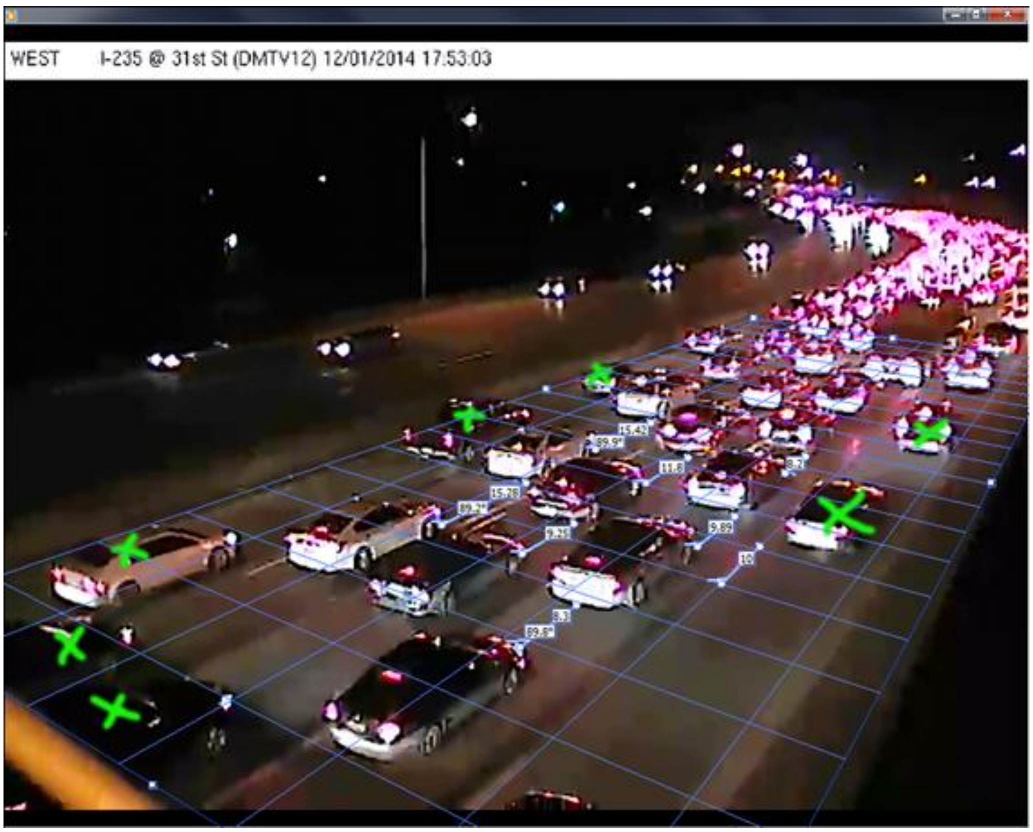

Figure 2. Example of an image processed for standstill distance measurements

Headway data collection

Several different data collection methods were explored with respect to collecting headway data. The criteria used to evaluate the usefulness of each of the methods investigated were as follows: (1) the ability to collect individual vehicle data; (2) the inclusion of lane, class, time of arrival, and speed in the data; (3) the accessibility and cost of the data or equipment; and (4) the reliability of the method. The types of data collection evaluated were manual collection, existing freeway loop detectors, laser-based detection, video processing, and radar-based detection.

The manual collection option was deemed to be too resource intensive because several other less resource intensive options were available. The loop detectors provided the necessary data, but they were sparsely located and it was not possible to easily set up at locations where data collection would be desired. No any laser-based detector options were found to meet the data requirements, so the majority of the investigative effort went into comparing the video and radar based methods. With the video based products, it was often difficult to determine the level of detail that was actually provided, even after inquiring with the company directly. One benefit of the video based method is that the video can also be used to validate the data by manually counting it for a short period and comparing the count with the automated results; a video-based product is thus all-in-one product. With the radar based methods, there were fewer options compared; however, information about the data they provide was more readily accessible, and the products were determined to provide the individual vehicle data necessary. Additionally, a major benefit of Wavetronix in particular was that IDOT has been installing Wavetronix radar detectors throughout its urban areas, particularly on freeways. This allowed for the possibility of connecting to IDOT’s existing sensors for additional accessibility and data.

After examining all of the options, it was determined that the Wavetronix SmartSensor HD detector was the best option. According to the Wavetronix website, “each individual vehicle is detected and its speed, duration, length and lane assignment is precisely measured” (Wavetronix 2006). Wavetronix’s accuracy has also been tested by a number of studies which Wavetronix references on its website to show its product’s reliability. In general, these studies showed around a 1 to 3 percent average error in volumes and about a 1 to 5 mph error in speeds (Wavetronix 2006). Figure 3 below shows an image of a SmartSensor HD.

Photo credit: wavetronix.com

Figure 3. Wavetronix SmartSensor HD

After selecting the device for collecting headway data, it was necessary to select locations from which to collect data. Because the purpose of the research is to compare parameters for different freeway scenarios and driver populations, it was important to collect data from different urban centers in Iowa, as well as rural locations when possible. Three urban areas in Iowa were selected: Des Moines, Council Bluffs, and the Quad

Cities (Davenport and Bettendorf in Iowa and Rock Island, Moline, and East Moline in Illinois). Additionally, one rural location a few miles outside of the Quad Cites was selected. While it would have been preferable to have more sites to compare, challenges with the data collection and time constraints prevented this.

In Des Moines, the Iowa DOT had not yet granted permission to use its already installed Wavetronix detectors, so a setup was created that could be installed on road signs temporarily to collect data for a few weeks at a time. This setup consisted of a metal pole on which the Wavetronix detector, a camera, and a solar panel were mounted. The solar panel charged batteries which were then used to power the camera and Wavetronix detector. Additionally, the Wavetronix detector was connected to the camera system, because the camera could be accessed through a cellular network. This also allowed for a live connection to the Wavetronix detector, which, in turn, allowed the data to be recorded. An example of this setup is provided in Figure 4.

This setup was installed at six locations in the Des Moines area over two separate periods of time. The first data collection period was from September 16 to 24, 2014, and the second period was from October 1 to 14, 2014. During the first data collection period, the temporary sensors were set up on Interstate 235 (I-235) just west of 73rd Street, directly across the interstate from one another. Each direction of traffic has three lanes of through traffic and one auxiliary lane for exit/entrance ramps. Though Wavetronix claims that its SmartSensors have the capability of observing up to 22 lanes, in order to test the accuracy of the sensors for the farther lanes, data from two sensors collecting the same data from opposite sides of the freeway from each other were compared. The other two locations during the first collection period were northbound on Interstate 35/80 (I-35/80) between an entrance ramp (from University Avenue) and an exit ramp (to Hickman Road). During the second data collection period, the two locations were on southbound I-35/80 in the same section of the roadway between Hickman Road and University Avenue. These locations are shown in Figure 5. The detector locations are shown in aerial photographs in Figure 6 through Figure 11, which also portray the lane configurations.

Figure 4. Example of Wavetronix detector and video setup

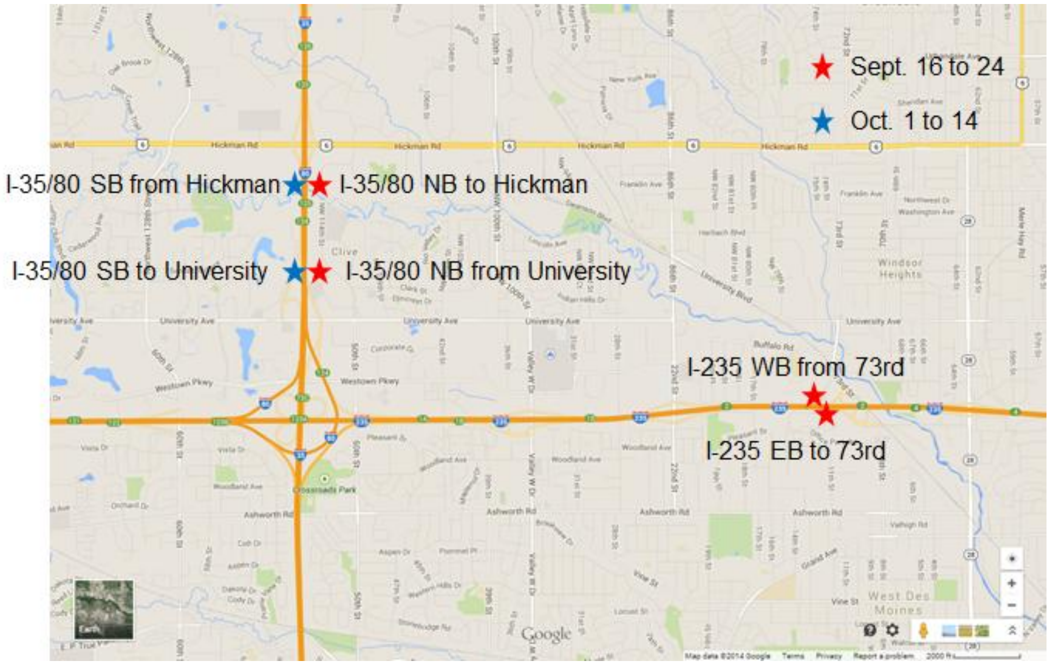

Figure 5. Des Moines data collection locations.

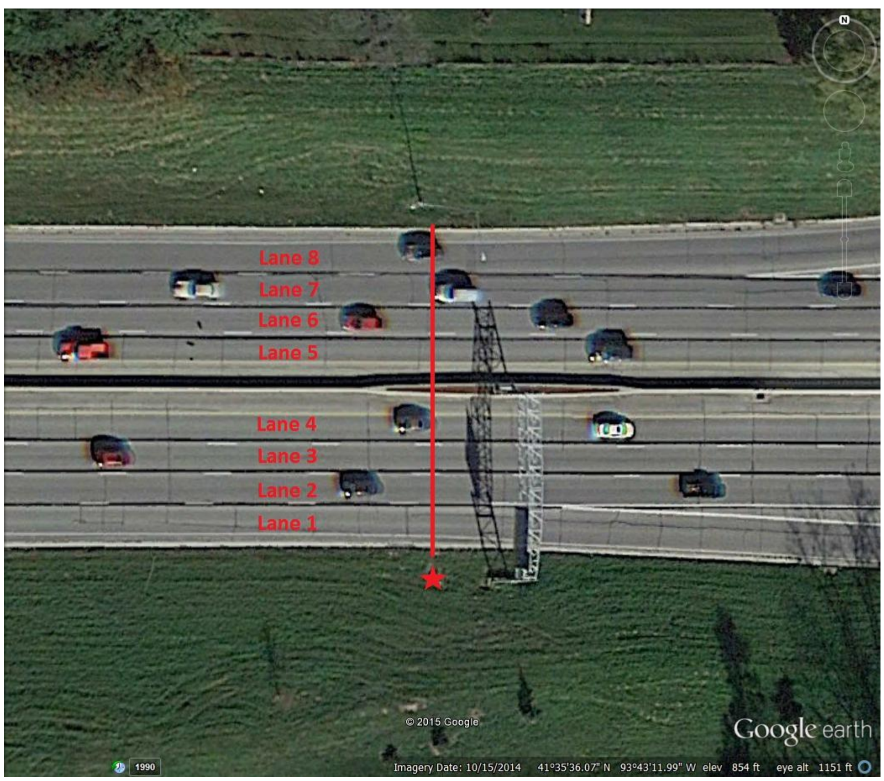

Figure 6. Wavetronix setup south of I-235 at 73rd St (Lane 1 is a weaving lane)

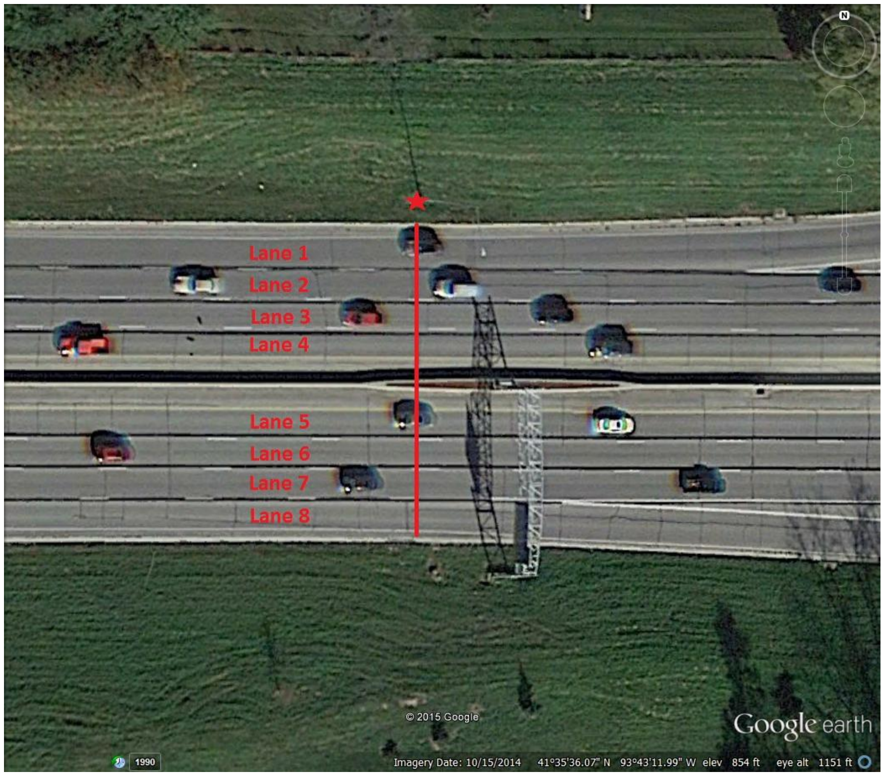

Figure 7. Wavetronix setup north of I-235 at 73rd St (Lane 1 is a weaving lane)

Figure 8. Wavetronix setup at I-80/35 NB at Hickman Rd (Lane 1 is an exit lane)

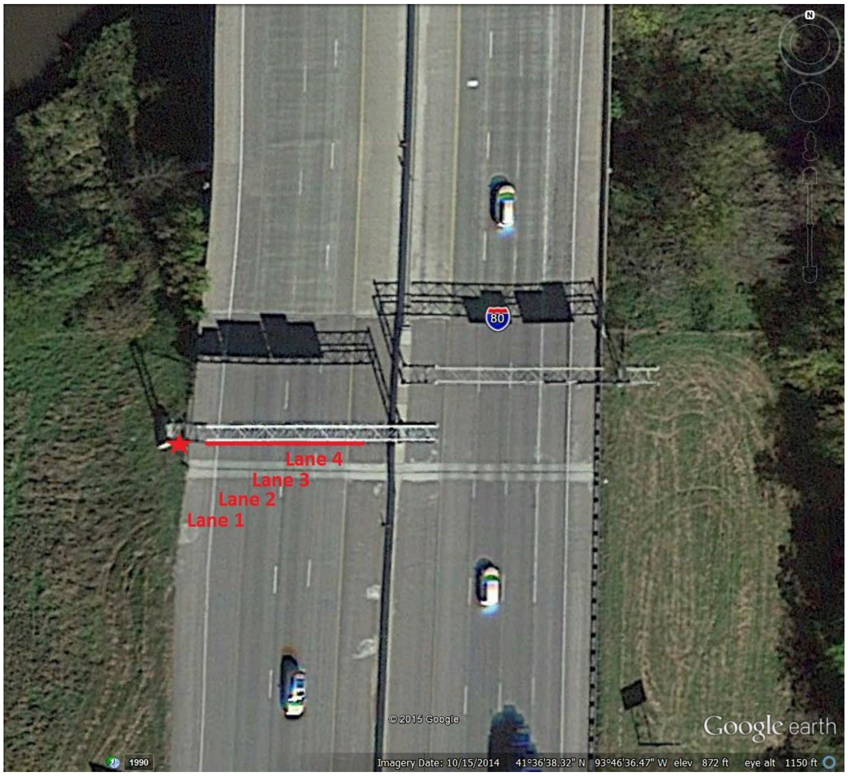

Figure 9. Wavetronix setup at I-80/35 SB at Hickman Rd (Lane 1 is a weaving lane)

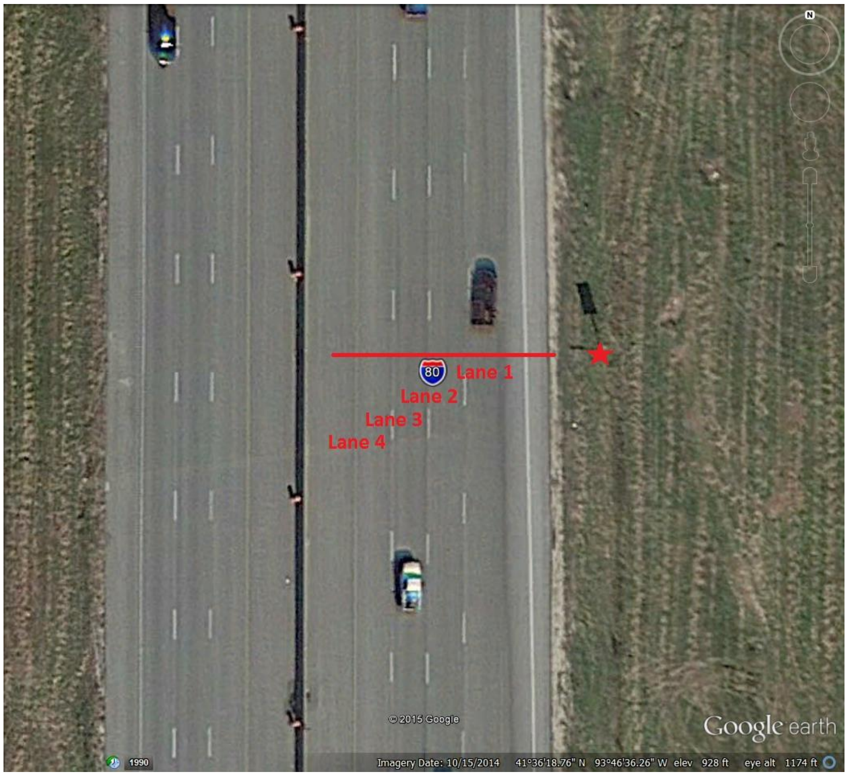

Figure 10. Wavetronix setup at I-80/35 NB at University Ave (Lane 1 is a merging lane)

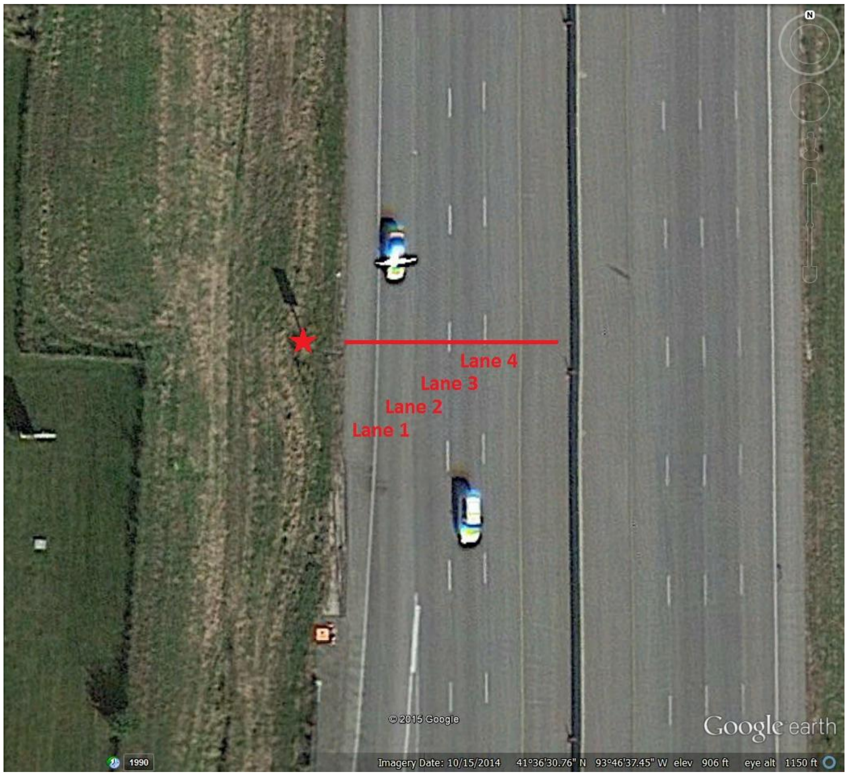

Figure 11. Wavetronix setup at I-80/35 SB at University Ave (Lane 1 is a weaving lane)

Because the data from the Des Moines locations were collected toward the end of the typical data collection season for Iowa, the rest of the data collection would have to wait for the summer of 2015. By that time, IDOT had granted permission to use its permanent sensors to obtain individual vehicle data, so at first that was the plan for the rest of the headway data collection. The permanent sensors were already accessible through an online data portal. However, this portal only provided aggregated data in a minimum increment of 20 seconds. The process of obtaining the individual vehicle data involved installing a small device made by Wavetronix called a Click301 in the communications cabinet where a Wavetronix detector was already installed. The Click301 receives power from the cabinet, connects to the existing Wavetronix setup, and connects to the IDOT network. This does not interfere with IDOT's data collection; it simply creates a copy of the stream of data. The Click301 does not record the data automatically, however. It has a unique IP address that allows it to be accessed remotely. Once the connection was established, the data recording was started manually, and, if the connection was lost, the recording was restarted manually. The data were stored in comma separated (CSV) files on the local computer used to access the Click301.

However, there were a number of issues with using the IDOT permanent sensors to collect individual vehicle data. The main issue was that most of the Wavetronix sensors were not compatible with this method of connecting to them. Additionally, the company that manages the sensors for IDOT, TransCore, did not maintain an up-to-date accounting of the sensors that would be compatible. Therefore, an inquiry had to be placed with TransCore about each individual sensor and its potential connectivity, and a response had to be awaited. Each inquiry about a group of sensors took at least a week, and after several rounds of communication only one sensor that would work with the Click301 could be located: a sensor on I-80 just east of the South Expressway entrance and exit ramps. Figure 12 shows the location and Figure 13 shows the aerial photograph. Once access to that sensor was gained, it recorded data off and on from May 30 to July 14, 2015. Interruptions to the recording were due to communication failures and malfunctions in the detector itself.

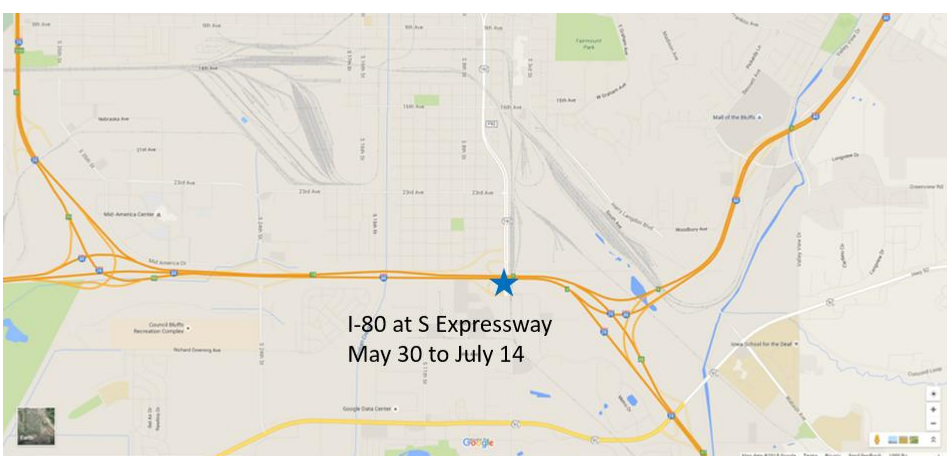

Figure 12. Council Bluffs detector location

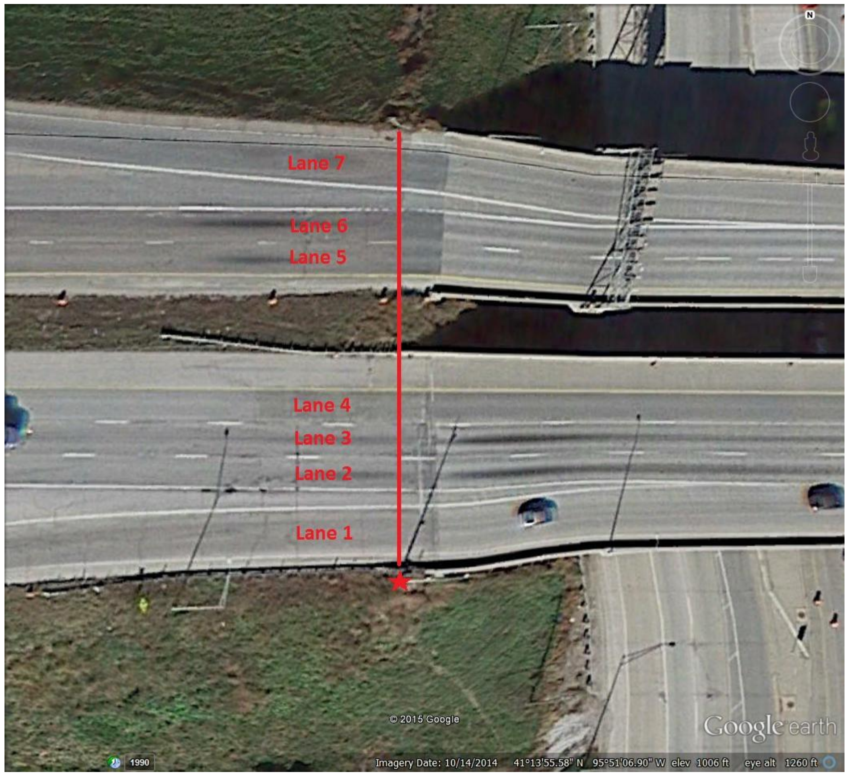

Figure 13. Wavetronix setup at I-80 at S Expressway (Lane 1 is an entrance, Lane 7 is an exit)

In the communication process with TransCore, it was discovered that none of the sensors in the Quad Cities were compatible with the chosen method of connecting to them. Because of this, the same temporary setup used in Des Moines was used at two urban locations in the Quad Cities and one rural location just outside of the Quad Cities. The urban locations were in the same section of Interstate 74 (I-74), with one just south of Spruce Hills Drive and the other just north of Middle Road. It should be mentioned that there was a major construction project on the I-74 bridge over the Mississippi River (south of these locations) which may have affected driver's behavior on that freeway, particularly in the southbound direction. Despite the construction, those sites were selected because I-74 is the only urban freeway in the Quad Cities that has heavy enough traffic to see a significant amount of car following. The rural location was on I-80 a few miles west of the Quad Cities. Figure 14 shows all three locations, and Figure 15 through

Figure 17 show the corresponding aerials. All sensors were collecting data off and on from July 17 to 31, 2015. The periods when the temporary setups were not collecting data were all due to communication errors or depleted batteries.

Figure 14. Quad Cities detector locations

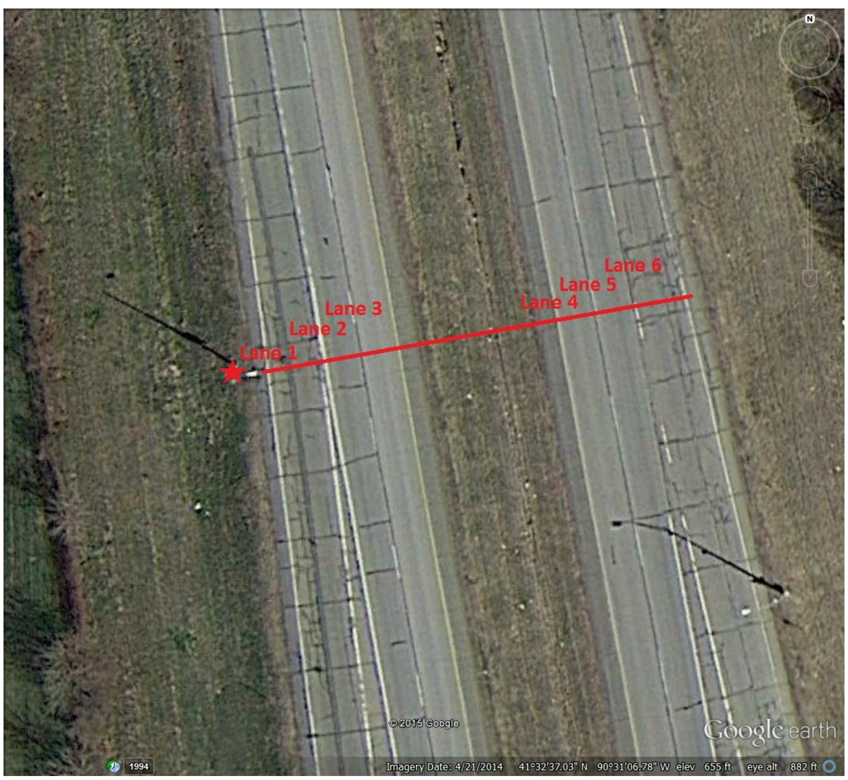

Figure 15. Wavetronix setup at I-74 at Middle Rd (Lane 1 is an exit, Lane 6 is an entrance)

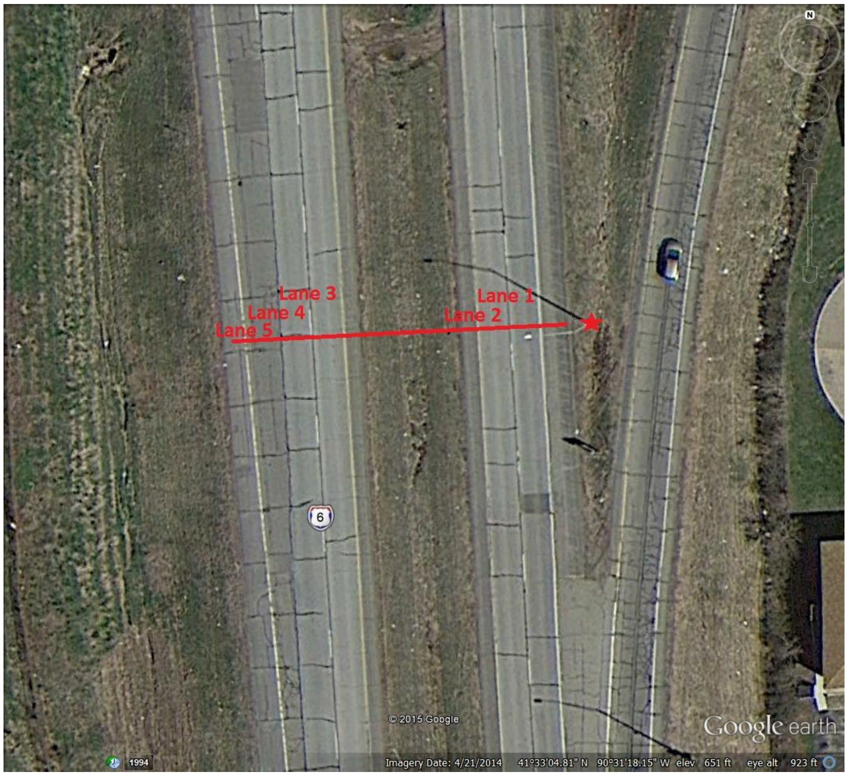

Figure 16. Wavetronix setup at I-74 at Spruce Hills Dr (Lane 5 is an entrance)

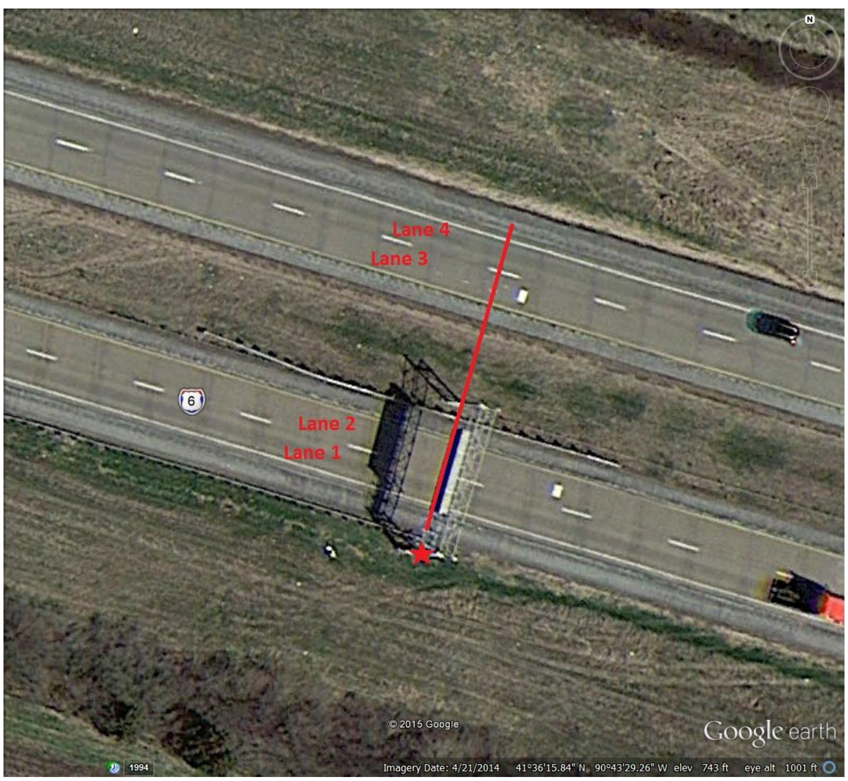

Figure 17. Wavetronix setup at I-74 west of the Quad Cities (Rural)

CHAPTER IV: DATA VALIDATION AND ANYLYSIS METHODOLOGY

Introduction

Once the headway and time gap data and standstill distances were collected, their quality was checked and they were analyzed using excel spreadsheets and the statistical software R. The accuracy of the standstill distance measurement process was validated, but due to the nature of the data, there was no ground truth to which the measurements themselves could be compared. The headway validation was achieved through comparison of the Wavetronix data to manual counts. The standstill distance data was analyzed by using t-tests between groups stratified by different variables. The headway data was analyzed separately for each detector location, and the results of those locations were compared using a practical significance threshold of 0.1 second. Additionally, statistical distributions were fitted to the headway data.

Data Validation

Introduction

Once the data were collected, their accuracy was evaluated. The accuracy of the standstill distances was validated by confirming lane line lengths with Google Earth and testing the accuracy of the Photoshop measuring tool. The headway data were validated through conducting a 30 minute manual count at each of the detector locations and comparing this count to what the temporary Wavetronix detector counted as well as what the closest IDOT-owned sensor counted (the research team had access to aggregated counts but not individual vehicle data for all IDOT sensors). In addition to the aggregated counts, relative vehicle class frequency and lane detection frequency were also compared. Finally, one 10 minute peak count was conducted during which the vehicle arrivals were recorded so headways could be calculated directly, and the average of those manually counted headways was compared to the average of the Wavetronix headways.

Standstill distance validation

Because the standstill distances were measured after the fact, it was not possible to directly validate the accuracy of the standstill distances by comparing the distances measured in Photoshop to the actual distances. However, the accuracy of the key assumption (the length of the lane line) and the accuracy of Photoshop's measuring capabilities were evaluated. The lengths of a number of lane lines in the areas of the stop-and-go traffic incidents were measured in Google Earth. All of the lane lines were within 0.9 feet of 10 feet, more than 93 percent of lines were within 0.6 feet, and the average error was 0.29 feet. There was also no observed trend of one city having longer or shorter lane lines than other cities. This supports the assumption that the lane lines measured 10 feet. In order to evaluate the accuracy of the Photoshop measurements, photos of a grid with known dimensions were taken from different angles and measured using the same method described in the methodology section. The average of the absolute relative error of these measurements was 1.2 percent. Additionally, the primary source of error appeared to be in determining the exact end points to be measured, which is limited by the image quality rather than the software. An example of one of these test images is shown below in Figure 18. To evaluate the human error, the undergraduate assistant and I processed the same video and found an average standstill distance approximately 0.5 feet off from each other. The remainder of the videos were all processed by the undergraduate assistant, so any potential bias should not have affected the comparisons within this dataset. With an accurate base measurement assumption and an accurate measuring method, the accuracy of the measurements overall can be reasonably be assumed.

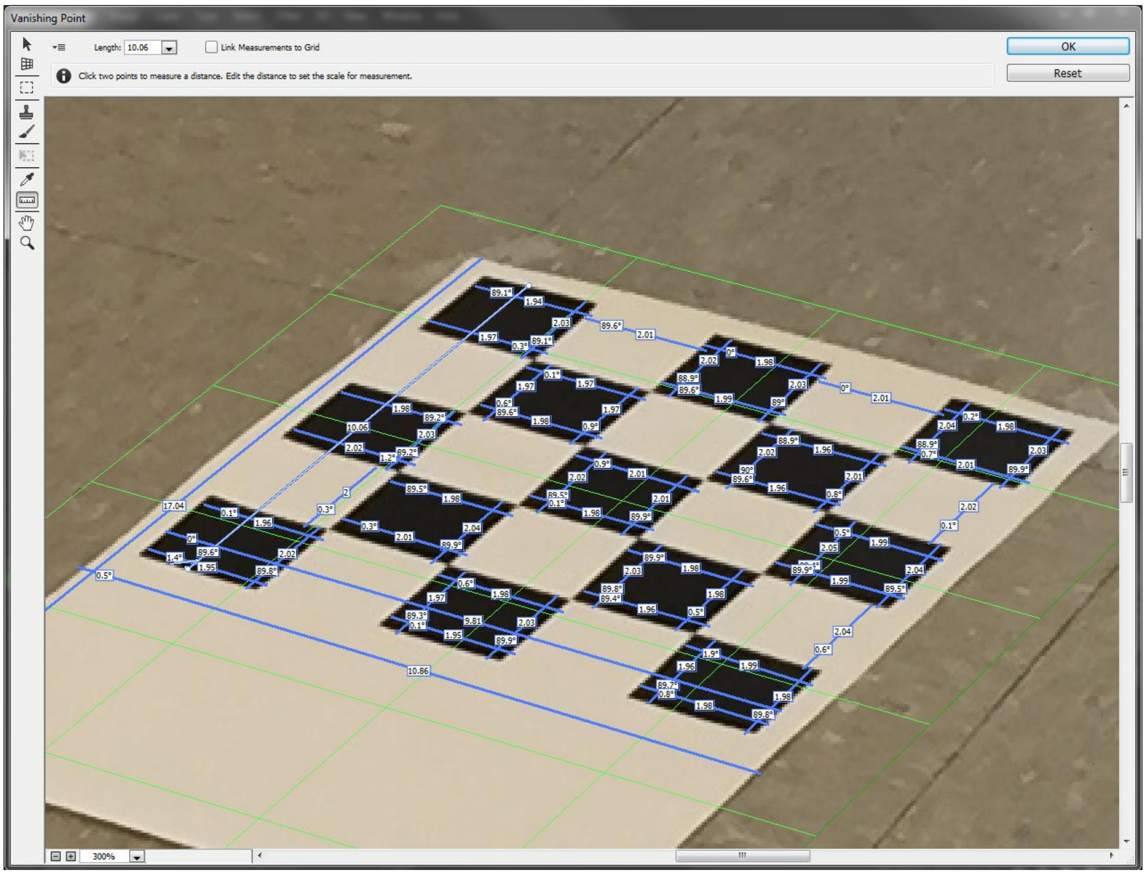

Figure 18. Example of Photoshop accuracy test using 2 in x 2 in squares in a grid pattern

Headway validation

Before the Wavetronix SmartSensor HD was selected, its accuracy was researched. There were a number of studies on the accuracy of Wavetronix's SmartSensor HD that had been completed. The SRF Consulting Group tested the accuracy of the detector's volume measurement in Minnesota as part of an on-ramp queue length measurement system and found the volume error was within 3 percent of the manual counts (MnDOT 2009). In South Korea, the accuracy of the Wavetronix detector was tested during different times of the day, and the study found a 95 percent volume accuracy and a 98 percent speed accuracy at all times (South Korea ITS Performance Test Institute 2008). In a study conducted at University of Maryland, a volume error of -3.6 to 2.7 percent, an average speed error of -1 to -2 mph, and an average absolute speed error of 2 to 5 mph were found (University of Maryland 2008). In a study conducted by Florida State University in association with the Florida DOT, 1 to 1.5 percent errors in daily volume and 2 to 9 percent (1 to 4 mph) errors in daily average speed were found (Florida State University 2007). In Denmark, the speed measurements were tested, and 1 to 5 percent average speed errors were observed (Hansen and Henneberg 2008). Finally, the speed measurements were compared to those of a highly calibrated piezo sensor system in West Virginia, and it was found that 92 percent of speed observations fell within 5 mph of the true speed, and that number increased to 98 percent when a 2 mph bias was removed (Wavetronix 2006). These past studies established that the Wavetronix detector should be accurate, but it was still important for the present research to validate each detector in case there was an error in setting up the system.

For each detector from which data was recorded, whether it was a temporary setup or a connection to an existing IDOT sensor, a 30 minute period of peak traffic was manually counted from video data. In the manual counting process, the lane, vehicle type, and the minute during which the vehicle arrived were recorded for each vehicle in a Microsoft Excel spreadsheet. This allowed for a comparison of the total counts, the minute-by-minute counts, the lane assignments, and the vehicle length measurement (vehicle class assignments).

The validation for the locations in Des Moines is summarized in Table 1 and Table 2. Overall, the total Wavetronix detector counts were within 1 to 9 percent of the manual video counts, with the exception of the detectors on I-235, which counted half as much traffic as was actually present. These differences could have been due to an issue with how the detectors were set up at those locations; whatever the cause, the I-235 locations were excluded from the analysis. The lane assignments and vehicle class assignments were also generally within 1 to 4 percent of reality, and often less than 1 percent off. Where temporary setups were used, nearby Iowa DOT sensors were also used to further validate the temporary setup counts. Examples of visual comparisons including all three data sources are shown in Figure 19 to Figure 21. Through counts were used when comparing the temporary setups to the IDOT Wavetronix detectors, because the data obtained from the IDOT detectors were only recorded for the through lanes.

Table 1. Des Moines Wavetronix detector accuracy summary (first collection period)

		Locations				
		I-235 EB (73rd)	**I-235 WB (73rd)**	**I-80/35 NB (Hickman)**	**I-80/35 NB (Hickman)**	**I-80/35 NB (University)**
	Time Observed	9/19/14 7:15-7:45	9/19/14 17:00-17:30	9/18/14 17:00-17:30	9/18/14 12:00-12:30	9/17/14 17:00-17:30
Error (in %)	**Count**	-50.71	-54.26	2.99	1.06	9.12
	Lane 1 %	-0.41	1.85	-0.23	-0.2	-0.68
	Lane 2 %	0.6	-0.35	0.04	0.33	-1.62
	Lane 3 %	-0.5	-0.44	-0.09	-0.44	-1.97
	Lane 4 %	0.3	-1.06	0.29	0.31	4.27
	Car %	0.7	9.4	1.9	3.3	1.72
	Truck %	-0.4	-9.2	-1.6	-3.2	1.72

Table 2. Des Moines Wavetronix accuracy summary (second collection period)

		Locations	
		I-80/35 SB (Hickman)	**I-80/35 SB (University)**
	Time Observed	10/8/14 17:00-17:30	10/6/14 17:00-17:30
Error (in %)	**Count**	1.01	7.27
	Lane 1 %	1.86	-1.59
	Lane 2 %	-3.1	-0.01
	Lane 3 %	1.05	-1.42
	Lane 4 %	0.18	3.02
	Car %	1	-0.42
	Truck %	-0.6	0.42

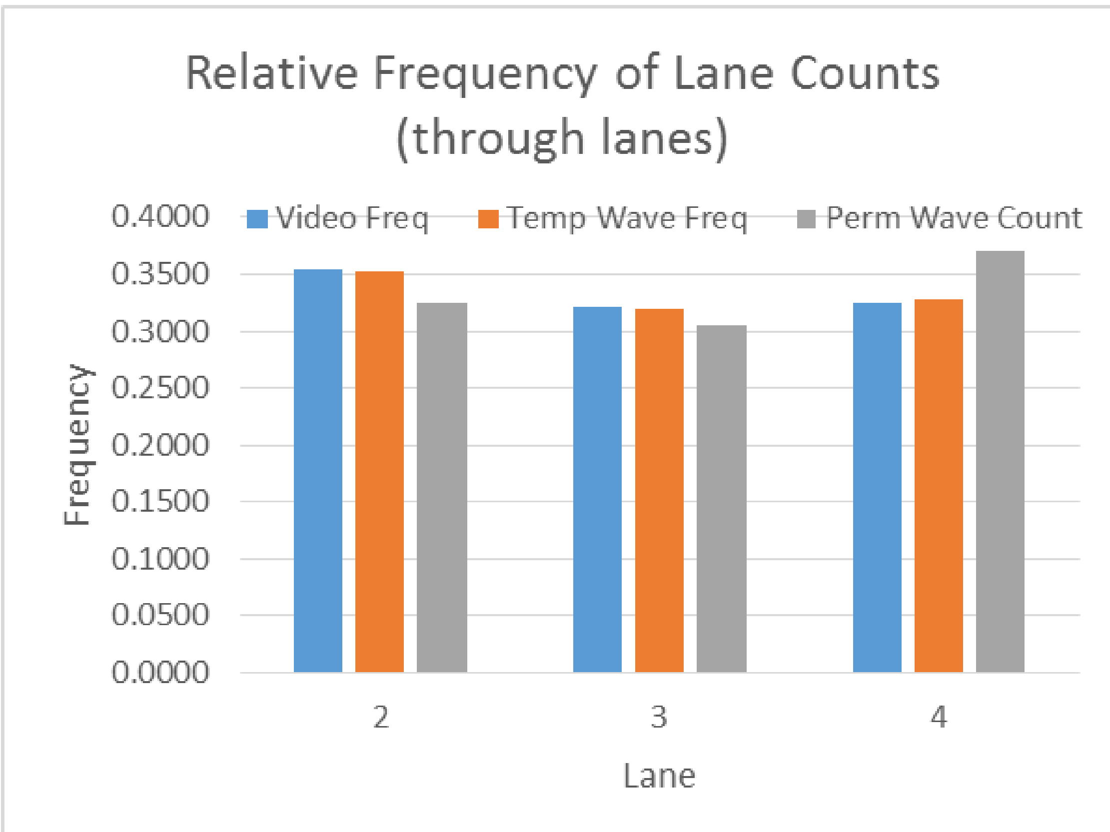

Figure 19. Example of visual comparison of lane proportions (NB I-35/80 at Hickman Sept 18 5 to 5:30 pm)

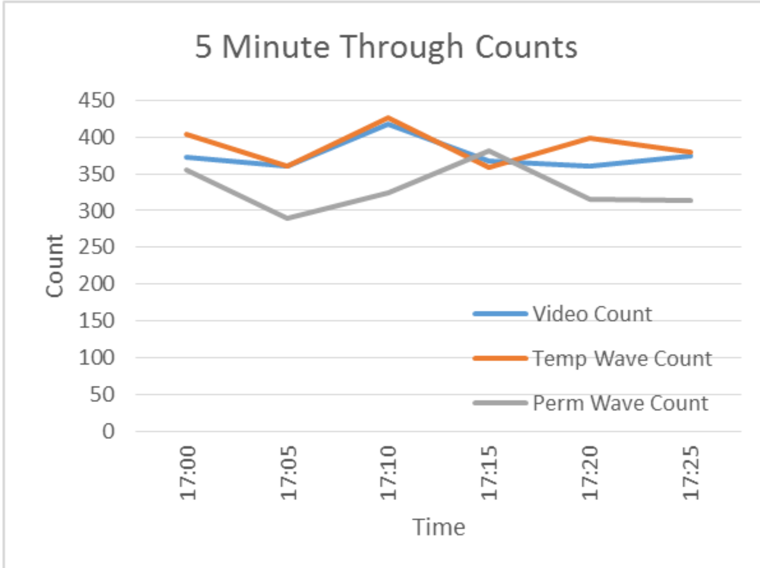

Figure 20. Example of visual comparison of 5 minute counts (NB I-35/80 at Hickman Sept 18 5 to 5:30 pm)

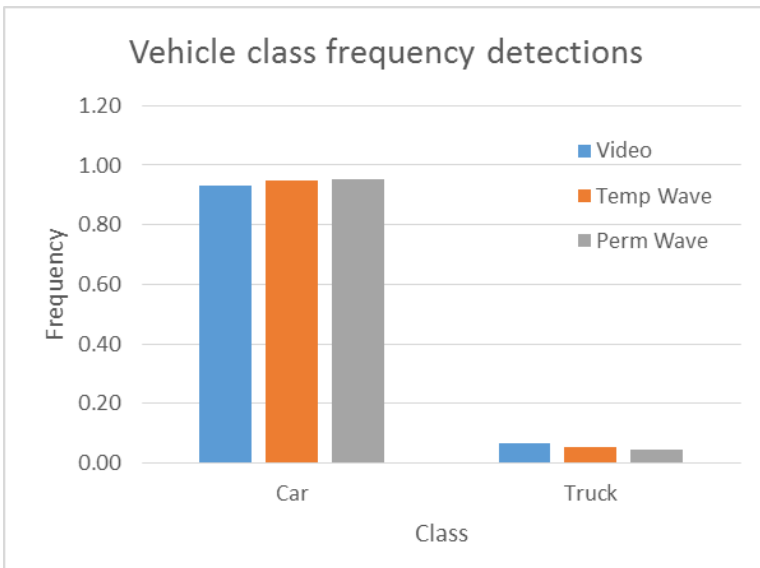

Figure 21. Example of visual comparison of lane proportions (NB I-35/80 at Hickman Sept 18 5 to 5:30 pm)

In addition to the 30 minute aggregate comparison, one 10 minute count was conducted at the I-80/35 northbound at Hickman Rd location, during which the vehicle arrival time was recorded in addition to the vehicle's lane and class. These individual vehicle arrival times were used to calculate individual headway and were compared with the individual vehicle headways of the Wavetronix detector. The average of the headways from the video was 2.7 seconds and the average of the headways from the Wavetronix was 2.73 seconds. In order to compare the distributions of the headways observed from the video and Wavetronix, a histogram of both the video's and the Wavetronix detector's individual headways is shown in Figure 22. This histogram shows visually that both distributions are quite similar.

As a follow-up to the visual comparison, the chi-square test was used to determine if there was a significant difference between the two distributions. The chi-square test compares between the observed and expected values within different categories or bins. In order to mitigate the differences at the extremes of the distribution (very small and large headways) having an excessively large impact on the chi-square statistic, several intervals were grouped together into larger bins. Headways of less than 2 seconds were grouped together as were headways greater than 6 seconds. Headways of less than 2 seconds were grouped together due to measurement error. This modified histogram is presented in Figure 23. This results in 10 groups (or 9 degrees of freedom), which corresponds to a critical value of 16.919 for a 95% level of confidence. The chi-square statistic for this distribution is 13.207 which leads to a p-value of 0.153, indicating that there is not a statistically significant difference observed between the two distributions.

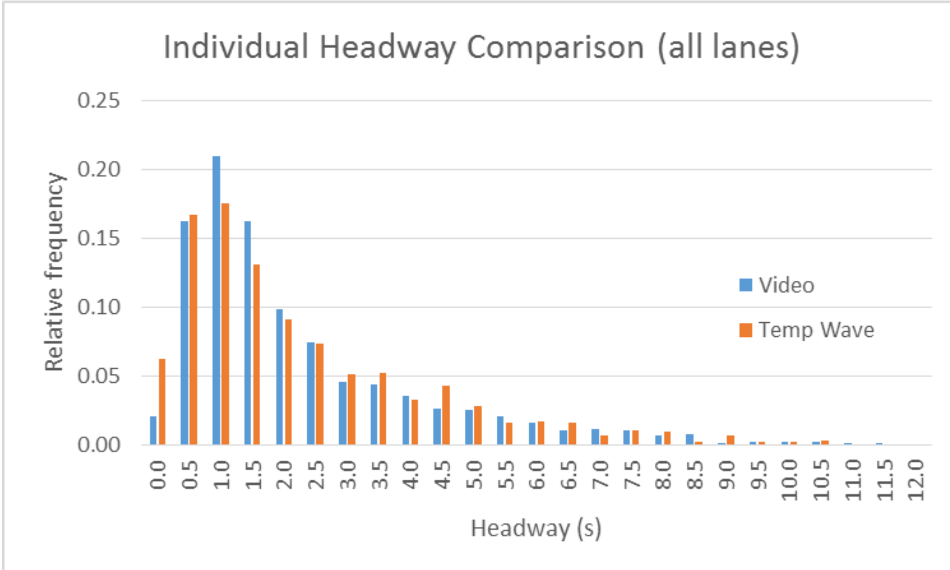

Figure 22. Histogram comparing individual vehicle headways from manual count and Wavetronix

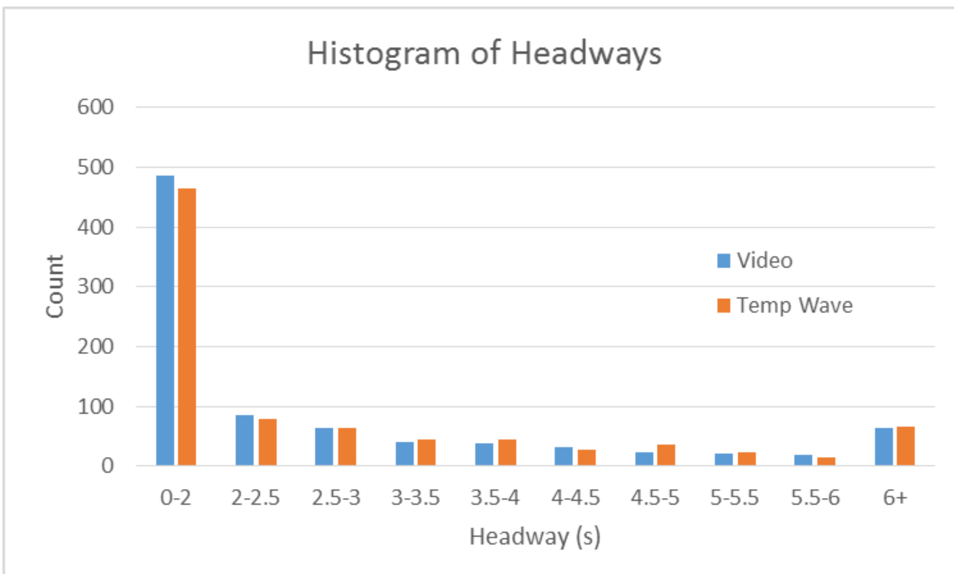

Figure 23. Modified histogram comparing individual vehicle headways from manual count and Wavetronix

The three temporary Wavetronix setups in the Quad Cities were not functioning for a large portion of their time in the field. In particular, the device at Spruce Hills Drive was only functioning for about seven hours on one day. However, the accuracy of all three detectors was validated using the same process that was used in Des Moines. In general, all detectors fairly accurate. The Wavetronix detector counts were all lower than the video counts: the rural I-80 location by 0.62 percent, the I-74 location at Middle Road by 2.44 percent, and the I-74 location at Spruce Hills Drive by 3.58 percent. The relative lane percentages were all off by less than 1.5 percent. The car and truck percentages were off by 1 to 3 percent. These results are summarized in Table 3 to Table 5.

Table 3. I-74 at Spruce Hills Drive Wavetronix detector accuracy

		I-74 Spruce Hills
	Time Observed	7/29/2015 9:00 to 9:30
Error (in %)	**Count**	-3.58
	Lane 1 %	0.42
	Lane 2 %	0.41
	Lane 3 %	0.1
	Lane 4 and 5 %	-0.93
	Car %	2.13
	Truck %	-2.13

Table 4. I-74 at Middle Road Wavetronix detector accuracy

		I-74 Middle Road
	Time Observed	7/23/15 17:00 to 17:30
Error (in %)	**Count**	-2.44
	Lane 1 %	-0.14
	Lane 2 %	0.45
	Lane 3 %	0.04
	Lane 4 %	0.05
	Lane 5 %	-0.76
	Lane 6 %	0.35
	Car %	1.25
	Truck %	-1.25

Table 5. Rural I-80 west of Quad Cities Wavetronix detector accuracy

		I-80 West of Quad Cities
	Time Observed	7/23/15 17:00 to 17:30
Error (in %)	**Count**	-0.62
	Lane 1 %	0.24
	Lane 2 %	1.37
	Lane 3 %	-1.13
	Lane 4 %	-0.48
	Car %	-3.21
	Truck %	3.21

In Council Bluffs, access to recording the individual vehicle data at one IDOT-owned Wavetronix detector was obtained, so the data used for the analysis were only compared to the manual count from the video (i.e., not an additional separate Wavetronix detector as well). Additionally, the video corresponding to the times the individual vehicle data were being recorded was mistakenly not downloaded. Therefore, the video was downloaded later and compared to the 20 second interval aggregated data obtained from the online data portal. Unfortunately, for some reason the vehicle class counts were not recorded in the aggregated data, so these data could not be compared. However, experience with the other detectors indicated that the class percentages are close, even if the counts are off, so it was assumed that the class percentages were reliable as well.

For the overall count, the detector counted 6 percent more vehicles than the video. The detector also appeared to be more accurate for the near lanes (eastbound) than the far lanes, with errors of 0.5 percent in the eastbound on ramp and 5 percent in the eastbound through lanes compared to 1 percent in the westbound exit ramp and 7 percent in the westbound trough lanes. These results are summarized in Table 6.

Table 6. Council Bluffs Wavetronix detector accuracy summary

		I-80 S Expressway
	Time Observed	8/24/15 17:00 to 17:30
Error (in %)	**Count**	6.03
	EB On Ramp %	-0.55
	EB Through %	-5.21
	WB Through %	7.10
	WB Exit Ramp %	-1.34

Analysis Methodology

Introduction

While the standstill distance data were compiled in Microsoft Excel, for reproducibility's sake the statistical software R was used for the analysis. For each stop-and-go incident, in addition to the standstill distance measurements, the conditions surrounding the incident were also recorded. These conditions included weather, presence of a curve, day or night conditions, the cause of the incident (if known), and the city in which it occurred. R was used to find sample statistics while stratifying the data in different ways. For example, the mean standstill distance was calculated for each of the incident types (accident, construction, slow traffic, stalled vehicle, and unknown). R was also used to plot the histogram of data to observe the distribution. Because it was a skewed distribution, it was transformed to make it more symmetric so t-tests could be used to compare the means of the different groups.

The analysis of the headway data was initially conducted using Microsoft Excel, then streamlined using Microsoft Access and R. It was discovered that working with individual vehicle data in Excel is unwieldy or even impossible due to the size of the dataset (hundreds of thousands of rows or more). Therefore, instead of using Excel, the raw data were imported into Microsoft Access so that each detector was in a table of its own. Then, using the built-in Windows program "Data Sources (ODBC)" a database was created using the Access file with all the detector data in it. Finally, the "RODBC" package in R was used to establish a connection to the ODBC database and import data from each detector into R. By doing this, the headway analysis was sped up considerably and was much more reproducible. This is because R is a statistical programming language, so the code can be written for one detector and tweaked slightly for other detectors.

The headway analysis itself consisted in part of calculating the headway and time gap for each vehicle pair, defining cars versus trucks, defining a maximum headway threshold for car following, filtering the data to congested conditions and following vehicles only, fitting statistical distributions to the data, finding summary statistics for the headway for each following combination for each site, and comparing all these results. It should be noted that headway and time gap are two different variables. Headway is the time between successive vehicles measured from the same point on each vehicle (the front bumper in the case of the Wavetronix data). Time gap is the time from the back bumper of the leading vehicle to the front bumper of the following vehicle as shown in Figure 24. Thus, a car can be following another car with the same headway as a car following a truck, but due to the length of the truck, the car following the truck will have the shorter time gap.

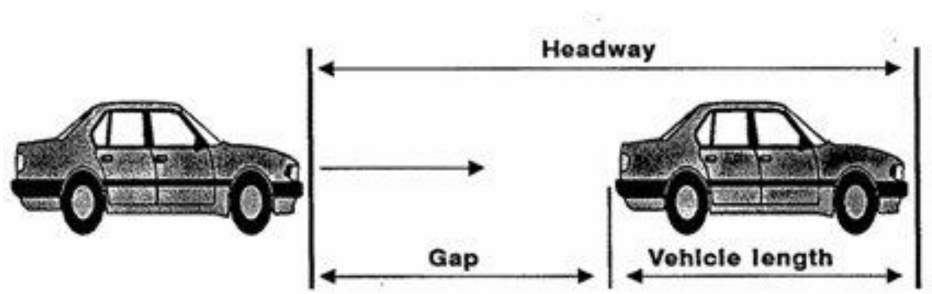

Photo credit: ops.fhwa.dot.gov

Figure 24. Difference between headway and time gap

Standstill distance analysis

While the collection of the standstill distance measurements was more time consuming and tedious than that of the headway data, analyzing the data was much more straightforward. The standstill distance measurements from each stop-and-go incident were compiled into one Excel file and saved as a CSV file. As mentioned above, in addition to the distance measurement and the vehicle pair type, the location, lighting, cause of incident, presence of curve, and weather conditions were recorded (if known) for each stop-and-go incident, as shown in Figure 25. Once everything was in one file, this file was imported into R using the read.csv() command. An overall histogram of the distances was created that revealed a skewed distribution and led to the exclusion of some outliers that did not fit with normal driving behavior (see Figure 26). It is apparent from this histogram that there are some excessively long measurements. From a visual inspection of the histogram, it was determined that measurements of longer than 25 feet fell outside typical standstill distances, and such measurements were excluded. These measurements could have been a result of vehicles stopping for reasons other than stopping for the vehicle in front of them (e.g., to perform a lane change maneuver).

Then, the mean, median, and standard deviations were calculated for different stratifications of the data to compare the standstill distance measurements across different groups. To compare between groups for statistically significant differences, t-tests were used. Rather than hypothesis tests for specific significance levels, p-values were used to get a better idea of the strength of the t-tests' conclusions. Because the standstill distances were a skewed distribution, they were transformed to be more symmetric before using the t-test to compare them. The distribution was right-skewed (long tail to the larger values), and the data were transformed by taking the square root of each observation. The entire process of collecting and analyzing the standstill distances is summarizes in the flowchart in Figure 27.

	A	B	C	D	E	F	G	H	I	J	K
1	ID	Video	City	File	Distance	PairType	Lighting	Cause	Curve?	Weather	Notes
2	1	AMTV1 Nov 21 1030 to 1130 Distances	Ames	10 33 47	17.88	CC	Day	Accident	No	Clear	
3	2	AMTV1 Nov 21 1030 to 1130 Distances	Ames	10 33 47	8.61	CC	Day	Accident	No	Clear	
4	3	AMTV1 Nov 21 1030 to 1130 Distances	Ames	10 33 47	14.22	CC	Day	Accident	No	Clear	

Figure 25. Sample standstill distance data

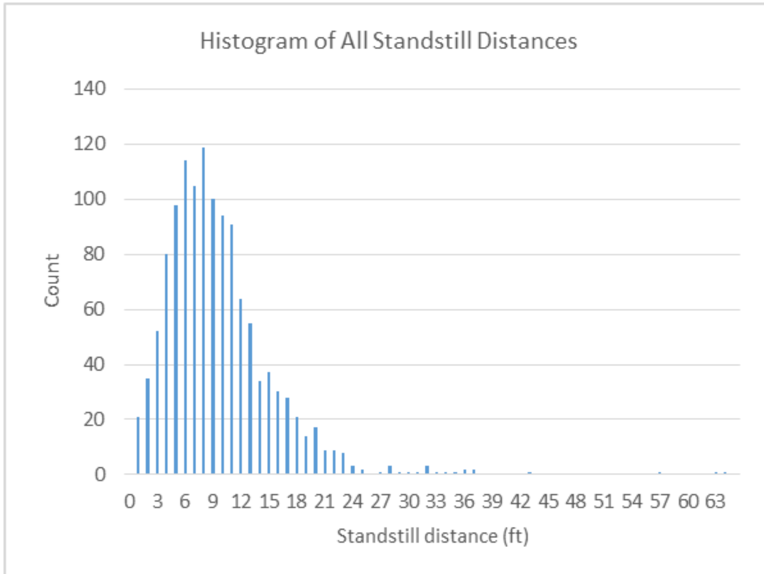

Figure 26. Histogram of unfiltered standstill distances

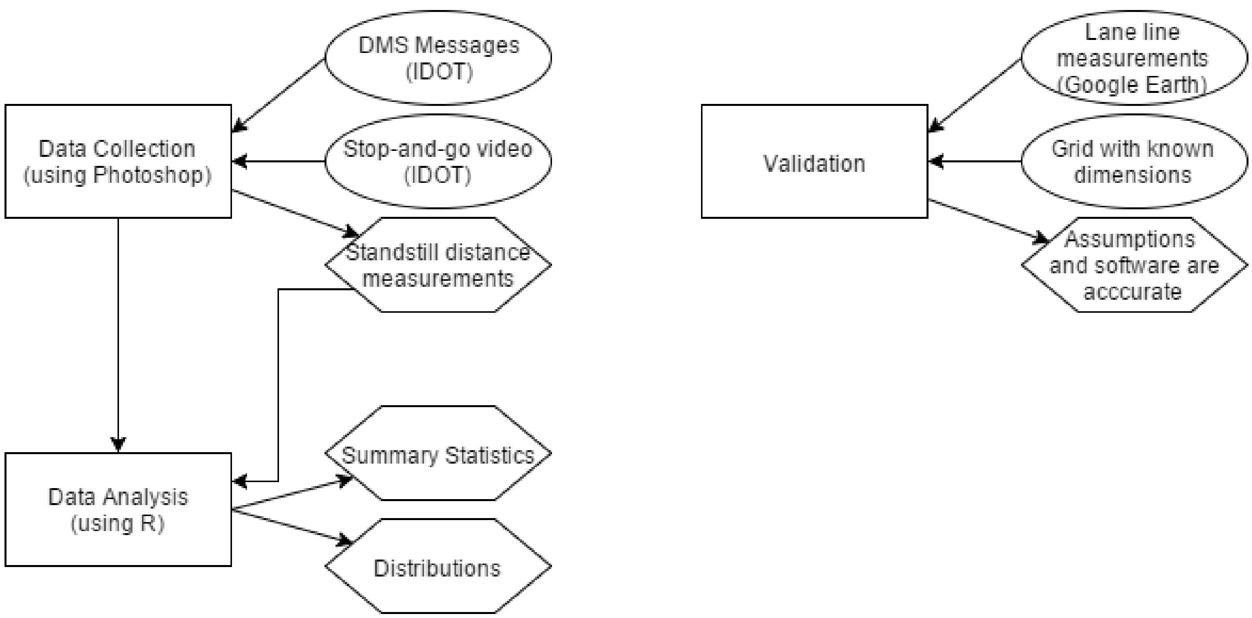

Figure 27. Flow chart for standstill distance analysis

Headway and time gap analysis

In the raw individual vehicle data from Wavetronix each vehicle was represented by a row in a CSV file. Each vehicle is assigned a lane, length (in feet), speed (in mph), vehicle class, range (distance from detector in feet), and time of detection. An example of this data is shown in Figure 28. One important note about the raw data is that not all vehicles are assigned a speed by the detector due to an internal quality control mechanism. Because the speed of the leading vehicle is required calculate the time gap, only vehicle pairs in which the leading vehicle had an assigned speed were used in the analysis. The analysis of these data is broken up into 4 main parts: headway and time gap calculation, vehicle class threshold determination, vehicle following determination/filtering, and headway and time gap distribution analysis. The overall process from data collection to analysis is shown in the flow chart in Figure 29.

2	LANE	LENGTH	(MPH)	CLASS	RANGE	YYYY-MM-DD HH:MM:SS.sss
3	LANE_01	19	71.7	2	32	9/16/2014 18:49:10.41
4	LANE_03	55	68.4	4	56	9/16/2014 18:49:11.10

Figure 28. Sample Wavetronix data

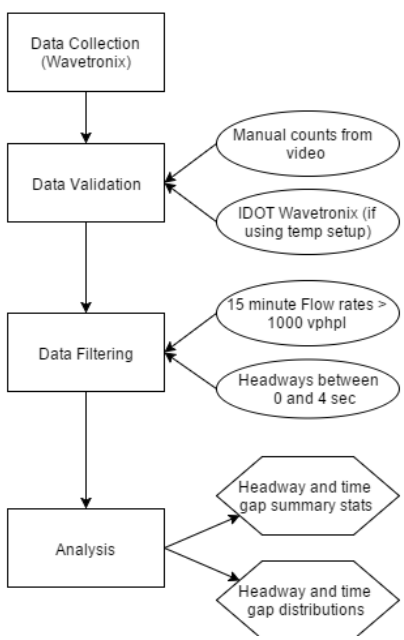

Figure 29. Flow chart for headway analysis

Headway and time gap calculation

To calculate the headway for each vehicle, the differences between successive vehicle arrival times in the same lanes were found. Once the data were imported into an Access database and that database was accessed through R, it became much faster to isolate the lanes and calculate the headways. Isolating lanes can be accomplished many ways in R, but this research used the filter() function in the "dplyr" package to assign each lane's data to a separate object in R. Then, a "for" loop was used to calculate the headway where the headway of the i^{th} vehicle was determined by subtracting the $(i-1)^{th}$ time of arrival from the i^{th} arrival (see Equation 1).

Equation 1

$$\boldsymbol{Headway_{ij} = t_{ij} - t_{(i-1)j}}$$

Where:

$Headway_{ij}$ = the headway of the i^{th} vehicle in the j^{th} lane (in seconds)

t_{ij} = time of arrival of the i^{th} vehicle in the j^{th} lane

$t_{(i-1)j}$ = time of arrival of the $(i-1)^{th}$ vehicle in the j^{th} lane

It should be noted that the time of arrival variable is stored as the number of days from the start of the year 1900, so January 1, 1900 is stored simply as "1" and times are stored as decimals, because they are fractions of days. So, to get the headway value in seconds, the difference is multiplied by 86400 (the number of seconds in a day).

Because the headways are measured from the front bumper of the leader to the front bumper of the follower, and the time gap is measured from the back bumper of the leader to the front bumper of the follower, the only difference between headway and time gap is that the time gap is shorter by the length of time it takes for the leading vehicle to clear the detector. That time can be calculated simply by dividing the length (in feet) of the leading vehicle by its speed (in feet per second). Then the time gap of the following vehicle is it's headway minus the time for the leading vehicle to clear the detector (see Equation 2). It is important to remember that because the time gap calculation introduces two more measurements than the headway calculation, there is a reduced level of confidence in each individual time gap measurement. However, as long the measurements are not biased in one direction or the other from the true measurement and there is a sufficiently large sample size, the sample average time gap should be close to the actual average gap time. Again, because all the vehicles did not have assigned speeds, only pairs where the leading vehicle had a speed could be used to calculate time gaps.

Equation 2

$$TimeGap_{ij} = Headway_{ij} - \frac{Length_{(i-1)j}}{Speed_{(i-1)j}}$$

Where:

$TimeGap_{ij}$ = the time gap of the i^{th} vehicle in the j^{th} lane (in seconds)

$Length_{(i-1)j}$ = the length of the $(i\text{-}1)^{th}$ vehicle in the j^{th} lane (in feet)

$Speed_{(i-1)j}$ = the speed of the $(i\text{-}1)^{th}$ vehicle in the j^{th} lane (in feet per second)

Vehicle Classification

While the Wavetronix detector assigns each vehicle to one of seven vehicle classes based on its length, this research was focused only on comparing passenger cars with trucks. It was therefore necessary to define a threshold length to distinguish between cars and trucks. With any chosen length, there is always some overlap in the types of vehicles and the capabilities included in each group. In particular, small trucks can sometimes behave as cars and other times as larger trucks. So, with that in mind, 35 feet was the length cutoff selected because IDOT uses four classes for its permanent sensors (0–10 feet, 10–19 feet, 19–35 feet, and 35–256 feet). Additionally, the distribution of vehicle lengths was observed through histograms. These revealed two distinct peaks, one for cars (which small trucks spill into) and a much smaller one for large trucks (see an example in Figure 30). By observing the distribution, it was clear that 19 feet should not be selected as the cutoff, because this would have split the cars into separate groups. The 35 foot mark also appeared to divide the long tail of the cars group (which represents small trucks) in half, which would cause the small trucks to be split fairly evenly between the car and truck groups.

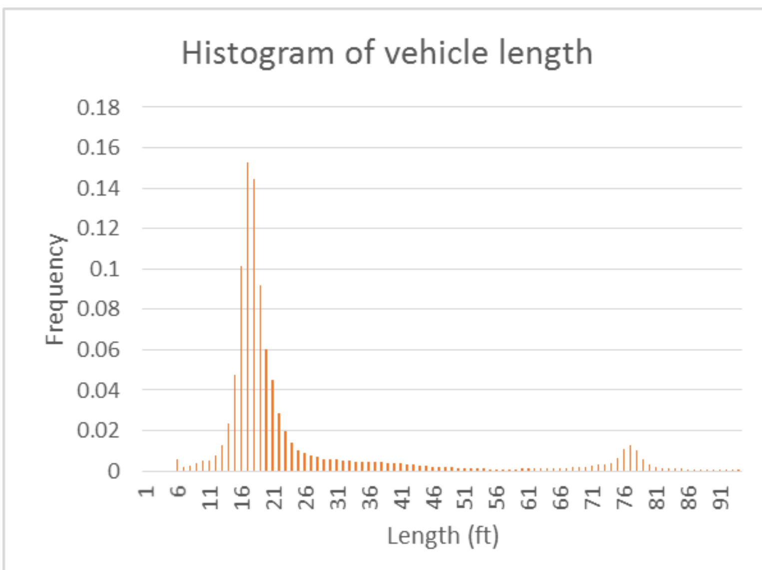

Figure 30. Example of a histogram of vehicle length (taken from I-80/35 NB at Hickman)

Vehicle following threshold and filtering

Another step in the analysis was determining the maximum headway at which the second vehicle could still be considered following the first vehicle. There have been a few efforts to establish this threshold in past studies, but these studies were mostly focused on rural two-lane roads. For example, the *HCM 2010* sets the threshold for rural two-lane roads to 3 seconds, but it does not offer any explanation for how this value was determined (TRB 2010). However, a study from Sweden outlined a process for determining which vehicles can be considered "free" by finding the correlation between leading and following vehicle speeds at different headway values (Vogel 2002). This methodology was applied with the opposite mentality in mind: which vehicles can be considered following? Thus, for the data from each of the detectors, the headways were rounded to the nearest second, and the Pearson correlation coefficient (see Equation 3) was calculated between the leading vehicle's speed and the following vehicle's speed (as long as both vehicles were assigned speeds) for each group of rounded headway data, and the results were plotted. An example of this for the detector on I-80 at S Expressway in Council Bluffs is shown in Figure 31. It can be seen from the graph that leading and following vehicle speeds are highly correlated at small values of headway, and that as the headway increases, the correlation decreases. This makes sense intuitively, but the advantage of this method is that it is possible to quantify how much of an influence the leading vehicle has at each headway value. The correlation drops from a peak of approximately 0.95 at a rounded headway of 0 seconds (i.e., headways of 0 to 0.5 seconds) to a baseline of approximately 0.65 to 0.7. It can also be seen in Figure 31 that the point of inflection of the graph (where it switches from concave down to concave up) is at a headway of approximately 4 seconds. This can be interpreted as the point where the influence of the lead vehicle begins to dissipate and the vehicles are more likely to select their own speeds. The correlation at 4 seconds is about 0.8, which is still quite high. Similar trends were observed for most of the other detectors. From Vogel's (2002) perspective (where truly free vehicles are always observed), the speed correlation levels off at around 6 or 7 seconds, which is what that study found as well. However, the focus of the present research is following vehicles, so the 4 second threshold was selected.

Equation 3

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Where:

r_{xy} = Pearson correlation coefficient

xi = the ith value of variable x

yi = the ith value of variable y

$\bar{x}$ = the mean value of variable x

$\bar{y}$ = the mean value of variable y

n = number of observations

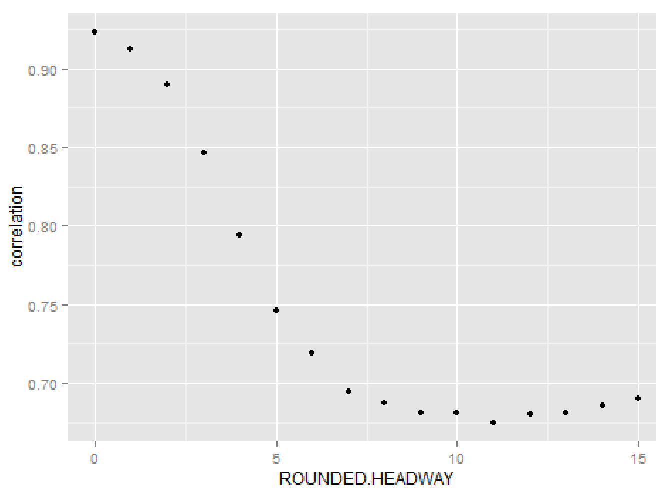

Figure 31. Correlation of leading and following vehicle speeds vs. headway (from I-80 at S Expressway)

The selection of this 4 second value is strengthened by Wasielewski's (1979) similar finding that the following vehicle distribution ranges from 0 to 4 seconds

(Wasielewksi 1979). In that study, the author measured 42,000 headways in one lane of an urban freeway over a variety of flow ranges. The study established that free flowing headways are exponentially distributed. The author thus looked for the smallest headway such that if an exponential distribution were fitted to the values higher than this headway, there would be no significant deviation from the exponential distribution above that headway, and there would be significant deviation within 0.5 seconds less than that headway. The results would indicate that traffic above this value is in free flow, while traffic below this value has enough car following to create a statistically significant difference from the free flow distribution.

In addition to filtering out headways larger than the 4 second cutoff, the individual vehicle data were also filtered to when the roadway was not in free flow. The purpose of this filtering was to limit the scope of the analysis to situations in which low headways are more likely representative of actual car following situations. If traffic on the road is minimal, then some low headway values are possibly the result of a following vehicle approaching a slower leading vehicle in the same lane and then passing the slower vehicle. This headway value would not necessarily be representative of the headway the driver may select in a following situation. The HCM 2010 suggests 1,000 passenger cars per hour per lane (pc/hr/ln) as the maximum value for which an uninterrupted flow facility can be said to be operating in free flow (TRB 2010).

Fifteen-minute flow rates were therefore calculated for each direction of travel using only the through lanes, giving an average flow rate in veh/hr/ln for each direction. Vehicles that arrived during a 15 minute period of less than 1,000 veh/hr/ln were excluded from the analysis. The flow rates were not converted to pc/hr/ln partly for simplicity and partly because 1,000 veh/hr/ln equates to more than 1,000 pc/hr/ln due to the presence of trucks, which slightly increases the level of congestion at the threshold flow rate (which is the point of eliminating free flow periods). The average headway was also observed to level out starting at 1,000 veh/hr/ln and higher traffic volumes. While a higher flow threshold would be beneficial for analysis, the traffic volumes throughout Iowa are not typically high enough that there would be enough data to analyze, particularly in the Quad Cities.

Additionally, in the filtering process, entrance and exit ramp lanes were excluded due to the different behavior of drivers in comparison with drivers in through traffic. Finally, consideration was given to including a filter for a speed difference of less than a certain threshold, extending the idea that following vehicles have similar speeds to the leading vehicles. However, the speed error of the Wavetronix detector would not make such a filter meaningful.

Headway and time gap distributions

In order to get a better understanding of the headway and time gap data, and as a point of comparison with past studies involving individual headway measurements, various statistical distributions were fit to the filtered data. This step of the analysis was only performed after the switch to R was made. The fitdistr() function in the R package "MASS" (which stands for Modern Applied Statistics with S) was used to fit these distributions using maximum likelihood estimation (MLE). There are 14 univariate distributions which are included in the fitdistr() function: beta, Cauchy, chi-squared, exponential, f, gamma, geometric, lognormal, logistic, negative binomial, normal, Poisson, t, and Weibull. Some of these distributions were not applicable to the headway data. For example, the Poisson and negative binomial distributions require the random variable (headway in this case) to be integers, and is most often used to model the probability of a specific number of random events occurring within a set period of time. The distributions which compared were Cauchy, exponential, gamma, lognormal, normal, and Weibull. They were compared based on each model's maximized log-likelihood as well as a visual comparison of each distribution's probability density function versus a histogram of the data. The three best models according to log-likelihoods were the lognormal, gamma, and Weibull distributions, and their log-likelihood values were close to each other. So their probability density functions were plotted against the histogram of the data (see Figure 32 for an example). From examining the distributions, it appears that the lognormal distribution fits better than the gamma and Weibull distributions (particularly the location of the peak), despite the log-likelihoods indicating slightly better performance for the gamma and Weibull distributions. This resulted in the selection of the lognormal distribution to represent the headway data, which is consistent which what past studies have found (see Chen et al. 2014, Zhang et al. 2007, and Ha et al. 2012). While other more complicated models have been shown to represent headway data marginally better, the lognormal distribution is usually cited as being the best univariate distribution for headway. Additionally, it appears that either the lognormal or gamma distribution would fit the data well, but for consistency with past literature, the lognormal was favored in this study.

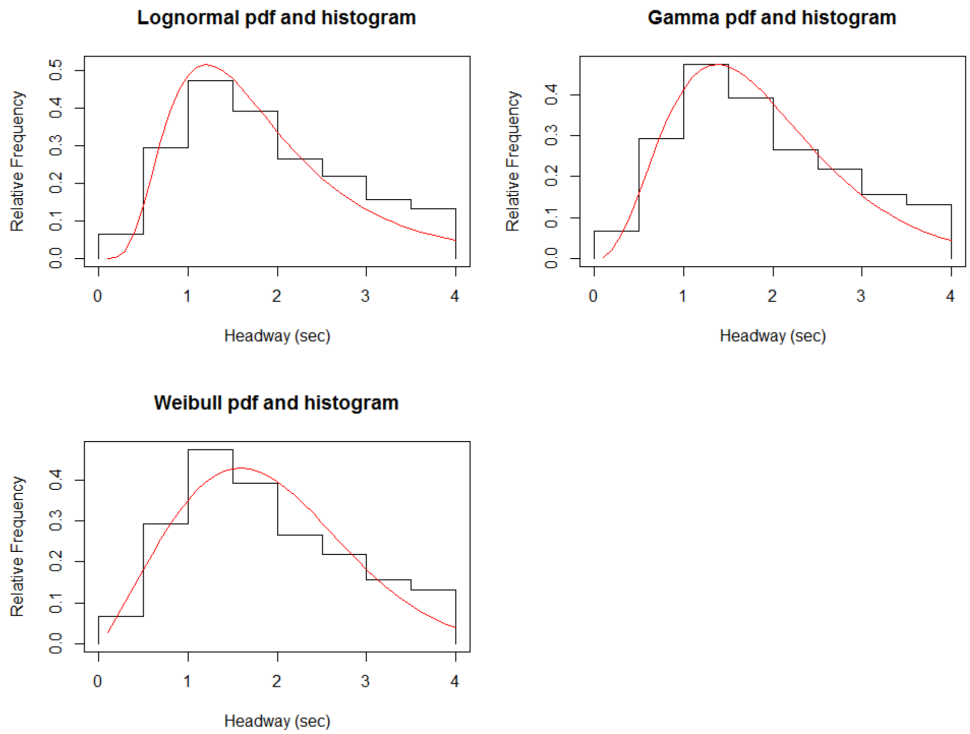

Figure 32. Comparison of lognormal, gamma, and Weibull distributions (data: I-80 in Council Bluffs)

Once the lognormal distribution was chosen to represent the headway distribution, it was also used to confirm that car following and truck following behaviors differ statistically. Lognormal distributions were fit to car following headways and truck following headways separately. Then, the likelihood ratio test was used to compare the effectiveness of modeling car and truck following separately (separate models) to modelling them both together (pooled model). The likelihood ratio test comparing two models (pooled model versus separate models) can be conducted by using the log likelihoods at convergence. The test statistic, which is chi-square distributed, is shown below in Equation 4. The null hypothesis (H_0) is that parameters for car following model and truck following model are the same. As shown in Table 7, the test statistics, χ^2, are all much greater than the critical value of the chi-square distribution at the 5% significance level (i.e. 5.99). This indicates that separate car and truck following headway models are significantly different from the headway model estimated with all data combined. This process was also performed for the time gaps and the same results were found, but since the time gaps are so closely related to the headways, those results are not presented here. Additionally, a similar analysis comparing car and truck following headways/time gaps to modeling each pair type (CC, CT, TC, and TT) separately, and it found that each pair type should be modeled separately.

Equation 4. Likelihood ratio test statistic

$$\chi^2 = -2(LL_R - LL_U)$$

Where:

LL_R –log likelihood for the pooled model

LL_U –sum of the log likelihood values for separate models

Table 7. Comparison of combined model to separate models for headway

Location	**Combined Model LL**	**Car Model LL**	**Truck Model LL**	χ^2
I-80 at S Expressway (CB)	-553499	-465081	-73134	30568
I-80/35 NB at Hickman (DM)	-84689	-71532	-11758	2797
I-80/35 NB at University (DM)	-22970	-19791	-2725	909
I-80/35 SB at University (DM)	-39141	-33762	-4436	1886
I-74 at Middle Road (QC)	-6141	-5990	-124	53.8
I-80 west of Quad Cities (rural)	-2274	-1786	-416	142

Finally, headway and time gap data can be summarized, which is relatively simple at this point. The vehicle count and the mean, median, and standard deviation are calculated. These calculations are repeated for the overall filtered data, then for car and truck following separately, then for all four different vehicle pair types (car-car, car-truck, truck-car, and truck-truck). Since the vehicle counts in these groups were quite large (for some locations), even extremely small differences in headway or time gaps between groups could be found to be statistically significant. Therefore, this research favors comparisons using a practical significance threshold. For this study, a practical significance threshold of 0.1 seconds was used. This 0.1 second threshold was selected based on the observed headway measurement error.

CHAPTER V: RESULTS

Results Introduction

Once the headway, time gap, and standstill distance data were collected and validated, they were analyzed separately using different techniques. Both datasets were subject to some filtering to create as uniform of data as possible for comparing different locations. For the standstill distance, observations of greater than 25 feet were excluded. A description of the reasoning in selecting these filters is given in the Analysis Methodology section. For the headway and time gap data, only observations in through lanes with headways of four seconds or less and which occurred during flow rates of greater than 1,000 veh/hr/ln were used. Though there will still be some free vehicles in this dataset, this should limit their numbers and impact on the means significantly. Additionally, the focus of this study is to compare headways for different driver populations, not find an extremely accurate following headway distribution, so limiting the data in the same way for all cities should achieve this goal, even if some vehicles measured are not following.

Standstill distances were measured at 47 stop-and-go traffic incidents on urban freeways in Iowa across 7 cities and a variety of conditions in accordance with the methodology laid out in the Data Collection section. This resulted in 1238 observations, of which 693 were from Des Moines. This imbalance was due to the abundance of dynamic message signs and cameras in Des Moines, as well as its relatively high traffic volumes. In order to attempt to deal with this, data summaries were often split by the city they occurred in or the data was summarized within that city for the variable of interest.

Microsoft Excel and “R” were also used for the analysis of standstill distances. With more typical sample sizes that were less likely to lead to the false rejections seen in the headway and time gap data, t-tests were used to compare mean standstill distances. The rest of this section will lay out the results for the both the headway and time gap analysis, as well as the standstill distance analysis.

An initial exploration of the headway and time gap data indicated a difference based on the type of vehicle pair. Due to this and the inherent differences in car and truck behavior and capabilities, it was determined that summary statistics should be reported for each vehicle pair type. Since different sites may have different vehicle compositions, comparing average headways and time gaps for the different vehicle types (rather than the average for all vehicles) can control for these differences in truck percentages. Microsoft Excel was used for initial headway and time gap data exploration, but the statistical software “R” was used for the analysis. In order to determine what a large difference was and what was not between the means of different cities, a practical threshold of 0.1 seconds was used. This was used rather than statistically significant differences, because the sample sizes of some of the sites would lead to extremely small mean differences (hundredths or thousandths of a second) being rejected as unequal by t-tests. Such small differences do not have a large impact on traffic operations and could be the result of many things other than actual differences in population means.

Standstill Distance Results

Introduction to standstill distance results

Standstill distance measurements for urban freeways in Iowa were collected from stop-and-go traffic incidents across the state. The process involved finding potential incidents, reviewing video of them, taking screenshots when vehicles were stopped in the video, and measuring the distances between these stopped vehicles using Photoshop. For a more complete description of this process, see the Data Collection section. This data collection process precluded data from rural locations because the required infrastructure was not installed at many locations that could be considered rural and because stop-and-go traffic was not observed at those locations where the infrastructure was present. Additionally, it was decided that measurements of greater than 25 feet would be excluded because they were deemed to be outside of normal behavior based on observations of vehicles during the data collection process as well as observations of the histogram of the measurements.

Along with the actual standstill distance measurement, the vehicle pair type (CC, CT, TC, or TT) was recorded for each observation. Additionally, a number of other attributes for each incident were recorded for each observation: the city in which the incident took place, the lighting at the time, the weather at the time, whether a curve was present, and the cause of the incident. Having this additional information allowed for the exploration of the potential influence of these data on standstill distances. This section will present the summary statistics and relevant distributions of standstill distances for the different levels of the variables recorded.

Standstill distance by city

Due to the distribution of cameras, sensors, dynamic message signs, and traffic in the state of Iowa, the majority of stop-and-go incidents that were processed came from Des Moines. Des Moines is the largest city in the state, so naturally it has the most traffic and receives the most attention from IDOT. The attention from IDOT in this case means that Des Moines has more cameras, sensors, and dynamic message signs than other urban area in the state, which provided many opportunities to capture stop-and-go incidents. Additionally, the traffic load in Des Moines, especially during the peak hours, is large enough that even relatively small disturbances (e.g., a stalled vehicle on the shoulder) can be enough to cause stop-and-go conditions. These factors all led to a much larger number of measurements being observed for Des Moines (a total of 693) than other cities. However, some data were collected for Ames, Cedar Rapids, Council Bluffs, Iowa City, the Quad Cities, and Sioux City. The top two cities other than Des Moines were the Quad Cities, with 277 observations, and Sioux City, with 126 observations.

The mean, median, and standard deviations for the standstill distance measurements for each city are reported in Table 8. It is interesting that the means are generally around 10 to 12.5 feet, except for those in Des Moines and Iowa City, but Iowa City only has 11 observations. However, the data from Ames, Cedar Rapids, and Council Bluffs were also limited. Therefore, the statistical analysis focused on only the top three cities: Des Moines, the Quad Cities, and Sioux City. The histograms for these three cities are presented in Figure 33. Because these distributions are skewed, they needed to be transformed in order to use the t-test for difference of means. By taking the square root of the distance, these distributions become more symmetric (see Figure 34), so that t-tests can be used to compare them. The t-test comparisons resulted in the p-values reported in Table 9. A p-value of less than 0.05 means there is a statistically significant difference between the means with 95% confidence. All three t-tests were highly statistically significant, but more investigation is necessary to determine whether this difference is due to different driver populations or differences in the circumstances of the stop-and-go traffic collected in each of the cities.

Table 8. Summary of standstill distance by city

City	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Ames	50	1	13	11.57	9.66	5.71
Cedar Rapids	59	3	11	11.17	10.65	5.45
Council Bluffs	22	1	10	12.33	11.05	4.16
Des Moines	693	25	153	8.59	7.95	4.37
Iowa City	11	2	3	9.69	9.98	5.11
Quad Cities	277	8	74	10.19	9.51	4.36
Sioux City	126	6	33	12.53	12.00	4.81

Table 9. P-values for t-tests for mean standstill distance comparisons by city

	Des Moines	Quad Cities	Sioux City
Des Moines	xxx	2.22e-08***	< 2.2e-16***
Quad Cities	2.22e-08***	xxx	3.32e-06***
Sioux City	< 2.2e-16***	3.32e-06***	xxx

Note: * for 95% confidence, ** for 99%, *** for 99.9%

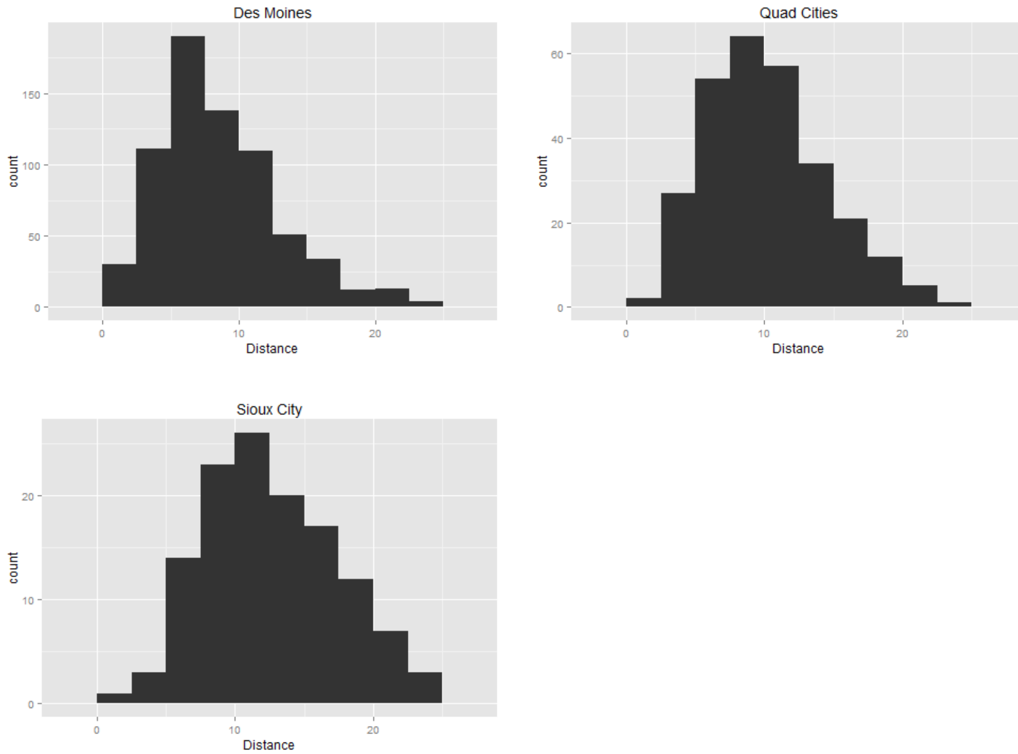

Figure 33. Standstill distance histograms by city

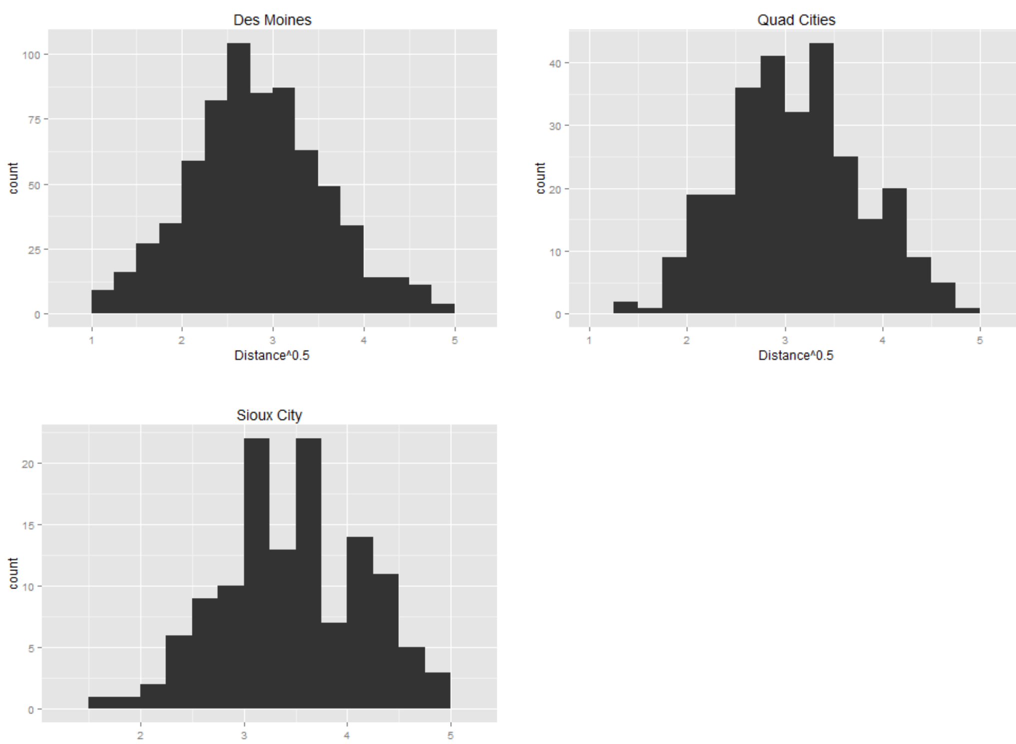

Figure 34. Square root of standstill distance histograms by city

If one set of conditions producing stop and go traffic is overrepresented in one city compared to the others, this can skew the results of the comparisons. For example, in Table 10, the number of incidents for each cause is reported for each city. It should be noted that all six incidents in Sioux City were the result of construction, whereas none of the incidents in Des Moines was the result of construction. To attempt to address this, the summary statistics for each cause type within each city are reported in Table 11 and Table 13 for Des Moines and the Quad Cities, respectively. These two cities were the only ones reported because they are the only ones with two or more cause types to compare and a sufficient number of observations. The p-values for the t-test comparisons for each incident cause combination in Des Moines and the Quad Cities are given in Table 12 and Table 14, respectively. In Des Moines, the only statistically significant result is the extremely marginally significant comparison (p-value of 0.0499) between slow traffic and unknown cause, which does not have a strong interpretation. The Quad Cities data, however, show a highly significant difference between construction and slow traffic and between construction and stalled vehicle.

Table 10. Number of incidents resulting from different causes by city

		No. of Incidents for each cause type				
City	**Total No. of Incidents**	**Accident**	**Construction**	**Slow Traffic**	**Stalled Traffic**	**Unknown**
Ames	1	1	0	0	0	0
Cedar Rapids	3	3	0	0	0	0
Council Bluffs	1	1	0	0	0	0
Des Moines	25	11	0	6	1	7
Iowa City	2	1	1	0	0	0
Quad Cities	8	0	3	4	1	0
Sioux City	6	0	6	0	0	0

Table 11. Summary of standstill distances for Des Moines by cause of incident

Cause	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Accident	252	11	56	8.63	8.00	4.41
Slow Traffic	162	6	37	8.12	7.66	4.28
Stalled Vehicle	32	1	12	7.72	7.31	3.29
Unknown	247	7	49	8.96	8.30	4.50

Table 12. P-values for t-tests for mean standstill distance comparisons by incident cause for Des Moines

	Accident	Slow Traffic	Stalled Vehicle	Unknown
Accident	xxx	0.238	0.285	0.377
Slow Traffic	0.238	xxx	0.754	0.0499*
Stalled Vehicle	0.285	0.754	xxx	0.120
Unknown	0.377	0.0499*	0.120	xxx

Note: * for 95% confidence, ** for 99%, *** for 99.9%

Table 13. Summary of standstill distances for the Quad Cities by cause of incident

Cause	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Construction	104	3	19	11.50	11.45	4.43
Slow Traffic	154	4	51	9.60	8.95	4.17
Stalled Vehicle	19	1	4	7.85	7.79	3.61

Table 14. P-values for t-tests for mean standstill distance comparisons by incident cause for the Quad Cities

	Construction	Slow Traffic	Stalled Vehicle
Construction	xxx	0.000565***	0.000812***
Slow Traffic	0.000565***	xxx	0.067
Stalled Vehicle	0.000812***	0.067	xxx

Note: * for 95% confidence, ** for 99%, *** for 99.9%

Standstill distance by vehicle pair type

The summary statistics for standstill distance for each vehicle pair type are presented in Table 15. An initial look at the summary statistics seems to indicate that vehicle pairs involving a truck tend to have larger standstill distances. To evaluate this impression statistically, t-tests were used. Though the smaller datasets for CT, TC, and TT make it difficult to assess the normality of their distributions, it can be seen in the overall distribution that the data are skewed and can be made more symmetric by taking the square root of the distances. Applying this procedure to the distances by vehicle pair type and conducting t-test comparisons yields the results presented in Table 16. The table shows that there is high confidence that CC standstill distances are significantly different than CT and TC standstill distances, but the sample size of TT pairs is too small to indicate that there is a difference between CC and TT. None of the other vehicle pair types were found to be significantly different.

Table 15. Summary of standstill distance by vehicle pair type

Pair Type (Lead-Follow)	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
CC	1140	45	287	9.41	8.80	4.54
CT	40	24	38	13.35	13.15	6.32
TC	48	26	41	12.37	11.20	5.78
TT	10	7	10	11.07	10.56	3.69

Table 16. P-values for t-tests for mean standstill distance comparisons by vehicle pair type

	CC	**CT**	**TC**	**TT**
CC	xxx	0.000421***	0.000861***	0.118
CT	0.000421***	xxx	0.515	0.276
TC	0.000861***	0.515	xxx	0.548
TT	0.118	0.276	0.548	xxx

Note: * for 95% confidence, ** for 99%, *** for 99.9%

These results were not likely to be overly influenced by the Des Moines data, as some of the other incident-based variables, because the data were fairly well spread out throughout the different cities, as seen in Table 17. Every city contributes something to each pair type, and Des Moines contributes fairly evenly across the different pair types. About 58 percent of CC observations are from Des Moines, and about 33 percent of CT and TC observations are from Des Moines. While these numbers are somewhat unbalanced, they are much more balanced than most of the incident-based variables described subsequently.

Table 17. Number of observations of each pair type in each city.

City	No. of CC observations	No. of CT observations	No. of TC observations	No. of TT observations	Total No. of Observations
Ames	42	2	3	3	50
Cedar Rapids	43	6	8	2	59
Council Bluffs	17	2	2	1	22
Des Moines	663	13	16	1	693
Iowa City	6	1	3	1	11
Quad Cities	264	6	6	1	277
Sioux City	105	10	10	1	126

Standstill distance by lighting

The summary statistics for standstill distance for each lighting condition (day or night) are shown in Table 18. While a t-test of the means shows that there is a statistically significant difference (p-value of 5.96e-05), there is not enough coverage within the data to make any conclusions with respect to the influence of lighting conditions. All of the night-time observations occurred in Des Moines as the result of one incident. This is not acceptable for analysis.

Table 18. Summary of standstill distance by lighting conditions

Lighting	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Day	1076	44	265	9.87	9.20	4.78
Night	159	1	31	8.33	7.44	4.21

Standstill distance by weather

The summary statistics for standstill distance in different weather conditions are provided in Table 19. The weather conditions were observed from the video of each incident. Due to the small sample size and the lack of a theoretical basis for a difference in standstill distance between clear and cloudy conditions, the data for these conditions were combined and compared to the rainy condition data using a t-test. Again, this t-test was conducted on the square roots of the measurements. This test resulted in a p-value of

0.0884, which means that there is not a statistically significant difference with a minimum confidence level of 95%. Even if the distance had been found to be significantly significant, it would have to be taken with a grain of salt, again because of the sample size and coverage. All three incidents involving rain occurred in Des Moines.

Table 19. Summary of standstill distance by weather conditions

Weather	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Clear/Cloudy	1202	43	287	9.70	9.02	4.75
Rainy	36	3	10	8.40	7.69	3.77

Standstill distance by curve presence

The summary statistics for standstill distance when a curve was or was not present are provided in Table 20. The presence of a curve was noted from watching the video of the incidents. A t-test comparison barely did not show a statistically significant difference at the 95% confidence level (p-value of 0.0564), and there was fairly good coverage of the data in this case, which lends support to this conclusion. The incidents for which data were recorded on a curve were found in Des Moines, Cedar Rapids, and the Quad Cities, and the causes of these incidents included an accident, slow traffic, construction, and an unknown cause.

Table 20. Summary of standstill distance by curve presence

Curve	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
No	915	40	214	9.83	9.15	4.83
Yes	323	7	83	9.19	8.57	4.41

Standstill distance by cause

The causes of the stop-and-go incidents were ascertained by noting the message displayed on the dynamic message board sign, which indicated the cause that led to the video being downloaded. For example, if a sign said ACCIDENT AHEAD PREPARE

TO STOP, that incident was coded as being caused by an accident. There is clearly some ambiguity involved in this method, because a sign warning of slow traffic does not necessarily mean that the incident was not caused by an accident. The message could simply mean that the traffic management center was unaware of the cause or that the message was displayed automatically because the IDOT detectors recorded the speed dropping below a certain threshold. Despite this ambiguity, this method was the best option because video rarely showed what caused the stop-and-go conditions directly.

The summary statistics for standstill distance for each cause are reported in Table 21. The summary statistics indicate that incidents caused by a stalled vehicle tended to have the smallest standstill distances, and incidents caused by construction tended to have the highest. While the stalled vehicle category has the fewest observations, its two incidents are from different cities (Des Moines and the Quad Cities) and have a similar number of observations (32 in Des Moines and 19 in the Quad Cities). The means of the standstill distances are 7.72 feet for the Des Moines incident and 7.85 for the Quad Cities incident (see Table 26). When this is tested using a t-test in a similar fashion to the preceding analyses, the p-value is 0.891, indicating no statistically significant difference. This does not completely validate the result, because the sample size is small and there could be interacting factors in both the observed and unobserved information about each site leading to this result, but the consistency between the two sites lends support to the conclusion that stalled vehicle incidents tend to have lower standstill distances.

Table 21.Summary of standstill distance by cause

Cause	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Accident	392	17	92	9.61	8.94	4.95
Construction	232	10	53	12.05	11.50	4.64
Slow Traffic	316	10	88	8.84	8.16	4.28
Stalled Vehicle	51	2	16	7.77	7.50	3.38
Unknown	247	7	49	8.96	8.30	4.50

The t-test p-values for each combination of incident type are reported below in Table 22. However, caution should be used when interpreting these results due to the unbalanced nature of the data. Because Des Moines yielded so many more observations than the other cities and appears to have consistently lower standstill distances than other cities, the influence on the means of other variables can be large.

Table 22. P-values for t-tests for mean standstill distance comparisons by the cause of the incident

	Accident	Construction	Slow Traffic	Stalled Vehicle	Unknown
Accident	xxx	3.60e-11***	0.0534	0.00465**	0.122
Construction	3.60e-11***	xxx	< 2.2e-16***	4.24e-10***	6.358e-14***
Slow Traffic	0.0534	< 2.2e-16***	xxx	0.0859	0.820
Stalled Vehicle	0.00465**	4.24e-10***	0.0859	xxx	0.0708
Unknown	0.122	6.358e-14***	0.820	0.0708	xxx

Note: * for 95% confidence, ** for 99%, *** for 99.9%

In general, there was a good spread of locations for each of the incident causes. For a simple breakdown of the locations of incidents for each cause, refer to Table 10. Every cause was present in at least two cities, except for the unknown cause, which was only present in Des Moines. To further break down the data, it is possible to look at the variation between cities for each specific incident type; these results are presented in Table 23 to Table 26 (unknown cause was not included because all incidents occurred in Des Moines, so there was no variation between cities for this cause). The most noticeable results that these tables show are that the means for Des Moines are typically lower than for other cities and that there is generally variation between cities for each cause.

Table 23. Summary of standstill distances for incidents caused by accidents for each city

City	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Ames	50	1	13	11.57	9.66	5.71
Cedar Rapids	59	3	11	11.17	10.65	5.45
Council Bluffs	22	1	10	12.33	11.05	4.16
Des Moines	252	11	56	8.63	8.00	4.41
Iowa City	9	1	2	9.40	8.08	5.66

Table 24. Summary of standstill distances for incidents caused by construction for each city

City	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Iowa City	2	1	1	11.02	11.02	0.71
Quad Cities	104	3	19	11.50	11.45	4.43
Sioux City	126	6	33	12.53	12.00	4.81

Table 25. Summary of standstill distances for incidents caused by slow traffic for each city

City	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Des Moines	162	6	37	8.12	7.66	4.28
Quad Cities	154	4	51	9.60	8.95	4.17

Table 26. Summary of standstill distances for incidents caused by a stalled vehicle for each city

City	Count	No. of Incidents	No. of Photos	Mean (ft)	Median (ft)	Std. Dev. (ft)
Des Moines	32	1	12	7.72	7.31	3.29
Quad Cities	19	1	4	7.85	7.79	3.61

Standstill distance results conclusion

Overall, it appears that mean standstill distance throughout Iowa is generally between 8 and 12 feet. Due to the way the data were collected, more than half of the observations were from Des Moines. Additionally, it appears that Des Moines had significantly lower standstill distances than the other cities. These two factors, along with some other imbalances in the data, made it difficult to assess the effects of other incident-based characteristics such as the cause of incident, the weather at the time, etc. The vehicle pair types were spread among the incidents and the locations fairly well, so the mean standstill distances could be reasonably tested for each of the vehicle pair types. It was found that the CC pair type had a significantly lower mean than the CT and TC pair types, while the data were not sufficient to reach the same conclusion for the TT pair type.

Headway and Time Gap Results

Introduction to headway and time gap results

Due to the data validation process as well as the filtering process, the data from several detectors were excluded from the analysis. Regarding data validation, both Wavetronix detectors placed on I-235 in Des Moines counted around half as many vehicles as were counted manually, so the data obtained from these detectors were not used. Additionally, it was determined that in order to filter the data to mostly following vehicles, only through lane headways of four seconds or less and headways observed during a period of time when the 15 minute through vehicle flow rate exceeded 1,000 veh/hr/ln would be used. This filtering completely eliminated the detector on I-74 at the Spruce Hills Drive exit in the Quad Cities because it was not operating for long and the flow rate never exceeded 1,000 veh/hr/ln. Limited high-traffic intervals also significantly reduced the sample size of the other two locations in the Quad Cities area (I-74 at the Middle Road exit and the rural location on I-80 west of the Quad Cities). Finally, for consistency's sake, the same data set was used for analyzing the headways and time gaps for each detector; this required that both the leading and following vehicle had a speed measured by the detector. For most detectors, at least 60 to 70 percent of vehicles had speeds, but the detector on I-80/35 southbound at the Hickman Road entrance only had speeds for 17 percent of vehicles (and only 2 percent of through vehicles). The percentages of vehicles with speeds and the detectors that were excluded are summarized in Table 27. While it would have been ideal to have multiple detectors in each city to check for consistent headway values within the same driving population, the locations in Des Moines and the two directions in Council Bluffs were used to check this assumption before comparing the different cities.

Table 27. Summary of detectors used in analysis

Detector	City	% vehicles with speed	Data used in analysis?	Reason for Exclusion
I-80 S Expressway	Council Bluffs	81.7	Yes	N/A
I-80/35 SB @ Hickman	Des Moines	16.9	No	% vehicles with speed (sample size)
I-80/35 SB @ University	Des Moines	89.5	Yes	N/A
I-80/35 NB @ Hickman	Des Moines	73.2	Yes	N/A
I-80/35 NB @ University	Des Moines	95.8	Yes	N/A
I-235 EB @ 73rd	Des Moines	N/A	No	Validation counts
I-235 WB @ 73rd	Des Moines	N/A	No	Validation counts
I-74 @ Middle Road	Quad Cities	81.0	Yes	N/A
I-74 @ Spruce Hills	Quad Cities	90.4	No	No flow rates > 1000 veh/hr/ln
I-80 West of Quad Cities	Quad Cities	62.8	Yes	N/A

Additionally, lower flow rate data excluded from the actual calculation of the summary statistics for headways and time gaps were still used to compare how average headways varied as the flow rate changed for the different detectors, as outlined in the methodology chapter above. This comparison showed that the average headways and time gaps measured by different detectors in different parts of the state were consistently similar for similar flow rates. Figure 35 and Figure 36 show this consistency: there is not one detector or one city for which flow rate was consistently higher or lower than that of any other (other than the rural location, which was expected to be different), and all were tightly grouped for almost all flow rates. These figures also demonstrate that headway (and, to a lesser extent, time gap) starts to level off at flow rates greater than 1,000 veh/hr/ln. This leveling off is even more pronounced at higher flow rates (approximately

1,300 veh/hr/ln), but this higher threshold would exclude even more sites and data, so 1,000 veh/hr/ln was used as the threshold for "congested" traffic. Similar graphs were produced for each vehicle pair type (e.g., car-truck), and the graphs show similar trends, but the graphs were shifted up or down based on the vehicle pair type.

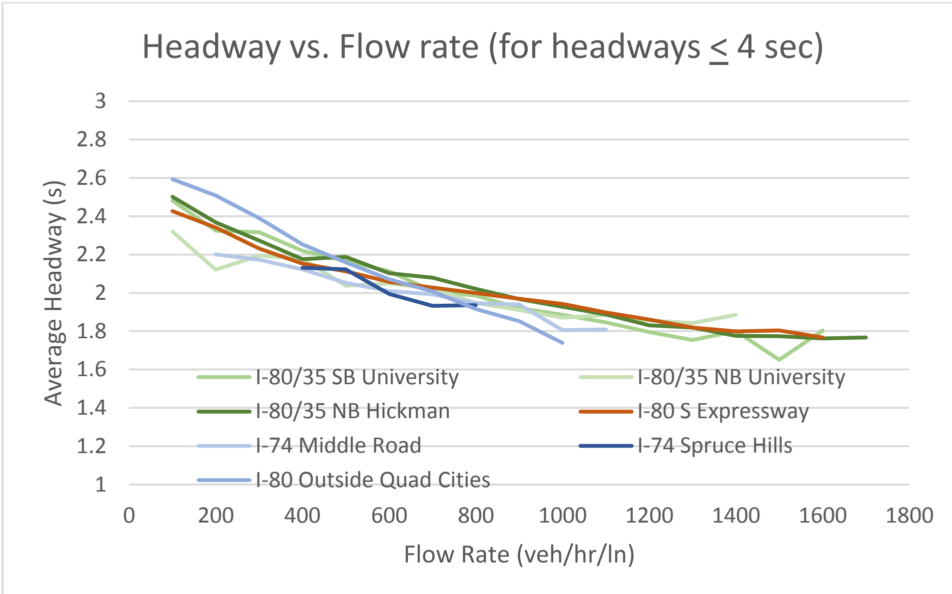

Figure 35. Average headway versus flow rates for headways less than 4 seconds

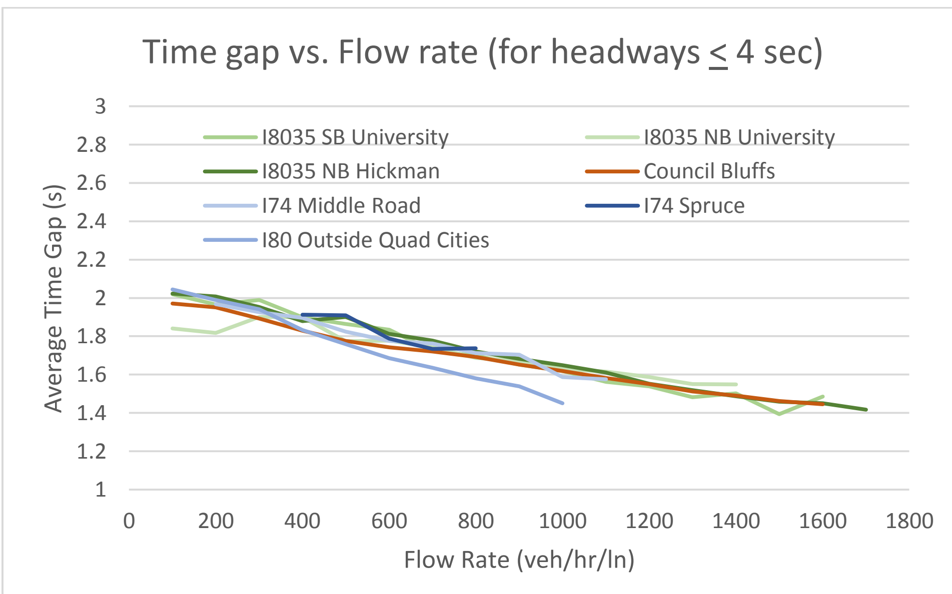

Figure 36. Average time gap versus flow rates for headways less than 4 seconds

Des Moines headway/time gap results

Three detectors from Des Moines produced data that could be used to analyze headways and time gaps: I-80/35 southbound at the University Avenue exit, I-80/35 northbound at the University Avenue entrance, and I-80/35 northbound at the Hickman Road exit (see Figure 5). These locations experienced the same driver population (mostly commuters to and from the center of Des Moines). Though the roadway segments were similar geometrically (all had three through lanes and one weaving/merge lane), the detector locations were different: two were located at exit ramps and one was at an entrance ramp. Additionally, in the southbound direction there was a weaving lane, whereas in the northbound direction the University Avenue entrance merged into four lanes and the fourth lane became an exit-only lane for the Hickman Road exit.

The means, medians, and standard deviations for each pair type for each detector in Des Moines are reported in Table 28 to Table 30. When comparing the means, t-tests were not used because the large sample sizes led to a situation where differences in headways that are within the error of the detector were rejected by the t-test as being significantly different. For example, the mean difference of headways from car leading car pairs (CC) between I-80/35 southbound at University Avenue (in Table 28) and I-80/35 northbound at University Avenue (in Table 30) is only 3 hundredths of a second, but the sample size leads the t-test to conclude that they are different, with a p-value of 2.586×10^{-5}. This is not always the case; because observations involving trucks are less frequent, this phenomena of “false rejections” is less common outside the CC observations. For the sake of consistency, however, all comparisons were made assuming that a practical difference of headway or time gap is one tenth (0.1) of a second. The selection of the 0.1 second value was based on the measurement error of the Wavetronix detector. While it was not possible to validate the accuracy of the time stamp for individual vehicles, the average of individual headways was found to differ by 0.03 seconds from a manual measurement for a 10 minute peak period at the I-80/35 northbound location at Hickman Road. Because this was a relatively short period, the true error of the detector could be higher or lower than this; to be conservative, this error was assumed to be 0.05 seconds. Therefore, two detectors that were off by that error in opposite directions would lead to a difference of 0.1 without there being a true difference in means.

For the mean headway, only 2 comparisons out of the 12 possible were outside 0.1 seconds: the average for truck leading truck pairs (TT) for I-80/35 southbound at University Avenue was 2.36 seconds, while the average for each of the other two sites was 2.25 seconds, a difference of 0.11 seconds, which was just outside the established threshold. For the median headway, 2 out of the 12 comparisons were outside 0.1 seconds: differences of 0.16 seconds for CC and 0.13 seconds for TT between I-80/35 southbound at University Avenue and I-80/35 northbound at University Avenue. The mean time gaps were even more consistent, with only one comparison outside the range: a difference of 0.11 seconds for car leading truck pairs (CT) between I-80/35 northbound at University Avenue and I-80/35 northbound at Hickman Road. The median time gaps were the worst, with 4 out of the 12 comparisons outside of 0.1 seconds: 0.17 seconds for CC and TT between I-80/35 southbound at University Avenue and I-80/35 northbound at University Avenue and 0.12 seconds for CC and 0.11 seconds for truck leading car (TC) between I-80/35 southbound at University Avenue and I-80/35 northbound at Hickman Road.

Table 28. Headway and time gaps summary statistics for I-80/35 SB at University

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	23472	47	1.72	1.58	0.92	1.51	1.31	0.92
CT	2832	47	2.30	2.14	0.85	2.09	1.95	0.85
TC	3278	47	1.82	1.73	0.88	1.09	0.88	0.88
TT	636	47	2.36	2.18	0.84	1.59	1.50	0.85

Table 29. Headway and time gaps summary statistics for I-80/35 NB at University

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	13069	26	1.79	1.74	0.94	1.58	1.48	0.94
CT	1678	26	2.34	2.18	0.87	2.13	1.97	0.87
TC	1920	26	1.86	1.78	0.89	1.12	0.91	0.88
TT	383	26	2.25	2.05	0.85	1.49	1.33	0.85

Table 30. Headway and time gaps summary statistics for I-80/35 NB at Hickman

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	46886	123	1.75	1.67	0.92	1.53	1.43	0.92
CT	6797	123	2.24	2.08	0.91	2.02	1.87	0.91
TC	8814	123	1.85	1.77	0.88	1.16	0.99	0.89
TT	1951	123	2.25	2.09	0.87	1.52	1.40	0.87

In all, 9 out of 48 comparisons fell outside 0.1 seconds of each other, and most of those were not far out of that range, which is summarized in Table 31. The maximum difference between any two values was only 0.17 seconds. The closeness of these summary statistics indicates that there are not practical differences in headway and time gap values at these sites with the same driver populations. This means that the preferred headway and time gap should be fairly consistent throughout Des Moines, unless there are significantly different roadway geometries or other factors. The summary statistics of the three Des Moines locations combined are presented in Table 32.

Table 31. Number of differences greater than 0.1 between summary statistics for Des Moines sites

	Number of differences greater than 0.1 sec (out of 3 possible)			
Pair Type (Lead-Follow)	**Mean Headway**	**Median Headway**	**Mean Time Gap**	**Median Time Gap**
CC	0	1	0	2
CT	0	0	1	0
TC	0	0	0	1
TT	2	1	0	1

Table 32. Summary statistics for headway and time gap data for Des Moines overall

Pair Type (Lead-Follow)	**Count**	**No. of Non Free-Flow Intervals**	**Mean Headway (s)**	**Median Headway (s)**	**Std. Dev. Headway (s)**	**Mean Time Gap (s)**	**Median Time Gap (s)**	**Std. Dev. Time Gap (s)**
CC	83427	170	1.74	1.66	0.92	1.53	1.42	0.92
CT	11307	170	2.27	2.10	0.89	2.05	1.91	0.89
TC	14012	170	1.85	1.77	0.88	1.14	0.95	0.88
TT	2970	170	2.27	2.11	0.86	1.53	1.40	0.87

Council Bluffs headway and time gap results

I-80 at the South Expressway entrance in Council Bluffs was the one location where an existing IDOT-owned sensor was used to collect the data. This allowed for the collection period to last much longer than it did for the other sites and allowed for an even larger sample size. The detector was recording data off and on for six weeks and detected over 2.5 million vehicles in total. It recorded data for both the eastbound and westbound directions. The eastbound direction has three through lanes and one auxiliary entrance ramp lane. The westbound direction has two through lanes and an exit ramp. Having the data for both directions allowed for the opportunity to support the finding in Des Moines that the observed headway and time gap values do not vary within the same driver population despite differences in geometry. The summary statistics for the eastbound and westbound directions are reported in Table 33 and Table 34, respectively. It is important to note that there were only 32 intervals when the 15 minute flow rate exceeded 1,000 veh/hr/ln for the eastbound direction compared to about 1,200 intervals for the westbound direction. This is mainly due to the fact that there are three through lanes eastbound and two through lanes westbound.

Despite this difference in traffic operation and sample size, the summary statistics are quite similar between the two directions. The differences in measurements were less than 0.1 seconds for the mean and median headway and time gap values for CC, TC, and TT. For CT, the mean headway was off by 0.14, the median headway was off by 0.38, the mean time gap was off by 0.16, and the median time gap was off by 0.38. It is interesting that the means are not much different than the 0.1, second threshold while the medians are both off by 0.38 seconds (lower for the westbound traffic). Overall, it appears that there is enough evidence to support the finding in Des Moines that headway and time gap values do not vary much when considering the same driver population faced with different geometries. This is especially true if the focus is narrowed to only mean values, because the median values have been shown to be more volatile when comparing sites.

Table 33. Summary statistics for headway and time gap data for I-80 at S Expressway EB

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	11315	32	1.82	1.75	0.91	1.63	1.55	0.91
CT	1095	32	2.52	2.60	0.83	2.33	2.40	0.83
TC	1655	32	1.90	1.80	0.86	1.11	0.93	0.85
TT	243	32	2.45	2.33	0.82	1.61	1.54	0.79

Table 34. Summary statistics for headway and time gap data for I-80 at S Expressway WB

Pair Type	Count	No. of	Mean	Median	Std. Dev.	Mean	Median	Std. Dev.

(Lead-Follow)		Non Free-Flow Intervals	Headway (s)	Headway (s)	Headway (s)	Time Gap (s)	Time Gap (s)	Time Gap (s)
CC	294480	1206	1.76	1.68	0.91	1.54	1.45	0.91
CT	40439	1206	2.38	2.22	0.86	2.17	2.02	0.86
TC	59207	1206	1.89	1.80	0.86	1.13	0.97	0.86
TT	14205	1189	2.44	2.24	0.80	1.63	1.51	0.82

The combined data for I-80 at S Expressway is presented in Table 35. Because the westbound traffic had so many more observations meeting the filtering criteria, the overall statistics are essentially the same as the westbound statistics. These values are similar to the values found in Des Moines, though a few measurement differences were outside of 0.1 seconds. For CT, the mean headway difference was 0.12 seconds, the median headway difference was 0.12 seconds, the mean time gap difference was 0.13 seconds, and the median time gap difference was 0.11 seconds. For TT, the mean headway difference was 0.17 seconds, the median headway difference was 0.13 seconds, and the median time gap difference was 0.11 seconds. While it seems like a lot of measurements are off, they are not substantially more than the threshold for the most part, and the biggest difference is still only 0.17 seconds.

Table 35. Summary statistics for headway and time gap data for I-80 at S Expressway overall

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	305795	1208	1.76	1.68	0.91	1.55	1.45	0.91
CT	41534	1208	2.39	2.22	0.86	2.18	2.02	0.86
TC	60862	1208	1.89	1.80	0.86	1.13	0.97	0.86
TT	14448	1194	2.44	2.24	0.80	1.63	1.51	0.82

Quad Cities headway and time gap results

In the Quad Cities, two temporary Wavetronix detectors were set up on an urban freeway (I-74 at Middle Road and I-74 at Spruce Hills Drive), and one temporary Wavetronix detector was set up on a rural freeway (I-80 west of the Quad Cities). All detectors recorded traffic flowing in both directions on and off for about two weeks. The I-74 Spruce Hills Drive location only functioned for about 8 hours and did not experience any 15 minute flow rates greater than 1,000 vehicles per hour. The I-74 Middle Road location only observed one 15 minute flow rate greater than 1,000 vehicles per hour in the southbound direction, compared to 16 intervals northbound. Due to these data collection limitations, the consistency found within driver populations in Des Moines and Council Bluffs could not be confirmed with data in the Quad Cities. The consistency was therefore assumed to hold true, and the data collected from the I-74 Middle Road location were deemed representative of the Quad Cities. Geometrically, the I-74 Middle Road location had two through lanes in both directions and an entrance ramp (northbound) and exit ramp (southbound) that did not have auxiliary lanes associated with them.

The summary statistics for the detector at I-74 at Middle Road are presented in Table 36. It should be noted that only 10 instances of TT pairs meeting the filtering criteria were observed by this detector, which is not a large enough sample size to judge its similarity to the other locations. However, it is still included in the table for the sake of consistency and completeness. The numbers of CT and TC pairs observed were also fairly low but were still substantial enough to get an idea of the true measurements. In order to be 95% confident that the true mean is within $\pm$0.2 seconds of the estimated mean, at least 78 observations are necessary, according to the sample size formula based on the normal distribution (Equation 5) and using a standard deviation of 0.9, which was observed at this site and others. While the headway and time gap distributions were not normal (they were somewhat skewed), this equation gives a low-end approximation. Though it would have been helpful to have more confidence in a narrower margin (such as the 0.1 second threshold), these observations provided a decent estimate of the mean. Despite this small sample size, there were still only a few measurements outside of the 0.1 second threshold. For the TC pairs, between the I-74 Middle Road detector and Council Bluffs detector the mean time gap difference was 0.19 seconds, and the median time gap difference was 0.17 seconds. Between the I-74 Middle Road detector and the Des Moines detectors, the mean time gap difference was 0.18 seconds, and the median time gap difference was 0.19 seconds. All of these differences were due to the measurements at the I-74 Middle Road location being consistently higher than the other two locations.

Equation 5. Sample size estimation formula

$$n = \left(\frac{Z_{\frac{\alpha}{2}} * \sigma}{E}\right)^2$$

Where:

n – number of observations needed

$Z_{\frac{\alpha}{2}}$ – the critical z-score for a significance level of $\frac{\alpha}{2}$

σ – sample standard deviation

E – acceptable error

This consistent difference in time gaps but not headways could be due to various reasons. First, it could just be a result of the small sample size. However, none of the other measurements differed much from the other two locations, and they were consistently off by about the same amount, so this does not appear to be the most likely explanation. Another potential explanation for the skewed results is that drivers may have tended to maintain a consistent headway rather than a consistent time gap and the average truck was shorter in the Middle Road data than at the other two sites. This would lead to the similar headway values observed, as well as the larger time gaps. When the average truck length was calculated at each of the locations, it was found that the average truck length in Des Moines and Council Bluffs was 61.8 feet, and the average length at the Middle Road location was 52.5 feet. This lends credence to the theory that drivers maintain consistent headway rather than a consistent time gap, and different vehicle compositions can therefore affect the average time gap value. Though the data support this theory, simulation software and numerous past studies have found that vehicles maintain a consistent time gap rather than headway, so this theory is not likely either. The most likely explanation is that there is a higher percentage of small trucks on I-74 compared to the locations in Des Moines and Council Bluffs. Cars may interact with these smaller trucks differently than they interact with 18 wheelers. If cars do interact differently with smaller trucks and these vehicles are present in different proportions, this could affect the overall average, because both smaller and larger trucks are considered trucks by this analysis. It appears that this could be the case, as evidenced by the histograms of vehicle length, in which the Quad Cities location clearly has more small trucks than the other locations (see Figure 37). Finally, the consistent difference in time gaps but not headways could just be an anomaly in the data collection process. Without more sites to investigate, it is not possible to make a strong assertion about the cause of this difference. Despite this difference in time gaps for TC, overall the data from the I-74

Middle Road detector in the Quad Cities matched the data from Des Moines and Council Bluffs fairly well.

Table 36. Summary statistics for headway and time gap data for I-74 at Middle Road

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	4508	16	1.79	1.72	0.92	1.58	1.50	0.92
CT	87	16	2.36	2.16	0.88	2.14	1.96	0.88
TC	143	16	1.93	1.84	0.87	1.32	1.14	0.89
TT	10	5	2.26	2.24	0.49	1.70	1.75	0.40

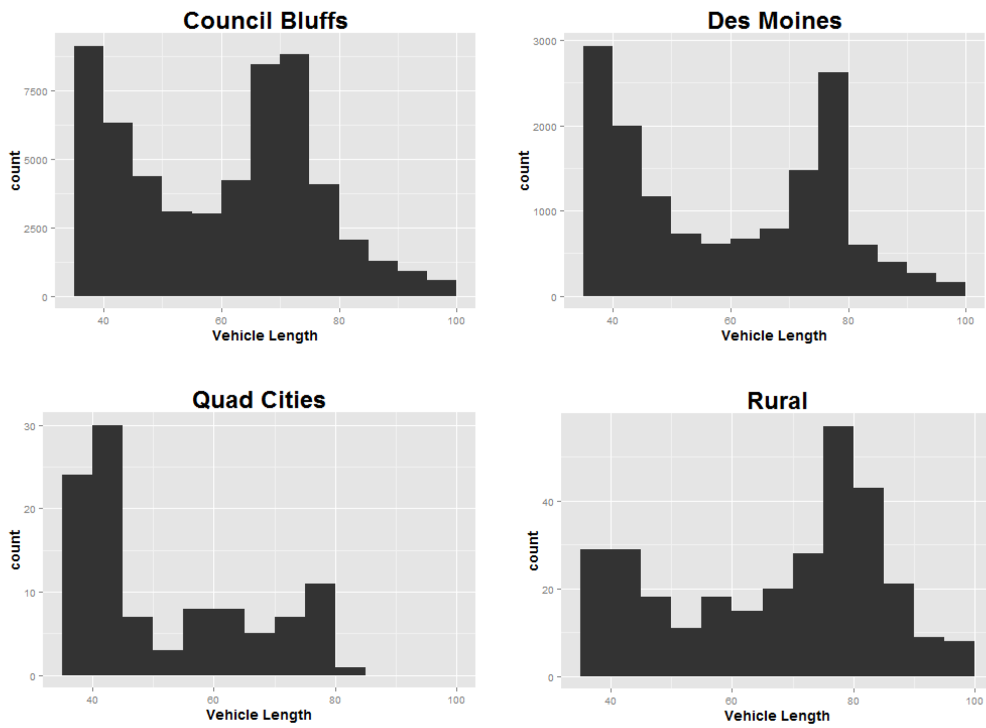

Figure 37. Truck length histograms for the three urban areas and one rural location

Rural Quad Cities headway and time gap results

Finally, a temporary Wavetronix detector was set up on I-80 approximately two miles west the corporate limits of the Quad Cities and the I-280 bypass interchange. Using the HCM definition of interchange density (the total number of interchanges three miles upstream or downstream of the location divided by total miles), the interchange density at this location is 0.33 interchanges per mile (HCM 2010). There is one interchange approximately three miles west of the location, and the I-280 interchange is about two miles east. This limits the interaction between entering and exiting traffic that is characteristic of urban freeways. There are two through lanes in each direction and obviously no entrance or exit lanes. The driver population is also somewhat different even from the I-74 Middle Road location nearby, because the I-74 Middle Road location is in the heart of town and is likely mostly commuters, whereas the rural location is likely to be passenger and freight vehicles making through trips. This difference in driver population is supported by only 3.4 percent of all vehicle detections at the I-74 Middle Road location being trucks, while that same value is 34.7 percent for the rural location. Even when limiting the scope to only the data which was used in the analysis (flow rates of greater than 1000 veh/hr/ln and headways of less than 4 sec), the truck percentages were still 2.0 percent and 18.6 percent for I-74 Middle Road location and the rural location, respectively. The truck percentage in the Des Moines data was 12.8 and in the Council Bluffs data was 13.2.

The reason for the inclusion of this site is to offer a point of comparison to evaluate the potential impacts of a rural freeway setting, since all the other locations were on urban freeways. The original intention of this research was to obtain more rural locations to compare with the urban locations, but, as was mentioned in the data collection portion, this was not possible due to a number of data collection setbacks in conjunction with time constraints. Since this data represents only one rural location, consistency among rural locations cannot be established as it was with the urban locations. Therefore, the results presented for this detector should not be generalized to represent all rural locations – it is simply a point of comparison which could be used as a starting point for future research.

The summary statistics for the rural location on I-80 west of the Quad Cities are presented in Table 37. These values appear to differ substantially from the values at the three urban locations. There are too many differences to point out individually in the text, so comparison are summarized in Table 38 to

Table 40 below. Most of the largest differences are in the CC and CT groups where the rural headway means were as much as 0.34 sec lower than the corresponding values for the urban locations and the time gap means were as much as 0.3 sec lower.

Table 37. Summary statistics for headway and time gap data for I-80 west of the Quad Cities

Pair Type (Lead-Follow)	Count	No. of Non Free-Flow Intervals	Mean Headway (s)	Median Headway (s)	Std. Dev. Headway (s)	Mean Time Gap (s)	Median Time Gap (s)	Std. Dev. Time Gap (s)
CC	1139	10	1.56	1.17	0.90	1.40	1.02	0.90
CT	224	10	2.04	1.94	0.87	1.88	1.77	0.87
TC	293	10	1.88	1.81	0.92	1.21	1.05	0.92
TT	104	10	2.33	2.14	0.84	1.62	1.47	0.84

Table 38. Summary of differences between the rural and Des Moines (rural minus Des Moines)

Pair Type (Lead-Follow)	Mean Headway (s)	Median Headway (s)	Mean Time Gap (s)	Median Time Gap (s)
CC	-0.18	-0.49	-0.13	-0.40
CT	-0.22	-0.17	-0.18	-0.13
TC	0.03	0.05	0.07	0.10
TT	0.06	0.03	0.09	0.07

Table 39. Summary of differences between the rural and Council Bluffs (rural minus Council Bluffs)

Pair Type (Lead-Follow)	Mean Headway (s)	Median Headway (s)	Mean Time Gap (s)	Median Time Gap (s)
CC	-0.20	-0.51	-0.15	-0.44
CT	-0.34	-0.29	-0.30	-0.25
TC	-0.01	0.01	0.08	0.08
TT	-0.11	-0.11	-0.01	-0.04

Table 40. Summary of differences between the rural and I-74 Middle Road (rural minus I-74 Middle Road)

Pair Type (Lead-Follow)	Mean Headway (s)	Median Headway (s)	Mean Time Gap (s)	Median Time Gap (s)
CC	-0.23	-0.55	-0.18	-0.48
CT	-0.32	-0.22	-0.26	-0.19
TC	-0.05	-0.03	-0.11	-0.09
TT	0.07	-0.10	-0.08	-0.28

This is an interesting result and one which could be due to a number of things. The most likely possibility is that people drive differently in a rural setting without the influence of entering or exiting traffic. Without those considerations, drivers are free to follow at closer distances. Another potential cause could be that there are bigger differences between drivers' desired speeds on rural freeways compared to urban freeway leading to more aggressive following when a slow vehicle is encountered. From personal experience on rural freeways, this can often occur when cars are passing trucks at too slow of a speed for the preferences of those behind them. With the higher truck percentages on rural freeways, this type of passing would occur more frequently than in urban settings. Additionally, in urban settings the drivers are more likely to anticipate heavy traffic and could be more willing to accept a following role, rather than looking to pass. Of course, it could just be a result of a fairly small sample size, as there were only ten intervals during which the 15 minute flow rate exceeded 1000 veh/hr/ln. If there was something unique or unusual about a few of those intervals, those intervals could have substantial affect the overall measurements. These are all just possible explanations, and without other locations, it is difficult to discern if there is a true difference between rural and urban locations.

Alternative to practical significance

There is another method of addressing the large sample sizes other than using a practical significance threshold. Effect size statistics are a group of statistics which adjust more traditional sample statistics (t-tests, etc.) based on their sample size to achieve comparable differences. One of the most commonly used effect size statistics is Cohen's *d* which is calculated from Equation 6. Commonly used thresholds for the magnitude of the effect size based on *d* are: 0 to 0.2 is negligible, 0.2 to 0.5 is small, 0.5 to 0.8 is medium, 0.8 to 1.3 is large, and greater than 1.3 is very large. These thresholds are fairly arbitrary, but have become widely accepted in interpreting effect sizes.

Equation 6. Cohen's *d* statistic

$$d = t\sqrt{\frac{n_1 + n_2}{n_1 n_2}}$$

Where:

d – Cohen's d

t – value of t-test statistic

n_1 – sample size of group 1

n_2 – sample size of group 2

The Cohen's *d* was calculated for the comparison of mean headways and time gaps for the different vehicle pair types at one site (see Table 41 and Table 42). Additionally, the Cohen's *d* was used to compare the mean headways and time gaps across two different sites for corresponding vehicles types (see Table 43). These tables indicate the same things that using the practical significance indicated. First, Table 41 shows that the difference in headways in CC and TC pairs (pairs where a car is following) is negligible. Table 41 also shows the same for CT and TT pairs (pairs where a truck is following). When comparing the time gaps in Table 42, it can be seen the difference is negligible between vehicle pairs where both are the same vehicle (CC and TT. Finally, Table 43 shows that the difference between the headway and time gap means is negligible between the NB and SB directions of I-80/35 at University.

Table 41. Cohen's *d* values for headways at I-80/35 SB at University

	CC	**CT**	**TC**	**TT**
CC	xxx	-0.68	-0.12	-0.76
CT	0.68	xxx	0.55	-0.07
TC	0.12	-0.55	xxx	-0.63
TT	0.76	0.07	0.63	xxx

Table 42. Cohen's *d* values for time gaps at I-80/35 SB at University

	CC	**CT**	**TC**	**TT**
CC	xxx	-0.68	0.48	-0.09
CT	0.68	xxx	1.16	0.60
TC	-0.48	-1.16	xxx	-0.58
TT	0.09	-0.60	0.58	xxx

Table 43. Cohen's *d* for comparing b/w I-80/35 SB and NB at University

	Headway d statistic	**Time gap d statistic**
CC	0.08	0.07
CT	0.05	0.04
TC	0.04	-0.03
TT	-0.13	-0.11

The fact that the conclusions drawn from the effect sizes are in agreement with the conclusions drawn from the practical significance threshold supports the use of a practical significance threshold. However, effect sizes were not used throughout this analysis for a few reasons. First, they are harder to explain and interpret than a practical significance threshold. Most people have not heard of effect size statistics, but most people can grasp that 0.1 sec is a small measure of time. Second, the accepted thresholds used for interpreting effect sizes are arbitrary and somewhat controversial, whereas the 0.1 sec threshold was derived directly from the measurement error observed in this research. Additionally, the effect size statistic can only be used to compare means, while the practical significance threshold can also be used to compare medians (which can be helpful, since the headway time gap distributions are skewed). Finally, it is the opinion of this researcher that when two methods provide similar conclusions, the simpler method should be favored. So for those reasons, it was determined that the practical significance threshold would be used for the analysis of headways and time gaps.

Comparison of predicted and observed capacity

In order to evaluate the effectiveness of this method of filtering data and measuring the average headway in predicting actual traffic operations, the capacity predicted from the mean headway observed was compared to the estimated capacity of the respective facility. The predicted capacity was estimated from Equation 7, which is derived from the definition of the inverse relationship between headway and flow rate. Typically, this equation overestimates capacity, because flow breakdown usually occurs before every vehicle is at its preferred following headway, so it is more of a theoretical maximum capacity of the roadway than the actual capacity (Hoogendoorn and Botma, 1996). Applying this equation with the measured headway implies the assumption that the mean observed headway from the filtered data is the same as the mean following headway. The actual capacity of the facility was estimated by observing a plot of average speeds versus flow rates and approximating the maximum flow rate which is typically reached before a speed drop occurred. This has customarily been used in traffic engineering to estimate the flow rate at which flow breakdown occurs (i.e. the capacity of the facility). The average speeds and flow rates were calculated for five minute intervals for the weekdays of three months around the time of the data collection for each site. This was possible because of the presence of IDOT owned Wavetronix sensors near each of the data collection sites. The IDOT sensors were used rather than the temporary sensors, because this allowed for a larger sample of the traffic and more congested intervals to be observed. It also allowed for a comparison to the capacity from a different data source, so any consistency is not just a result of using the same data for calculating both the headway and the capacity.

Equation 7. Predicted capacity

$$C_{pred} = \frac{3600}{\bar{h}_{following}}$$

Where:

C_{pred} – Predicted capacity (veh/hr/ln)

$\bar{h}_{following}$ – Mean following headway (seconds)

Unfortunately, it became evident from observing the speed-flow graphs that only the Des Moines sites experienced flow breakdown due to demand exceeding capacity (as opposed to an accident, construction, or other external stressor). This means that the predicted and observed capacities can only be compared for the NB and SB directions of I-80/35 at Hickman. For the NB direction, the overall mean headway is 1.83 sec which leads to a predicted capacity of 1967 veh/hr/ln (3600/1.83). The speed-flow graph of the NB direction shows a capacity of approximately 1800 to 1900 veh/hr/ln (see Figure 38). This is consistent with the past finding that predicted capacity would exceed the actual capacity, but the prediction is actually fairly close to the actual value. There are actually a number of intervals which reach or slightly exceed the predicted capacity.

For the SB direction, the overall mean headway is 1.80 sec which leads to a predicted capacity of 2000 veh/hr/ln. The speed-flow graph of the SB direction shows a wider range of potential capacities than the NB direction – approximately 1700 to 1900 veh/hr/ln (see Figure 39). This larger disparity in potential capacities is due to a large scatter of data, which is likely due to more complicated weaving behavior in the SB direction compared to the NB direction. The weaving behavior can lead to flow breakdowns at a wider variety of flow rates due to its unpredictability. Again, the predicted capacity exceeded the actual capacity from the speed-flow graph, but not by much. And again, there were a few intervals in which the flow rate matched or exceeded the predicted capacity.

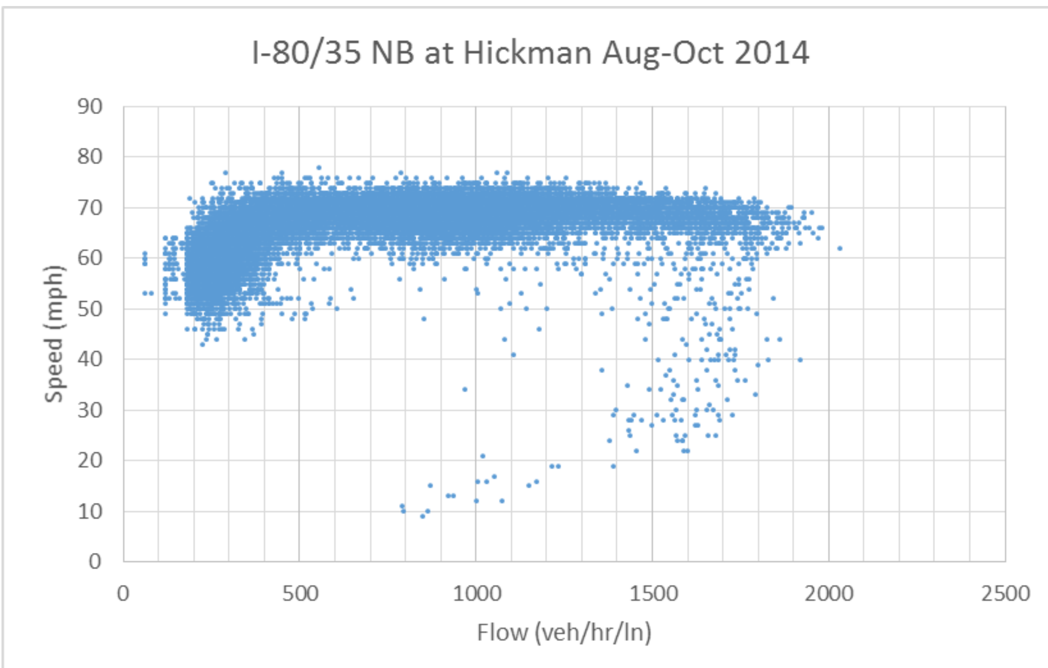

Figure 38. Speed vs flow for 5 min intervals for I-80/35 NB at Hickman

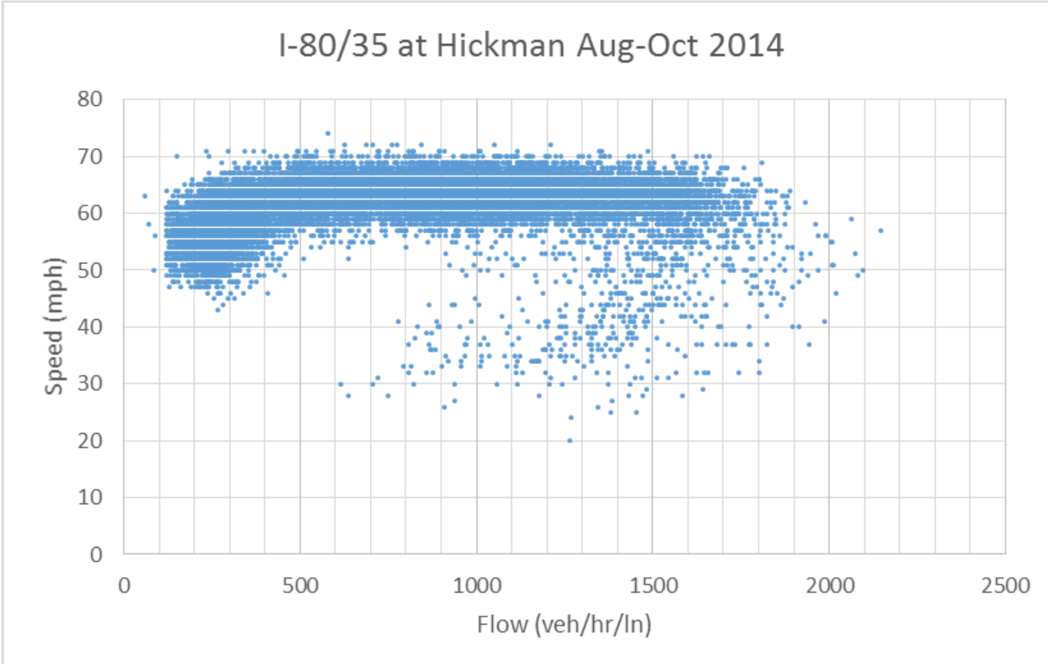

Figure 39. Speed vs flow for 5 min intervals for I-80/35 SB at Hickman

Conclusion of headway and time gap results

Overall, it was observed through the data collected in three different regions of Iowa that headway and time gap measurements are largely similar within the same driver population as well as across different driver populations, provided that the environment is generally the same (urban conditions and somewhat similar geometries). These regions were compared mostly on the similarities of the mean, median, and standard deviation values, as well as on a visual examination of the distributions. The data were filtered to include only observations that occurred during intervals that exceeded a 15 minute flow rate of 1,000 veh/hr/ln and observations with headways of 4 seconds or less.

In the central portion of Iowa, data were analyzed from threes sites in close proximity with one another on I-80/35 between the University Avenue and Hickman Road interchanges. The sites had fairly similar geometries but slightly different lane configures, and it was found that they had similar measurements for headway and time gap. This finding established the consistency of measurements within the same driver population. In the western portion of Iowa, data were collected from Council Bluffs on I-80 eastbound and westbound at the interchange for South Expressway. The eastbound and westbound directions had similar headway and time gap measurements, further supporting the finding that the same driver population produces similar headways and time gaps. In the eastern portion of Iowa, data were used from one site in the Quad Cities located on I-74 at the Middle Road interchange. It was not possible to further confirm consistency within driver population in the Quad Cities, because the other detector set up nearby did not provide any data meeting the filtering criteria and the amount of data in both directions of I-74 at Middle Road was not sufficient to compare. However, the data from Des Moines, Council Bluffs, and the Quad Cities were compared, and it was found that the headways and time gaps across the different cities were similar.

Additionally, data from a rural location on I-80 west of the Quad Cities were used as a point of comparison for the three urban locations. Not as much data met the filtering criteria as at some of the other sites, but this was expected due to the site’s lower flow rates. The usable data indicated a substantial difference in the rural location compared to the three urban locations. This difference can be seen visually in Figure 35 and Figure 36, which show average headway versus flow rate for each site, and it appears that the rural location follows a slightly different pattern from the others. This is particularly pronounced for the time gap data in Figure 36, where the rural location consistently shows time gaps lower than those of the urban locations.

Results Conclusion

The preceding section laid out the results of analysis of standstill distance, headway, and time gap data. A total 1238 standstill distance measurements were recorded from 47 stop-and-go traffic incidents in seven cities in Iowa. This data was heavily skewed toward Des Moines (which had 693 observations) due to the data collection process as well as the distribution of traffic in Iowa. This made it difficult to draw many definitive conclusions about site-based variables such as the cause of the incident, weather, curve presence, etc. However there were two main takeaways: first, Des Moines appears to have generally lower standstill distances than other parts of the state, and, second, the CC vehicle pair type tends to have lower standstill distances as well.

Separately from the standstill distance data, headway and time gap data were obtained for three urban freeways in different parts of the state of Iowa, as well as one rural urban location. These data were summarized for each site and each vehicle pair type (CC, CT, TC, and TC). It was found that in general, average headways and time gaps varied across the different vehicle types, but this difference is more pronounced for the time gaps. For the headways, the average when a car was following a car was usually similar to a car following a truck, and the average when a truck was following a car was usually similar to a truck following a truck. The headway averages for when a car was following compared to when a truck was following were markedly different, however. For the time gaps, all four vehicle pair types were distinct, due to the differing lengths of cars and trucks. When comparing averages using common vehicle types across the different locations, it was found that headways and time gaps were fairly consistent across the different urban freeway locations, but were different for the rural location.

As has been mentioned, standstill distances and headways/time gaps are strongly connected to the macroscopic traffic measures jam density and capacity, respectively. The smaller the average standstill distance, the higher the jam density will be, and the smaller the headway/time gap, the higher the capacity will be. Using average vehicle lengths which ranged from 20.6 feet to 23.9 feet for the urban areas and an average standstill distance of approximately 10 feet yields an estimated of jam density ranging from 156 veh/mi/ln to 173 veh/mi/ln. This can be compared to the assumed HCM 2010 value of 190 pc/mi/ln (TRB 2010). It is important to note that the HCM 2010 value is in pc/mi/ln, not veh/mi/ln, so if it were adjusted to include the impact of large trucks, it would be even closer to the range estimated from the standstill distance measurements.

Additionally, the average headway ranged from 1.81 sec to 1.86 sec. This corresponds to an estimated capacity of 1935 veh/hr/ln to 1989 veh/hr/ln which is compared to the HCM 2010 values of 2,250 pc/hr/ln at a free flow speed of 55 mph and 2,400 pc/hr/ln at a free flow speed of 70 mph for ideal conditions (HCM 2010). Though, when these values are adjusted for the presence of trucks (assumed to be 12 %, which is in line with the actual percentages in Des Moines and Council Bluffs locations) and either level or rolling conditions, the range of capacities drops to 1907 veh/hr/ln to 2264 veh/hr/ln. It is not surprising that the capacity estimated from the average headway values is on the low side of the capacities listed in the HCM for three main reasons. First, drivers in Iowa may not be as comfortable following as closely as the drivers on which the HCM was calibrated. Second, not all of the vehicles in the data set in this research were following, so more vehicles could still be added to the road before causing flow breakdown. And third, the HCM capacities quoted here are for basic freeway segments, which does not consider weaving behavior, which reduces the capacity. In addition to comparing to the HCM capacity, the predicted capacity was also compared to the capacity observed from speed-flow graphs of five minute intervals. The speed-flow graphs indicated a capacity of around 1700 to 1900 veh/hr/ln which is slightly lower than the predicted capacity.

The jam density and capacity are two of the most important traffic operations parameters, and they have a large impact on simulation models as well as roadway design. Capacity is frequently used on the planning and operational levels to diagnose potential bottlenecks and problem areas. Therefore, reliable capacity outputs from simulation models are critical in roadway planning and design. Additionally, jam density is one of the major determinants of the extent of a queue when a flow breakdown occurs on the facility. If a queue was estimated to be one mile long with a jam density of 190 pc/mi/ln, and the jam density was actually 160 pc/mi/ln, the queue would actually stretch nearly 1000 feet further than the estimate, which could hinder operations more than expected. Clearly, these parameters are important to understanding the traffic operations of an urban freeway.

CHAPTER VI: DISCUSSION AND CONCLUSION

Summary of Key Findings

Standstill distance findings

This study found that the average standstill distances in Iowa are generally between 8 and 12 feet. The average appeared to be lowest in the Des Moines area with a mean of 8.59 feet. After Des Moines, the Quad Cities had the most observations and had a significantly different mean of 10.19 feet (p-value of 2.22×10^{-8}). The city with the third most observations, Sioux City, had a mean of 12.53 that was significantly different than Des Moines (p-value of 2.69×10^{-17}) and the Quad Cities (p-value of 3.32×10^{-6}). However, it is worth noting that all the observations in the Quad Cities were the result of construction, so that could have some effect on the result. It was observed that construction standstill distances were consistently higher than other causes in each city where multiple causes were present. The cities which did not have as many stop-and-go incidents as the top three cities tended to have average standstill distances in the range of 10 to 12 feet.

The vehicle pair types were spread among the incidents and locations fairly well, so the mean standstill distances could be reasonably tested for each of the vehicle pair types without much concern for the data imbalances. It was found that the CC pair type had a significantly lower mean, 9.41 feet, than the CT and TC pair types, 13.35 feet and 12.37 feet, respectively. The level of confidence in these conclusions was quite high, with a p-value of 0.000421 between CC and CT pairs and a p-value of 0.000861 between CC and TC. There was not enough not enough data to reach the same statistically significant conclusion for the TT pair type, but its mean was more than 1.5 feet larger than the CC mean and it is believed that a statistically significant difference would arise with more data.

Headway and time gap findings

Both headways and time gaps were found to be fairly consistent within driver populations as well as on similar urban freeways with different driver populations. The one rural location examined was not consistent with the data observed from the three urban areas. These comparisons were made separately for each vehicle pair type (CC, CT, TC, and TT), because it was observed in the headway and time gap data that each of the vehicle pair types exhibit different following behavior. There was some variation of headways based on pair type, but the following vehicle was much more influential in the average headway values than the leading vehicle. The conclusion that the vehicle pair types should be treated separately was made by observing the summary statistics, but was confirmed by fitting log-normal distributions to the overall data and then to the individual groups and comparing the two methods using the likelihood ratio test.

For CC pairs in urban areas, the mean headway ranged from 1.72 to 1.82 sec, but were typically within 1.74 to 1.79 sec. The mean time gaps ranged from 1.51 to 1.63, but were usually within 1.53 to 1.58 sec. These values are substantially larger than those observed at the rural location of 1.56 sec headway and 1.40 sec and time gap. When a truck was leading a car (TC), the car tended to follow with a slightly higher headway by about 0.1 sec, but due to the length of the trucks this resulted in a much smaller time gap, typically about 0.4 sec shorter. When a truck was following, it typically resulted in about

0.5 to 0.6 sec longer headways and time gaps than the corresponding scenario for car following, but this relationship had more variability, presumably due to smaller sample sizes when trucks were involved (particularly TT). At the rural location, trucks were observed following cars at 0.2 to 0.3 second closer for both headways and time gaps compared to the urban locations.

In addition to the consistency displayed in the summary statistics in the urban areas, these locations were shown graphically to have similar mean headways and time gaps for the range of traffic volumes they experienced (see Figure 35 and Figure 36). These graphs also showed that average headway decreased with increasing traffic flow (as one would expect), but this decrease in headway tended to level off at flow rates of higher than 1000 veh/hr/ln. This illustrates the presence of a minimum headway which is accepted by each individual driver.

Discussion

Standstill distance discussion

A total of 1238 standstill distances were observed as a result of 47 stop-and-go incidents spread across 7 different cities in Iowa. However, 693 of these observations came from Des Moines. This imbalance was due to a large number of dynamic messaging signs (DMS), cameras, and detectors in Des Moines as well as the higher traffic volumes observed in Des Moines compared to the rest of the state. The higher traffic volumes made it easier for stop-and-go traffic conditions to develop, and the increased coverage from DMS and cameras made it easier to observe this stop-and-go traffic. There were some other imbalances in the data as well. For example, of the top three cities (by number of observations) only Des Moines had incidents caused by accidents, and only the Quad Cities and Sioux City had incidents caused by construction. These data imbalances and others made it somewhat difficult to discern which differences were due to incident cause or other conditions and which were the results of different driver populations across the state.

In addition to the imbalances in the data, another limitation of this study was the amount of data, specifically within certain groups. There were only 10 observations of trucks following of trucks (TT), and only about 40 observations for CT and TC pairs. There were also several cities which only had data from three or fewer stop-and-go incidents. Additionally, some conditions (e.g. night time, rainy, caused by stalled vehicle) were only observed in a few incidents.

Despite these limitations, there are some conclusions which can be made from the analysis. The first main conclusion is that standstill distance varies by site across Iowa, and it ranges from approximately 8 to 12 feet. Even the smallest mean standstill distance measurements were greater than one microsimulation program's (VISSIM's) default value of 4.92 feet (PTV 2011). So, it can be reasonably concluded that the default parameter does not reflect driving behavior in Iowa. Additionally, it makes sense that the standstill distance varies some from site to site, particularly that it is lower for Des Moines. Stop-and-go incidents in Des Moines are more common than anywhere else in the state. This likely helps condition the drivers in Des Moines to maintain smaller standstill distances to prevent other vehicles from changing lanes in front of them.

The second main conclusion is that the standstill distance for cars following cars is less than when a car and truck interact, and probably less than when a truck is following a truck. This result makes sense, because car drivers tend to be more comfortable when following other cars and give trucks a little bit wider berth. Truck drivers also tend to be more cautious with their spacing, since they understand the limitations of their vehicle's acceleration and deceleration capabilities as well as its blind spots.

Finally, it was found that standstill distances vary considerably even within the same stop-and-go incident. Some drivers leave as little as one or two feet, while others leave well more than one car length. Some microsimulation models do not allow for standstill distances (or their equivalent) to vary, instead treating it as a constant for all vehicles.

Headway and time gap discussion

The headway and time gap data were collected using side-fired radar detectors (Wavetronix) at sites in three urban areas and one rural location. In total, the detector data from three detectors in Des Moines, one detector in Council Bluffs, and one detector in the Quad Cities were used in the analysis of headway and time gap data. It is regrettable that more locations could not be observed in each city, in other cities, and in more rural locations, but due to data collection issues and time constraints, data from more locations could not be obtained. The locations where data was collected, however, produced a large amount of data. The three Des Moines detectors which were used in analysis observed over 600,000 vehicles total, the Council Bluffs detector observed over 2,100,000, the Quad Cities detector observed over 100,000, and the rural detector observed over 150,000. While these are large sample sizes on their own, it is important remember that only a small portion of these vehicles are actually in a following situation, which led to a small sample size for a few vehicle pair groups.

In order to address the issue that not all vehicles in the traffic stream are in a following situation, two filters were applied to the data. The first filter was that only headways which were observed during 15 minute flow rates of greater than 1000 veh/hr/ln would be used in the analysis. This value corresponds (roughly) to the HCM 2010 value for the free flow threshold of 1000 pc/hr/ln (TRB 2010). Additionally, it was shown that while headway decreases with increasing flow rate, this effect begins to level off around 1000 veh/hr/ln. The second filter was that headways of greater than 4 seconds would be excluded. This value was selected for two reasons. The first was that there was a study which found this was the threshold where statistically significant differences from the free flow distribution of headways began occurring (Wasielewski, 1979). Additionally, a variation on a methodology used to identify free vehicles (Vogel, 2002) indicated that headways of 4 seconds or less have high correlation between leading and following vehicle speeds, increasing the likelihood of following vehicles being observed.

These filters are not perfect – they will not guarantee all the vehicles in the data set are following, there will undoubtedly be free vehicles in the mix as well. They could be improved if a larger flow rate was used as the minimum. While 1000 veh/hr/ln is essentially the minimum for vehicles started to impede others, the majority of vehicles are still free. This is supported by the fact that even though the average headway begins to level off at 1000 veh/hr/ln, it becomes essentially constant after approximately 1300 veh/hr/ln. Unfortunately, this value could not be used, because it would entirely exclude the Quad Cities and rural locations. Additionally, the 4 second headway threshold is not cut and dry either. Past research has done everything from arbitrarily assigning a threshold to involved statistical analyses, with little agreement among them. Finally, a speed difference threshold would have been useful. By definition, following vehicle have speeds close to that of the leading vehicles, so if a threshold could be established which indicated small speed differences, this would help further narrow the focus to only following vehicles.

Despite these limitations, several conclusions can be drawn from analyzing this headway and time gap data. Since the focus of this paper is comparing headways and time gaps across different cities, even though some of the vehicles are not following, as long as they are all treated with the same filters, the results should be comparable. The main finding is that headways and time gaps are consistent for the same vehicle pair types (considering only cars and trucks) within the same driver population and across different urban areas. This result makes some sense, because experience indicates that there is not a large of difference in the way people drive in Iowa on one side of the state versus the other. However, it seems to go against the finding that standstill distances vary by city. This could be because the standstill distance data is unbalanced and that leads to the appearance of it varying by city, when the true value does not. It is completely possible that standstill distances vary by driver population, but the headways and time gaps do not, since they pertain to two separate traffic flow regimes and, therefore, two different driver behaviors.

Another major finding is that the headways and time gaps are distinct for the four different vehicle pairs, though the following vehicle has much more influence than the leading vehicle on headways. The mean headways and time gaps for the different pair types confirm the result which other stuff have shown – drivers adjust their headways to the capabilities of their vehicle and the surrounding vehicles. Cars are willing to follow trucks more closely than other cars because the car drivers know they can brake more effectively than trucks. Likewise, trucks follow cars further than they follow other trucks in order to allow room for their relatively poor braking abilities. One factor which was not examined in this study was how small trucks interact with cars and larger trucks. Despite these clear differences in vehicle behaviors, some microsimulation models do not allow for headway/time gap values to vary by vehicle class.

Finally, this paper confirmed what other studies had found that headway and time gap distribution can be adequately modelled by a fitted lognormal distribution. Additionally, the parameters for the fitted distributions are different for the different vehicle pair types. Again, despite this variation, some microsimulation models treat headways/time gaps as constant for all drivers rather than following a distribution.

Conclusion

This paper has established that standstill distance and headway (or time gap) are two of the most important parameters for microsimulation calibration, both in theory and in practice. These parameters have a large impact on microsimulation, which can then have a large impact on roadway design projects using that microsimulation model as a decision making tool. Projects which use microsimulation are often large and expensive, so accurate simulation results are essential to selecting the best design for the traffic dynamics and the most cost effective design. This paper also showed several gaps in the past literature. Most past microsimulation calibration efforts did not collect data on the parameters themselves in order to provide a physical basis for the final parameter selections. Additionally, there only a few studies which collected standstill distance data, and no studies which did so for a freeway setting. Finally, there were not many studies which compared headways or time gaps for different cities within a region.

The paper proceeded to outline a repeatable methodology for collecting standstill distance and headway/time gap data. This collection process relies on manual processing of video for standstill distances and individual vehicle data from radar detectors for the headways/time gaps. This data was then validated and analyzed using Microsoft Excel and the statistical software R. Standstill distance analysis consisted of comparison of group means for different variables using t-tests and examining the distribution of the data. The headway/time gap analysis consisted of a filtering process to limit the data to mostly following vehicles, comparisons of summary statistics of those data sets within the same city and across different cities, and fitting statistical distributions to the data.

There were a number of major findings as a result of this analysis. The fact that freeway standstill distances were collected at all is a significant contribution, since, to the best of the author's knowledge, it had not been done before. The standstill distance was found to vary from city to city and from CC following to when a truck was involved. Headways and time gaps tended to be consistent for the same vehicle pair types for the same driver population and for different driver populations, as long as the conditions were similar (i.e. urban freeway). The headways and time gaps for the rural location collected did not have the same consistency as the urban locations, giving a preliminary indication that driver behavior varies from the urban to the rural setting. Both standstill distance and headways/time gaps followed relatively disperse skewed distributions. The average standstill distance and headway parameters were found to be significantly larger than the default values in one microsimulation program, VISSIM. In particular, the headway was found to be about half a second larger than the default, which significantly reduces the simulated capacity of the roadway.

The findings summarized above have significant implications for future microsimulation models. They demonstrate the need to allow standstill distances and headways/time gaps to be treated as distributions. Additionally, headways/time gaps should be set separately for different vehicle classes. The consistency of the headway/time gaps in different cities within Iowa indicates that they can serve as a starting point based on physical evidence for future microsimulation calibration efforts on urban freeways in Iowa. The standstill distances found in this study can also be used as a starting range as well.

Future research in the area of estimating microsimulation parameters based on point measurements could collect data on more sites and cities in Iowa or another state or region to confirm its findings. In particular, standstill distances could be collected in a more balanced way in order to produce more easily interpretable results. Additionally, more rural and other dissimilar locations could be studied to compare to the urban locations. Other research could focus on refining the distribution models for standstill distance and following headway.

Additional future research with respect to the microsimulation parameters collected in this paper, as well as other microsimulation parameters, could focus on using instrumented vehicles. Instrumented vehicles are more suited for such data collection, because they can measure the microsimulation parameters directly for different drivers.

The SHRP 2 Naturalistic Driving Study data could be good for this purpose (TRB 2014). Its breadth of different drivers, traffic conditions, and regional locations would allow for the standstill distance, headway, and other parameters to be examined in a variety of conditions. Differences in how individual people drive in these different situations could be examined as a part of a comprehensive analysis of which factors contribute to how people drive (for example, the factors impacting the preferred headway of drivers).

REFERENCES

ÅF - Hansen & Henneberg — "Framework Agreement for the Supply of the Detectors: Result of the Detector Test." Farum, Denmark. Jul. 1-6, 2008. (Freeway)

Chen, Xiqun, Li Li, and Qixin Shi. "A Markov Model Based on Headway/Spacing Distributions." Stochastic Evolutions of Dynamic Traffic Flow (2014): 49-79.

Elefteriadou, Lily. An Introduction to Traffic Flow Theory. New York: Springer, 2014. Springer.com. Springer. Web.

Federal Highway Administration. Office of Operations. Traffic Analysis Toolbox Volume III: Guidelines for Applying Traffic Microsimulation Modeling Software. By Richard Dowling, Skabardonis Alexander, and Vassili Alexiadis. 2004.

Federal Highway Administration. Office of Operations. Traffic Analysis Toolbox Volume IV: Guidelines for Applying CORSIM Microsimulation Modeling Software. By Peter Holm, Daniel Tomich, Jaimie Sloboden, and Cheryl Lowrance. 2007.

FAMU-FSU / Florida DOT —"Evaluation of Wavetronix Traffic Sensor." Tallahassee, Florida. May 21 2007, Jul. 28-Aug. 28, 2007.

Gomes, Gabriel, Adolf May, and Roberto Horowitz. "Congested Freeway Microsimulation Model Using VISSIM." Transportation Research Record 1876.1 (2004): 71-81.

Ha, Duy-Hung, Maurice Aron, and Simon Cohen. "Time Headway Variable and Probabilistic Modeling." Transportation Research Part C: Emerging Technologies 25 (2012): 181-201.

Hoogendoorn, S., & H. Botma. 1996. On the estimation of parameters in headway distributions. Proceedings of the 2nd TRAIL Ph.D. Congress, Delft.

Hourdkais, J., P.G Michalopoulos, and J. Kottommannil, (2003), "Practical Procedure for Calibrating Microscopic Traffic Simulation Models", Transportation Research Record 1852, Paper No. 03-4167

ITS Performance Test Institute — "Certificate of Test ." Korea. Nov. 3-5, 2008

Long, Gary. "Intervehicle spacings and queue characteristics." Transportation Research Record: Journal of the Transportation Research Board 1796 (2002): 86-96.

Lowens, Nicholas E., and Randy B. Machemehl. "Sensitivity of Simulated Capacity to Modification of VISSIM Driver Behavior Parameters." Transportation Research Record 1988 (2006): 102-10.

Manjunatha, Pruthvi, Peter Vortisch, and Tom V. Mathew. Methodology for the Calibration of VISSIM in Mixed Traffic. Transportation Research Board, Proc. of Transportation Research Board 92nd Annual Meeting, Washington DC. 2013

Krbálek, Milan. "Inter-vehicle gap statistics on signal-controlled crossroads." Journal of Physics A: Mathematical and Theoretical 41.20 (2008): 205004.

Menneni, Sandeep, Carlos Sun, and Peter Vortisch. "Microsimulation Calibration Using Speed-Flow Relationships." Transportation Research Record 2088 (2008): 1-9.

Park, Byungkyu, and J. Schneeberger. "Microscopic Simulation Model Calibration and Validation: Case Study of VISSIM Simulation Model for a Coordinated Actuated Signal System." Transportation Research Record 1856.1 (2003): 185-92.

Paz, Alexander, Victor Molano, and Alauddin Khan. Calibration of Micro-simulation traffic-flow models considering all parameters simultaneously. Proc. of Transportation Research Board 93rd Annual Meeting, Washington DC. 2014.

PTV AG (2011). VISSIM User Manual ver. 5.30

Rahman, Mizanur, Mashrur Chowdhury, Taufiquar Khan, and Parth Bhavsar. A Parameter Estimation and Calibration Method for Car-Following Models. Proc. of Transportation Research Board 93rd Annual Meeting, Washington DC. 2014.

Sbayti, Hayssam and David Roden. Best Practices in the Use of Microsimulation Models. Statewideplanning.org. AASHTO, Mar. 2010. Web. 29 Apr. 2014.

SRF Consulting Group / Minnesota DOT — “Mn/DOT ITS Innovative Ideas Program: Ramp Queue Detection.” St. Paul, Minnesota. Aug. 13, 2008.

Traffic Safety and Operations Lab, U. of Maryland — “Performance Evaluation of the Detection System Provided by ASTI.” Salisbury, Maryland. Apr. 2008.

Transportation Research Board of the National Academies, Highway Capacity Manual (2010).

Transportation Research Board of the National Academies, SHRP2 safety facts sheet: New safety data available to tackle tough safety issues, improve highway safety (2014).

United States. Federal Highway Administration. Office of Infrastructure. *Interstate System Access Informational Guide*. *Fhwa.dot.gov*. Aug. 2010. Web.

Wang, Yuhao, and Panos D. Prevedouros, Ph.D. "Synopsis of Traffic Simulation Models." Thesis. University of Hawaii at Manoa, 1996. July 1996. Web. 29 Apr. 2014.

Wasielewski, Paul. "Car-Following Headways On Freeways Interpreted By The Semi-Poisson Headway Distribution Model." Transportation Science 13.1 (1979): 36. Academic Search Alumni Edition. Web. 10 Oct. 2015.

Wavetronix —"SmartSensor HD Performance Test Results." Richmond, Virginia. Jul. 17, 2006.

Woody, Tony. CALIBRATING FREEWAY SIMULATION MODELS IN VISSIM. 2006. Web.

Wu, X., & Liu, H. X. The Uncertainty of Drivers' Gap Selection and its Impact on the Fundamental Diagram. Procedia-Social and Behavioral Sciences,80, 901-921. 2013.

VISSIM Standards Project. Delaware Valley Regional Planning Commission, June 2013. Web.

Vogel, K. What characterizes a "free vehicle" in an urban area? Transporation Research Part F (2002), 5, 313-327.

Zhang, Guohui, Yinhai Wang, Heng Wei, and Yanyan Chen. "Examining Headway Distribution Models with Urban Freeway Loop Event Data." Transportation Research Record 1999.1 (2007): 141-49.

APPENDIX: SAMPLE R-CODE

````
---
title: "Headway - I80 S Expressway (CB)"
author: "Andrew Houchin"
date: "Monday, August 24, 2015"
output: html_document
---

```{r}
#Establish connection and download dataset
library(compute.es)
library(RODBC)
library(plyr)
library(dplyr)
library(ggplot2)
library(MASS)
I80_CB_data <- odbcConnect("I80 S Expressway - CB")
I80_CB_data
I80_CBDS14 <- sqlFetch(I80_CB_data, "CBDS14_May_30_Complete_TimeAsNum")
colnames(I80_CBDS14)[c(4,7)] <- c("SPEED", "TIME")
I80_CBDS14 <- I80_CBDS14[order(I80_CBDS14$ID),]

close(I80_CB_data)
```

```{r}
# Establish thresholds

#Assign car/truck threshold in feet
cartruck<-35

#Flow rate interval in minutes
interval <- 15

#Establish what interval to round headways in seconds
round.headway <- 1

#Headway threshold in seconds
head_thresh <- 4
```

```{r}
# % vehicles w/speed
(nrow(I80_CBDS14)-sum(is.na(I80_CBDS14$SPEED)))/nrow(I80_CBDS14)
```

```{r}
#Add follow type to each vehicle
I80_CBDS14$FOLLOW_TYPE <- NA
I80_CBDS14$FOLLOW_TYPE[I80_CBDS14$LENGTH<=cartruck] <- "C"
I80_CBDS14$FOLLOW_TYPE[I80_CBDS14$LENGTH>cartruck] <- "T"
```
````

```{r}
#Create free flow keys
I80_CBDS14$START_INT <- round_any(I80_CBDS14$TIME,1/(24*60/interval),floor)

FF_Key_EB <- I80_CBDS14 %>% filter(LANE == "EB_CL" | LANE == "EB_LL" | LANE == "EB_RL")
%>% group_by(START_INT) %>% summarise(thru_flow = length(START_INT)*(60/interval)/3, free =
thru_flow<1000)
nrow(FF_Key_EB)-sum(FF_Key_EB$free)

FF_Key_WB <- I80_CBDS14 %>% filter(LANE == "WB_LL" | LANE == "WB_RL") %>%
group_by(START_INT) %>% summarise(thru_flow = length(START_INT)*(60/interval)/2, free =
thru_flow<1000)
nrow(FF_Key_WB)-sum(FF_Key_WB$free)
```

```{r}
#Isolate lanes
# ER is "exit ramp", OR is "on ramp"
EB_CL <- filter(I80_CBDS14, LANE == "EB_CL")
EB_LL <- filter(I80_CBDS14, LANE == "EB_LL")
EB_RL <- filter(I80_CBDS14, LANE == "EB_RL")
WB_ER <- filter(I80_CBDS14, LANE == "WB_ER")
WB_LL <- filter(I80_CBDS14, LANE == "WB_LL")
WB_OR <- filter(I80_CBDS14, LANE == "WB_OR")
WB_RL <- filter(I80_CBDS14, LANE == "WB_RL")
```

Adding variables function

```{r}
add.variables <- function(input,FF_Key,round.headway){
  n <- nrow(input)
  input$HEADWAY <- NA
  input$LEADINGLENGTH <- NA
  input$LEADINGSPEED <- NA
  input$LEAD_TYPE <- NA
  input$PAIRTYPE <- NA
  input$TIMEGAP <- NA
  res.HEADWAY <- NA
  res.LEADINGLENGTH <- NA
  res.LEADINGSPEED <- NA
  for (i in 2:n){
    res.HEADWAY[i] <- input$TIME[i]-input$TIME[i-1]
    res.LEADINGLENGTH[i] <- input$LENGTH[i-1]
    res.LEADINGSPEED[i] <- input$SPEED[i-1]
  }
  input$HEADWAY <- (res.HEADWAY)*86400
  input$LEADINGLENGTH <- res.LEADINGLENGTH
  input$LEADINGSPEED <- res.LEADINGSPEED
  input$LEAD_TYPE[input$LEADINGLENGTH<=35] <- "C"
  input$LEAD_TYPE[input$LEADINGLENGTH>35] <- "T"
  input$PAIRTYPE <- paste(input$LEAD_TYPE,input$FOLLOW_TYPE,sep="")
  input$TIMEGAP <- input$HEADWAY-
input$LEADINGLENGTH/(input$LEADINGSPEED*1.4666667)
  input$ROUNDED.HEADWAY <- round_any(input$HEADWAY,round.headway)
  input$SPDDIFF <- input$SPEED-input$LEADINGSPEED
  input <- input %>% filter(TIMEGAP > 0)
```

````
  input <- merge(input, FF_Key, by='START_INT')
  return(input)
}
```

Add variables
```{r}
# Add headways, leading speeds to all lanes
ptm <- proc.time()
EB_CL <- add.variables(EB_CL,FF_Key_EB,round.headway)
proc.time() - ptm

ptm <- proc.time()
EB_LL <- add.variables(EB_LL,FF_Key_EB,round.headway)
proc.time() - ptm

ptm <- proc.time()
EB_RL <- add.variables(EB_RL,FF_Key_EB,round.headway)
proc.time() - ptm

ptm <- proc.time()
WB_ER <- add.variables(WB_ER,FF_Key_WB,round.headway)
proc.time() - ptm

ptm <- proc.time()
WB_LL <- add.variables(WB_LL,FF_Key_WB,round.headway)
proc.time() - ptm

ptm <- proc.time()
WB_OR <- add.variables(WB_OR,FF_Key_WB,round.headway)
proc.time() - ptm

ptm <- proc.time()
WB_RL <- add.variables(WB_RL,FF_Key_WB,round.headway)
proc.time() - ptm
```

Counts by lane
```{r}
count_EB_CL <- EB_CL %>% group_by(START_INT) %>% summarise(count=length(START_INT))
count_EB_LL <- EB_LL %>% group_by(START_INT) %>% summarise(count=length(START_INT))
count_EB_RL <- EB_RL %>% group_by(START_INT) %>% summarise(count=length(START_INT))
count_WB_ER <- WB_ER %>% group_by(START_INT) %>% summarise(count=length(START_INT))
count_WB_LL <- WB_LL %>% group_by(START_INT) %>% summarise(count=length(START_INT))
count_WB_OR <- WB_OR %>% group_by(START_INT) %>% summarise(count=length(START_INT))
count_WB_RL <- WB_RL %>% group_by(START_INT) %>% summarise(count=length(START_INT))
```

Merge lanes back into one data frame
```{r}
# Reorder columns so ID is first
EB_CL <-
EB_CL[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED","
CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE",
"PAIRTYPE","START_INT","thru_flow","free")]
````

````
EB_RL <-
EB_RL[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED","
CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE",
"PAIRTYPE","START_INT","thru_flow","free")]
EB_LL <-
EB_LL[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED","
CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE",
"PAIRTYPE","START_INT","thru_flow","free")]
WB_ER <-
WB_ER[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED","
CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE",
"PAIRTYPE","START_INT","thru_flow","free")]
WB_LL <-
WB_LL[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED","
CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE",
"PAIRTYPE","START_INT","thru_flow","free")]
WB_OR <-
WB_OR[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED",
"CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE
","PAIRTYPE","START_INT","thru_flow","free")]
WB_RL <-
WB_RL[c("ID","LANE","LENGTH","SPEED","SPDDIFF","LEADINGLENGTH","LEADINGSPEED","
CLASS","RANGE","TIME","HEADWAY","TIMEGAP","ROUNDED.HEADWAY","FOLLOW_TYPE",
"PAIRTYPE","START_INT","thru_flow","free")]

# Merge lanes into original data frame
I80_CBDS14 <- rbind(EB_CL, EB_RL, EB_LL, WB_ER, WB_LL, WB_OR, WB_RL)
I80_CBDS14 <- I80_CBDS14[order(I80_CBDS14$ID),]

# Filter by congested conditions and different lane configurations
I80_CBDS14_cong <- I80_CBDS14 %>% filter(HEADWAY<=head_thresh & free==F)
I80_CBDS14_cong_thru <- I80_CBDS14_cong %>% filter(LANE!="WB_ER" & LANE!="WB_OR")
I80_CBDS14_cong_thru_EB <- I80_CBDS14_cong %>% filter(LANE=="EB_CL" | LANE=="EB_LL" |
LANE=="EB_RL")
I80_CBDS14_cong_thru_WB <- I80_CBDS14_cong %>% filter(LANE=="WB_LL" |
LANE=="WB_RL")
```

```{r}
# Confirm the threshold for car following
correlation <- I80_CBDS14 %>% group_by(ROUNDED.HEADWAY) %>% summarize(
  count=length(LANE)-sum(is.na(SPEED-LEADINGSPEED)),
  correlation=cor(LEADINGSPEED,SPEED,use="pairwise.complete.obs"))

qplot(ROUNDED.HEADWAY,correlation,data=correlation[1:16,])
```

Analyze overall and by following type for both directions
```{r}
# Summary overall through
summ_I80_CBDS14 <- I80_CBDS14_cong_thru %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14
````

````
summ_I80_CBDS14_follow <- I80_CBDS14_cong_thru %>% group_by(FOLLOW_TYPE) %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_follow
summ_I80_CBDS14_pair <- I80_CBDS14_cong_thru %>% group_by(PAIRTYPE) %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_pair

# Summary EB through
summ_I80_CBDS14_EB <- I80_CBDS14_cong_thru_EB %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_EB
summ_I80_CBDS14_follow_EB <- I80_CBDS14_cong_thru_EB %>% group_by(FOLLOW_TYPE)
%>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_follow_EB
summ_I80_CBDS14_pair_EB <- I80_CBDS14_cong_thru_EB %>% group_by(PAIRTYPE) %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_pair_EB

# Summary WB through
summ_I80_CBDS14_WB <- I80_CBDS14_cong_thru_WB %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_WB
summ_I80_CBDS14_follow_WB <- I80_CBDS14_cong_thru_WB %>% group_by(FOLLOW_TYPE)
%>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_follow_WB
summ_I80_CBDS14_pair_WB <- I80_CBDS14_cong_thru_WB %>% group_by(PAIRTYPE) %>%
summarise(count_headway=length(free),num_cong_int=length(unique(START_INT)),mean_headway=me
an(HEADWAY),med_headway=median(HEADWAY),sd_headway=sd(HEADWAY),mean_timegap=mea
n(TIMEGAP,na.rm=T),med_timegap=median(TIMEGAP,na.rm=T),sd_timegap=sd(TIMEGAP,na.rm=T))
summ_I80_CBDS14_pair_WB
```

```{r}
# Writing results
write.csv(summ_I80_CBDS14,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research Projects\\Iowa
DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14.csv")
write.csv(summ_I80_CBDS14_follow,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_follow.csv")
write.csv(summ_I80_CBDS14_pair,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_pair.csv")
````

```
write.csv(summ_I80_CBDS14_EB,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_EB.csv")
write.csv(summ_I80_CBDS14_follow_EB,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_follow_EB.csv")
write.csv(summ_I80_CBDS14_pair_EB,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_pair_EB.csv")
write.csv(summ_I80_CBDS14_WB,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_WB.csv")
write.csv(summ_I80_CBDS14_follow_WB,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_follow_WB.csv")
write.csv(summ_I80_CBDS14_pair_WB,file="S:\\(S) SHARE\\_project CTRE\\1_Active Research
Projects\\Iowa DOT VISSIM Calibration\\Headway Results\\summ_I80_CBDS14_pair_WB.csv")
```

Fitting distributions for headway for overall and both directions

```{r}
# Fitting distributions and plotting histograms with pdfs for overall through
fit.thru <- fitdistr(I80_CBDS14_cong_thru$HEADWAY,"lognormal")
fit.thru
fit.thru$loglik
h <- hist(I80_CBDS14_cong_thru$HEADWAY,breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru$estimate["meanlog"],sdlog=fit.thru$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic")
lines(xfit,yfit, col="red")

fit.thru.car <-
fitdistr(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$FOLLOW_TYPE=="C"],"lognor
mal")
fit.thru.car
fit.thru.car$loglik
h <-
hist(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$FOLLOW_TYPE=="C"],breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.car$estimate["meanlog"],sdlog=fit.thru.car$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - car following")
lines(xfit,yfit, col="red")

fit.thru.truck <-
fitdistr(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$FOLLOW_TYPE=="T"],"lognor
mal")
fit.thru.truck
fit.thru.truck$loglik
h <-
hist(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$FOLLOW_TYPE=="T"],breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
```

```
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.truck$estimate["meanlog"],sdlog=fit.thru.truck$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - truck following")
lines(xfit,yfit, col="red")

# Is it more effective to model car and truck following separately?
fit.thru$loglik
fit.thru.car$loglik
fit.thru.truck$loglik
-2*(fit.thru$loglik - (fit.thru.car$loglik + fit.thru.truck$loglik))




# Fitting distributions and plotting histograms with pdfs for EB through
fit.thru.EB <- fitdistr(I80_CBDS14_cong_thru_EB$HEADWAY,"lognormal")
fit.thru.EB
fit.thru.EB$loglik
h <- hist(I80_CBDS14_cong_thru_EB$HEADWAY,breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_EB$HEADWAY),max(I80_CBDS14_cong_thru_EB$HEADWAY),len
gth=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.EB$estimate["meanlog"],sdlog=fit.thru.EB$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nEB thru traffic")
lines(xfit,yfit, col="red")

fit.thru.EB.car <-
fitdistr(I80_CBDS14_cong_thru_EB$HEADWAY[I80_CBDS14_cong_thru_EB$FOLLOW_TYPE=="C"]
,"lognormal")
fit.thru.EB.car
fit.thru.EB.car$loglik
h <-
hist(I80_CBDS14_cong_thru_EB$HEADWAY[I80_CBDS14_cong_thru_EB$FOLLOW_TYPE=="C"],br
eaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_EB$HEADWAY),max(I80_CBDS14_cong_thru_EB$HEADWAY),len
gth=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.EB.car$estimate["meanlog"],sdlog=fit.thru.EB.car$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nEB thru traffic - car following")
lines(xfit,yfit, col="red")

fit.thru.EB.truck <-
fitdistr(I80_CBDS14_cong_thru_EB$HEADWAY[I80_CBDS14_cong_thru_EB$FOLLOW_TYPE=="T"]
,"lognormal")
fit.thru.EB.truck
fit.thru.EB.truck$loglik
```

```
h <-
hist(I80_CBDS14_cong_thru_EB$HEADWAY[I80_CBDS14_cong_thru_EB$FOLLOW_TYPE=="T"],br
eaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_EB$HEADWAY),max(I80_CBDS14_cong_thru_EB$HEADWAY),len
gth=40)
yfit<-
dlnorm(xfit,meanlog=fit.thru.EB.truck$estimate["meanlog"],sdlog=fit.thru.EB.truck$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nEB thru traffic - truck following")
lines(xfit,yfit, col="red")

# Is it more effective to model car and truck following separately?
-2*(fit.thru.EB$loglik - (fit.thru.EB.car$loglik + fit.thru.EB.truck$loglik))




# Fitting distributions and plotting histograms with pdfs for WB through
fit.thru.WB <- fitdistr(I80_CBDS14_cong_thru_WB$HEADWAY,"lognormal")
fit.thru.WB
fit.thru.WB$loglik
h <- hist(I80_CBDS14_cong_thru_WB$HEADWAY,breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_WB$HEADWAY),max(I80_CBDS14_cong_thru_WB$HEADWAY),le
ngth=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.WB$estimate["meanlog"],sdlog=fit.thru.WB$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nWB thru traffic")
lines(xfit,yfit, col="red")

fit.thru.WB.car <-
fitdistr(I80_CBDS14_cong_thru_WB$HEADWAY[I80_CBDS14_cong_thru_WB$FOLLOW_TYPE=="C
"],"lognormal")
fit.thru.WB.car
fit.thru.WB.car$loglik
h <-
hist(I80_CBDS14_cong_thru_WB$HEADWAY[I80_CBDS14_cong_thru_WB$FOLLOW_TYPE=="C"],
breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_WB$HEADWAY),max(I80_CBDS14_cong_thru_WB$HEADWAY),le
ngth=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.WB.car$estimate["meanlog"],sdlog=fit.thru.WB.car$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nWB thru traffic - car following")
lines(xfit,yfit, col="red")
```

fit.thru.WB.truck <-
fitdistr(I80_CBDS14_cong_thru_WB$HEADWAY[I80_CBDS14_cong_thru_WB$FOLLOW_TYPE=="T
"],"lognormal")
fit.thru.WB.truck
fit.thru.WB.truck$loglik
h <-
hist(I80_CBDS14_cong_thru_WB$HEADWAY[I80_CBDS14_cong_thru_WB$FOLLOW_TYPE=="T"],
breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_WB$HEADWAY),max(I80_CBDS14_cong_thru_WB$HEADWAY),le
ngth=40)
yfit<-
dlnorm(xfit,meanlog=fit.thru.WB.truck$estimate["meanlog"],sdlog=fit.thru.WB.truck$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nWB thru traffic - truck following")
lines(xfit,yfit, col="red")

Is it more effective to model car and truck following separately?
-2*(fit.thru.WB$loglik - (fit.thru.WB.car$loglik + fit.thru.WB.truck$loglik))
```

```{r}
# Filter only those that have time gap values
I80_CBDS14_cong_TG <- I80_CBDS14_cong %>% filter(TIMEGAP>0 & is.na(TIMEGAP)==F)
I80_CBDS14_cong_thru_TG <- I80_CBDS14_cong_TG %>% filter(LANE!="WB_ER" &
LANE!="WB_OR")
I80_CBDS14_cong_thru_EB_TG <- I80_CBDS14_cong_TG %>% filter(LANE=="EB_CL" |
LANE=="EB_LL" | LANE=="EB_RL")
I80_CBDS14_cong_thru_WB_TG <- I80_CBDS14_cong_TG %>% filter(LANE=="WB_LL" |
LANE=="WB_RL")
```

Fitting distributions for timegap for overall and both directions
```{r}
# Fitting distributions and plotting histograms with pdfs for overall through
fit.thru <- fitdistr(I80_CBDS14_cong_thru_TG$TIMEGAP,"lognormal")
fit.thru
fit.thru$loglik
h <- hist(I80_CBDS14_cong_thru_TG$TIMEGAP,breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_TG$TIMEGAP),max(I80_CBDS14_cong_thru_TG$TIMEGAP),length
=40)
yfit<-dlnorm(xfit,meanlog=fit.thru$estimate["meanlog"],sdlog=fit.thru$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic")
lines(xfit,yfit, col="red")

fit.thru.car <-
fitdistr(I80_CBDS14_cong_thru_TG$TIMEGAP[I80_CBDS14_cong_thru_TG$FOLLOW_TYPE=="C"],
"lognormal")
fit.thru.car
fit.thru.car$loglik

```
h <-
hist(I80_CBDS14_cong_thru_TG$TIMEGAP[I80_CBDS14_cong_thru_TG$FOLLOW_TYPE=="C"],bre
aks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_TG$TIMEGAP),max(I80_CBDS14_cong_thru_TG$TIMEGAP),length
=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.car$estimate["meanlog"],sdlog=fit.thru.car$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - car following")
lines(xfit,yfit, col="red")

fit.thru.truck <-
fitdistr(I80_CBDS14_cong_thru_TG$TIMEGAP[I80_CBDS14_cong_thru_TG$FOLLOW_TYPE=="T"],"
lognormal")
fit.thru.truck
fit.thru.truck$loglik
h <-
hist(I80_CBDS14_cong_thru_TG$TIMEGAP[I80_CBDS14_cong_thru_TG$FOLLOW_TYPE=="T"],bre
aks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_TG$TIMEGAP),max(I80_CBDS14_cong_thru_TG$TIMEGAP),length
=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.truck$estimate["meanlog"],sdlog=fit.thru.truck$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - truck following")
lines(xfit,yfit, col="red")

# Is it more effective to model car and truck following separately?
-2*(fit.thru$loglik - (fit.thru.car$loglik + fit.thru.truck$loglik))




# Fitting distributions and plotting histograms with pdfs for EB through
fit.thru.EB <- fitdistr(I80_CBDS14_cong_thru_EB_TG$TIMEGAP,"lognormal")
fit.thru.EB
fit.thru.EB$loglik
h <- hist(I80_CBDS14_cong_thru_EB_TG$TIMEGAP,breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_EB_TG$TIMEGAP),max(I80_CBDS14_cong_thru_EB_TG$TIMEGA
P),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.EB$estimate["meanlog"],sdlog=fit.thru.EB$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nEB thru traffic")
lines(xfit,yfit, col="red")

fit.thru.EB.car <-
fitdistr(I80_CBDS14_cong_thru_EB_TG$TIMEGAP[I80_CBDS14_cong_thru_EB_TG$FOLLOW_TYPE
=="C"],"lognormal")
```

```
fit.thru.EB.car
fit.thru.EB.car$loglik
h <-
hist(I80_CBDS14_cong_thru_EB_TG$TIMEGAP[I80_CBDS14_cong_thru_EB_TG$FOLLOW_TYPE==
"C"],breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_EB_TG$TIMEGAP),max(I80_CBDS14_cong_thru_EB_TG$TIMEGA
P),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.EB.car$estimate["meanlog"],sdlog=fit.thru.EB.car$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nEB thru traffic - car following")
lines(xfit,yfit, col="red")

fit.thru.EB.truck <-
fitdistr(I80_CBDS14_cong_thru_EB_TG$TIMEGAP[I80_CBDS14_cong_thru_EB_TG$FOLLOW_TYPE
=="T"],"lognormal")
fit.thru.EB.truck
fit.thru.EB.truck$loglik
h <-
hist(I80_CBDS14_cong_thru_EB_TG$TIMEGAP[I80_CBDS14_cong_thru_EB_TG$FOLLOW_TYPE==
"T"],breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_EB_TG$TIMEGAP),max(I80_CBDS14_cong_thru_EB_TG$TIMEGA
P),length=40)
yfit<-
dlnorm(xfit,meanlog=fit.thru.EB.truck$estimate["meanlog"],sdlog=fit.thru.EB.truck$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nEB thru traffic - truck following")
lines(xfit,yfit, col="red")

# Is it more effective to model car and truck following separately?
-2*(fit.thru.EB$loglik - (fit.thru.EB.car$loglik + fit.thru.EB.truck$loglik))




# Fitting distributions and plotting histograms with pdfs for WB through
fit.thru.WB <- fitdistr(I80_CBDS14_cong_thru_WB_TG$TIMEGAP,"lognormal")
fit.thru.WB
fit.thru.WB$loglik
h <- hist(I80_CBDS14_cong_thru_WB_TG$TIMEGAP,breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_WB_TG$TIMEGAP),max(I80_CBDS14_cong_thru_WB_TG$TIMEG
AP),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.WB$estimate["meanlog"],sdlog=fit.thru.WB$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nWB thru traffic")
lines(xfit,yfit, col="red")
```

```
fit.thru.WB.car <-
fitdistr(I80_CBDS14_cong_thru_WB_TG$TIMEGAP[I80_CBDS14_cong_thru_WB_TG$FOLLOW_TY
PE=="C"],"lognormal")
fit.thru.WB.car
fit.thru.WB.car$loglik
h <-
hist(I80_CBDS14_cong_thru_WB_TG$TIMEGAP[I80_CBDS14_cong_thru_WB_TG$FOLLOW_TYPE=
="C"],breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_WB_TG$TIMEGAP),max(I80_CBDS14_cong_thru_WB_TG$TIMEG
AP),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.WB.car$estimate["meanlog"],sdlog=fit.thru.WB.car$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nWB thru traffic - car following")
lines(xfit,yfit, col="red")

fit.thru.WB.truck <-
fitdistr(I80_CBDS14_cong_thru_WB_TG$TIMEGAP[I80_CBDS14_cong_thru_WB_TG$FOLLOW_TY
PE=="T"],"lognormal")
fit.thru.WB.truck
fit.thru.WB.truck$loglik
h <-
hist(I80_CBDS14_cong_thru_WB_TG$TIMEGAP[I80_CBDS14_cong_thru_WB_TG$FOLLOW_TYPE=
="T"],breaks=seq(from=-0.2,to=6.8,by=0.5))
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru_WB_TG$TIMEGAP),max(I80_CBDS14_cong_thru_WB_TG$TIMEG
AP),length=40)
yfit<-
dlnorm(xfit,meanlog=fit.thru.WB.truck$estimate["meanlog"],sdlog=fit.thru.WB.truck$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nWB thru traffic - truck following")
lines(xfit,yfit, col="red")

# Is it more effective to model car and truck following separately?
-2*(fit.thru.WB$loglik - (fit.thru.WB.car$loglik + fit.thru.WB.truck$loglik))
```

```{r}
#Pair type comparison


fit.thru.CC <-
fitdistr(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="CC"],"lognormal"
)
fit.thru.CC
fit.thru.CC$loglik
h <-
hist(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="CC"],breaks=12)
xhist<-c(min(h$breaks),h$breaks)
```

```
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.CC$estimate["meanlog"],sdlog=fit.thru.CC$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - CC following")
lines(xfit,yfit, col="red")

fit.thru.CT <-
fitdistr(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="CT"],"lognormal"
)
fit.thru.CT
fit.thru.CT$loglik
h <-
hist(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="CT"],breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.CT$estimate["meanlog"],sdlog=fit.thru.CT$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - CT following")
lines(xfit,yfit, col="red")

fit.thru.TC <-
fitdistr(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="TC"],"lognormal"
)
fit.thru.TC
fit.thru.TC$loglik
h <-
hist(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="TC"],breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.TC$estimate["meanlog"],sdlog=fit.thru.TC$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - TC following")
lines(xfit,yfit, col="red")

fit.thru.TT <-
fitdistr(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="TT"],"lognormal"
)
fit.thru.TT
fit.thru.TT$loglik
h <-
hist(I80_CBDS14_cong_thru$HEADWAY[I80_CBDS14_cong_thru$PAIRTYPE=="TT"],breaks=12)
xhist<-c(min(h$breaks),h$breaks)
yhist<-c(0,h$density,0)
xfit<-
seq(min(I80_CBDS14_cong_thru$HEADWAY),max(I80_CBDS14_cong_thru$HEADWAY),length=40)
yfit<-dlnorm(xfit,meanlog=fit.thru.TT$estimate["meanlog"],sdlog=fit.thru.TT$estimate["sdlog"])
plot(xhist,yhist,type="s",xlim=c(0,6),ylim=c(0,max(yhist,yfit)), main="Lognormal pdf and histogram
for:\nall thru traffic - TT following")
lines(xfit,yfit, col="red")
```

```
# Is it more effective to model CC, CT, TC, and TT following separately?
fit.thru.CC$loglik
fit.thru.CT$loglik
fit.thru.TC$loglik
fit.thru.TT$loglik
-2*(fit.thru.car$loglik + fit.thru.truck$loglik - (fit.thru.CC$loglik + fit.thru.CT$loglik + fit.thru.TC$loglik + fit.thru.TT$loglik))
```