Project ID: NTC2015-MU-R-07

IDENTIFICATION OF POTENTIALLY HAZARDOUS ROADWAY NETWORK LOCATIONS USING MICROSCOPIC OBSERVATIONAL VEHICLE DATA AND MACROSCOPICALLY MODELED REACTION TIME

Final Report

by

Ishtiak Ahmed
Graduate Research Assistant
Institute for Transportation Research and Education
North Carolina State University

Billy M. Williams, Ph.D., P.E.
Professor, North Carolina State University,
Centennial Campus, Box 8601, Raleigh, NC 27695-8601,
Phone: (919) 515-7813; Fax: (919)515-7908; Email: billy_williams@ncsu.edu
North Carolina State University

for

National Transportation Center at Maryland (NTC@Maryland)
1124 Glenn Martin Hall
University of Maryland
College Park, MD 20742

March 2018

ACKNOWLEDGMENTS

This project was funded by the National Transportation Center @ Maryland (NTC@Maryland), one of the five National Centers that were selected in this nationwide competition, by the Office of the Assistant Secretary for Research and Technology (OST-R), U.S. Department of Transportation (US DOT).

DISCLAIMER

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

EXECUTIVE SUMMARY

The car following model developed by Gazis, Herman, and Rothery (known as the GHR model) can yield an analytical estimate of driver reaction time required for asymptotic stability as a function of density. Furthermore, equivalent steady-state macroscopic flow models have been derived for this car-following model. This study exploits this bridge between the macroscopic and microscopic version of the GHR model and hypothesizes that this reaction time and its change with the change of traffic state will demonstrate a correlation with observed crash rates that can be exploited in the USDOT's *Highway Safety Manual* methodology, thereby providing improved estimation of road segment safety performance. One of the challenges to investigate this hypothesis is to identify the transition state and to determine the change in reaction time during this state. To tackle this issue, the format of the GHR model developed in this study is characterized by the discontinuous traffic flow model. It enables the modeling of two traffic states for the density region where the traffic stream may be either in an uncongested or congested state and may transition from one regime to the other. When the driver reaction time required for asymptotic stability is estimated for the two regimes, the discontinuous model also reveals the drop in driver reaction time during transition. The primary objective of this study is to investigate the relationship between freeway crash rates and this drop in driver reaction time during transition of traffic states.

As the proposed method of model fitting is data-driven, two algorithms are applied to macroscopic traffic data to remove outliers such as inconsistent and mixed state observations and instrumental errors. The first algorithm is applied to the raw data which removes low speed associated with low-density observations, and those having a first difference of speed greater than 10 mph. The second one is a modified robust regression technique that removes outliers iteratively by judging their standard error.

The key modification to the GHR macroscopic model proposed in this study is the introduction of a transition regime based on fitted density breakpoint values. Each transition regime observation is initially modeled as both uncongested and congested regime and the one that gives the lower error of flow rate is finally selected. A non-linear optimization tool in MATLAB is used to fit the model. The fitted flow-density models yield the so-called inverse lambda shaped curve. Two additional constraints justified by the Highway Capacity Manual is incorporated to reasonably fit the observations.

As the estimated required driver reaction times for asymptotic stability are plotted against different density values, two drops in reaction times are identified that are associated with the transition of traffic states. The larger one (kb_1) is the difference in reaction time in the uncongested and congested state at the breakpoint density of the congested regime. The relatively smaller drop (kb_2) is associated with the breakpoint density of the uncongested regime. The hypothesis is that a higher value of these drops will be associated with a higher expected frequency of target crashes. In this study, the rear-end crashes were selected as the target crash type as it mostly occurs during car-following situation.

The proposed macroscopic model was fitted to 21 roadside sensor data collected from three Interstates near the Triangle Area of North Carolina for the calendar year of 2013. The fitted transition states range from 33 pc/mi/ln to 48 pc/mi/ln. Site-specific breakpoints and capacity values and their deviation from the defaults specified in the Highway Capacity Manual also revealed the limitation of using national average values for fitting macroscopic traffic stream models. All the parameters having physical interpretation had reasonable values expect for jam density of the congested regime which spanned within an odd range from 229 to 4207 pc/mi/ln.

Crash data were collected for a three-year period (2011-2013) from the archive maintained by North Carolina Department of Transportation. Crash rates were estimated for the road segment associated with each sensor and it varied from 17 to 173 crashes per 100 million vehicles-miles traveled. The percentages of rear-end crash rates diverse from 14% to 75%. The similarity between rear-end and peak-hour crash rates led to a correlation analysis which revealed that a moderately strong correlation exists between these two crash types.

Regression analysis between crash rates and both drops in reaction time revealed that Δt_{r2} does not have any significant correlation with either total, rear-end, or peak hour crash rates. The reason could be that the reaction time at uncongested regime is not that stable. Δt_{r1} showed a positive correlation with rear-end crash rates with an adjusted R square of 0.34. The coefficient of Δt_{r1} is significant at a level of 0.01.

The positive correlation between crash rates and Δt_{r1} obtained from this study opens the window of future research on the mechanism of rear-end crashes and the role of traffic stream characteristics behind it. However, the study suffered from several limitations. A small sample size of locations used in this study is one of them. In addition, the selection of thresholds for applying the outlier detection algorithms should be calibrated. Furthermore, the issue with unusual estimates of jam density, which is probably due to the lack of observations near jammed condition should be addressed in the future studies.

1.0 INTRODUCTION

1.1 BACKGROUND

Research on the factors associated with roadway crashes is not new in the field of transportation, as prominent tools like Haddon Matrix was developed as early as in 1971 (Haddon & Kelley, 1971). Identifying these factors assists planners and engineers to predict the crash propensity of a road or a vehicular property, or a human behavior, and to develop appropriate countermeasures. However, the factors behind a crash can be so complicated and interconnected to each other that finding their genuine contribution to the crash can be very challenging. Nonetheless, continuous efforts of researchers and engineers from around the world have made it possible to develop different tools to determine the role of different factors behind roadway crashes.

It is an intuitive and well-known fact that the time required by a driver to safely react to an event on the roadway can impact the potential of the event leading to an accident. However, the exact relationship between the required reaction time and the chance of crash occurrence is difficult to find since the reaction time is not easy to measure. Microscopic traffic flow theory has established the conditions for maintaining asymptotic stability in stimulus-response car-following models as a function of the driver reaction time and the modeled follower-response sensitivity (May, 1990). Furthermore, equivalent steady-state macroscopic flow models have been derived for the well-known Gazis-Herman-Rothery's (GHR) car-following model (Gazis, et al., 1961). Therefore, based on the GHR model the driver reaction time required for maintaining asymptotic stability at any traffic state (flow, speed, and density) can be analytically derived using a fitted GHR macroscopic model. The relationship between the number of crashes or crash rates at a road segment with this analytically derived reaction time may reveal the contribution of traffic characteristics to crash propensity, a topic that is yet to be investigated according to the authors' knowledge.

1.2 PROBLEM STATEMENTS, OBJECTIVES, AND SCOPE

Since the establishment of traffic flow theory, several car-following models had been developed such as, GHR model (Skabardonis, et al., 2003), Gipps model (Gipps, 1981), Cellular Automation model (Nagel & Schreckenberg, 1992) etc. The GHR model, which is very popular in literature is selected in this study to model the car-following behavior of a traffic stream because of the existence of equivalent macroscopic and microscopic model forms. As the mathematical formulation of driver reaction time required for asymptotic stability using the parameters of GHR model has already been established, the primary challenge lies in identifying the transition state and determining the change in reaction time during this state. Traffic condition when the flow reaches the capacity of a road section is defined as the boundary between two states by past studies (Transportation Research Board, 2010). However, it does not give the site-specific density breakpoint value since the HCM model uses a fixed value of 45 pc/mi/ln, which is based on national average values. In addition, data-driven model fitting approaches require robust filtration of inconsistent, mixed state observations and instrumental errors. Therefore, research is warranted to formulate a model with steady-state observations that is capable of revealing the site-specific transition regime.

The primary objective of this study is to investigate the relationship between crash rates and analytically derived drop in driver reaction time required for asymptotic stability of traffic stream. From the above discussion, it is apparent that prior to investigating the crash data, a macroscopic model is necessary that can imitate the transition state between uncongested and congested traffic condition with sufficient accuracy.

The format of the GHR model developed in this study is characterized by the discontinuous traffic flow model or the inverse lambda shape flow-density model (Wong & Wong, 2002). The advantage of this method is that it enables the modeling of two traffic states for the density region where the traffic stream may be either in an uncongested or congested state and may transition from one regime to the other. When the driver reaction time required for asymptotic stability is estimated for the two regimes, the discontinuous model also reveals the drop in driver reaction between the two regimes.

The drops in driver reaction time at the beginning and the end of the transitioning state are considered to be of particular interest in this study. Large magnitude of these drops indicates that the drivers have to adapt themselves quickly to this abrupt change in state. These drops in reaction time are mostly associated with queue growing from downstream bottleneck and rear-end crashes are more likely to occur at this condition. Moreover, this phenomenon is more frequently observed during the peak hours.

This study exploits this derivation and hypothesizes that this reaction time representing a traffic state can predict the potential of crashes that are more likely to occur at that traffic state. Moreover, it aims at particular crash types that can be associated with the car-following maneuver. It is assumed that the target crash types are likely to occur when traffic state transfers from the uncongested to the congested regime. As the mathematical formulation of the study is based on a car-following model, its scope is kept within the road segment types where the least number of lane changing behavior is expected. Thus, the scope of this study is narrowed down to estimating the reaction time associated with the transition of state at basic freeway segments (as defined in the HCM) and finding the potential of occurrence of the target crashes.

1.3 ORGANIZATION OF THE REPORT

This report is divided into five chapters. The current chapter presents the background information, motivation, problem statements, objectives, and scopes of the study. Chapter 2 provides a detailed review of the past studies related to macroscopic model development and to the link between crashes and traffic stream characteristics. Chapter 3 describes the development of the proposed macroscopic model and the calculation of the changes in driver reaction time. Chapter 4 presents the application of the proposed model and investigation of the relationship between crash rates and the change in reaction time. Chapter 5 summarizes the main conclusions of the analysis as well as recommendations for future study.

2.0 LITERATURE REVIEW

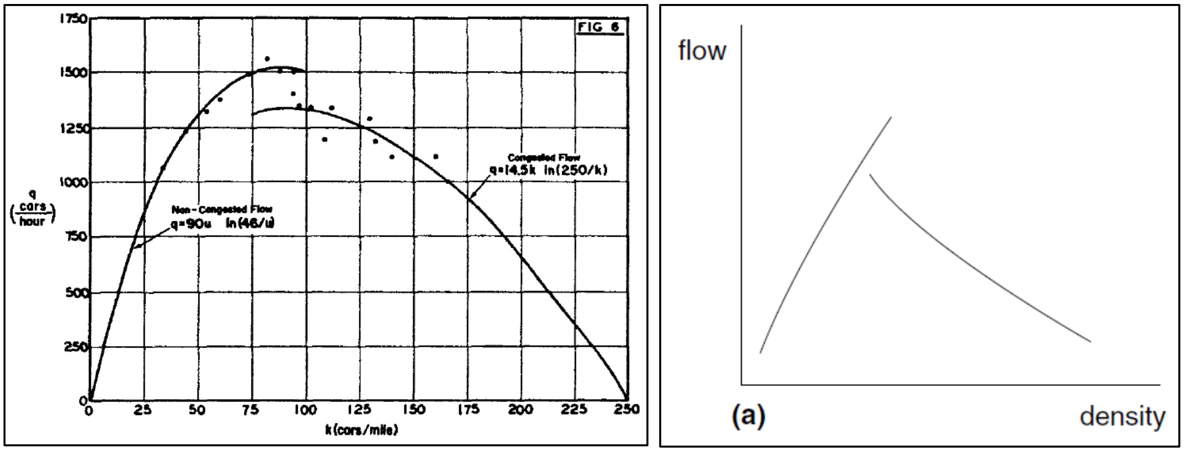
This chapter presents a review of past studies related to macroscopic traffic stream models and the relationship between freeway safety and traffic characteristics. It is divided into two major sections- each on one of the topics mentioned above. The section on macroscopic traffic stream models is further divided into two subsections. The relationship between freeway safety and traffic characteristics is subdivided based on different traffic parameters used to establish the relationship.

2.1 LITERATURE ON MACROSCOPIC TRAFFIC STREAM MODELS

The history of macroscopic traffic stream model goes back to 1934 when Greenshields first proposed his speed-density relationship based on an aerial photographic study (May, 1990). Greenshields concluded that the speed-density relationship is linear and depends on two parameters namely free-flow speed and jam density. Later in 1959, Greenberg proposed a non-linear speed-density relationship using tunnel data in a hydrodynamic analogy (Greenberg, 1959). From then on, numerous non-linear speed-density models have been developed. All these models can be divided into two categories. The first category has some sort of discontinuity and the second one is continuous over the entire regime of observation.

2.1.1 Discontinuous Model

Edie (1961) developed a complementary theory for steady-state conditions by combining the models proposed by Greenberg (1959)and Underwood (1960). It was the first study to observe a sharp speed drop in a small density range in some observed speed–density phase plots and proposed a two-regime phase diagram to model it. The frequent appearance of this discontinuity in several cases was mentioned in this study, which evidenced that the discontinuity did not appear because of random factors or by circumstances upstream from a bottleneck. Based on these evidence, Edie used the Underwood model for the free flow regime and the Greenberg model for the congested regime. The Lincoln Tunnel data used by Greenberg (1959) and Gazis et al. (1961) was fitted in this study discontinuously to the proposed model and showed that the regression lines have better fits than the continuous model. The fitted flow-density plot is shown in Figure 2-1 (a) below.



(a) (b)

FIGURE 2-1: (a) Flow-density model proposed by Edie fitted to the Lincoln tunnel data (b) Inverted lambda shaped flow-density model proposed by Koshi et al. (1983)

Based on the plot, the discontinuity is suggested between 70 to 100 veh/miles of density. Within this range, the two fitted models also overlapped on one another with the non-congested and congested models capturing the higher and lower flow observations respectively. This overlap was found of particular importance in later studies that proposed an inverted-lambda shaped flow-density plot.

Drake & Schofer (1966)) conducted statistical analysis to compare the fit and model parameters obtained by seven popular macroscopic models. Among them, Edie's model was the only discontinuous one, and it resulted the best estimates of fundamental parameters. Also, its standard error was lowest among all hypothesized models. Optimal density for multi-regime analysis was selected by a maximum likelihood approach in this study.

Kerner (1998) proposed a three-phase traffic model that includes a free flow phase and two parts of congested phase namely: synchronized flow and wide moving jam. Synchronized flow occur when the location of the front of the queue remains same, and the phase lasts for a long time. Wide moving jams are characterized by moving front of the queue at a constant speed, and it lasts for a short period. Although this paper did not discuss the discontinuity between the free flow and the synchronized flow phase, it did point out the fact that the maximum flow at the synchronized flow phase is substantially lower than the capacity of a freeway section (max flow of the free flow phase). Thus, the resulting flow-density diagram represents a three-phase discontinuous diagram.

The term "inverted Lambda shaped flow-density curve" was first used by Koshi et al. (1983) who analyzed traffic data from the Tokyo Expressway and acknowledged the presence of the discontinuity. Koshi reported that capacity drop phenomena is a common feature for congested flow condition and consequently, advocated the discontinuity in the fundamental diagram.

More recently, Zhang & Kim (2005) proposed a modified car-following theory to model capacity drop and traffic hysteresis phenomena. The Pipe's theory of car-following model (1966) was modified and four different models were proposed. By introducing a transition region of density in one of the models, a fundamental diagram that resembles the mirrored image of the reversed lambda was obtained. The model was developed by incorporating the concept varying “gap-time” which is defined as the time required for the follower to travel a gap-distance.

2.1.2 Continuous Model

Although many studies evidenced the presence of discontinuity in the fundamental diagrams, several researchers yet advocated the existence of a continuous relationship among the parameters. The speed-flow relationship for base conditions described in the 2010 version of Highway Capacity Manual (Transportation Research Board, 2010) can be interpreted as a three-regime continuous model. The free flow regime that spans from zero to a breakpoint density constitutes a straight-line relationship between flow and density. The transition regime represents a parabolic curve that spans from the breakpoint to a density of 45 pc/mi/ln. The congested regime (from 45 pc/mi/ln to jam density) represents a straight-line flow-density relationship. According to Chapter 10 and Chapter 25 of HCM, the flow-density curve for the oversaturated regime (3rd regime) assumes a simplified linear relationship. Others have argued against the existence of a discontinuity in the fundamental diagrams as a permanent character and advocated the continuous flow-density relationship (Li & Zhang, 2013; Cassidy, 1998).

Cassidy (1998) showed that the flow-density scatter plot forms a reproducible continuous curve if there is no downstream queue and the data is near-stationary. The study has a remarkable contribution in demonstrating the adverse effect of non-stationary observations for developing a bivariate relationship which is described in a later section. Near-stationary observations filtered from loop detectors placed at different locations of two freeway sites showed that the discontinuity in flow-density curve evidenced by past studies mainly forms because of the presence of a downstream queue.

Li and Zhang (2013), while modeling the time-space inhomogeneities with the kinematic wave theory of traffic flow demonstrated how the discontinuity in the fundamental diagram can theoretically violate the boundary conditions. Although the research stated that the discontinuity is frequently observed in the fundamental diagrams, it showed that a discontinuous fundamental diagram is non-differentiable at the discontinuous point and results in infinite characteristic wave speeds.

2.2 LITERATURE ON FREEWAY SAFETY AND TRAFFIC CHARACTERISTICS

The effect of detailed traffic characteristics on roadway safety is not that direct and has been a subject of important research since 1960’s (Gwynn, 1967; Turner, 1986; Frantzeskakis, 1987) More recent studies have shown that the safety of a road segment can be related to its degree of concentration of volume, level of service, and V/C ratio during peak hours (2010; Zhou & Sisiopiku, 1997; Persaud & Nguyen, 1998). Also, certain traffic characteristics that were found to be recurrently existed before crashes and thus, may have a causal relationship with crashes at any road segment were subjected to the investigation as well. Golob and Rucker (Golob & Recker, 2003) studied in detail the relationship between crash characteristics and real-time traffic flow variables, road lighting, and weather condition using nonparametric canonical correlation analysis. Lee et al. (2003) analyzed crash risks regarding traffic density, speed, and other geometric characteristics of roads. Several studies recognized the traffic state as a critical variable and developed their models for a target crash type, while most of the studies did not draw any border based on traffic states or crash types. In this section, detailed of the studies that investigated the relationship between freeway safety and traffic characteristics are discussed. The discussion is divided based on the traffic parameters used by the studies.

2.2.1 Density Based Traffic States as Predictors

Xu et al. (2012) evaluated the relationship between crash risks and traffic states defined by the occupancy of the freeway segment. It addressed several limitations of past studies on the same topic such as the arbitrary classification of traffic states, ignoring nearby traffic condition and effect of different confounding factors. To resolve these issues, the study considered the traffic state observations from four detectors for each crash, two on each direction from the crash occurrence location. The effect of confounding factors was addressed by the inclusion of three parameters namely the time, season, and location of the crash occurrence. Also, only day-time crashes were considered assuming that driver error or lighting condition cause most of the nighttime crashes.

To define different traffic states combining the traffic data obtained from four detectors for each recorded crash, the study exploited the technique of K-means clustering. This technique generates K groups by minimizing the intragroup distance and maximizing the intergroup distance. Next, a case-control strategy was used to relate the crashes with the traffic states 30 minutes prior to each crash. Here, the traffic states 30 minutes before a crash is considered as the "case". For each case, traffic characteristics data during the same clock time, season, and the location was randomly selected for four non-crash days. These observations are termed as "controls" (i.e. case: control =1:4). Such case-control studies are common in the epidemiological observational study.

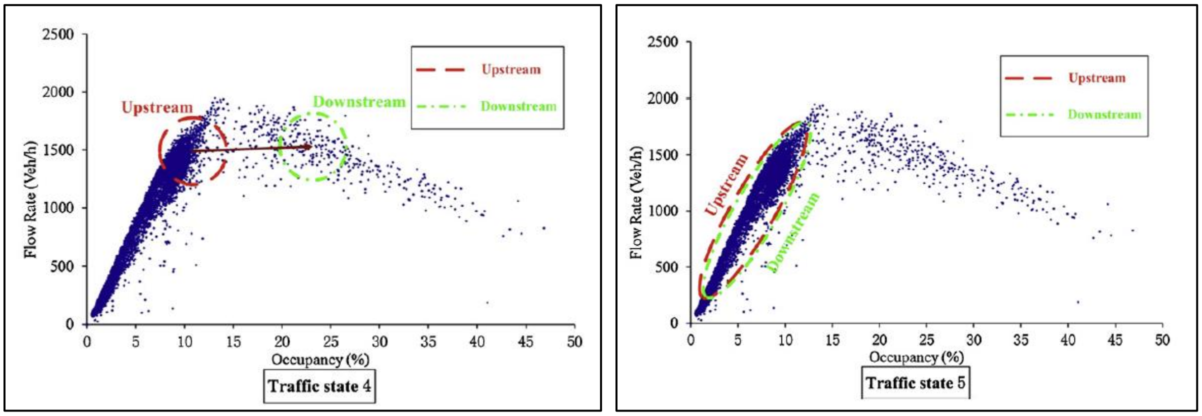
The final task of this study involves establishing a probabilistic model using logistic regression analysis. The probability of a traffic state being a "case" is expressed as-

$$P(Y = 1) = 1/\{1 + \exp[-(\alpha + \sum_{i=1}^{p} \beta_i x_i)]\} \quad (1)$$

Where,

Y=1,0 if a traffic state is a case or a control respectively;
α = Effect of matching variables (e.g. ,time, season, and location),
x_i = Value of explanatory variable i and β_i being its coefficient. Although this general form is writted to consider a vector of $i = 1,2,3, \dots .p$ variables, only the corresponding traffic state was used as the explanatory variable.
Traffic data from PeMS (2003) collected for the specific sites exhibited three traffic regimes namely free flow, congested, and transition regime. Combining the traffic regimes of four sensors 30 minutes prior to a crash, five traffic states (k=5 was obtained by maximizing the Spearman correlation parameter) were defined. Two out of these five states are demonstrated in Figure 2-2 (a) (Traffic state-4) and Figure 2-2 (b) (Traffic state-5). Three other traffic states,

although not shown here, were made up of different combinations of upstream and downstream states.



(a) (b)

Figure 2-2: (a) Traffic state-4 consists of congested condition at downstream and free flow condition at upstream of a crash, (b) Traffic state-5 consists of free flow condition at both upstream and downstream of a crash (Courtesy: Xu et al. (2012))

The relative odd-ratio of a traffic state that can signal a crash occurrence estimated from the logistic regression model showed that traffic state-4 and traffic state-5 has the largest and smallest value. It indicates that the crash risk was highest when the upstream detectors had free flow condition and the downstream detectors had congested condition. The crash risk was lowest when free flow condition prevailed at both upstream and downstream detectors. The authors subjectively defended the outcome of the models by pointing to the fact that Traffic state-4 indicates a positive speed difference between upstream and downstream detector before a crash. This fact was established in earlier studies (Lee, et al., 2003; Hossain & Muromachi, 2010). Also, another supposition backed by past studies is that a higher speed variance upstream of a crash location prior to the crash occurrence may increase the crash risk. This conjecture was validated in this study as the speed variance was found greatest in Traffic state-4.

2.2.2 Traffic Speed as a Predictor

Zheng et al. (2010) investigated the impact of traffic oscillations during congested condition on freeway crash occurrences. Since this study was focused on crash risk during congested condition, the outcomes are different than other studies. This study used high-resolution loop detector data fused with crash data collected for four years during the PM peak hours. The selected road segment experienced recurrent congestion during the PM peak hours which was verified by observing the speed contour profile.

The study applied a matched case-control strategy to relate crash risk with traffic characteristic of a certain period prior to each crash. By investigating the oblique curves of cumulative time mean speed, the study suggested that an optimum time duration of 10 minutes prior to a crash traffic data should be collected. In the proposed case-control analysis method, traffic characteristics (oscillation, average speed, count, and average occupancy) are considered as the cases and traffic characteristics at the same location weather condition on a non-crash day were selected as controls. A conditional logistic regression model is used to estimate the probability of a traffic characteristic being to be the prior situation of a crash. The model evaluation technique was fine-tuned since a repetitive sampling process of the control samples was implemented. The results revealed that in 90% of the repetitions, traffic oscillation was found significant. Other parameters representing average traffic conditions were significant only for 5% to 45% of the repetitions. Thus, the study established traffic oscillation as the single most significant predictor of crash risks during congested condition in freeways. Despite proposing a sound methodology and having concrete findings, the study consider crash type in its model. 85% of the recorded crashes were rear-ended. The result could be different if other types of crashes were present significantly in the dataset.

Song and Yeo (2012) used speed data collected from both upstream and downstream detectors of a crash location to develop relationships between crash rates and traffic speed related variables. The difference of upstream and downstream detector speed when a crash occurred was used to define four traffic states. A "back of queue (BQ)" condition was declared when the upstream detector had free flow, and the downstream detector had congested condition. The reverse situation was termed as a "Bottleneck (BN)". Free flow (FF) and congested (CT) conditions were identified as well using an arbitrarily selected speed threshold of 50 mph.

The statistical models developed for every four states expressed crash rates for target crash types (rear-end, sideswipe, and others) as a quadratic function of the speed parameter. This form of the equation was selected by observing the trend of crash rates against these variables. Two types of speed parameters were used separately: average speed and the difference in speed for the upstream and downstream detectors. These two variables were binned by 5 mph interval and crash rates for each bin were calculated using three years of crash data.

The regression models showed that read-end and sideswipe crashes were strongly related to the speed difference while other crash types were better related to the average speed. Collision potential was found to be highest in case for congested condition and lowest for free flow condition. Despite generating logical and statistically significant results, the study divided traffic states merely by an arbitrarily selected speed threshold. Besides, the study did not mention anything about the validation of the proposed models.

2.2.3 Peak Hour Traffic Volume as a Predictor

Chapter 18 and Appendix B of the Highway Safety Manual (HSM) (American Association of State Highway and Transportation Officials, 2010) discusses two methods of predicting crashes on freeway segments: Predictive method and Empirical Bayes (EB) method. The predictive method provides a regression model of the so-called Safety Performance Function or SPF (it represents the predicted average crash frequency for a year) as a function of the Annual Average Daily Traffic (AADT) and the length of a segment. Several Crash Modification Factors (CMF) have been developed by regression analysis to reflect different site characteristics and is applied with the SPF in a multiplicative form. Only one of the CMFs are related to traffic characteristics (high volume during peak hours) where all others are associated with the geometric characteristics of the site. According to HSM, the change in traffic volume can influence the crash frequency, type, or severity for both single and multi-vehicle crashes. The proportion of AADT occurring during the peak hours of an average day is used as a surrogate for the degree of traffic volume concentration.

The general format of estimating the predicted average crash frequency according to HSM is shown in Equation 2.

$$N_{p,fs,n,y,z} = C_{fs,ac,y,z} * N_{spf,fs,n,y,z} * \prod_{m=1}^{m} CMF_{m,fs,ac,y,z} \quad (2)$$

Where, $N_{p,fs,n,y,z} =$ predicted average crash frequency of a freeway segment with n lanes, crash type y, and severity z (crashes/yr);
$N_{spf,fs,n,y,z} =$ predicted average crash frequency of a freeway segment with base conditions, n lanes, crash type y, and severity z (crashes/yr);
$CMF_{m,fs,ac,y,z} =$ crash modification factor for a freeway segment with any cross-section ac, feature m, crash type y, and severity z; and
$C_{fs,ac,y,z} =$ calibration factor for freeway segments with any cross-section ac, crash type y, and severity z.
The crash modification factor representing the effect of peaking is shown in Equation 3.

$$CMF_{peak,fs,ac,y,z} = exp(a \times P_{hv}) \quad (3)$$

Where, $P_{hv} =$ proportion of AADT during hours where volume exceeds 1,000 veh/h/ln.
$a =$ regression coefficient. The value of this coefficient is positive for multi-vehicle crashes and negative for single vehicle crashes. It implies that as the concentration of traffic peaking rises, multi-vehicle crash frequency increases while single vehicle crash frequency decreases. Here, the threshold volume of 1,000 veh/h/ln is chosen because traffic stream speed tends to drop as volume increased beyond this value. Although this CMF does not include any detailed information regarding other traffic characteristics when a road segment is recurrently congested, it applies a simple yet rational model to capture the effect of peaking on freeway safety.

2.2.4 Volume and Density as Predictors

Lord et al. (2005) aimed to determine the statistical relationship using commonly applied predictive models between crashes and hourly traffic flow characteristics. Single and multi-vehicle crash records along with corresponding traffic data from loop detectors were collected for a 5-year period. Moreover, the model was developed separately for two site types (urban and rural).

The proposed method started with the assumption that the number of crashes at the i-th segment and t-th period, Y_{it} has a poisson distribution with a mean of μ_{it}. μ_{it} can be expressed in terms of traffic flow characteristics and segment length (f(.)). The model error is gamma distributed with a mean of 1 and a variance of $1/\varphi$, where φ is called the dispersion parameter. With this condition given, it can be showed that Y_{it}, conditional on f(.) and φ is distributed as a negative binomial random variable with a mean f(·) and a variance f(·)(1 + f(·)/φ) respectively.

Three forms of functions was used for f(·) in this study: the function of hourly volume, the function of hourly volume and density, and function of hourly volume and V/C ratio. Pearson statistics, deviance, and Cumulative Residual plots were used to evaluate the developed models. The model with the only volume as predictor yielded some unusual results such as single vehicle and total crashes increases with the increase in volume. This did not contradict intuition only but states the opposite of what was found from the linear trend of the crash data against the hourly volume. However, the inclusion of density or V/C yielded statistically sound model with logical outcomes.

2.2.5 Speed and Occupancy as Predictors

Abdel-Aty et al. (2004) used matched case-control technique for developing a real-time crash prediction model as a function of traffic characteristics prior to the crash. Although the case study application suffered from small crash data sample (375 crashes recorded over a period of 8 months), the methodology was sound as it addressed some key issues associated with the case-control technique.

Like other studies that used case-control strategy, it considered the traffic characteristic prior to a crash as a case and traffic characteristic on the same location, day of the week, and season were treated as control. Individual logistic regression models were developed with five traffic-related variables for each loop detector. Based on their respective "hazard ratio" (a parameter that quantifies the significance of each variable in predicting crash risk), the average occupancy of the upstream detector and coefficient of variation of the speed of the downstream detector were selected as the most significant variables. The descriptive statistics of the coefficient of variation of speed in the downstream detector for crash and no-crash conditions were described as shown in Figure 2-3. It clearly shows that the coefficient of variation of speed associated with the crashes has a higher mean and range compared to the non-crash observations.

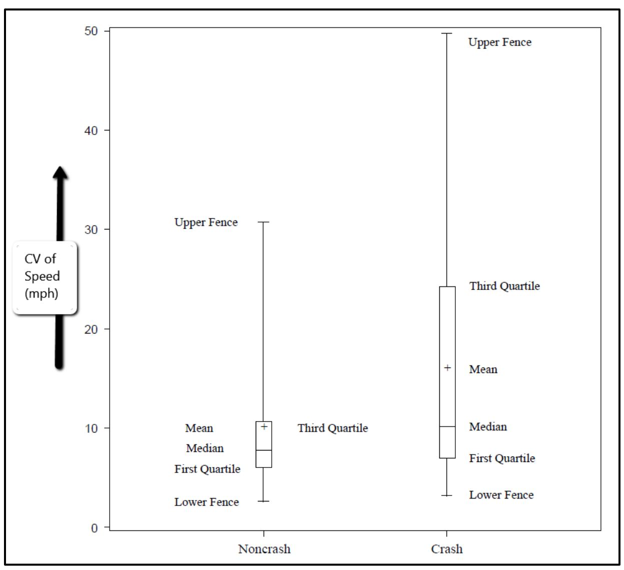


Figure 2-3: Box plot of coefficient of variation of the speed of the downstream detector Courtesy: Abdel-Aty et al. (2004)

Upon developing the logistic regression model, the odd-ratio for each traffic characteristic observation was calculated. A threshold value of 1 for the odd-ratio was selected to signal a crash. Application of this threshold identified about 69% of the crashes in the dataset (i.e., about 31% false negative). Given the applicability of the model in real-time, this result seems promising.

Pande and Abdel-Aty (2006) employed traffic surveillance data from freeway loop detectors to identify traffic conditions that are prone to rear-end crashes. Traffic speed, variation of speed, and average occupancy for both crashes (collected for a 5-year period) and no-crash scenarios were analyzed. Exploration of these variables using a Variable Importance Measure (VIM) revealed that rear-end collisions are associated with two types of traffic condition. The first type is associated with the congested condition, which can be signaled by the average occupancy and variation of speed of the detector closest to the crash location. The second type, which occurs when relatively free flow condition prevails 5-10 minutes before the crash can be predicted using the average occupancy and variation of speed data from a downstream detector. One of the key aspects of this study was that the proposed model was implemented to a different set of crash data. Despite the model identified about 75% of the crashes, the author acknowledged that the methodology induced a significant number of false alarms.

2.3 SUMMARY

The review of past studies on developing macroscopic traffic stream model revealed that the discontinuity in the steady state fundamental diagrams had been recognized from as early as the 1960s. Nonetheless, modern computational tools can improve the development of such models by using data collected over an extended period. In contrast, a few studies challenged the existence of fundamental diagram discontinuity.

The review on the link between roadway safety and traffic stream characteristics demonstrated that several studies applied probabilistic modeling technique to signal crash occurrence with traffic data. However, not many studies focused on the recurrent traffic characteristics that can be emerged and applied as a Crash Modification Factor in safety studies. Instead, the studies developed models using instantaneous traffic data almost immediately prior to a crash. Such models are challenging to implement in reality. Thus, it implies that further research is needed to find the link between roadway safety and persisting traffic stream characteristics.

3.0 METHODOLOGY

This chapter presents the mathematical formulation for developing a two-regime macroscopic model and the driver reaction time estimation for different density condition. From the algorithms for outlier detection to the estimation of driver reaction time, the overall process is described in a step-by-step manner. A description of the study area and the available data sources are presented toward the end of this chapter.

3.1 FITTING A MACROSCOPIC GHR TRAFFIC STREAM MODEL

3.1.1 Initial Cleaning of Data

For developing the proposed discontinuous steady-state model, traffic stream data from fixed-location sensors are needed for a long span of time (preferably one year). Since the model development is data-driven, inconsistent and mixed state data were removed in two stages of the process. This section presents the first stage, which involves applying three thresholds that are introduced in a prior study (Xu, et al., 2013). In this study, the thresholds are modified and applied in a slightly different manner to conform to the study objectives.

The first two thresholds applied to detect inconsistent observations are known as the Critical Speed Threshold (CST) and Critical Density Threshold (CDT). The combined application of these two thresholds removes the low-speed observations associated with low volume. Such observations exist due to inclement weather, work zones, incidents, or any other sort of capacity drop phenomena as well as from observations that include both congested and uncongested flow. CST is used to identify low speed (congested) observations. Analysis of traffic data reveals that 10 mph below the speed limit is a reasonable threshold to determine congested observations. CDT is used to identify low volume observations. A density of 35 pc/mi/ln, the threshold that separates LOS of D and E for urban freeways in the HCM is defined as the CDT threshold. Observations that fall below both CST and CST thresholds are tagged as inconsistent (low-speed and low-volume) data points. Figure 3-1 shows speed-flow data obtained from Traffic.com for calendar year 2013 and for sensor ID 23774 WB. The inconsistent data points are shown in Figure 3-1(a) bounded by the CST and CDT threshold lines (green points inside the large triangle).

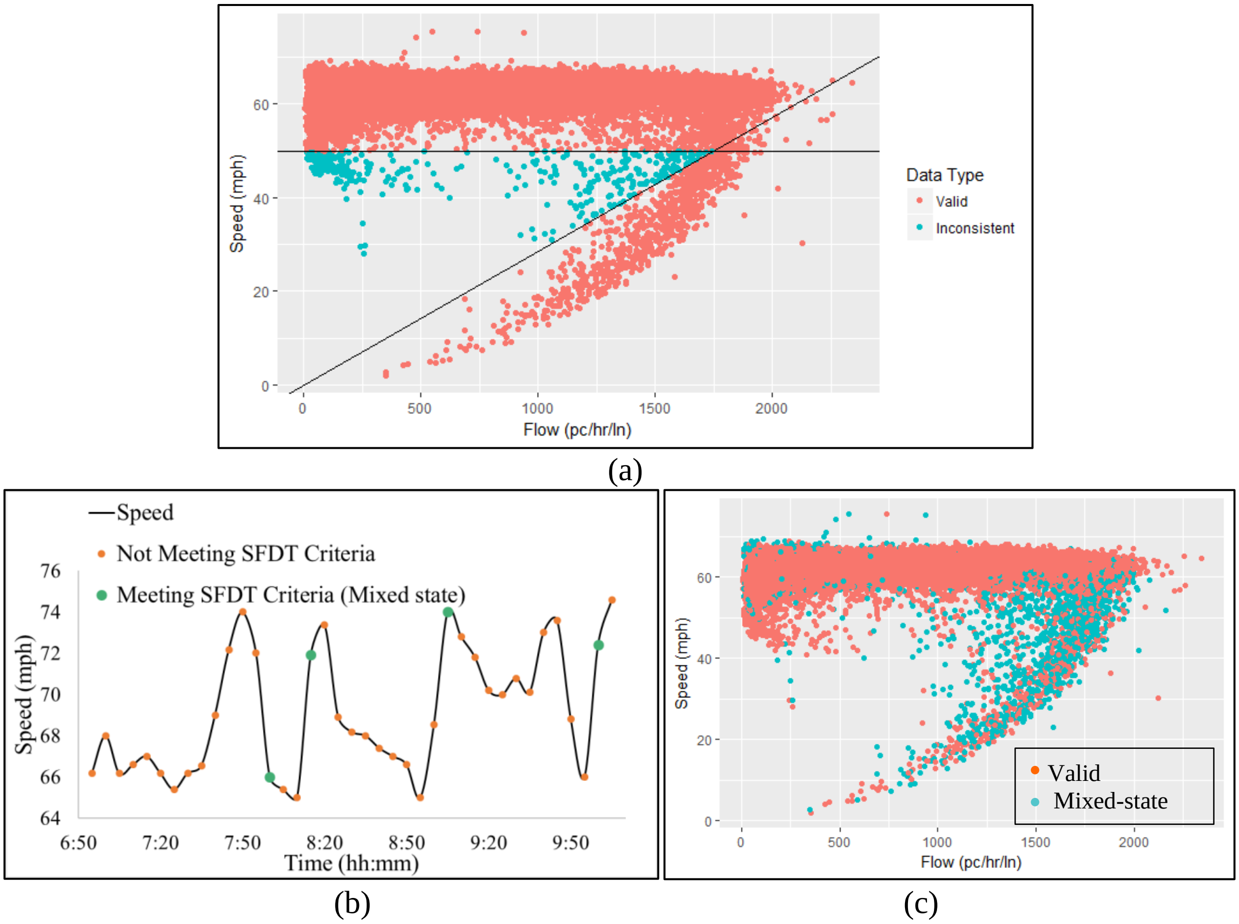


Figure 3-1: Inconsistent data filtered by CST and CDT (b) Mixed state observation identification by SFDT (c) Mixed state observations in the speed-flow diagram (23774WB)

To detect time interval (in this case, 5 minutes) data points that represent non-stationary traffic conditions, the third threshold termed Speed First Difference Threshold (SFDT) is applied. SFDT excludes observations whose speeds differ from the previous 5-minute observation by more than 10 mph as shown in Figure 3-1 (b). These non-stationary observations for sensor 23774W is shown in Figure 3-1 (c) by the green points. It should be noted that outliers generating from system measurement error may still exist even after applying these thresholds.

3.1.2 Fitting a Discontinuous Traffic Stream Model with an Overlap

The process of fitting the proposed steady-state model is explained in this section. The starting point is the flow density relationship for the single regime full GHR model (Gazis, et al., 1961) as shown in Equation 4.

$$q_{M,i} = k_i * u_f * \left(1 - \left(\frac{k_i}{k_j}\right)^{l-1}\right)^{\frac{1}{1-m}} \tag{4}$$

Where**,**

i=1,2,3 . . . Observation index

$q_{M,i}$= Model flow for observation I (pc/hr/ln)
u_f= Free flow speed (mph)
k= Observed density (pc/mi/ln)
k_j=Jam density (pc/mi/ln)
l= Distance headway exponent
m= Speed difference exponent

In a two-regime (uncongested and congested regime) GHR model, the same form of Equation 4 is used for each regime, and the regime boundary is determined by a density break point. In the proposed model, a transition range is introduced by allowing the two regimes to overlap each other which produces the recognized inverted lambda shape when plotted in the flow-density domain. The major challenge is to model this transition range. Empirical observations of traffic stream data depict that the data points in the transition range follow either the uncongested or the congested regime's characteristics. In other words, the transition range (defined in terms of density) includes a mixture of observations from both regimes. In this study, it is proposed to model each data point within the overlap by the regime model, either uncongested and congested, that results in the smallest absolute error.

The algorithm for modeling the uncongested and congested regimes along with the overlap range is defined by Equation 5, 6, and 7. The subscripts with the parameters represent the corresponding regime number

Observed Density	Formula for model flow	
$k_i \leq kb_2$	$q_{M,i,r=1} = k_i * u_{f1} * \left(1 - \left(\frac{k_i}{k_{j1}}\right)^{l_1-1}\right)^{\frac{1}{1-m_1}}$	(5)
$k_i \geq kb_1$	$q_{M,i,r=2} = k_i * u_{f2} * \left(1 - \left(\frac{k_i}{k_{j2}}\right)^{l_2-1}\right)^{\frac{1}{1-m_2}}$	(6)
$kb_2 < k_i < kb_1$	$q_{M,i,r} = \begin{cases} q_{M,i,r=1} \; if \; \lvert q_{M1,i} - q_i \rvert < \lvert q_{M2,i} - q_i \rvert \\ q_{M,i,r=2} \; Otherwise \end{cases}$	(7)

Where,
r= Regime index (1, 2)
kb_1= Density breakpoint for regime 1 (uncongested regime)
kb_2= Density breakpoint for regime 2 (congested regime)

In Equation (2) through (4), the two density breakpoints (kb_1 and kb_2) define the upper and lower limit of the transition range respectively.

3.1.3 Requirement for Additional Constraints

Before discussing the constraints added to fit the proposed model, it is important to explain the limitations of distinguishing the two regimes and defining the overlap based on empirical data.

All data in this study are from HERE roadside sensors (formerly Traffic.com) in the Raleigh, North Carolina urban area. Figure 3-2 shows an example application of the algorithm described above to clarify the limitations. In this figure, the observed data are for HERE sensor ID 23771 WB for the calendar year of 2013 aggregated at a 5-minute interval. The model was fitted with MATLAB software using a nonlinear optimization tool with an objective function of minimizing the sum of squared error of flow. The red points show the fitted model with no additional constraints, while the green ones show the same but with the inclusion of several constraints described later.

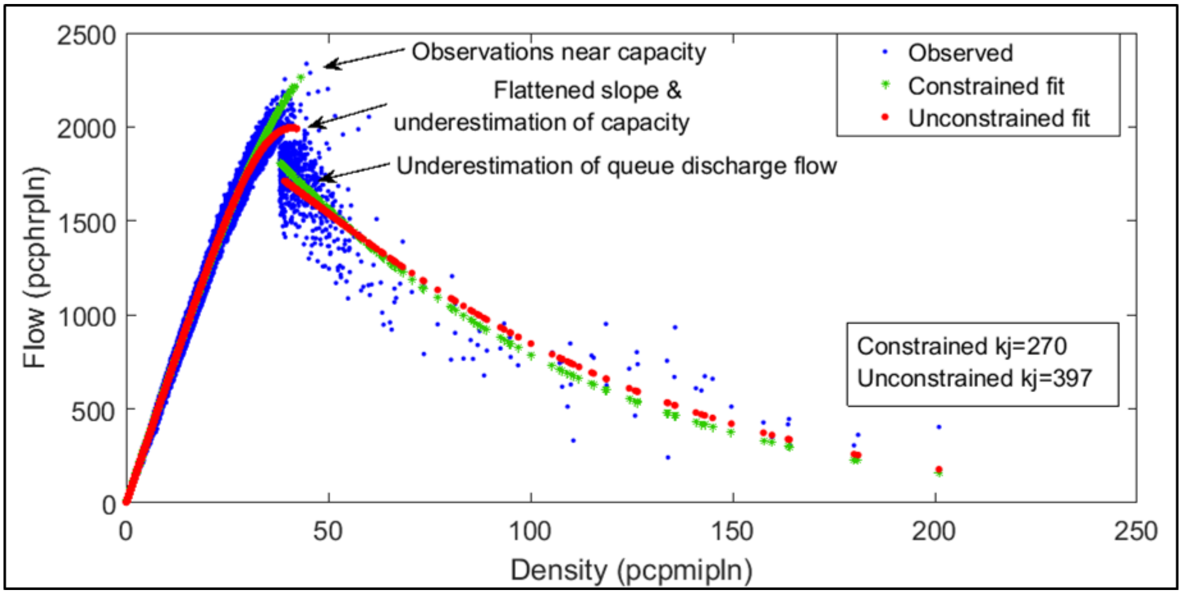


Figure 3-2: An example of the proposed flow-density model with and without any constraint

Several important aspects of the unconstrained fitted model shown in Figure 3-2 are noted below:

- The uncongested curve of the unconstrained model (red line) does not capture a group of data points with high flow (<2,000 pc/hr/ln). Instead, the slope of the uncongested regime curve flattens out at a flow value near 2,000 pc/hr/ln. It is apparent that this model poorly fits several important and valid observations. Assuming a free flow speed of 65 mph (posted speed limit at this location was 60 mph), the capacity of the road section should be 2350 pc/hr/ln according to HCM. Therefore, the fit of the uncongested curve can be improved by constraining the slope at capacity using this information.
- The queue discharge flow rate (or post-breakdown flow rate) is appeared to be 1,700 pc/hr/ln from Figure 3-2. Approximating the pre-breakdown flow rate from the observed data as 2,350 pc/hr/ln, it is found that the drop in post-breakdown flow rate is about 28% of the pre-breakdown flow rate. This percentage is significantly higher than the usual range mentioned in HCM (2%-20%).

The observations highlighted above are also present in the datasets for the other traffic monitoring sensors used in this study. The essence of the above discussion is that fitting the proposed empirical model without additional constraints may result in unusual and infeasible values that deviate from the findings of past research cited in the HCM. To resolve these issues, several constraints were implemented to properly fit the model. Although these constraints are selected empirically, they are justified by the guidelines provided in HCM.

3.1.3.1 Constraint on Queue Discharge Flow Rate

The post-breakdown flow rate is usually lower than the pre-breakdown flow rate, resulting in a significant loss of freeway throughput during congestion. According to a past study cited in Chapter 10 and Chapter 26 of HCM (2010), the average difference between the post-breakdown and the pre-breakdown flow rates vary widely from as little as 2% to as much as 20%. In the absence of local information, a default value of 7% is recommended. In light of this information, a maximum limit of 20% and a minimum limit of 2% of the pre-breakdown flow rate is imposed for the reduction in the queue discharge flow rate. The mathematical expressions for pre-breakdown and post-breakdown flow rates are shown in Eq. (8) and Eq. (9).

$$q_{pre} = k_{b1} * u_{f1} * \left(1 - \left(\frac{k_{b1}}{k_{j1}}\right)^{l_1-1}\right)^{\frac{1}{1-m_1}} \tag{8}$$

$$q_{post} = k_{b2} * u_{f2} * \left(1 - \left(\frac{k_{b2}}{k_{j2}}\right)^{l_2-1}\right)^{\frac{1}{1-m_2}} \tag{9}$$

$$\textit{Constraint}\text{: } q_{pre} * 0.98 \geq q_{post} \geq q_{pre} * 0.8$$

3.1.3.2 Constraint on Slope at Capacity

It is shown earlier in Figure 3-2 that the capacity is underestimated by the unconstrained model as the slope of the uncongested regime flattens out, and the regime continues to a high-density value. The slope of the flow-density curve at any point of the uncongested regime can be obtained by differentiating the model flow equation with respect to density.

$$\frac{dq_{M,i,r=1}}{dk_i} = u_f(\left(1 - \left(\frac{k_i}{k_{j1}}\right)^{l_1-1}\right)^{\frac{1}{1-m_1}} - k_i\left(\frac{1}{1-m_1} * \left(1 - \left(\frac{k_i}{k_{j1}}\right)^{l_1-1}\right)^{\frac{m_1}{1-m_1}} * \frac{l_1-1}{k_{j1}} * \left(\frac{k_i}{k_{j1}}\right)^{l_1-2}\right)) \tag{10}$$

Substituting $k_i = kb_1$ in Equation (10) yields the slope at capacity. To resolve the issue of decreasing slope at capacity, the slope at the capacity value in HCM is used as a guideline. According to Chapter 10 and Chapter 25 of HCM, the equations for flow in a basic freeway segment at different density regimes are shown in Equation (11).

$$q_{HCM,i} = \begin{cases} k_i * u_f \quad \forall\, k_i \in [0, k_b] \\ q_b + \dfrac{\sqrt{1 - 4 * \beta * q_b * k_i + 4 * \beta * u_f * k_i^2 - 1)}}{2 * \beta * k_i} \quad \forall\, k_i \in (k_b, 45] \\ q_c * \left(1 - \dfrac{k_i - 45}{k_j - 45}\right) \quad \forall\, k \in (45, k_j] \end{cases} \tag{11}$$

Where,
$q_{HCM,i}$=HCM model flow for observation *i*
β= A coefficient defined as a function of free flow speed
q_b= Flow breakdown (pc/hr/ln) as a function of free flow speed
q_c= Capacity of the segment (pc/hr/ln) as a function of free flow speed
k_b =Density at breakpoint (pc/mi/ln)

To obtain the slope at capacity equation, Equation (11) for density range 45 to $k_b$is differentiated with respect to the observed density and the following expression is obtained

$$\frac{dq_{HCM,i}}{dk_i} = \frac{1 + \dfrac{k_i(-4q_b\beta + 8\beta u_f k_i)}{2\sqrt{1 - 4q_b\beta k_i + 4\beta u_f k_i^2}} - \sqrt{1 - 4q_b\beta k_i + 4\beta u_f k_i^2}}{2\beta k_i^2} \tag{12}$$

In Equation (12), plugging $k_i = 45$ gives the slope at capacity value according to the HCM model. The proposed constraint here is that the slope at capacity according to the proposed model must be greater than or equal to the HCM slope at capacity value.

3.1.4 Robust Regression

It is mentioned earlier that the use of robust regression techniques to remove outlier bias is gaining popularity. Although the initial stage of model development process involves removing clearly inconsistent and non-stationary data from the raw datasets, outlying observations are likely to remain in the data after the stage one filters. Since the method described above is data-driven, it is imperative to fit the model through valid observations only. A customized application of robust regression method is proposed in this study. According to this approach, the standard error for each data point is estimated by fitting the model. Data points with a standard error higher than a certain threshold are removed from the original dataset. Then, the model is fitted again with the updated dataset. The process is continued until the maximum standard error becomes lower than the threshold.

In this regard, the selection of threshold is critical. First, there is a significant difference in the number of data points in the uncongested and congested regime. Therefore, it is essential to distinguish the two regimes and estimate the standard error for each of them separately. The estimation of the standard error for each regime is shown in Eq. (13)

$$SE_{i,r} = (q_{M,i,r} - q_{M,i,r})/Std_r \quad (13)$$

Where,
r= *1 or 2* for uncongested and congested regimes respectively
$q_{M,i,r}$=Modeled flow for observation i in regime r
$q_{i,r}$=Observed flow in regime r
Std_r= Standard deviation of flow error for regime r

The second issue with robust regression is that there are two types of observations that need to be removed: mixed state observations and observations with extreme measurement error. The mixed state data points are valid observations that represent time intervals in which there was a switching between congested and uncongested regimes. Following the initial cleaning described above, such observations that remain are expected to be more prevalent in the congested regime than in the uncongested regime as illustrated in Figure 3-1 (b). The distribution of the mixed state observations will be asymmetric around the flow-density curve of the congested regime (lying on the left side of the curve) Because these observations do not result from measurement error, they are likely to be more prevalent than measurement errors. For example, for every flow breakdown, there is likely to be two pronounced mixed state observations, one for the time interval when the queue formation shock wave passed by and one for the time interval when the queue clearance shock wave passed by. Considering these facts, a symmetric threshold for the standard error of ±3.5 is applied for the uncongested regime. This symmetric threshold is expected to remove the observations with high measurement error in the uncongested regime. For the congested regime, an asymmetric threshold of +2 is applied for all but the final step of the robust regression. This threshold is expected to remove the remaining mixed state data points and the measurement error outliers on the left side of the flow-density curve. In the final step of robust regression, a threshold of -3.5 is applied to exclude any remaining outliers from the congested regime on the right side of the flow-density curve.

The symmetric threshold of ±3.5 standard error applied on the uncongested regime observations represents a confidence interval of approximately 99.95% (assuming that the errors have Gaussian distribution). On the other hand, a less conservative threshold is applied to the congested regime due to the potential presence of mixed state observations. With these set of thresholds, the robust regression process is expected to converge without removing excessive data points.

3.2 ESTIMATION OF DROP IN DRIVER REACTION TIME

3.2.1 Estimating Driver Reaction Time for Different Traffic States

To estimate the driver reaction time required for asymptotic stability, the macroscopic form of the GHR model in Eq. (1) needs to be converted to a microscopic form. The microscopic form of the fifth and final GHR car-following model is expressed in Eq. 2. Here, the acceleration of the *(n+1)th* vehicle in a traffic stream at a time $(t+\Delta t)$ (termed as $x''_{n+1}(t+\Delta t)$) in response to the relative speed between the *nth* and *(n+1)th* vehicle at time *t* is expressed as the product of the relative speed between the two vehicles and the sensitivity term.

$$x''_{n+1}(t+\Delta t) = \alpha \frac{[x'_{n+1}(t+\Delta t)]^m}{[x_n(t)-x_{n+1}(t)]^l} * [x'_n(t) - x'_{n+1}(t)] \quad (14)$$

Where,

n= Position of a driver in a traffic stream. (n=0 is the most downstream driver)

x_n= Location of the nth driver with respect to a reference point.

$x'_n(t)$= Speed of the nth driver at time *t*

This format of the model is also called the stimulus-response model. According to May (1990), the parameter α can be expressed as shown in Eq (15).

$$\alpha = \frac{(l-1)u_f^{1-m}}{(1-m)k_j^{l-1}} \quad (15)$$

Thus, the sensitivity factor is equivalent to what is shown in Eq. (16)

$$Sensitivity\ factor = \frac{(l-1)u_f^{1-m}}{(1-m)k_j^{l-1}} * \frac{[x'_{n+1}(t+\Delta t)]^m}{[x_n(t)-x_{n+1}(t)]^l} \quad (16)$$

For steady-state observation, individual vehicle speed represents the speed of the traffic stream and the spacing between two successive vehicles represents the inverse of the traffic density. Thus, Eq. (17) can be written as

$$Sensitivity\ factor = \frac{(l-1)u_f^{1-m}}{(1-m)k_j^{l-1}} * \frac{u^m}{\left(\frac{1}{k}\right)^l} \quad (17)$$

Now, according to May (1990), for a traffic stream to be asymptotically stable, the product of the reaction time and sensitivity must be less than or equal to 0.5. Hence, the expression for the required reaction time for asymptotic stability can be derived as:

$$t_i \leq \frac{(l-1)u_f^{1-m}}{2(1-m)k_j^{l-1}} * \frac{u_i^m}{\left(\frac{1}{k_i}\right)^l} \quad (18)$$

Eq. (18) gives the formula to estimate the driver reaction time required for stability for each observation of flow, speed, and density.

3.2.2 Drop in Drivers Reaction Time Between Two Regimes

Fitting a discontinuous flow-density model enabled the research team to investigate the drop in driver reaction time when the traffic state moves from the free flow (regime 1) to the congested flow regime (regime 2). If the reaction time described in Eq. (18) is estimated for a series of density values, the following curves are obtained.

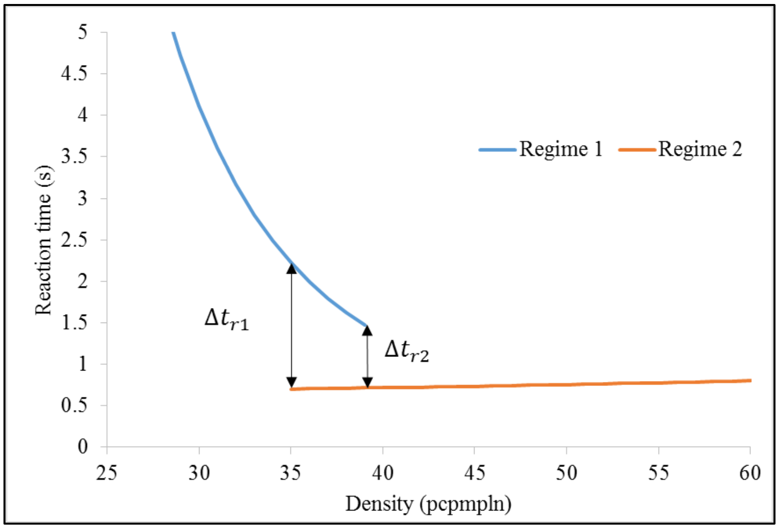


Figure 3-3: A typical driver reaction time vs. density plot

Figure 3-3 shows a typical reaction time vs. density plots. To investigate the relationship between crash rates and driver reaction time, selecting the appropriate reaction time is critical and the logic behind the selection is described here.

Plotting this reaction time vs. density diagram for different locations revealed that the reaction time for the congested regime does not vary significantly. Moreover, the reaction time for a very low-density value (such as less than 20) may not have any significance as traffic stream barely follows the car-following model**.** On the other hand, if the transition regime is focused here (bounded by the vertical arrows in Figure 3-3), there are two reaction times for each density point within this regime. As each vehicle transfers from the free flow to the congested regime, the required driver reaction time also drops to get adapted to the change in traffic state. The research team hypothesizes that higher the drop in reaction time, higher the risk of getting involved in the target accident type.

The two drops in reaction time shown in Figure 3-3 are considered as of particular interest. These two drops termed as Δt_{r1} *and* Δt_{r2} are the drops in reaction time at the beginning and the end of the transition regime respectively. However, since the end of transition regime is sensitive to the break point density for regime 2 (38 pc/mi/ln in Figure 3-3), the primary focus is given on the drop at the beginning of the transition regime (Δt_{r1}).

3.3 TARGET CRASH TYPE

Not all the freeway crashes can be related to the traffic characteristics or the reaction time derived from a traffic flow model (e.g., animal or debris related crashes). Moreover, some crash occurrence cannot be related to the car-following behavior of drivers (side-swipes). Since the estimated driver reaction time is based on the asymptotic car-following stability, rear-end crashes were selected as the target crash type. Moreover, given that the transition between flow regimes is more likely during heavy peak-hour travel, peak-hour crash rates were also analyzed. Peak-hour crashes are defined empirically in this study as the crashes that occur either from 6 am to 9 am or from 4 pm to 7 pm. However, it is likely that the rear-end and peak-hour crashes are correlated.

3.4 STUDY AREA AND DATA COLLECTION

To test the relationship between crash rates and driver reaction time, side-fire radar data for different freeway locations near the Triangle Area of Raleigh, North Carolina is collected. The collected dataset contains the flow, speed, and lane occupancy for both directions of a freeway section for the year of 2013 in a time resolution of 5 minutes. Crash data for the segments surrounding the sensor locations from 2011 to 2013 is extracted from NCDOT's accident archive and crash analysis tool called "Traffic Engineering Accident Analysis System (TEAAS)" (North Carolina Department of Transportation, 1988).

To select the side-fire radar sensors and the associated segments for crash rate analysis, following conditions were considered. To classify the segments of the selected freeway locations, Highway Capacity Manual's (Transportation Research Board, 2010) segmentation rule is applied.

- To avoid mandatory lane changing phenomena of vehicles, only sensors that are within a basic freeway segment are selected
- Google Earth's (Google Inc., 2017) historical imagery tool was used to make sure that the selected locations did not experience any significant change in road geometry and surrounding conditions from 2011 to 2013.
- The selected sensors' dataset does not contain a significant number of outliers and hence, enables to fit the traffic model properly.
- Each segment should be homogenous in terms of the number of lanes throughout their entire length.

After filtering according to these criteria, in total 21 directional segments were selected. The locations of these 21 sensors are shown in Figure 3-4 below. The numbers show the tag of each sensor.

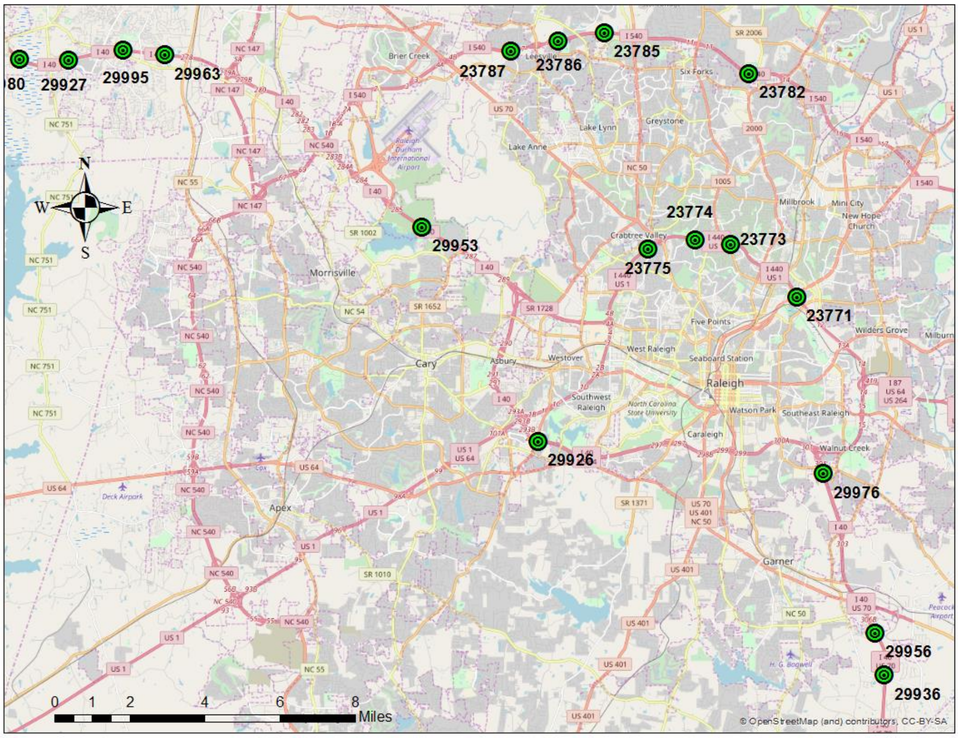


Figure 3-4: Location of the Sensors

4.0 ANALYSIS AND RESULTS

4.1 FITTING A MACROSCOPIC MODEL TO SENSOR DATA

4.1.1 Application of CST, CDT, and SFDT

This section summarizes the two stages of cleaning inconsistent, mixed state, and outlying observations. Figure 4-1 shows the percentage of data reduced by the initial thresholds (CST, CDT, and SFDT) and by robust regression.

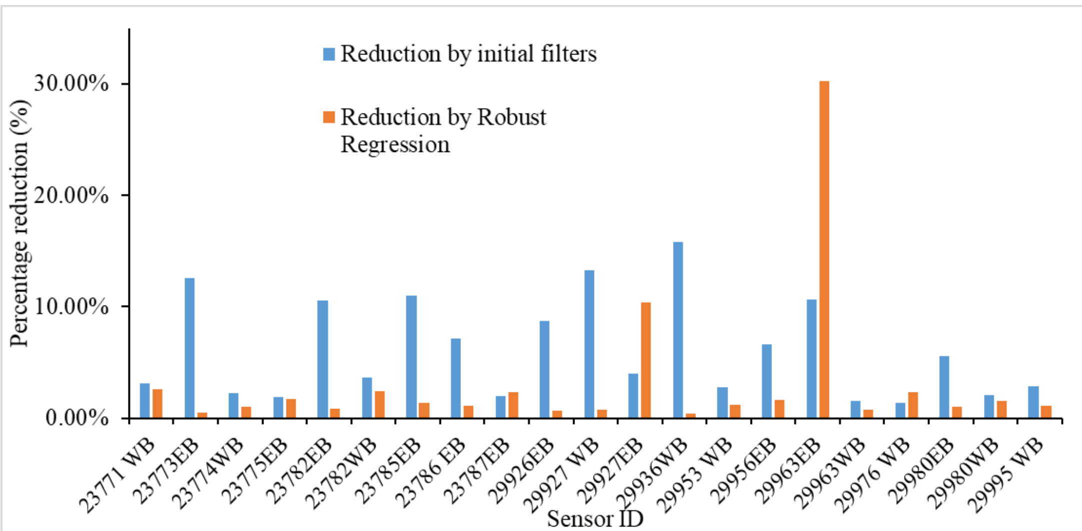


Figure 4-1: Percentage of data reduction in different sensors

The most striking observation from Figure 4-1 is the higher percentage of data reduced by the robust regression for sensor 29963EB (~30%) compared to that for other sensors. Data exclusion by robust regression is lower than that by the initial filters for all other sensors. The highest proportion of data reduced by the initial thresholds is about 16% (29936WB). The proportion of outlier detection may also depend upon the health of the detectors.

4.1.2 Fitted Two-Regime GHR Traffic Stream Models

A non-linear optimization tool embedded in a MATLAB (MathWorks, 1984) environment is used to develop the proposed model for the 21 sensors. The fundamental diagrams for randomly selected five sensor locations are shown in Figure 4-2 below. The final parameter values used to plot the flow-density and speed-flow fundamental diagrams shown in Figure 4-2 are the parameters resulting from convergence of the robust regression algorithm. The estimated values along with their standard deviations are listed in Table 4-1. Parameter values for all 21 sensors are provided in Appendix B.

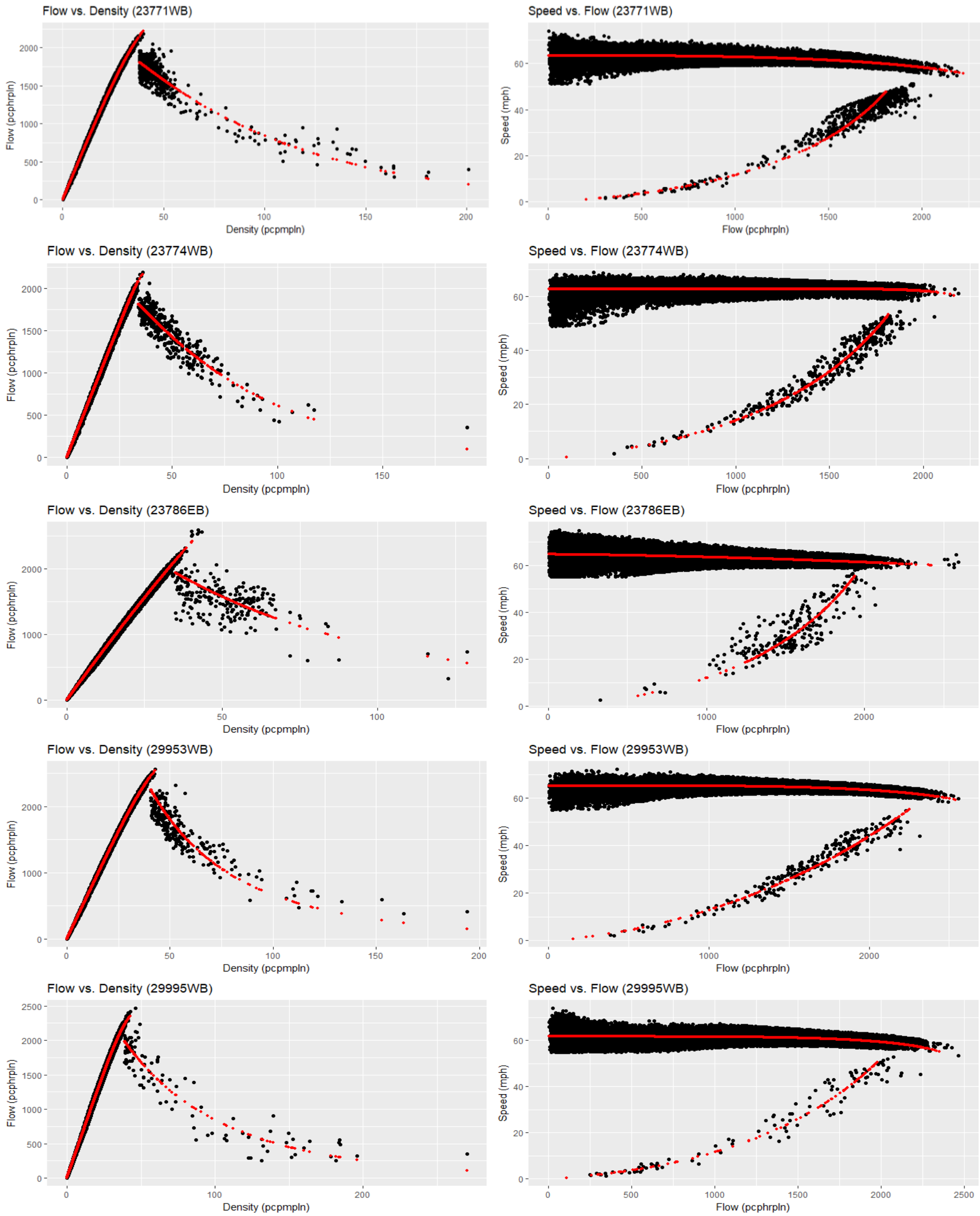


Figure 4-2: Flow-density and Speed-flow diagrams for 5 sensors

Table 4-1: Parameter estimates for different sensors obtained from the fitted models for five selected sensor locations

Sensor ID	*uf_1 (mph)*	*kb_1 (pc/m/l)*	*kj_1 (pc/mi/ln)*	*l_1*	*m_1*	*uf_2 (mph)*	*kb_2 (pc/mi/ln)*	*kj_2 (pc/mi/ln)*	*l_2*	*m_2*
23771 WB	63	39.9	606	3.8464	0.9967	474662	35.5	510	1.0374	0.7415
23774 WB	63	37.2	69	11.2073	0.9707	537130	33.2	345	1.0347	0.7222
23786 EB	65	40.4	251	2.4451	0.0056	100000	34.9	1001	1.0976	0.8298
29953 WB	65	48.0	572	4.8622	0.9995	1447754	40.0	1800	1.1102	0.8943
29995 WB	62	43.7	256	5.4517	0.9970	100000	38.7	1160	1.1101	0.8463

In general, the plots in Figure 4-2 illustrate that the fitted models reasonably follow through the steady-state observations. Here, the value of the parameters that have physical interpretation needs to be discussed. The distance head exponent (*l*) and speed exponent (*m*) for both regimes do not have any straightforward physical meaning. Same statement applies to the jam density and free flow speed for the uncongested and congested regime respectively. From the fitted speed-flow plots in Figure 4-2, the free flow speed for the uncongested regime (ffs_1) is evident. The density breakpoint for the uncongested regime (kb_1) is less compared to the HCM default of 45 pc/mi/ln for all but one (29953W). This parameter along with ffs_1 are of importance for estimating the capacity of the segment at the sensor locations. According to the HCM default method, the capacity of a basic roadway segment varies between 2,320 to 2,350 pc/hr/ln for the given range of ffs_1 (62-65 mph). However, the site-specific capacities appear from the plots in Figure 4-2 are apparently less for sensor 23771 WB and 23774 WB and more for sensor 23786 EB and 29953 WB than the HCM defaults. The fundamental diagrams for all segments that are provided in Appendix A reveal that the model capacity deviates from HCM specified value for 13 out of 21 sensors. The close agreement of HCM default capacity to only 13 sensors' data underscores that the national average capacity values provided by HCM need to be calibrated with field data if high fidelity analysis is desired.

The difference between the two density break points represents the overlap range. This range appears to be a unique characteristic for the five locations based on the proposed approach modeled. The fitted overlap in density varies from 2.5 to 8 pc/mi/ln across all sensors. This site-specific modeling of overlap along with the drop-in capacity has many potential applications such as estimating pre-breakdown and queue discharge flow rate, identifying recurring freeway bottlenecks, etc.

Jam density across all sensors varies within a wide range of about 229 to 4207 pc/mi/ln. These unusual estimates of jam density are resulted because of the absence of sufficient observations near jammed condition. Such unusually large jam density values interpret that the average vehicle length in the traffic stream can be as small as 1.26 ft., which is not physically possible. Although the reasonable limit of jam density specified in the HCM (150-270 pc/mi/ln) could have been applied to the model as a constraint, it may impact the estimate of other parameters due to the lack of jammed observations.

4.1.3 Reaction Time Calculation

Driver reaction time required to maintain stability is calculated using Equation 18. The parameter values obtained from fitting the model is applied to different density values and the two drops in reaction time are estimated as demonstrated in Figure 4-3. Similar plots are generated for the five sensors in Table 4-1. Values of the two drops in reaction time for all 21 sensors are provided in Appendix B.

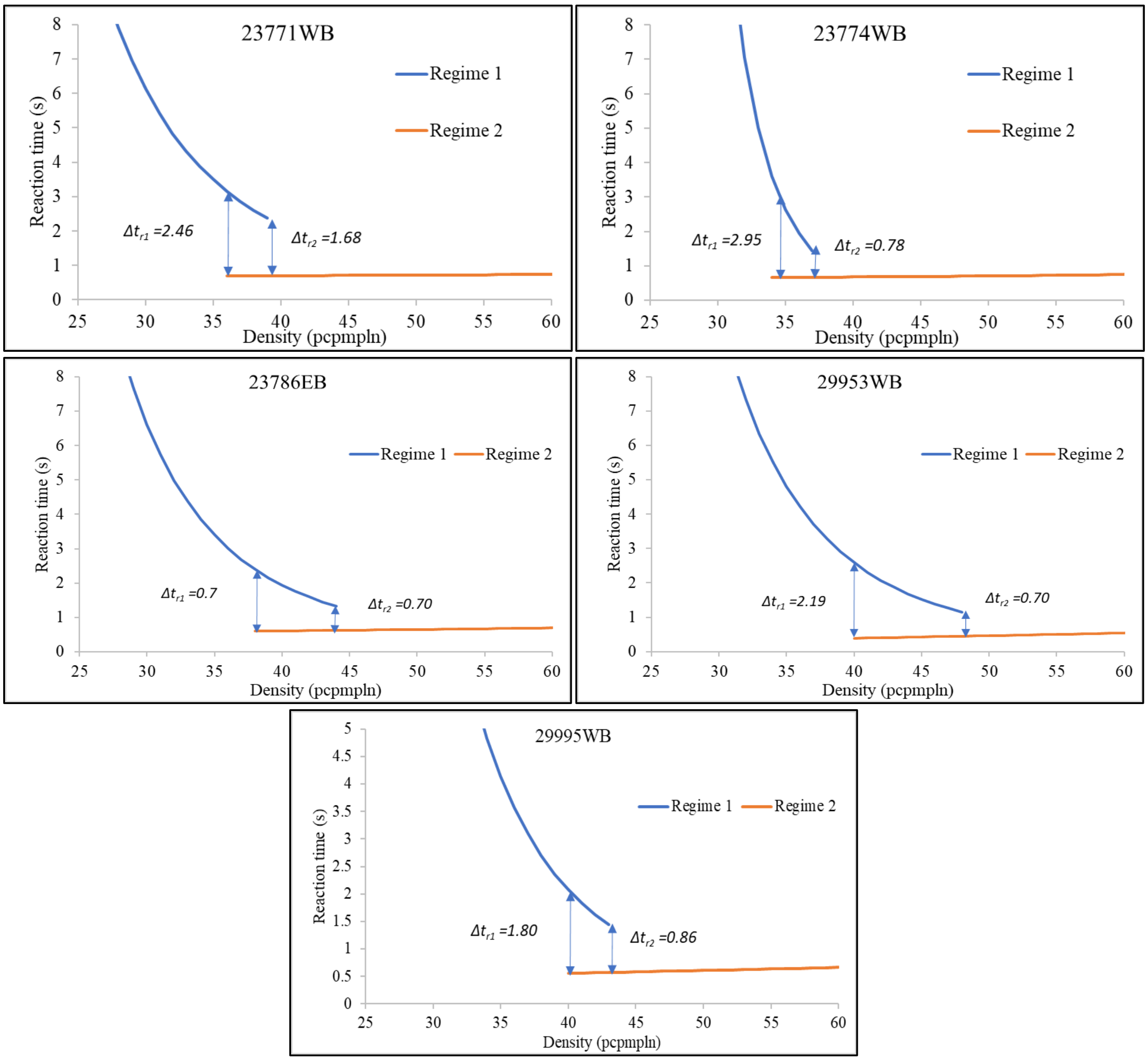


Figure 4-3: Reaction time and its drops during transition for five sensors

It is apparent from these plots that the value of the two drops in reaction time is sensitive to the two density breakpoint values. Therefore, it is important to fit these breakpoints as well as the overlap properly. Δt_{r1} which represents the difference in reaction time between the two regimes at a density equal to kb_2 varies from 1.22 to 2.95 seconds. Δt_{r2} is the difference in reaction time between the two regimes at a density equal to kb_1, and it ranges from 0.68 to 1.68 seconds. As the drop in required reaction time is estimated for all sensors, the remaining task is to analyze the crash data at these locations and to investigate the correlation of crash rates with these drops.

4.2 ANALYSIS OF CRASH DATA

Figure 4-4 shows the total, rear-end, and peak-hour crash rates for all the selected segments. It is clear that for some segments, rear-end or peak-hour crashes are the major crash type (23774WB, 29927EB, 29927WB etc.). The lowest percentage of rear-end crash rate relative to the total crash rate is 14% (segment at sensor 23782EB).

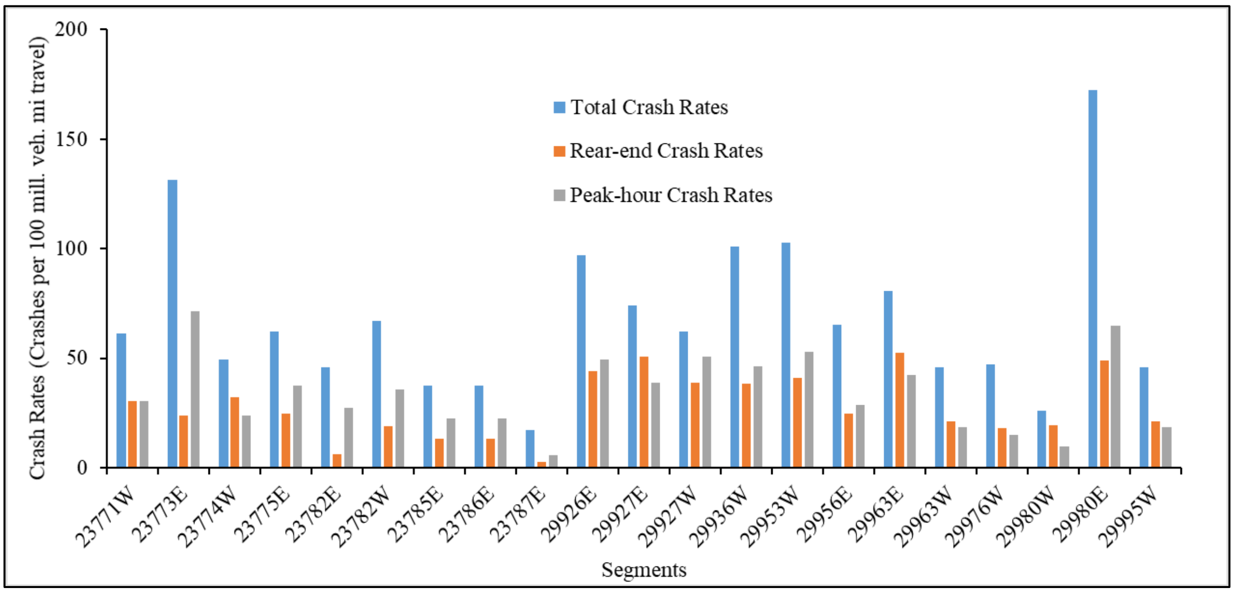
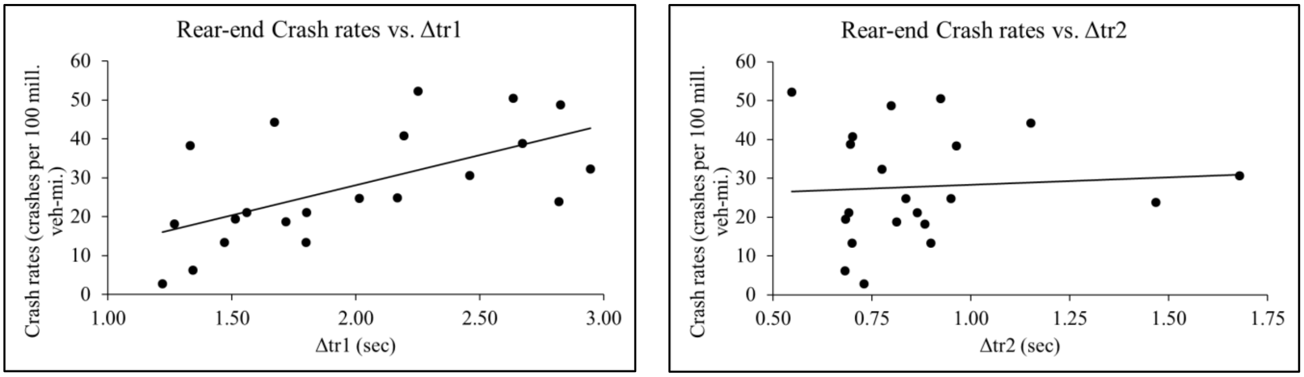


Figure 4-4: Total, rear-end, and peak-hour crash rates for different segments

Another important observation from Figure 4-4 is that for many segments, the rear-end and peak-hour crash rates are close to each other. However, for some segments, they are quite different (23773EB, 23782EB, 23782WB etc.). To check whether these two crash types are correlated or not, a statistical hypothesis test is conducted. The null hypothesis of the test is that the true correlation between these two parameters is equal to zero and the alternate hypothesis is that it is not. The test result gives a *t-statistics=3.845, p-value=0.001, and Pearson Correlation Coefficient=0.66.* The small *p-value* rejects the null hypothesis and the *Pearson Coefficient* value proves that there is a moderately strong correlation exists between these two crash rates.

4.2.1 Crash Rates vs. Drop in Reaction Time

Figure 4-5 demonstrates the relationship between the rear-end crash rates and the drop in driver reaction times. The straight lines show the linear trend between these two parameters.



(a) (b)

Figure 4-5: Rear-end crash rates vs. (a) Δtr_1, (b) Δtr_2

From visual observation, it is evident that there is no correlation exists between the rear-end crash rates and Δt_{r2}. It could be because this drop in reaction time is sensitive to the selection of breakpoint density for the uncongested regime, which makes this predictor less robust to the outliers. On the other hand, a positive trend is seen in the rear-end crash rates vs. Δt_{r1} plot. To find the details of this trend, a linear regression model is fitted and the results are shown in Table 4-2 below.

Table 4-2: Regression analysis for rear-end crash rates and $\Delta tr1$

Residuals	Minimum	1st Quartile	Median	3rd Quantile	Maximum
	-16.856	-6.504	-3.485	7.946	21.229
Coefficients		**Estimate Std.**	**Error**	**t-statistics**	**Probability (>|t|)**
	Intercept	-2.685	9.485	-0.283	0.7802
	Δt_{r1}	15.408	4.605	3.346	0.00339 **
Residual Standard Error		11.68 on 19 degrees of freedom			
Multiple R-squared		0.3708			
Adjusted R-squared		0.3376			
F-statistics		11.2 on 19 Degree of freedom			
p-value		0.0033			

From the regression analysis, it is evident that there is a strong correlation exists between the rear-end crash rates and Δt_{r1}. The coefficient of Δt_{r1} is significant at a level of 0.01, where the intercept does not have much significance. However, the interpretation of the intercept term here does not have any meaning here since $\Delta tr1 > 0$ is true for any location.
Figure 4-6 shows the variation of peak-hour crash rates with the drop in driver reaction times.

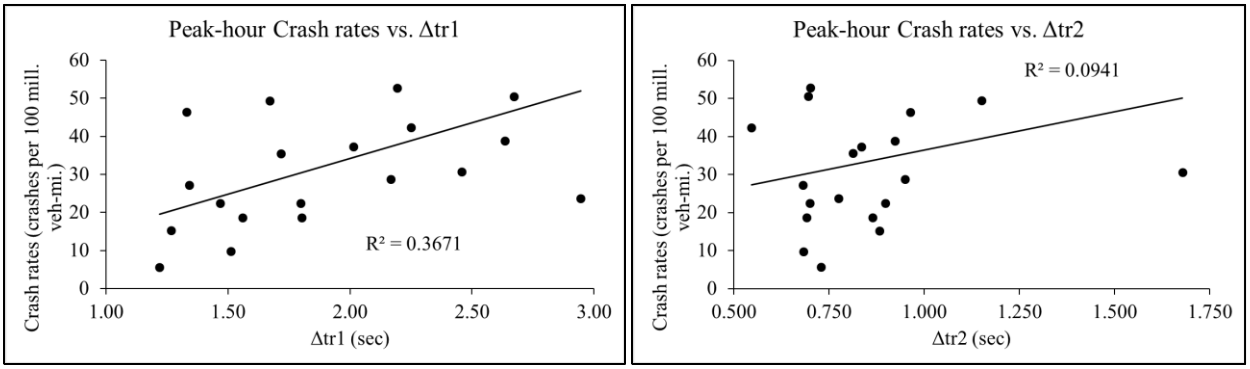


Figure 4-6: Peak-hour crash rates vs. (a) Δtr_1, (b) Δtr_2

Visual observation and the R-square values shown in Figure 4-6(a) and 4-6(b) suggests that the variation of peak hour crash rates with Δtr_1 and Δtr_2 is similar to what is observed for rear-end crash rates. As it is shown above that peak-hour crashes are correlated with rear-end crashes, the correlation of peak-hour crash rates with Δtr_2 serves as an alternate validation of the hypothesis that rear-end crash rates are correlated with Δtr_2.

4.3 SUMMARY OF THE FINDINGS

The application of the proposed two-regime macroscopic model to the sensor data showed reasonable fits. The fitted parameters values deviated significantly from the values specified in the Highway Capacity Manual, which evidences the necessity of estimating site-specific values for free flow speed, density breakpoints, and capacity. The positive correlation of rear-end crash rates on the freeway segments with the drop in reaction time (Δtr_1) signals that the change in traffic state may play a significant role behind the mechanism of rear-end crashes.

5.0 CONCLUSIONS AND RECOMMENDATIONS

Based on the analysis of the results obtained and discussed in Chapter 4, the key findings obtained from this study are summarized here. The first section of this chapter discusses the findings on fitting a two-regime GHR traffic stream model. Next, key results from the estimation of drop in driver reaction time and its correlation with crash rates are discussed. The limitations of the study and recommendations for future research are presented in the last section.

5.1 RESULTS FROM FITTING THE MACROSCOPIC MODEL

The study proposes a method for fitting the well-known Gazis-Herman-Rothery (GHR) macroscopic model. A discontinuity between the transition of the uncongested and congested regime is introduced which based the formation of the so-called inversed lambda shaped flow-density diagram. Such discontinuity enabled the estimation of the drop in driver reaction time during the transition from uncongested to congested regime.

Macroscopic speed and flow data were collected for the calendar year of 2013 from 21 road-side sensors located at three different Interstates near Raleigh, North Carolina. Mixed state observations are removed by applying thresholds on critical speed and density and a time series-based threshold on speed data prior to fit the model. An iterative robust regression method is applied to remove remaining outliers. A non-linear optimization tool with an objective function of minimizing the sum squared error of flow was applied to the observed data. Two constraints justified by the Highway Capacity Manual (Transportation Research Board, 2010) were applied to fit the models properly. The transition regime observations were modeled based by contrasting their conformity with congested and uncongested regime model. The key results obtained from the model fitting part are-

- The reduction of data by the initial filters varied across different sensors from 1.4% to 15.8%. Although the percentage reduction by robust regression is comparatively less than that for 20 sensors (0.4-10.4%), the percentage is unusually high for the sensor 29956EB (30.3%). These percentages may depend on the health of the sensors.
- The fundamental diagrams resulting from fitting the two-regime GHR model showed reasonable fits. The estimated free flow speed across 21 locations varied from a little over 56 mph (29927EB) to 68 mph (23782W). The fitted transition states varied from 33 pc/mi/ln to 48 pc/mi/ln, while the overlap or the difference between these ranged from 2.5 (23787EB) to 8 (29953W) pc/mi/ln. With all these parameters having values within reasonable ranges, the jam density at the congested regime showed an odd characteristic. It spanned from 229 to 4207 pc/mi/ln which led to the interpretation that the average vehicle length in the traffic stream is too small. Other parameters (free flow speed of congested regime, jam density of uncongested regime, *m*, *l*), despite being important in fitting the models, do not have any physical interpretation.

- The calculation of driver reaction time for different density values revealed that the two drops in reaction time during the transition (Δtr_1 and Δtr_2) are sensitive to the two density breakpoint values. The value of Δt_{r1} and Δtr_2 varied from 1.22 to 2.95 seconds and 0.68 to 1.68 seconds respectively.

5.2 RESULTS FROM ANALYSIS OF CRASH DATA

Crash data for the segments surrounding the 21 sensor locations were collected for a three-year (2011-2013) period. Crash rates were calculated in 100 million vehicles-miles traveled. Key findings from analyzing the crash data and its correlation with the drops in driver reaction time are listed below-

- Crash rates of the selected segments varied from 17 to 173 crashes per 100 million vehicles-miles traveled. The percentages of rear-end and peak hour crash rates diverse from 14% to 75% and from 33% to 82% respectively. The similarity of these two crash types led to a correlation analysis which revealed that a moderately strong correlation exists between them.
- The drop in reaction time at a density equal to kb_1 (Δt_{r2}) did not exhibit any correlation with either total, rear-end, or peak hour crash rates. The reason could be that the reaction time at uncongested regime is not that stable and sensitive to the parameter estimates.
- The drop in reaction time at a density equal to kb_2 (Δt_{r1}) showed a positive correlation with rear-end crash rates. A linear regression analysis evidenced that there is a correlation exists (adjusted R square of 0.34) between the rear-end crash rates and Δt_{r1}. The coefficient of Δt_{r1} is significant at a level of 0.01. The intercept did not have much significance since $\Delta tr1$ is non-zero for any location.

5.3 RECOMMENDATIONS FOR FUTURE STUDY

The positive correlation between crash rates and Δt_{r1} obtained from this study cracked the window of future research on crash occurrence and traffic stream characteristics. However, the study has several limitations that need to be addressed prior to advancing research on this topic. The authors recommend addressing the following major issues for future research-

- It is recognized by the authors that the first-stage removal of unwanted data points recommended in this study may remove some valid data points. Moreover, the selection of the threshold for this filter was somewhat arbitrary. In future, attempts will be taken to combine the two-stage data reduction process into a single robust process that is more efficient and consistent.
- The jam density values obtained from fitting the model did not fall in a reasonable boundary and implied that the average vehicle length is too small. The authors believe

that this is occurring due to lack of data at a very high-density condition. Further research is required to resolve this issue.

- The authors recommend conducting future studies on the roadway safety using a larger sample size of observation. A more diverse sample of freeway sites should be included in the analysis. As the first effort to rigorously and jointly estimate discontinuous traffic stream models, this study naturally focused on a limited sample of freeway sites to address the elements of complexity involved in the model fitting.
- The nonlinear optimization technique used in this study may require an extensive period for computation if traditional license-based (e.g., Excel) or open source (e.g., R) software is used. Therefore, it is suggested to use MATLAB for this purpose given that a licensed version is available to the user.

6.0 REFERENCES

Abdel-Aty, M. et al., 2004. Predicting freeway crashes from loop detector data by matched case-control logistic regression. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 1897, pp. 88-95.

American Association of State Highway and Transportation Officials, 2010. *Highway safety manual.* Washington, DC: AASHTO.

Cassidy, M., 1998. Bivariate relations in nearly stationary highway traffic. *Transportation Research Part B: Methodological,* Volume 32(1), pp. 49-59.

Chen, C., 2003. *Freeway Performance Measurement System (PeMS),* BERKELEY: California PATH Research Report.

Drake, J. & Schofer, J., 1966. A statistical analysis of speed density hypothesis. . *Highway Research Record* , Volume 154, pp. 53-87.

Edie, L., 1961. Car-following and steady-state theory for noncongested traffic. *Operations research,* Volume 9(1), pp. 66-76.

Frantzeskakis, J. I. D., 1987. Volume-to-capacity ratio and traffic accidents on interurban four-lane highways in Greece. *Transport Research Record 1112,* pp. 29-38.

Gazis, D. C., Herman, R. & Rothery, R., 1961. Nonlinear follow-the-leader models of traffic flow. *Operations research,* Volume 4, pp. 545-567.

Gazis, D. C., Herman, R. & Rothery, R. W., 1961. Nonlinear follow-the-leader models of traffic flow.. *Operations research,* Volume 9(4), pp. 545-567.

Gipps, P. G., 1981. A behavioural car-following model for computer simulation.. *Transportation Research Part B: Methodological,* Volume 15(2), pp. 105-111..

Golob, T. F. & Recker, W. W., 2003. Relationships among urban freeway accidents, traffic flow, weather, and lighting conditions. *Journal of transportation engineering,* Volume 129(4), pp. 342-353..

Google Inc., 2017. *Google Earth Pro.* [Online]
Available at: https://www.google.com/earth/download/gep/agree.html

Greenberg, H., 1959. An analysis of traffic flow. *Operations research,* pp. 79-85.

Gwynn, D., 1967. Relationship of accident rates and accident involvements. *Traffic Quarter,* p. 407–418.

Haddon, J. W. & Kelley, A., 1971. Reducing highway losses: a logical framework for categorizing highway safety phenomena and activity. *ABA Sec. Ins. Negl. & Comp. L. Proc.,* p. 54.

Hossain, M. & Muromachi, Y., 2010. *Evaluating location of placement and spacing of detectors for real-time crash prediction on urban expressways.* s.l., s.n., pp. 10-1069.

Kerner, B. S., 1998. Experimental features of self-organization in traffic flow. *Physical review letters,* p. 3797.

Koshi, M. I. M. &. O. I., 1983. *Some findings and an overview on vehicular flow characteristics.*. Toronto, University of Toronto, pp. 403-426.

Lee, C., Hellinga, B. & Saccomanno, F., 2003. Real-time crash prediction model for application to crash prevention in freeway traffic. *Transportation Research Record: Journal of the Transportation Research Board, (1840),,* pp. 67-77.

Li, J. & Zhang, H., 2013. Modeling space–time inhomogeneities with the kinematic wave theory. *Transportation Research Part B: Methodological,* Volume 54, pp. 113-125.

Lord, D., Manar, A. & Vizioli, A., 2005. Modeling crash-flow-density and crash-flow-V/C ratio relationships for rural and urban freeway segments. *Accident Analysis & Prevention,* Volume 37(1), pp. 185-199.

MathWorks, 1984. *MATLAB and Statistics Toolbox.* Natick(Massachusetts): The MathWorks, Inc..

May, A. D., 1990. *Traffic flow fundamentals.*. s.l.:Prentice-Hall Inc..

Nagel, K. & Schreckenberg, M., 1992. A cellular automaton model for freeway traffic.. *Journal de physique,* Volume 2(12), pp. 2221-2229..

North Carolina Department of Transportation, 1988. *Traffic Engineering Accident Analysis System (TEAAS).* [Online]
Available at: https://connect.ncdot.gov/resources/safety/Pages/TEAAS-Crash-Data-System.aspx
[Accessed 29 October 2016].

Pande, A. & Abdel-Aty, M., 2006. Comprehensive analysis of the relationship between real-time traffic surveillance data and rear-end crashes on freeways. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 1953, pp. 31-40..

Persaud, B. & Nguyen, T., 1998. *Safety consideration and capacity analysis.* Copenhagen, Denmark, s.n., p. 817–831..

Pipes, L., 1966. Car Following Models and The Fundamental Diagram of Road Traffic. *Transportation Research/UK,* Volume 1, pp. 21-29.

Skabardonis, A., Varaiya, P. & Petty, K. F., 2003. Measuring Recurrent and Nonrecurrent Traffic Congestion. *Transportation Research Record 1856, Paper No. 03-4261.*

Song, S. & Yeo, H., 2012. Method for estimating highway collision rate that considers state of traffic flow. *Transportation Research Record: Journal of the Transportation Research Board,* Volume 2318, pp. 52-62.

Transportation Research Board, 2010. *Highway Capacity Manual.* Washington, DC.: Transportation Research Board, National Research Council.

Turner, D. T. R., 1986. Motorway accidents: an examination of accident totals, rates and severity and their relationship with traffic flow. *Traffic Eng. Contr. 27 (78),* p. 377–387.

Underwood, R., 1960. Speed, volume, and density relationships. *Transportation Research Board.*

Wong, G. C. K. & Wong, S. C., 2002. A multi-class traffic flow model–an extension of LWR model with heterogeneous drivers.. *Transportation Research Part A: Policy and Practice,* Volume 36 (9), pp. 827-841.

Xu, C., Liu, P., Wang, W. & Li, Z., 2012. Evaluation of the impacts of traffic states on crash risks on freeways. *Accident Analysis & Prevention,* Volume 47, pp. 162-171.

Xu, Y., Williams, B., Rouphail, N. & Chase, R., 2013. Development of an Oversaturated Speed-Flow Model Based on the Highway Capacity Manual. *Transportation Research Record: Journal of the Transportation Research Board,* pp. 41-48.

Zhang, H. & Kim, T., 2005. A car-following theory for multiphase vehicular traffic flow. *Transportation Research Part B: Methodological,* Volume 39(5), pp. 385-399.

Zheng, Z., Ahn, S. & Monsere, C. M., 2010. Impact of traffic oscillations on freeway crash occurrences. *Accident Analysis & Prevention,* Volume 42(2), pp. 626-636.

Zhou, M. & Sisiopiku, V., 1997. Relationship between volume-to-capacity ratios and accident rates. *Transport. Res. Record 1581,* p. 47–52.

APPENDICES

APPENDIX A. FUNDAMENTAL DIAGRAMS OBTAINED FROM FITTING THE MACROSCOPIC MODEL

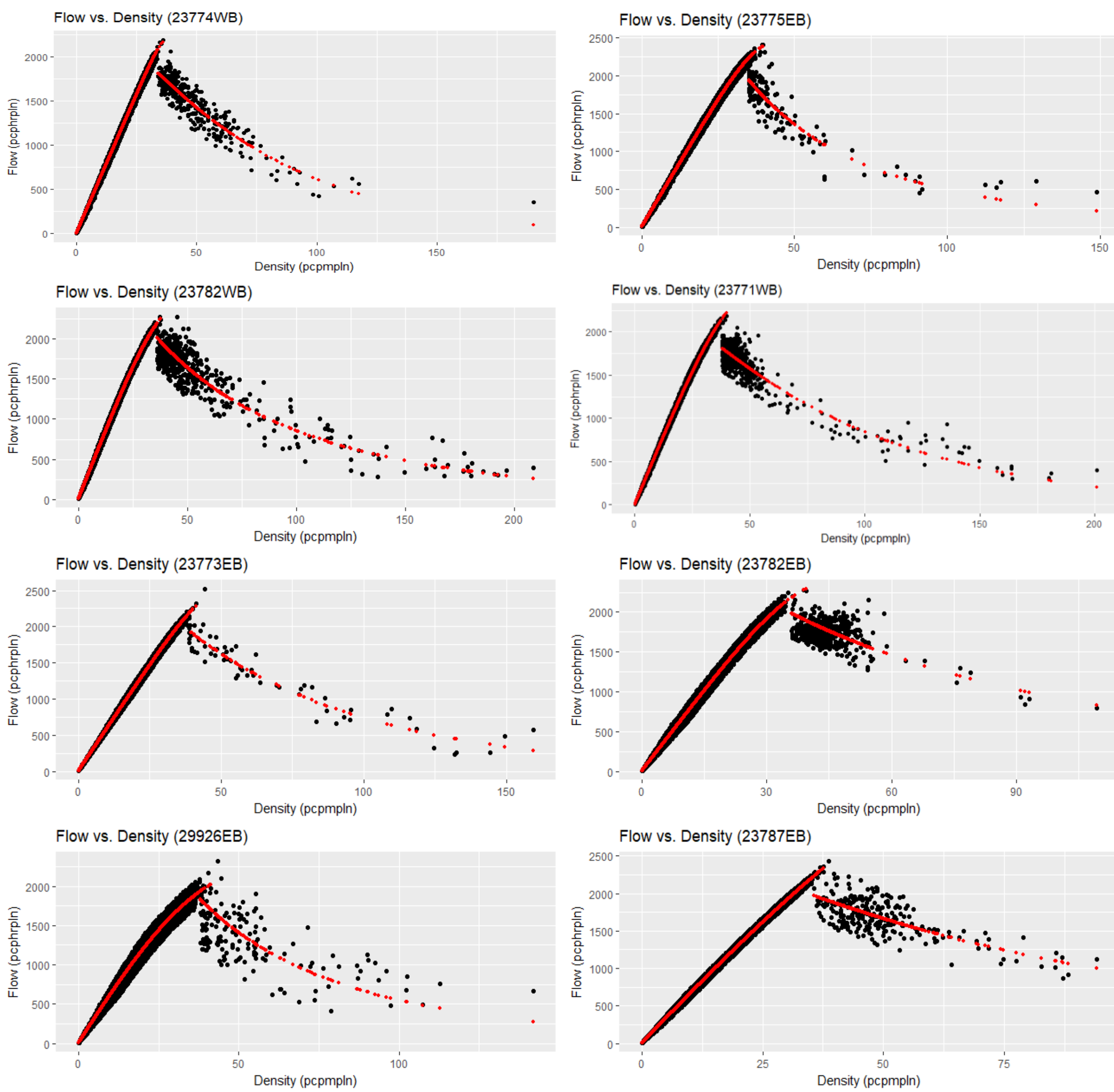


Figure A - 1: Fitted flow-density diagrams

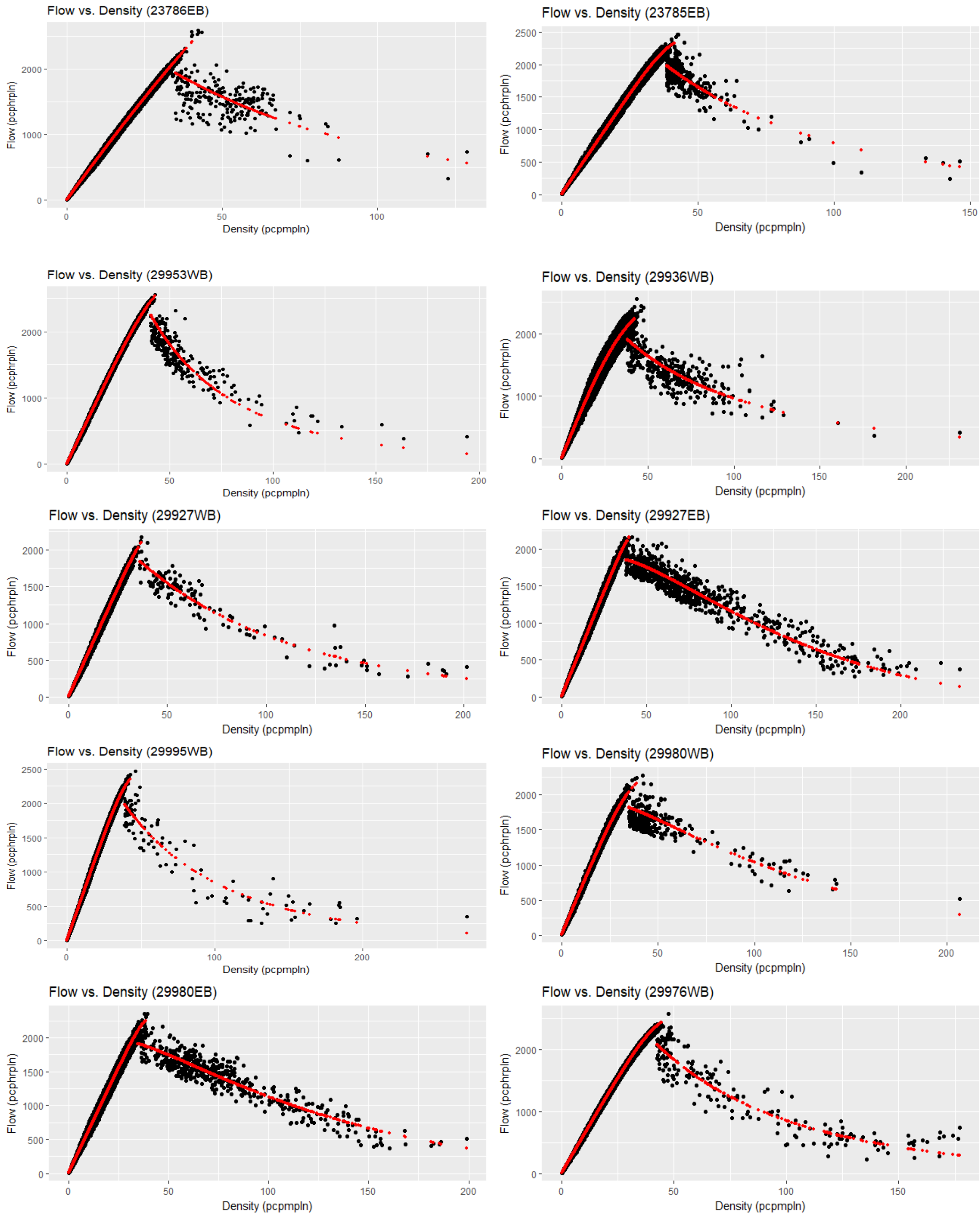


Figure A - 1: Fitted flow-density diagrams (Continued)

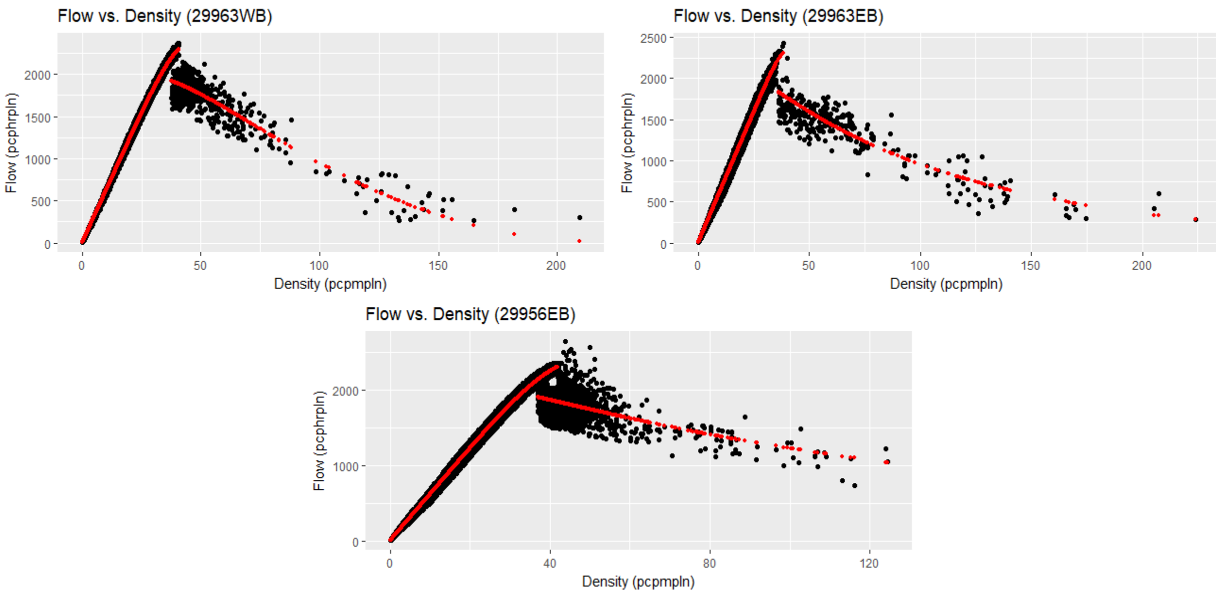


Figure A - 1: Fitted flow-density diagrams (Continued)

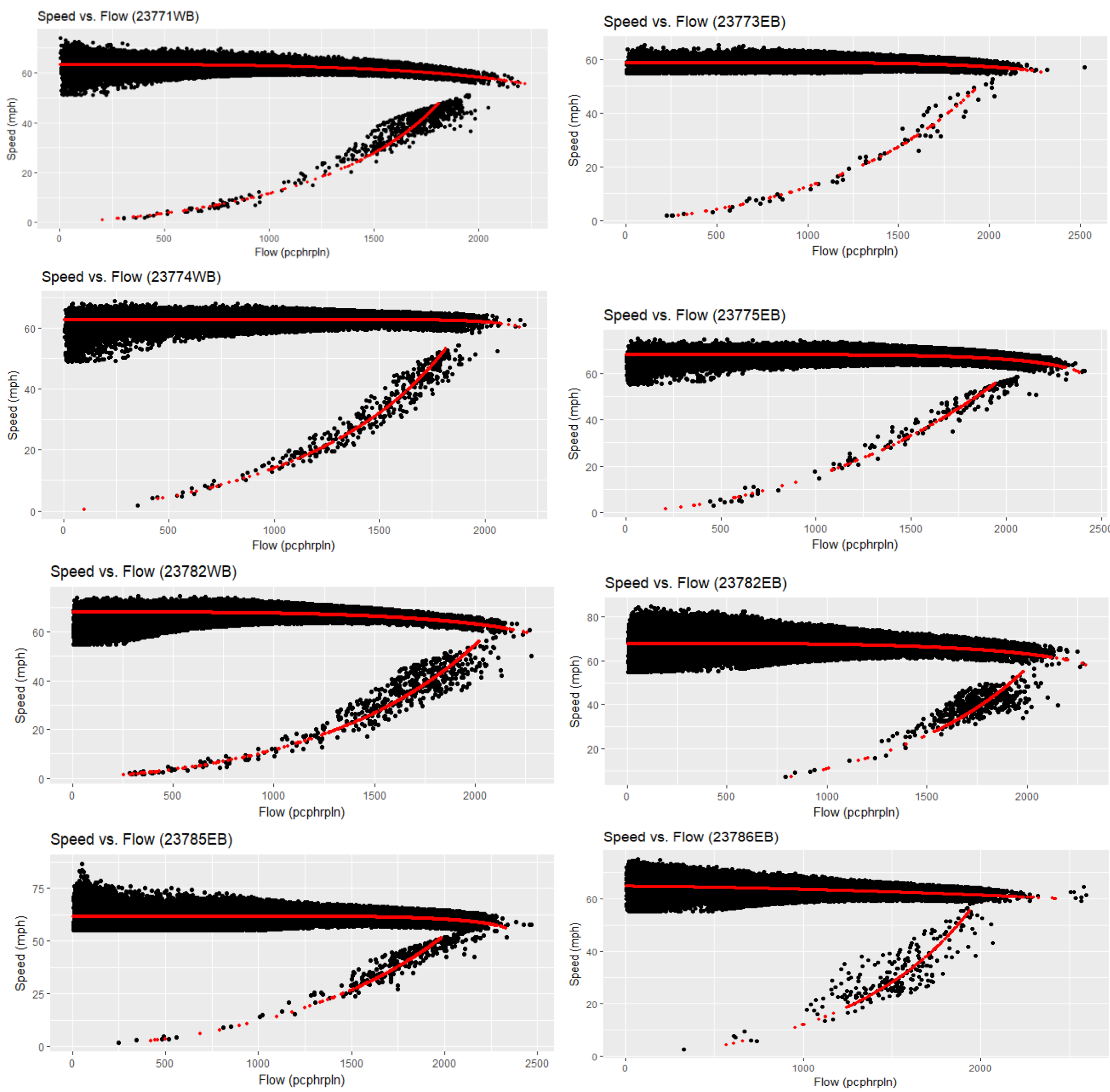


Figure A - 2: Fitted speed-flow diagrams

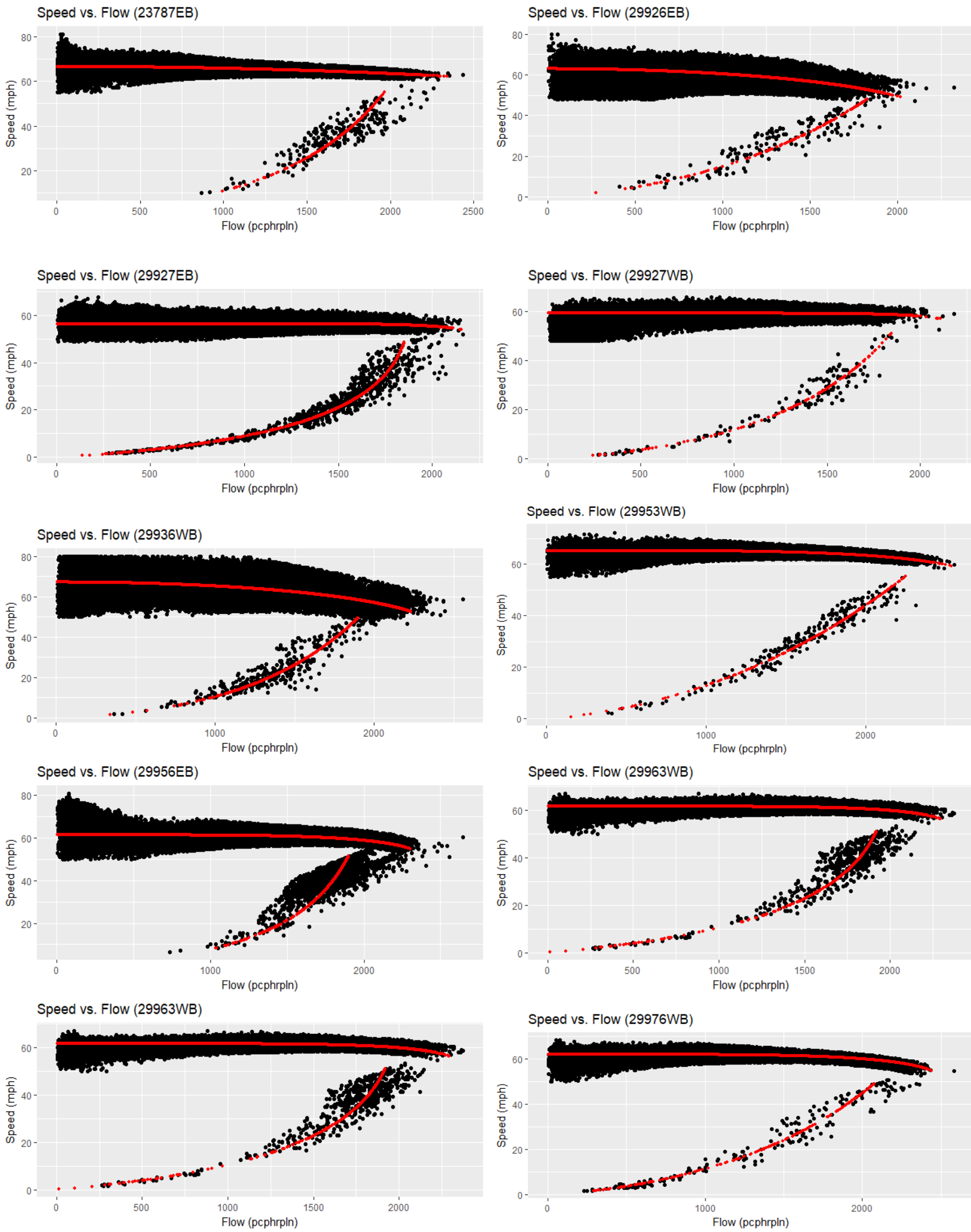


Figure A - 2: Fitted speed-flow diagrams

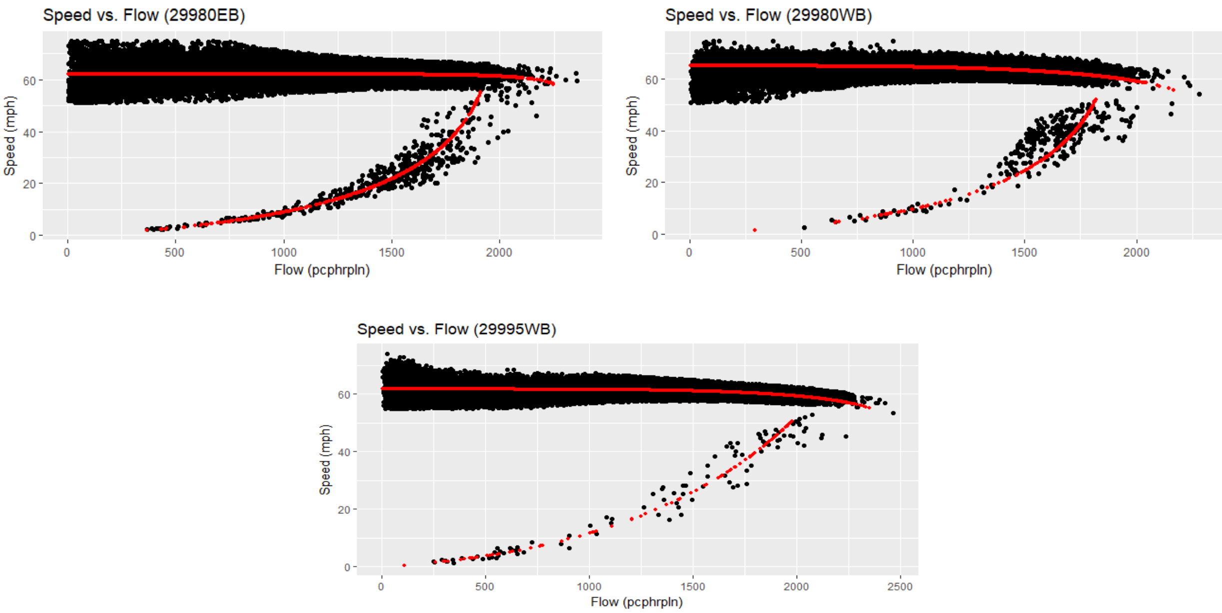


Figure A - 2: Fitted speed-flow diagrams

APPENDIX B. FITTED PARAMETER VALUES

Table B - 1: Fitted parameter values for all 21 stations

Station/Sensor ID	Regime 2					Regime 1					Δt_{r1}	Δt_{r2}
	u_f	k_b	k_j	l	m	u_f	k_b	k_j	l	m		
23771WB	63.3	39.9	606.2	3.85	0.99	4.7E+05	35.5	510.1	1.04	0.74	2.46	1.68
23773EB	58.8	44.3	246.0	5.55	0.99	1.0E+05	39.3	608.6	1.09	0.80	2.82	1.47
23774WB	62.8	37.2	69.4	11.21	0.97	5.4E+05	33.2	344.7	1.03	0.72	2.95	0.78
23775EB	67.8	39.6	196.0	6.09	0.99	1.0E+05	34.6	1808.3	1.17	0.91	2.01	0.84
23782EB	67.7	41.0	473.1	4.40	0.99	1.6E+06	36.0	1426.3	1.05	0.83	1.34	0.68
23782WB	68.1	41.0	447.7	4.19	0.99	6.8E+06	36.0	1289.3	1.04	0.83	1.72	0.81
23785EB	61.8	41.6	289.3	6.95	0.99	1.0E+05	38.6	1044.9	1.11	0.85	1.47	0.90
23786EB	63.8	44.0	427.7	4.51	0.99	4.3E+05	38.0	716.3	1.06	0.80	1.80	0.70
23787EB	66.0	38.5	79.7	5.83	0.75	4.6E+05	36.0	416.5	1.03	0.70	1.22	0.73
29926EB	62.9	42.0	399.9	2.97	0.95	3.4E+05	38.0	1306.6	1.12	0.88	1.67	1.15
29927EB	56.4	41.7	129.8	9.33	0.99	1.2E+07	38.0	337.4	1.00	0.59	2.63	0.92
29927WB	59.2	39.0	237.5	8.22	0.99	4.4E+05	35.0	641.3	1.04	0.76	2.67	0.70
29936WB	67.1	42.4	183.0	3.06	0.79	2.9E+06	38.4	4207.0	1.07	0.88	1.33	0.96
29953WB	65.3	48.0	572.0	4.86	0.99	1.4E+06	40.0	1800.1	1.11	0.89	2.19	0.70
29956EB	62.4	40.2	1099.7	1.97	0.20	1.0E+05	36.2	1812.2	1.09	0.84	2.17	0.95
29963EB	63.3	39.0	76.7	11.02	0.98	5.7E+06	36.0	918.0	1.03	0.78	2.25	0.55
29963WB	61.7	41.5	146.0	6.72	0.99	6.2E+06	37.5	228.9	1.00	0.54	1.56	0.69
29976WB	61.9	44.6	342.1	5.61	0.99	1.0E+05	41.6	1043.3	1.12	0.85	1.27	0.88
29980WB	65.0	39.5	88.9	4.42	0.61	4.7E+05	35.0	465.8	1.02	0.68	1.51	0.68
29980EB	62.0	38.6	81.5	10.37	0.99	4.5E+05	34.6	467.0	1.02	0.68	2.82	0.80
29995WB	61.7	43.6	209.6	5.47	0.99	1.0E+05	38.6	2132.9	1.13	0.88	1.80	0.86