Project No. _____ COPY NO. _____

DEVELOPMENT OF ANALYSIS METHODS USING RECENT DATA

FINAL REPORT

Prepared for
The Strategic Highway Research Program 2
Transportation Research Board
of
The National Academies

TRANSPORTATION RESEARCH BOARD
OF THE NATIONAL ACADEMIES
PRIVILGED DOCUMENT

This report, not released for publication, is furnished only for review to members of or participants in the work of the CRP. This report is to be regarded as fully privileged, and dissemination of the information included herein must be approved by the CRP.

Gary Davis & John Hourdos
University of Minnesota
Minneapolis, MN

July 2009

DRAFT

ACKNOWLEDGEMENT OF SPONSORSHIP

This work was sponsored by the Federal Highway Administration in cooperation with the American Association of State Highway and Transportation Officials, and was conducted in the Strategic Highway Research Program, which is administered by the Transportation Research Board of the National Academies.

DISCLAIMER

This report describes independent research conducted by the contractor. The opinions and conclusions expressed or implied in the report are those of the research agency. They are not necessarily those of the Transportation Research Board, the National Academies, or the program sponsors.

DRAFT

DRAFT

Project No. _____ COPY NO. _____

DEVELOPMENT OF ANALYSIS METHODS USING RECENT DATA

FINAL REPORT

Prepared for
The Strategic Highway Research Program 2
Transportation Research Board
of
The National Academies

Gary Davis & John Hourdos
University of Minnesota
Minneapolis, MN

July 2009

DRAFT

DRAFT

Table of Contents

AUTHOR ACKNOWLEDGMENTS III

ABSTRACT V

EXECUTIVE SUMMARY 1

1 . BACKGROUND AND PROJECT OBJECTIVES 5
1.1 BACKGROUND 5
1.2 HIGHWAY SAFETY MANUAL 5
1.3 SHRP 2 SAFETY PROGRAM 6
1.4 STRUCTURAL MODELING OF CRASH-RELATED EVENTS 7
1.5 CRASH SURROGATES 9
1.6 SUMMARY 11

2 . OVERVIEW OF ANALYTIC METHOD 13
2.1 STATE-SPACE MODEL 13
2.2 ILLUSTRATIVE EXAMPLE 15

3 . ANALYSES USING VEHICLE-BASED DATA 19
3.1 DATA ACQUISITION AND PREPARATION 19
3.2 CASE NUMBER 99540 20
3.3 CASE NUMBER 104119 21
3.4 CASE NUMBER 73082 22
3.5 CASE NUMBER 104851 23
3.6 CASE NUMBER 104283 24
3.7 CASE NUMBER 60289 24
3.8 CASE NUMBER 92660 25

4 . ANALYSES USING SITE-BASED VIDEO DATA 27
4.1 DATA ACQUISITION AND PREPARATION 27
4.2 I-94 CASE 1. 27
4.3 I-94 CASE 2. 28
4.4 I-94 CASE 3. 28

DRAFT

4.5 I-94 CASE 4. 29

5 . ANALYSES USING CICAS SITE-BASED SYSTEM.....31

5.1 GENERAL METHODOLOGY 31

5.2 CICAS NORTH CAROLINA US-74E NEAR CRASH CASE: 12:11PM, APRIL 25th, 2007 34

6 . CONCLUSIONS AND RECOMMENDATIONS 39

6.1 SUMMARY..... 39

6.2 CONCLUSIONS..... 40

6.3 RECOMMENDATIONS FOR FUTURE WORK 41

REFERENCES.....43

APPENDIX A: ANALYSIS TOOLS DEVELOPED IN THIS PROJECT

APPENDIX B: THE CICAS SITE-BASED SYSTEM

APPENDIX C: OUTLINE OF A CAUSAL THEORY OF TRAFFIC CONFLICTS AND COLLISIONS

DRAFT

AUTHOR ACKNOWLEDGMENTS

The research described in this report was performed under SHRP 2 Project S01 (A) by the Center for Transportation Studies, the Minnesota Traffic Observatory, and the Department of Civil Engineering at the University of Minnesota.

Dr. Gary A. Davis, Professor of Civil Engineering, was the Principal Investigator. Other authors of this report are Dr. John Hourdos, Director of the Minnesota Traffic Observatory and co-Principal Investigator, Ted Morris, Manager of the Minnesota Traffic Observatory, Hui Xiong, Graduate Assistant in the Department of Civil Engineering, and Indrajit Chatterjee, Graduate Assistant in the Department of Civil Engineering.

DRAFT

DRAFT

ABSTRACT

This report documents the second phase of a two-phase project SHRP 2 S01(A). The main goal was to develop and test analytic methods for using data collected in the Strategic Highway Research Program 2 (SHRP 2) field studies to support microscopic modeling of crashes and near-crashes. After illustrating the analysis tools developed for this project, we describe reconstruction of individual crashes and near-crashes using vehicle-based data from the 100-car study, site-based video data from the Minnesota Traffic Observatory, and site-based Doppler shift data from Cooperative Intersection Collision Avoidance Systems (CICAS) project. From the vehicle-based data and site-based data we reconstructed several events involving crashes or conflicts between a following and leading vehicle. In each case plausible estimates of acceleration histories for the leading and following driver reaction time and the situation at the start of the reaction phase were obtained. For the CICAS site-based data, we first developed data-mining procedures to identify possible near-crash events, but found that manual post-processing was needed to obtain data in a useable form. Nonetheless, we were able to illustrate the potential of this data collection system to provide data on intersection near-crashes, given some technical modifications.

DRAFT

DRAFT

EXECUTIVE SUMMARY

This report describes work done during Phase 2 of SHRP 2 Project S01 (A): Development of Analysis Methods Using Recent Data. In the report submitted at the end of Phase 1 we identified three research problems on which progress was needed. The first of these was identification of an appropriate class of structural models describing how crash and near-crash events developed, together with analytic tools for fitting these models to data expected from the vehicle-based and site-based field studies. The second problem involved counterfactual screening of supposed near-crash events to determine their similarity to crashes. The third problem involved developing plausible models of how drivers select evasive actions as functions of the situations in which they find themselves. Solutions to the second and third research problems are contingent on solution of the first, so the bulk of our effort during Phase 2 has been devoted to structural modeling of crash and near-crash events, using data from the 100-car vehicle-based field study, from site-based video data collected by the Minnesota Traffic Observatory (MTO), and from site-based Doppler shift data obtained from the Cooperative Intersection Collision Avoidance Systems (CICAS) project. Our focus has been on crashes and near-crashes involving more than one vehicle, of the sort that occur at intersections.

DRAFT

Background

Chapter 1 of this report outlines the context within which this research has taken place, that in the United States two substantial national efforts related to road traffic safety are underway. The first is development of the first edition of a *Highway Safety Manual* (HSM), the second is the design and execution of the SHRP 2 safety focus field studies. We point out that it may be possible for the data collected in the SHRP 2 field studies to support development of microscopic (i.e., individual event) crash models, which can be incorporated into traffic simulation models to supplement or replace the macroscopic regression methods used in the first edition of the HSM. For this to occur, though, it will be necessary to identify and fit plausible microscopic models of crash-related events using the SHRP 2 field data. We illustrate how this might be accomplished using a simple braking-to-stop model applied to trajectory data extracted from site-based video, and then illustrate how once a fitted model is at hand, it is possible to quantify the expected number of crashes in a set of non-crash events.

Chapter 2 takes up the problem of extending these ideas to more complicated situations and we propose a modeling strategy were driver behavior is treated as a piece-wise constant sequence of acceleration changes. Given such an acceleration history and initial values for a vehicle's location and speed, it is logical to move toward a system of ordinary differential equations to get predicted histories of the vehicle's speed and position. Fitting such a model then involves identifying the appropriate break points in the acceleration profile, the corresponding acceleration levels, and the initial conditions that best fit observed trajectory data. We illustrate model identification and estimation

using speedometer, radar range and radar range-rate data for a near-crash event from the 100-car vehicle-based study. We also illustrate a what-if counterfactual analysis, where the final deceleration of the following vehicle is varied over a range of values, and for each we compute the probability, other things being equal, that a collision would have resulted is computed.

Findings

Chapter 3 describes our work with data from the 100-car study. We obtained data for 33 crash and near-crash events, each consisting of approximately 30 seconds of forward video and 30 seconds of instrumentation measures at 10 Hz. The instrumentation measures included speeds from the instrumented vehicle's speedometer, range and range-rate for objects ahead of the instrumented vehicle from its forward radar, accelerometer measures, GPS positions, heading, and yaw measures, and indicators of brake, accelerator, and turn signal use. After reviewing the data, we determined that seven of the 33 events had data sufficient to attempt modeling, and we describe model identification, estimation, and goodness-of-fit evaluation using speedometer, range, and range-rate data. Six of the events each involved a leading vehicle and following vehicle decelerating in the same lane of traffic, and it was possible to identify plausible acceleration profiles for the leading and following drivers. From these measures it was then possible to estimate the following driver's reaction time, together with measures describing the situation at the time his or her reaction phase began. We also reconstructed one event involving a swerving maneuver by the following driver, where the forward radar data were limited.

Chapter 4 describes our work using site-based video data from the University of Minnesota's Beholder system. This is a set of video cameras, computers, and wireless communication equipment positioned to overlook a section of Interstate 94 in downtown Minneapolis. Vehicle positions, extracted from video recordings and rectified to correct for camera position effects, provided the raw data for our analyses. A total of four rear-end crash events occurring in July-October 2008 were analyzed, with two of these events also including trajectories for vehicles not involved in the crash. In each case it was possible to identify plausible acceleration profiles for each of the involved drivers, which in turn provided information on reaction times and conditions at the start of the reaction phase.

Chapter 5 describes pilot work using site-based Doppler shift data collected by the University of Minnesota's Intelligent Vehicles Laboratory as part of the Cooperative Intersection Collision Avoidance Systems (CICAS) project. This configuration consists of a coordinated set of radar units collecting information on the positions and speeds of major approach vehicles approaching a two-way stop controlled intersection, and LIDAR units collecting information on the position of minor approach vehicles. Since the instrumentation configuration was designed to support a prototype driver information system and not to collect data on vehicle trajectories, data acquisition and preparation was not as straightforward as with our other data sets, but with some data mining and post processing, it was possible to identify one event suggesting braking on the part of a

DRAFT

major approach vehicle in response to minor approach crossing. It was then possible to identify and fit plausible acceleration profiles for both drivers.

Chapter 6 presents our conclusions and recommendations for further research.

Conclusions

1. At least for situations where direction of travel is roughly constant, trajectory-based reconstruction of crash-related events, where trajectory data are used to fit parsimonious models of driver behavior, is feasible using both vehicle-based and site-based data.
2. It is possible to extend the methods of counterfactual analysis to more complicated structural models involving differential equations.
3. At least for rear-ending events, there is some limited evidence that the distributions of evasive actions for crashes and near-crashes share some overlap, so that it should be possible to find near-crash events that are similar in other respects to crashes.
4. Although the CICAS system as currently configured was not designed to collect and process crash and near-crash trajectory data, with technical modifications it could support site-based field research, at least at lower-volume intersections.
5. The usefulness of the data produced by the SHRP 2 vehicle-based field study will be strongly dependent on the ability to calibrate and maintain the data collection systems.

Recommendations:

1. The modeling methods presented in this report should be extended to handle two-directional trajectories.
2. When attempting to include crash events in a microscopic traffic simulation, plausible models that close the feedback loop between existing conditions and driver actions will be necessary. This issue should be pursued using the data from the SHRP 2 field studies.
3. The trajectory modeling methods described in our report should be enhanced to allow for possible serial correlation in trajectory data.
4. Compiling data on gap-selection and other intersection-related events will require a data setup different from the SHRP 2 vehicle-based field study, so the SHRP 2 vehicle-based field study should be complemented with site-based research.
5. Clear descriptions of data collection and processing, and associated metadata, should be required in future major data collection efforts.

DRAFT

DRAFT

CHAPTER 1

1. BACKGROUND AND PROJECT OBJECTIVES

1.1 BACKGROUND

In its essence, rational planning involves using predicted consequences to guide selection from among a set of possible actions. In road safety engineering, this requires being able to predict the frequency and/or severity of crashes that are expected to result from a given design or operational configuration. Hauer (*1*) has argued persuasively for developing a scientifically justifiable methodology for making these predictions, and the last 10 years have seen two major initiatives in the United States related to this issue. One is the development of the *Highway Safety Manual* (HSM), the first edition of which is scheduled for release in the next year. The other is the safety focus of the SHRP 2 research program.

1.2 HIGHWAY SAFETY MANUAL

The goal of the HSM is to provide highway professionals with tools for explicitly considering the safety impacts of engineering actions. The dominant methodology used to develop the HSM is statistical analysis of crash-frequency data, where the basic units of analysis are either sections of highway or intersections, and where the dependent variables are crash frequencies observed over one or more years. In some cases these frequencies can be broken down by crash type or severity. Generalized linear models are used to describe baseline associations between crash frequency and observable road features, while the effects of changes from these baseline conditions are captured through empirically determined crash-modification factors. The effect of an intervention is then predicted by first using the base model to predict crash frequency under the prevailing conditions, and then multiplying this expected frequency by a crash-modification factor which reflects the effect of the change of interest. Ideally, the crash-modification factor was estimated from a well-conducted before/after study that controlled for selection bias effects. The strong reliance of the HSM's first edition on this type of statistical modeling is in large part due to historical developments, where the integrated crash and roadway databases maintained by the Highway Safety Information System (HSIS) and several individual states had attained a useful degree of completeness, while the development of empirical Bayes and hierarchical Bayes methods during the 1980s and 1990s brought the supporting statistical tools to a useful degree of maturity.

Regression analyses of aggregated observational data have well-understood limits as to their ability to discover and describe underlying causal processes (*2*). In 2006 a day-long workshop was held during the annual meeting of the Transportation Research Board (TRB), which focused on elucidating these limits and discussing alternative methods. At that workshop Bonneson and Lord (*3*) pointed out an interesting analogy with the development of the *Highway Capacity Manual*, where first generation regression

DRAFT

models for predicting traffic signal delay using naïve specifications of independent variables were later replaced by regression models where the form of the independent variables was justified theoretically. These in turn were replaced by models where the functional forms relating traffic flow, capacity, and signal timing to delay were justified theoretically. The macroscopic methods developed for the *Highway Capacity Manual* were then later supplemented and in some instances replaced by microscopic traffic simulation models. Hope was expressed that a similar evolution might take place for the HSM, with first-edition regression models being supplemented or replaced by structural models that explicitly describe the mechanisms underlying crash occurrence. A better understanding of crash mechanisms could also support using microscopic traffic simulation models to predict the safety consequences of engineering decisions, similar to how microscopic models are now used to predict operational consequences. As those researchers working in this area are acutely aware, however, a major obstacle to progress is the lack of good microscopic data regarding crash occurrence, and driver behavior more generally.

1.3 SHRP 2 SAFETY PROGRAM

As originally conceived, the safety focus area of the SHRP 2 research program was to consist of two major field studies, a vehicle-based study involving "volunteer drivers and a sophisticated instrumentation package installed in the volunteers' vehicles" and a site-based study involving video-recording of vehicle movements at specific locations. These studies were "…intended to support a comprehensive safety assessment of how driver behavior and performance interact with roadway, environmental, vehicular, and human factors and the influence of these factors and their interactions on collision risk…" (*4*). At the present time, the vehicle-based study is going forward while the site-based study has been limited to preliminary design, with further work dependent on the availability of additional funding. The SHRP 2 S01 projects are to develop and apply analytic methods relevant to these field studies, by identifying salient research questions and then attempting to answer these using existing data of the type expected from the field studies. The S01 request for proposal explicitly identified as important "application of crash surrogates" and "the formulation of analytic methods to quantify the relationship of human factors, driver behavior, vehicle, roadway, and environmental factors to collision risk." Special attention was to be directed to roadway departure and intersection crashes.

There is little doubt that the SHRP 2 vehicle-based study should produce a rich and unprecedented source of information concerning driver behavior in normal and crash situations, and that this should support development and evaluation of vehicle-based safety technologies. The study should also support traditional statistical investigations seeking to identify associations between roadway conditions and crash occurrence. In our view the SHRP 2 field studies could also provide data supporting the development and application of microscopic crash models, similar to how existing crash record and roadway databases support the development and application of regression-based

DRAFT

approaches. For this to occur, however, we will need analytic tools that can fit and test microscopic models using field study data, and that can extract the sort of measurements needed to quantify driver behavior in crash-related conditions. Our S01 project thus has two inter-related objectives. The first is to develop analytic tools and demonstrate how these can be used to carry out structural model development, using the sort of data expected from the SHRP 2 field studies. The second is to develop a rigorous method for characterizing near-crashes, so that observations of near-crashes might serve as useful surrogates for actual crashes. The approach we take can be called trajectory-based reconstruction of crash-related events. That is, we use time history data of vehicle positions and/or speeds to estimate values for variables describing drivers' actions and characterizing the conditions leading to crash-related events. The focus of our project is on crashes involving more than one vehicle, resulting from car-following or gap-selection behavior of the types often occurring at intersections.

1.4 STRUCTURAL MODELING OF CRASH-RELATED EVENTS

Our starting point is Pearl's (*5*) notion of a causal model, which in the abstract consists of a set of exogenous variables, a set of endogenous variables, and for each endogenous variable, a structural equation describing how that variable responds to changes in other model variables. A causal model can be represented qualitatively using a directed graph, with the nodes of the graph representing variables and directed arrows indicating direct causal dependencies.

Figure 1.1 displays a simple graphical model for a generic crash. The node u, possibly vector-valued, denotes variables describing background conditions. The node x denotes the variable describing an evasive action, and the node y denotes the crash condition. The crash condition y is assumed to be a *deterministic* function of u and x, such that

y(u,x) = 0, if the values for u and x do not produce a crash,
y(u,x) = 1, if u and x produce a crash.

To make these ideas a bit more concrete, consider a simple two-vehicle rear-ending collision model (*6-7*). Such an event might be observed in the field, or it could arise within a microscopic traffic simulation. The initial speed and braking deceleration of the leading vehicle are denoted by v_1 and a_1 respectively, v_2 and a_2 denote the initial speed and braking deceleration of the following vehicle, and h_2 and r_2 denote the following driver's headway and reaction time. A collision occurs when the stopping distance available to the following driver is less than that needed to stop without colliding with the lead vehicle. Using simple physics this can be expressed as

$$v_2 r_2 - \frac{v_2^2}{2a_2} > h_2 v_2 - \frac{v_1^2}{2a_1} \tag{1.1}$$

DRAFT

If the following driver's deceleration is taken as the avoidance action then for the rear-end collision, the variables (v_1,a_1,v_2,r_2,h_2) are components of u, a2 is the evasive action, and the collision function is

$$y(u,x) = \begin{cases} 0, \ if \ v_2 r_2 - \dfrac{v_2^2}{2a_2} \le h_2 v_2 - \dfrac{v_1^2}{2a_1} \\ 1, \ if \ v_2 r_2 - \dfrac{v_2^2}{2a_2} > h_2 v_2 - \dfrac{v_1^2}{2a_1} \end{cases} \tag{1.2}$$

In this treatment an event consists of a specification of values for each of the model variables. This specification plus structural equation (1.2) is sufficient to determine whether or not that event leads to a crash. As they are currently implemented, microscopic simulation models are restricted to combinations of values that do not lead to crashes, and an open research question involves determining realistic relative frequencies for those combinations which do.

To help illustrate the usefulness of this approach, we can use an example originally presented by Davis and Swenson (*7*). Figure 1.2 shows trajectories for a platoon of seven vehicles successively braking to stops while traveling in the same lane of a freeway. The left-most vehicle was the first vehicle in the platoon, the right-most vehicle was the last, and a collision occurred between the two right-most vehicles. These trajectories were constructed from a video recording of the event by first digitizing each vehicle's position on successive video frames and then using standard photogrammetry methods to convert from image coordinates to ground coordinates. Applying the simple braking model described above, each of these trajectories can be described by the physical model

$$y_k(t) = \begin{cases} v_k t, \quad t \le t0_k \\ v_k t0_k + a_k (t - t0_k)^2 / 2, \quad t0_k < t \le t0_k + v_k / (-a_k) \\ v_k t0_k - v_k^2 / 2a_k, \quad t > t0_k + v_k / (-a_k) \end{cases} \tag{1.3}$$

where $y_k(t)$ denotes the (one-dimensional) position of vehicle k at time t, v_k denotes the initial speed of driver k, a_k denotes his or her braking acceleration and $t0_k$ denotes the time at which braking began. This model can be connected to the rear-end collision model described above by noting that the reaction time of driver k is simply

$$r_k = t0_k - t0_{k-1} \tag{1.4}$$

while the initial following headway between vehicles k and k-1 when driver k-1 began braking is

DRAFT

$$h_k = (y_k(t0_{k-1}) - y_{k-1}(t0_{k-1}))/v_k \quad (1.5)$$

At least for simple event types like this, if it is then possible to generate estimates of initial speeds, braking rates and times of braking initiation by fitting equation (1.3) to observed trajectories. Davis and Swenson (*7*) described how Bayes estimates could be computed from trajectory data of the type displayed in Figure 1.2 and Table 1.1 summarizes the resulting estimates for these data.

In principle then, trajectory-based reconstruction can be used to fit structural models and estimate important features of crash-related events. One of our objectives is to extend these methods to handle more complicated situations, and to exploit the type of data expected from the SHRP 2 in-vehicle field study.

1.5 CRASH SURROGATES

The second objective of our project is to develop a quantitative method for characterizing and identifying events that can serve as useful crash surrogates. This is because crashes, especially severe crashes, tend to be rare so that, if one could identify near-crashes or other surrogate events that carry information about how crashes occur, the value of both in-vehicle and site-based studies would increase.

Roughly speaking there are two ways that near-crash events might be used as crash surrogates. On the one hand, one might carry out an intensive study of how individual near-crashes occurred, with the goal of identifying causal factors for each of these events. This would be similar to using investigation and reconstruction of actual crashes to gain insight into how and why crashes occur. On the other hand, one might use counts of near-crashes as a dependent variable and then look to see how these are associated with roadway and/or driver characteristics. This would be similar to carrying out a statistical study of crash frequency. In either case, though, the starting point is a set of non-crash events and the need to determine the extent to which each could be regarded as a near-crash.

Returning to the literature, it is possible to find two related, but different, approaches to defining crash surrogates. One is the definition of conflict as put forward by the International Calibration Study of Traffic Conflict Techniques (ICSTCT): "A traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged (*8*)." However, it turned out that when attempting to find empirical associations between conflict and crash frequencies it was helpful if conflicts could be graded as to their seriousness or severity. This distinction is included in the definition of near crash used in the 100-car study, which can be regarded as a pilot for the SHRP 2 vehicle-based field study: "Any circumstance that requires a rapid evasive maneuver by the subject vehicle, or any other vehicle, pedestrian, cyclist, or animal to avoid a crash. A rapid, evasive maneuver is defined as a steering, braking, accelerating or any combination of control inputs that approaches the limit of the vehicle capabilities (*9*)."

DRAFT

Both of these definitions assume a counterfactual definition of the surrogate event, where had an evasive action not been performed a crash would have occurred. The 100-car study definition places an additional condition, however, that the magnitude of the evasive action should be in some sense extreme. To see the value of this additional condition, it is helpful to return to Table 1.1. The table gives estimates of initial speeds, headways, reaction times, and the actual and minimum successful decelerations estimated from the vehicle trajectories. First, consider the interaction between vehicles 1 and 2. The minimum deceleration by vehicle 2 needed to avoid collision was about -6.2 feet/sec^2, while the actual deceleration was about -6.5 feet/sec^2. Had the deceleration been slightly less, other things being equal, a crash would have occurred so arguably this event satisfies the ICSTCT definition of a conflict. Most would agree, though, that it does not satisfy the 100-car study condition that the evasive action be extreme. For the interaction between vehicles 5 and 6, the minimum successful deceleration was about -17.1 feet/sec^2 and the actual deceleration was about -17.3 feet/sec^2. This event also satisfies the ICSTCT definition of conflict but comes closer to satisfying the 100-car study condition as well. We will use causal models to construct a quantitative measure that captures this difference.

To start, Figure 1.3 shows the probability of collision as a function of the following vehicle's braking deceleration for both vehicle pairs. The analysis is probabilistic because the values of the important event variables are not known with certainty, but rather only up to their posterior distributions given the trajectory data. Figure 1.3 was prepared by setting the follower's deceleration to each of a set of target values, and then using Monte Carlo simulation to compute the probability of a crash.

Figure 1.3 shows that for vehicles 1 and 2 decelerations greater than about -7 feet/sec^2 prevents a crash with high probability, while for vehicles 5 and 6 decelerations greater than about -20 feet/sec^2 are needed for a similar degree of certainty. The latter seems qualitatively close to the definition of near-crash used in the 100-car study, but to quantify this degree of closeness we need to specify what is meant by an evasive action that approaches the limit of the vehicle's capabilities. One (but not necessarily the only) way to do this is to apply the results of the emergency braking study carried out by Fambro et al. (*10*), where the distribution of braking decelerations used by drivers confronted with a surprise braking situation had a mean of about -20.3 feet/sec^2 and a standard deviation of about 2.6 feet/sec^2.

Figure 1.4 adds to Figure 1.3 a normal distribution with the above mean and standard deviation. Roughly speaking, the degree to which a conflict qualifies as a near-crash is determined by how much of this extreme braking distribution lies to the left of the crash probability curve. More formally, the probability that a conflict could have been a crash is found by integrating the crash probability curve with respect to the extreme braking distribution. Although analytically intractable, this computation is readily carried out using Monte Carlo methods. Table 1.2 gives these results for each of the non-colliding vehicle pairs from Table 1.1, and the sum of these probabilities can be taken as the expected number of crashes in this set of conflict events had the evading drivers taken their decelerations from the above distribution.

DRAFT

1.6 SUMMARY

To summarize, we have introduced an approach to microscopic modeling of crash-related events, where driver actions together with initial speeds and vehicle locations are treated as inputs to a physical model describing vehicle motion. Our choice of this modeling strategy is rooted in the fact that models of this sort are needed if realistic crash processes are to be included in microscopic traffic simulation models. We have illustrated how a simple version of a trajectory model can be used to estimate features of crash and near-crash events, such as driver reaction times, following headways, and deceleration rates, from trajectory data of the sort produced from a site-based field study. Given sufficiently large samples of crash and near-crash events, this method could be used to compile distributions for these inputs, which could in turn be used in traffic simulation models. Finally, we have also illustrated how a trajectory model together with estimates of input variables can quantify the degree to which a non-crash event could have been a crash. One potential application of this technique would be to process a set of non-crash events produced either by a single driver or at a single location to produce an expected number of crashes in this set. This expected number of crashes could then serve as a dependent variable in a study of driver or site features believed to be related to safety.

The remainder of this report describes our effort to extend these ideas to more complicated scenarios, using data produced by both vehicle-based and site-based field studies. Chapter 2 outlines the analytic procedures and tools developed for this project and illustrates their use. Chapter 3 presents our analyses of data obtained from the 100-car vehicle-based field study. Chapter 4 describes analyses of data from site-based video on Interstate 94, while Chapter 5 describes work with site-based radar data from the Cooperative Intersection Collision Avoidance Systems (CICAS) intersection in North Carolina. Chapter 6 presents our conclusions and recommendations.

DRAFT

DRAFT

CHAPTER 2

2. OVERVIEW OF ANALYTIC METHOD

2.1 STATE-SPACE MODEL

In Chapter 1 we illustrated our trajectory-based approach with simple braking-to-stop models estimated from data extracted from site-based video. An important objective of our project was to develop a common analytic framework that could be applied to data from either site-based or vehicle-based sensor configurations and could accommodate events more complicated than constant speed followed by simple braking-to-stop. Ultimately our goal is model vehicle trajectories in two-dimensions, where both braking and steering could serve as evasive actions. This report however will focus on the simpler problem of modeling vehicle motion in one direction.

Our basic idea is to model driver behavior as a piecewise-constant series of acceleration changes, which are then treated as inputs into a dynamic trajectory model. The vehicle's state at a given time is its location and velocity, and the trajectory model takes the acceleration input sequence and numerically integrates the associated differential or difference equations to produce time histories of vehicle locations and speeds. For discrete-time data, the trajectory model can be conveniently represented using the generic linear state-space form

$$\begin{aligned} x(t+1) &= Ax(t) + Ba(t) \\ y(t) &= Cx(t) \end{aligned} \tag{2.1}$$

where x(t) is vector of state variables (position and velocity), a(t) is a vector input variables (accelerations), and y(t) is the vector of observed variables. A, B, and C in equation (2.1) stand for matrices of coefficients.

The nature of A,B,C, x(t) and y(t) will vary depending on the class of events being modeled and the sort of data that is available. For two vehicles following on a straight road, the simplest trajectory model would consist of two state variables for each vehicle, its location and speed, with linear acceleration values as inputs. That is, if Δ denotes the basic time interval of the data, then the deterministic progression for a leading and following vehicle can be captured by the linear equation

$$\begin{bmatrix} x1(t+1) \\ x2(t+1) \\ v1(t+1) \\ v2(t+1) \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta & 0 \\ 0 & 1 & 0 & \Delta \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x1(t) \\ x2(t) \\ v1(t) \\ v2(t) \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ \Delta & 0 \\ 0 & \Delta \end{bmatrix} \begin{bmatrix} a1(t) \\ a2(t) \end{bmatrix} \tag{2.2}$$

DRAFT

Here x1(t) and x2(t) give the locations of the leading and following vehicles at time t, v1(t), v2(t) are the corresponding speeds, and a1(t), a2(t) are accelerations. For motion in two directions a similar structure can be used but with state and input variables for each direction. Given initial values for the state variables and the time history for the inputs, the trajectories of both vehicles can be replicated.

Since the above state equation describes vehicle motion irrespective of the data collection scenario, the primary difference between the site and vehicle-based scenarios will be the observation equation. For vehicle trajectories extracted from video, the observations consist of measurements of position for each vehicle, leading to an observation equation of the form

$$\begin{bmatrix} y1(t) \\ y2(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} x1(t) \\ x2(t) \\ v1(t) \\ v2(t) \end{bmatrix} \tag{2.3}$$

A standard calculation shows that this system is observable, and in fact the two-vehicle system decomposes into two one-vehicle systems.

For the vehicle-based data from the 100-car study, the primary observations are the speed of the following vehicle obtained from its speedometer, and the range and range-rate for the lead vehicle relative to the follower, obtained from the forward radar. This leads to an observation equation of the form

$$\begin{bmatrix} y1(t) \\ y2(t) \\ y3(t) \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} x1(t) \\ x2(t) \\ v1(t) \\ v2(t) \end{bmatrix} \tag{2.4}$$

This system is not observable. For instance it is not possible to obtain estimates of absolute position for each of the vehicles. It also does not decompose into two independent subsystems, so that estimation and inference will generally require working with both vehicles together.

For site-based data from the CICAS radar system, observations consist of both position and speed for individual vehicles, leading to an observation equation of the form

$$\begin{bmatrix} y1(t) \\ y2(t) \\ y3(t) \\ y4(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} x1(t) \\ x2(t) \\ v1(t) \\ v2(t) \end{bmatrix} \tag{2.5}$$

DRAFT

This also decomposes into two separate subsystems, one for each vehicle.

Given estimates of a driver's initial speed, the times at which he or she changed acceleration and the corresponding accelerations, the differential equations can be solved to give predicted time histories of that vehicle's position and speed and predicted values for the observations. Fitting a trajectory model then involves searching plausible combinations of values for these input quantities to find those that best account for the data. The counterfactual simulation needed to assess the degree to which a near-crash might have been a crash, in turn, involves using probability distributions characterizing the residual uncertainty in the model parameters as input to a Monte Carlo simulation, where the state equation is integrated using random draws from this distribution and the occurrence or non-occurrence of a collision recorded. In the early stages of this study we experimented with several different approaches to implementing these steps, including modeling the differential equation model using response-surface approximations (*11*), nonlinear least-squares estimation using asymptotic normal approximations to characterize posterior uncertainty (*12*), and Bayesian analysis using Markov Chain Monte Carlo (MCMC) simulation. The WinBUGS software (*13*) can be used to implement MCMC estimation for a variety of relatively complicated models, but our initial attempts to implement this by directly coding the differential equations in WinBUGS led to estimation runs with excessively long time demands. This problem was circumvented by taking advantage of the WinBUGS differential equation interface. This provides compiled procedures that can be included in a WinBUGS model specification, which numerically solves ordinary differential equations using Runge-Kutta methods. As is our standard practice in working the MCMC estimation, we first conducted exploratory analyses using frequentist methods, in this case nonlinear least-squares, implemented using either MATLAB (*14*) or R (*15*). This was done in order to understand the complexity of the acceleration model suggested by a given data set and to get reasonable starting values for the MCMC simulation. Bayes estimates for model parameters were then computed using WinBUGS, and counterfactual simulation carried out using the MCMC sample of the posterior distribution for these parameters.

DRAFT

2.2 ILLUSTRATIVE EXAMPLE

We will illustrate our general approach using case 104119 from the 100-car study. This event was a potential near-crash that involved a lead vehicle and a following vehicle successively braking to a stop, with the follower stopping short of collision. This was a vehicle-based study, with the following vehicle being the instrumented vehicle. The data employed in this analysis were the speedometer-measured speeds for the instrumented vehicle, and range and range-rate for the leading vehicle, obtained from the follower's forward radar.

Step 1: Graphical inspection of data.

Figure 2.1 shows the time history of the speedometer-measured speeds for the following vehicle converted to units of feet/sec. The piece-wise constant nature to this relation was characteristic of all speedometer data we obtained for 100-car study cases. Inspection of Figure 2.1 suggested a two-phase model where the following driver was initially traveling at about 21 feet/sec and accelerating until about 11 seconds from the start of the data series. He or she then decelerated at a roughly constant rate until coming to a stop.

Figure 2.2 shows the time history of the range and range-rate data provided by the instrumented vehicle's forward radar. The discontinuities in the time-series are due to periods of missing data, when the forward radar apparently lost the leading vehicle as a target. In order to get an initial sense of the actions of the leading driver, we often used an approximate speed profile for the leading vehicle obtained by adding the range-rate to the following vehicle's speedometer data. A plot of these approximate speeds is shown in Figure 2.3. This suggests a three-phase model where a period of constant acceleration is followed by a period of roughly constant speed, which in turn is followed by a period of constant deceleration leading to a stop.

Step 2: Nonlinear least-squares estimation of proposed models

For the following vehicle, we need to estimate its initial speed, its acceleration during the first phase, the time at which deceleration began, and deceleration characteristic of the second phase. To estimate these, a MATLAB script was written which took trial values for these parameters as inputs, and simulated the following vehicle's position and speed over time by solving the differential equations using a simple Euler's method. The difference between the simulated speeds and the speedometer speeds was computed for each time interval, and the squares of these differences were summed to produce a measure of fit between the model and the speedometer data. This script was then embedded in a numerical search procedure to find the parameter values that minimized the sum-of-squares. The resulting estimates and initial approximate standard errors are shown in Table 2.1.

For the lead vehicle, the three-phase model was fit to the approximate speed data shown in Figure 2.3. The resulting nonlinear least-squares estimates and approximate standard errors are given in Table 2.2.

Step 3: Bayes estimation of vehicle models.

Final Bayes estimates were computed using the MCMC software WinBUGS. In essence, WinBUGS generates a simulated realization of a Markov chain whose stationary distribution is the same as the Bayesian posterior distribution of the model parameters

DRAFT

given the data. For this, the data were the speed profile for the follower and the range and range-rate profiles for the leader. Model parameters consisted of acceleration profiles and initial speeds for each vehicle, and the parameters for both vehicles were estimated in combination simultaneously. In the WinBUGS model, predicted values for the follower's speed and the range and range-rate for the leader were computed by numerically solving the differential equations using ordinary differential equation interface. For this case, a 10,000 iteration burn-in period followed by a 70,000 iteration MCMC sample produced acceptable convergence. Table 2.3 summarizes the Bayes estimates for both leader and follower.

Figure 2.4 shows the speedometer speeds for the following vehicle together with the posterior means estimated from the MCMC sample. Figure 2.5 shows similar results for the range data from the forward radar. Both cases appear to provide reasonable approximations to the observed information.

Finally, Figure 2.6 shows the probability of a rear-end collision between the two vehicles as a function of counterfactual final decelerations by the following driver. These probabilities were computed using Balke and Pearl's (*16*) Twin Network method, where the follower's deceleration is set at a target value and then for each outcome of the MCMC sample the values for the remaining parameters are used as inputs to solve the differential equations. This describes what would have happened had the event involved those parameter values and the counterfactual follower's deceleration. Simulated range values less than zero are taken to indicate a collision, and the fraction of the MCMC sample values that lead to collision is an estimate of the collision probability. In this case, had the follower braked at less than about -10 feet/sec^2 a collision would probably have resulted.

DRAFT

DRAFT

CHAPTER 3

3. ANALYSES USING VEHICLE-BASED DATA

3.1 DATA ACQUISITION AND PREPARATION

In May and June 2008, the project team received from the Virginia Tech Transportation Institute (VTTI) data for 33 events observed during the 100-car study. These consisted of time series of measurements from the in-vehicle sensors, along with the videos from the forward camera, for about 30 seconds preceding and including the crash or near-crash event. Table 3.1 lists the variables obtained.

The first task after receiving these data was to review them for completeness and accuracy, and flag those events for which the data were either incomplete or problematic. For this task, a MATLAB-based graphical browsing tool was developed that loaded the data from the VTTI-provided text file, allowing the analyst to plot the history of selected measurements and to perform two initial simulations of the trajectory of the instrumented vehicle, using either speed-based measurements or acceleration-based measurements as input. Inconsistencies between the outputs of these two models would then indicate problems with either the speed or acceleration data (or possibly both) for that event.

For five of the events, measurements from the forward radar were not available, so for these it was not possible to reconstruct the actions of the leading vehicle. For several more of the events, there was a clear discrepancy between the trajectory of the instrumented vehicle as indicated by the speedometer, heading, and yaw measurements, and the trajectory reconstructed from the acceleration measurements, indicating that caution should be used when using these data.

For the instrumented vehicle, the data available consist of speedometer output, lateral and longitudinal accelerations, yaw, heading, and indications of the status of the turn signal, the brake, and the accelerator, recorded at 10 Hz. For the lead vehicle, the available data consist of range, range-rate, and azimuth obtained from the forward-viewing radar, also recorded at 10 Hz. Latitude and longitude values from the GPS receiver were available, but in all cases these values were essentially constant throughout the duration of the event.

After initial examination of the data obtained for each of the 33 cases, a summary was prepared (see Table 3.2) identifying the potential cases that could be examined for this research, based on the quality and the completeness of the available data. Several of the events involved lane-changing, swerving, or merging on the part of one of the involved vehicles so that the forward radar of the instrumented vehicle provided only limited information about the leading vehicle. In total seven of the 33 events were analyzed. Those seven cases are 104119, 104283, 104851, 60289, 73082, 92660, and 99540.

In what follows, we will describe our analysis of each of these seven events. We would like to point out, though, that the accuracy of the conclusions depends critically on

DRAFT

the accuracy of the data on which they are based. Although we are confident that our results are consistent with the data provided, we present these analyses as examples of what could potentially be accomplished using vehicle-based field data, rather than as final determinations of what truly happened in these events.

3.2 CASE NUMBER 99540

Description from video: Figures 3.1-3.3 illustrate that in this event the instrumented (i.e., following) vehicle is traveling in the right-hand lane of a multi-lane highway, and exits this highway to the right. The exit ramp connects to another multi-lane highway and the leading vehicle slows and then comes to a stop at the merge point. The following vehicle collides with the lead vehicle and the leading vehicle moves forward and then stops on the roadway's shoulder.

Approximately 35 seconds of data were available from the instrumented vehicle at 10 Hz. These data included speeds from the instrumented vehicle's speedometer, and range and range-rate from its forward radar, as shown in Figure 3.4.

In addition to the original data, approximate speeds for the leading vehicle were computed by adding the instrumented vehicle's speedometer speed to the forward radar's range-rate. The speeds for the two vehicles are displayed in Figure 3.5.

Exploratory modeling for both vehicles was carried out using MATLAB. For the following (instrumented) vehicle, we fit a three-stage model, where a period of initial deceleration lasting about 2 seconds was followed by a period of gentler deceleration, which was followed by a short period of much stronger deceleration starting less than 1 second before the collision. For the leading vehicle, we fit a two-stage model where roughly 2 seconds of deceleration was followed by a period of being stopped, which lasted until the collision. Bayes estimates for each vehicle's initial speed, the time points at which each driver changed acceleration, and the corresponding accelerations were computed using WinBUGS, and these results are displayed in Table 3.3. For this case, because the data acquisition began after the leading driver had begun his or her final deceleration, we computed the follower's reaction time as the difference between when the lead vehicle came to a stop and when the following driver initiated the final deceleration.

At the time the forward radar acquired the leading vehicle, the driver of the instrumented vehicle was traveling at about 14.75 feet/sec while the driver of the leading vehicle was traveling at about 8.9 feet/sec. The leading driver was decelerating at about -3.7 feet/sec^2 and came to a stop about 2.4 seconds after the acquisition. The driver of the following vehicle was initially decelerating at about -2.9 feet/sec^2, but after about 1.9 seconds eased up to about -0.5 feet/sec^2. About 2.37 seconds after the leading vehicle came to a stop, the driver of the following vehicle began braking at about -12.6 feet/sec^2

DRAFT

but this was not sufficient to prevent a collision. At the time the lead vehicle came to a stop the following vehicle was about 20.7 feet behind the leader and traveling at about 9 feet/sec.

Figure 3.6 shows the speedometer speeds for the instrumented vehicle along with the speed history predicted by our model. Allowing for the piece-wise constant nature of the recorded speedometer output, the model gives a reasonable representation of the follower's speed profile.

Figure 3.7 shows the radar-measured range between the leading and following vehicle, along with the range predicted by our model. Again, the model gives a reasonable representation of the data.

To assess the avoidability of this crash, probabilities of collision were computed as a function of counterfactual final decelerations on the part of the following driver, and this relationship is displayed in Figure 3.8. For this event, because the following driver did not initiate evasive action until close to collision, even fairly high counterfactual decelerations are not sufficient to prevent the collision.

3.3 CASE NUMBER 104119

Description from video: (see Figures 3.9-3.11) In this event the instrumented (i.e., following) vehicle was traveling on a signalized roadway, and turned left at an intersection following the lead vehicle. The leader braked to a stop, as did the follower without colliding.

Approximately 35 seconds of data were available at 10 Hz, including speedometer-measured speeds for the instrumented vehicle and range and range-rate from the follower's forward radar. It was possible to reliably identify the lead vehicle in the radar data for about 16 seconds, and the speed, range, and range-rate for the period are displayed in Figure 3.12.

Figure 3.13 compares the speed trajectories of the leading and following vehicles.

Exploratory modeling for both vehicles was carried out using MATLAB. For the following (instrumented) vehicle, we fit a two-stage model, where a period of initial acceleration was followed by deceleration to a stop. For the leading vehicle, we fit a 3-stage model where a first stage of acceleration was followed by a stage of gentler acceleration, which was then followed by decelerating to a stop. Bayes estimates for each vehicle's initial speed, the time points at which each driver changed acceleration, the corresponding accelerations were computed using WinBUGS, and these results are displayed in Table 3.4. The following driver's reaction time was determined as the difference between when the follower began final deceleration and when the leader began final deceleration.

In this case, at the time of the start of the data series, the leading vehicle was traveling at about 18.9 feet/sec and accelerating at about 3.34 feet/sec^2, while the follower was traveling at about 20.6 feet/sec and accelerating at about 1.45 feet/sec^2. After about 3.46 seconds, the leader eased his or her acceleration to about 0.62 feet/sec^2, and after about 10.6 seconds from the start of the data series, the leader began decelerating at about

DRAFT

-11.94 feet/sec^2, which continued until the vehicle stopped. About 11.15 seconds from the start of the data series, the follower began decelerating at about -9.47 feet/sec^2. At the time when the leader began final deceleration, the follower was about 20.7 feet behind and traveling at about 36 feet/sec. The follower's reaction time was fairly quick, about 0.53 seconds.

Figure 3.14 compares the following vehicle's speed as given by its speedometer and the speed as predicted by the fitted model. Figure 3.15 shows a similar comparison of the range between the follower and leader as given by the forward radar and the range as predicted by the trajectory models. In both cases, we have a plausible reconstruction of the data series.

Finally, Figure 3.16 displays the probability of a rear-end crash occurring as a function of counterfactual final decelerations on the part of the following driver.

3.4 CASE NUMBER 73082

Description from video: (Figures 3.17-3.19) In this event both the leading and following vehicles are initially stopped at a signalized intersection. The lead vehicle accelerates, then the follower accelerates and leader pulls away from the follower. The leader then brakes to a stop, as does the follower, and the follower stops short of the leader without colliding.

Figures 3.20 and 3.21 show the speed, range, and range-rate plots of the following and leading vehicle.

Exploratory modeling suggested a three-stage model for the following vehicle and a four-stage model for the leader. Bayes estimates for each vehicle's initial speed, the time points at which each driver changed acceleration and the corresponding accelerations were computed using WinBUGS, and these results are displayed in Table 3.5. The following driver's reaction time was determined as the difference between the time when the follower's final deceleration began and the time when the leader's final deceleration began.

For this case, the follower's initial speed was taken to be known and equal to 0 feet/sec, while at the start of the data sequence the leader was traveling at about 2.15 feet/sec and accelerating at 10.1 feet/sec^2. After about 1.8 seconds, the leader reduced acceleration to about 4.1 feet/sec^2. After about 8.8 seconds the leader began decelerating at -3.9 feet/sec^2, and after about 10.2 seconds the leader increased the deceleration rate to about -13.5 feet/sec^2. The following driver initially accelerated at about 5.4 feet/sec^2, and after about 6.65 seconds eased to about 2.0 feet/sec^2. After about 12 seconds, the follower began braking at the relatively high rate of -18.3 feet/sec^2. The follower does not appear to have obviously responded to the leader's initial deceleration.

Figure 3.22 compares the following vehicle's speed as given by its speedometer and the speed as predicted by the fitted model. Figure 3.23 shows a similar comparison of the range between the follower and leader as given by the forward radar and the range as predicted by the trajectory models. In both cases, we have a plausible reconstruction of the data series.

DRAFT

3.5 CASE NUMBER 104851

In this case the instrumented vehicle (i.e., the following vehicle) at the start of the video (see Figure 3.24) took a right turn and continued to follow the leading vehicle. But then the leading vehicle decelerated and came to a complete stop. This forced the following vehicle also to decelerate, resulting in a near- crash. However, the following driver's deceleration was sufficient to enable the vehicle to come to a complete stop without any collision. Although the total length of the video was 19 seconds, the event actually happened within first 8 seconds. For the remaining period of time, both the vehicles were in stopped condition.

The leading vehicle came to a complete stop, indicated by the brake lights, which almost resulted in a rear-end collision.

Figure 3.25 shows the speed trajectories of the leading and the following vehicles. The blue line indicates the speed of the following vehicle and red line indicates the speed of the leading vehicle in feet/sec.

The speed of the following vehicle was obtained directly from the speedometer of the instrumented car. The approximate speed of the leading vehicle was calculated by adding the speed of the following vehicle and the range-rate data obtained from radar. A similar approach was adopted for all the remaining cases discussed in this section. Figure 3.26 shows the range and range-rate data obtained for this event.

After initial estimates of the change points and accelerations were obtained from MATLAB, the trajectory model was fitted in WinBUGS for final estimates. In this case a three-stage model was developed for the following vehicle, with initial acceleration followed by two different declaration stages. Table 3.6 gives the final MCMC simulation estimates of the parameters.

At the time the radar acquired the leading vehicle the initial speeds of the following and leading vehicle were 25.66 feet/sec and 26.07 feet/sec respectively. The leading vehicle decelerated in three different stages. The first two deceleration stages were characterized by mild deceleration followed by a very steep deceleration (-24.29 feet/sec^2) bringing the leading vehicle to a complete stop. Subsequently, the following vehicle initially was accelerating for 2.626 seconds, and then it decelerated at -21.76 feet/sec^2 followed by a third deceleration of -2.87feet/sec^2. The predicted piece-wise acceleration model was compared by fitting the observed data. The range and speed of the following vehicle was fitted as shown Figures 3.27 and 3.28. Table 3.6 lists the estimates obtained from the WinBUGS output.

Similar to the previous cases, a counterfactual model based on different deceleration rates for the following vehicle was simulated, and for each deceleration, a probability of crash was computed. Figure 3.29 shows how the chance of crash varies over different counterfactual deceleration values.

DRAFT

3.6 CASE NUMBER 104283

This event, as shown in Figures 3.30-3.32, is a near-crash. The instrumented vehicle was traveling in the right-most lane of an arterial and continued to travel until it was forced to a complete stop to avoid a rear-end collision with the leading vehicle, which was waiting for a gap to change lanes at a merging section of the arterial. The total duration of the video was 35 seconds and the event occurred at about 23 seconds.

Respective speed trajectories for the leading and following vehicles were plotted in Figure 3.33. The radar could only mange to capture the leading vehicle's information for about 5 seconds. Also the range and range-rate data were collected as shown in Figure 3.34.

Initial speed of the following vehicle was 60 feet/sec compared to the initial speed of 30.93 feet/sec for the leading vehicle. This speed is the estimated speed of the leading vehicle when for the first time the radar captured information about the leading vehicle. A two-stage model was proposed for the following vehicle where in the first stage the vehicle decelerated at -7.57 feet/sec^2 until 3.92 seconds then it shifted to a stronger deceleration rate of -16.16 feet/sec^2. The leading vehicle's trajectory was also fitted with a two-stage model with -10feet /sec^2 of deceleration in the first stage, for 1.714 seconds, and -4.332 feet/sc^2 deceleration in the second stage.

The piece-wise model seems to be plausible as shown in Figures 3.35 and 3.36 fitting the observed speed of the following vehicle and range data.

Also, a counterfactual model (see Figure 3.7) was developed which shows that if everything else remained constant but the deceleration in the last stage for the following vehicle was stronger than -5.7 feet/sec^2 then the probability of crash is essentially zero, while on the other hand if the deceleration value was weaker than about -4.2 feet/sec^2 a crash is a near certainty.

3.7 CASE NUMBER 60289

As shown in Figure 3.38, in this event the two vehicles were closely following each other. The leading vehicle accelerated and then traveled at uniform speed before it decelerated to almost zero speed. The following vehicle follows the same pattern as the leading vehicle, shown in Figure 3.39. Range and range data obtained from the radar is shown in Figure 3.40.

A four-stage model was constructed to fit the following vehicle's speed trajectory. Initially the following vehicle was traveling at 11.21 feet/sec and then accelerated at 3.15feet/sec^2 for 5.36 seconds. Then it traveled at almost constant speed for another 5 seconds before decelerating at -2.419 feet/sec^2 for 4.3 seconds followed by a strong deceleration of -10.74 feet/sec^2 and finally came to a stop at 16.47 seconds. A similar pattern was observed for the leading vehicle, which had an initial acceleration stage of 8.16feet/sec^2 for 1.764 seconds, followed by a period of 8.277 seconds of almost constant speed and then two deceleration stage with the final deceleration rate as high as

DRAFT

-9.502 feet/sec^2. The similar speed profile of the two vehicles seems reasonable as they were following closely each other in this case. Table 3.8 shows the WinBUGS estimates.

Predicted versus observed speed and range values (see Figures 3.41 and 3.42) were plotted and look quite reasonably represented by the four-stage model.

A Similar counterfactual model (see Figure 3.43) was developed to show the probability of crash for different deceleration values in the final stage of the following vehicle.

3.8 CASE NUMBER 92660

In this event a vehicle was closely following another vehicle on a two-lane two-way highway as shown in Figure 3.44. The leading vehicle suddenly stopped and the following vehicle had to swerve to avoid the collision as shown in Figure 3.45.

The speed trajectory (see Figure 3.46) shows that the leading vehicle initially traveled at mild acceleration for a considerable period of time followed by a step deceleration which was more intensified in the last or final phase. On the other hand, the following (instrumented) vehicle started with a gentle acceleration, and then moved at almost constant speed for some time, followed by the final deceleration as a reaction to the leading vehicle's behavior. Range and range data is shown in Figure 3.47.

The video shows that the following vehicle had to swerve around the lead vehicle to avoid a crash which suggests a two-dimensional analysis. But here we only focus on one dimensional approach and try to extract as much information we can.

Table 3.9 below lists the WinBUGS estimates for the parameters. A three-stage model was developed for both the leading and the following vehicles. The most highlighted result we can see from the estimates is very strong deceleration in the final stage. The leading vehicle had an initial speed of 40 feet/sec and it accelerated gently at 0.5 feet/sec^2 for 7 seconds before decelerating at -14.76 feet/sec^2 for approximately 2.3 seconds, followed by more intense deceleration of -23 feet/sec^2 and came to a stop. The following vehicle with an initial speed of 33.75 feet/sec accelerated at 2.48 feet/sec^2 for 4.23 seconds and then moved at almost zero acceleration for another 4.5 seconds before final deceleration of -15.57feet/sec^2. Figure 3.48 shows the predicted and observed range values for 92660.

The counterfactual model for this case study is not shown because the following vehicle, in addition to deceleration, actually swerved to avoid the crash.

DRAFT

DRAFT

CHAPTER 4

4. ANALYSES USING SITE-BASED VIDEO DATA

4.1 DATA ACQUISITION AND PREPARATION

The locations selected for analyses of site-based video data were two freeway segments on westbound Interstate Highway 94 near downtown Minneapolis. One is about 500 feet long, located between Minnesota State Highway 65 and Portland Ave., and the other is about 600 feet long, located between Portland Ave. and Park Ave. S. Figure 4.1 gives an overhead view of these two sites.

The traffic traveling on these two segments was recorded using two cameras installed by the Minnesota Traffic Observatory (MTO) on the roof of a 121-foot-high building near 3rd Ave. The videos were transferred to the MTO and saved at 640 x 480 resolution and 10 frame/sec, from 11 a.m to 8 p.m daily. Given a target vehicle, its time-series of positions expressed in screen coordinates were manually extracted frame-by-frame using the VideoPoint software (see Figure 4.2).

The screen coordinates obtained from Videopoint were then converted to ground coordinates by first matching several reference points on the video images to corresponding points on high-resolution satellite photos, and then applying standard photogrammetric methods.

DRAFT

4.2 I-94 CASE 1.

Description from the video (Figures 4.3 and 4.4): In this event the following vehicle (#2) and leading vehicle (#1) are traveling in the middle lane. Vehicle 2 collided with vehicle 1.

Trajectory data following the collision were not collected. Approximately 15 seconds of data were available from the video. Inspection of these data indicated that both vehicles were traveling at constant speeds approximately 9 seconds before the collision. Only the trajectory data from the last 6 seconds were used and these are displayed in Figure 4.5.

Exploratory modeling for both vehicles was carried out using the R software. For vehicle 1, we fit a two-stage model, where a gentle acceleration lasted for about 2.7 seconds, and was followed by a stronger deceleration. For vehicle 2, we fit a one-stage model, where a gentle acceleration lasted until the collision. Bayes estimates for each vehicle's initial speed, the time points at which accelerations changed, and the accelerations in all stages were computed using WinBUGS. These results are displayed in Table 4.1.

At the beginning of study period, driver 2 was traveling at 24.21 feet/sec, while driver 1 was traveling at 27.5 feet/sec. Driver 1 accelerated at 1.6 feet/sec^2 for about 2.7 seconds and then decelerated at -8.779 feet/sec^2 for about 3.1 seconds. Driver 2 was

accelerating at about 1.16 feet/sec^2 until the collision occurred. No evidence from video or trajectory data showed that driver 2 decelerated to prevent a collision.
Figures 4.6 and 4.7 show the distance trajectories of vehicles 1 and 2 along with the distance trajectory predicted by our models.

4.3 I-94 CASE 2.

Description from the video (Figures 4.8 and 4.9): In this event the following vehicle (#2) and leading vehicle (#1) were traveling in the middle lane. Vehicle 2 collided with vehicle 1.

Approximately 4 seconds of data were available from the video. The trajectory data used for analysis are displayed in Figure 4.10.

Exploratory modeling for both vehicles was carried out using R. For vehicle 1, we fit a two-stage model, where a strong deceleration lasted for about 1.5 seconds, which was followed by a less strong deceleration. For vehicle 2, we fit a one-stage model, where a gentle acceleration lasted until the collision. Bayes estimates are displayed in Table 4.2.

At the beginning of the study period, driver 2 was traveling at 33.02 feet/sec while driver 1 was traveling at 33.12 feet/sec. Driver 1 was decelerating at -10.52 feet/sec^2 for about 1.5 seconds and then decelerated at -5.436 feet/sec^2 for about 2.3 seconds. Driver 2 was decelerating at -2.512 feet/sec^2 but was not able to prevent the collision. Figure 4.11 and 4.12 show the distance trajectories of vehicles 1 and 2 along with the distance trajectory predicted by our models.

4.4 I-94 CASE 3.

Description from the video (Figures 4.13 and 4.14): In this event three vehicles are traveling in the right lane of study segment. Vehicle 3 collides with vehicle 2.

Approximately 7 seconds of data were available from the video. The trajectory data used for analysis are displayed in Figure 4.15.

Exploratory modeling for each vehicle was carried out using R. For vehicle 1, we fit a three-stage model, where a gentle acceleration lasting for about 2.3 seconds was followed by a 2.4 second stronger deceleration and then by a 2.4 second gentler deceleration. For vehicle 2, we fit a two-stage model, where a gentle deceleration lasting for about 3.5 seconds was followed by a 3.2 second stronger deceleration. For vehicle 3, we fit a three-stage model. The trajectory modeling showed that driver 3 decelerated in all three stages. Bayes estimates are displayed in Table 4.3. The reaction time of driver 2 was calculated as the time difference between driver 2's first change point and driver 1's first change point. The reaction time of drivers 3 was calculated as the time difference between driver 3's second change point and driver 2's first change point.

At the beginning of study period, drivers 1, 2, and 3 were traveling at 21.33 feet/sec, 33.42 feet/sec, and 51.29 feet/sec respectively. Driver 1 accelerated at 1.048 feet/sec^2 for about 2.2 seconds first, then decelerated at -8.8 feet/sec^2 for about 2.3 seconds, and then decelerated at -0.8968 feet/sec^2 for about 2.4 seconds. Driver 2 was

DRAFT

decelerating at -0.8968 feet/sec^2 which lasted for about 3.5 seconds and then decelerated at -8.0 feet/sec^2 for about 3.2 seconds. Vehicle 3 was apparently initially traveling at a much higher speed (51.29 feet/sec) than the other two vehicles. Although driver 3 noticed that the traffic in front was slowing down and started decelerating before entering the data collection segment, driver 3 was still not able to avoid collision with vehicle 2 after decelerating at -10.24 feet/sec^2 for about 2.3 seconds in its last stage.

Figures 4.16, 4.17, and 4.18 show the distance trajectories of vehicles 1, 2, and 3 along with the distance trajectory predicted by our models.

In this case, to assess the avoidability of collision between vehicles 2 and 3, probabilities of collision between these vehicles were computed as a function of counterfactual final decelerations (the third stage) of vehicle 3. This relationship is displayed in Figure 4.19.

We also evaluated the probability of collision between vehicles 1 and 2, assuming that driver 2 was decelerating with different rates in the last stage. This relationship is displayed in Figure 4.20. In this case, since the relative speed between the vehicles 1 and 2 in the last stages is small, for most counter-factual deceleration rates of vehicle 2, the corresponding probabilities of collision are either 1 or 0.

DRAFT

4.5 I-94 CASE 4.

Description from the video (Figures 4.21 and 4.22): In this event three vehicles were traveling in the right lane of study segment. Vehicle 3 collided with vehicle 2.

Approximately 10 seconds of data were available from the video. The trajectory data used for analysis are displayed in Figure 4.23.

Exploratory modeling for both vehicles was carried out using R. We fit three-stage models to each vehicle trajectory. Vehicle 1 had an acceleration for about 2.6 seconds, which was followed by 4.1 second strong deceleration first, and then by a 3.7 seconds of gentle deceleration. The behavior of drivers 2 and 3 was almost the same as that of driver 1 but with stronger deceleration in their last stages. The WinBUGS estimates are shown in Table 4.4. The reaction times of drivers 2 and 3 were calculated as the time difference of last change points between vehicles 2 and 1, and between vehicles 3 and 2.

At the beginning of the study period, drivers 1, 2, and 3 were traveling at about 29.41 feet/sec, 25.67 feet/sec, and 21.94 feet/sec respectively. Driver 1 was accelerating at 2.247 feet/sec^2 for about 2.6 seconds first, followed by a decelerations of -0.7111 feet/sec^2 for about 4.1 seconds and then of -7.312 feet/sec^2 for about 3.7 seconds. Driver 2 was accelerating at 2.514 feet/sec^2 for about 3.8 seconds, decelerated at -0.2509 feet/sec^2 for about 3.8 seconds, and then decelerating at -10.48 feet/sec^2 to avoid a collision with vehicle 1. Driver 3 noticed the strong deceleration of vehicle 2 in its last stage. Although driver 3 had the strongest final deceleration (at -12.89 feet/sec^2) for about 1.3 seconds, this was not sufficient to prevent the collision.

Figures 4.24, 4.25, and 4.26 show the distance trajectories of vehicles 1, 2, and 3 along with the distance trajectory predicted by our models.

In this case, the relationship between a counterfactual final deceleration of the vehicle 3 and the probability of collision with vehicle 2 is displayed in Figure 4.27. We also evaluated the probability of collision between vehicle 1 and 2, assuming driver 2 was decelerating at a different rate in the last stage. This relationship is displayed in Figure 4.28.

DRAFT

CHAPTER 5

5. Analyses Using CICAS Site-Based System

5.1 GENERAL METHODOLOGY

As indicated in our Phase II research plan, one of the goals for this project was to explore the utility and usability of existing data in traffic safety analysis, specifically in the investigation of surrogate measures of safety. Our approach requires vehicle trajectories. The data collected from the Cooperative Intersection Collision Avoidance Systems (CICAS) project intersections, at least at first glance, looked like a promising source of information. This infrastructure was designed for the purposes of the CICAS project however, and certain features constrain the usefulness of these data for more general research purposes. In adherence to the principal objective of the SHRP2 S01, effort was devoted to assessing the potential utility this extensive data source. This was done both in order to test our core methodology, and to identify possible future improvements to the data collection infrastructure. In this chapter we summarize the findings of the analysis, problems encountered along with proposed solutions, and conclude with the exploratory analysis of one near-crash event.

During the course of this project we explored all CICAS project sites as potential data sources. The Minnesota Highway 52 site, although it has more recorded crash cases, is also the one that deploys the older technology, while the North Carolina US 74E site has both better sensors, and corrects some errors in the collection and post-processing of data. These were the main reasons for focusing our efforts on the data originating from the North Carolina site. The site is illustrated in Figure 5.1, with the US 74 serving as the main road and Strawberry Blvd., being the side street. Our goal was to attempt to isolate near-crash events and utilize our structural model methodology to investigate the background and input conditions characterizing a crash or near-crash event, and their relation to the occurrence or non-occurrence of a collision. The data harvesting procedure involved the extraction of possible cases from the CICAS database, the use of filters for the removal of false positives, and the visual inspection of the resulting trajectories. It is important to note that in contrast to the other data sources utilized in this project, the magnitude of the data available from the CICAS site made necessary the development of automated methods for event data extraction and the creation of utilities accelerating manual inspection of the extracted information.

The CICAS vehicle trajectory database structure and available information is described in Appendix B. In summary, the information available for each vehicle includes X, Y coordinates in the state plane system, speed, and acceleration measurements, all at 10 Hz. It is vital to understand that these are not the raw collected measurements from the site but are derived from the raw measurements provided by the instrumentation during a data reduction stage. From the detection hardware utilized, we know that the sensor radar-based measurements for main approach vehicles are Range, Range-rate, and Azimuth at 10 Hz, with similar information collected from the LIDAR

DRAFT

sensors for the side road vehicle (although it is unclear at what interval). The post-processing of the raw data includes filtering as well as projections to known roadway features that introduce some difficulties for our study. For example, the position of the vehicle is always projected onto the lane centerline, resulting in some loss of information regarding the transition during lane changes, but more importantly during evasive actions taken by the drivers to avoid collisions. In addition, after closer inspection of the data, we concluded that the side road vehicle position is not updated at 10 Hz, since the provided data generate a step wise trajectory profile. Speed and acceleration for the side road vehicle display apparently unrealistic behavior, so we excluded them from this investigation.

Figure 5.1 describes the North Carolina site for the CICAS project. The CICAS database codes events based on side street vehicle maneuvers along with the gap duration from the main road vehicle movements. This database post-processing makes possible the harvesting of completed cross-main-line maneuvers where the accepted gaps of the side road crossing vehicle may have been small. This involves finding a plausible time envelope of small accepted lag times during the cross-main-line maneuver over the complete lifetime of the tracked vehicle. We separated the search by direction for the two cross-main-road maneuvers (one for each direction). The technique is illustrated for one of the maneuvers. Essentially, it involves two inner join steps. The first step is the inner join of all tracked targets associated with small gap events for the given maneuver within the observation lifetime of the tracked target that completed the maneuver (i.e., when the target was first and last seen by the system):

```
CREATE VIEW closecalls AS SELECT targetid, maneuver, veh_class,
sub1_acpt_lag_zone_exit_date, sub1_acpt_lag_zone_exit_time
FROM vehicle_acpt_lag
WHERE  (maneuver = 1) AND  (( 3.0 > ANY(sub1_acpt_lags))  OR ( 3.0 >
ANY(sub2_acpt_lags)));

CREATE VIEW lifetime AS SELECT vt.targetid, vt.zone_first_seen,
vt.date_first_seen, vt.time_first_seen, vt.zone_last_seen, vt.date_last_seen,
vt.time_last_seen
FROM vehicle_time AS vt
INNER JOIN closecalls AS cc ON (vt.targetid = cc.targetid);
```

The second step uses the lifetime of this target to harvest all tracked targets that were present during the evolution of the accepted lag times for the maneuver.

```
SELECT *
FROM tracked_targets AS tt
INNER JOIN lifetime AS ltm ON
(tt.date = ltm.date_first_seen) AND
(tt.time >= ltm.time_first_seen) AND
(tt.time <= ltm.time_last_seen);
```


DRAFT

The above procedure returned close to 3,000 cases. These are possible near-crash events. To try to reduce the false positives, a filtering program was developed, which will be described next.

The first step in the filtering program is to determine the predominate direction of travel for all the tracked vehicles. Essentially, this can be achieved by keeping track of the lane number as the target progresses in time. Second, since we are primarily interested in events representing near-crash cases between a side road vehicle and main road vehicle, after the first sort, the side road vehicles are examined. For each side road vehicle, main road vehicles are scanned at the same time step and run through a rejection process. In this process, if any test fails to return at least one possible conflicting main road vehicle, the time step is advanced. The process is summarized in the following steps:

Step 1: If no mainline vehicles are present for the given time step, then the time step is advanced.
Step 2: A distance between the main road vehicles and side road vehicle must lie below a specified threshold.
Step 3: The speed of the side road vehicle must be above a specified minimum threshold.
Step 4: The speed of the main road vehicle must be above a specified minimum threshold.

The rejection at each step can be stored in a table that contains the date/time and target IDs. The level of elimination is dependent on the specified speed/distance parameters of the trajectories. The result of the filtering program contained 297 cases of near-crashes similar to a recorded crash case.

A sample trajectory of a main road tracked vehicle can be seen in Figure 5.2a for one of the resulting maneuvers. Note that as long as the lateral position relative to the center lane is within a certain margin, the target is snapped to the lane. Generally what we have determined thus far is that the main road tracked vehicle trajectories seem plausible. Unfortunately, this is not the case with the trajectories of the side road vehicle. Although when the complete maneuver is plotted in space trajectories seem plausible (Figure 5.2b), when animated in time, it is clear that peculiar tracking errors exist at critical locations during the maneuver. Figure 5.3 illustrates one such case, which has been typical of what we have observed thus far.

The side road tracked vehicle completed a northbound cross maneuver over US 74 (aligned E-W in Figure 5.2b). We can reasonably assume that the vehicle never reverses course during the maneuver. Under such circumstances, the displacement curve from the start of the maneuver to the last tracked location should always monotonically increase. As can be seen in Figure 5.3, particularly at points A, B, and C, this is not the case.

The side vehicle seems to make jumps backward, which is unrealistic. More important, these jumps do not seem to be isolated points on the overall trajectory since progression seems realistic following these course corrections. This last element renders difficult the correction of these trajectories even manually. Following closer inspection of many cases, we have formed a number of hypotheses on the nature of these discrepancies.

DRAFT

Specifically, the errors seem to concentrate in areas where the side road vehicle is stationary or a main road vehicle is in close proximity. For the portable site deployments such as North Carolina, side road vehicles are essentially tracked through the intersection utilizing horizontal LIDAR units placed at each approach of the side road and at a location in the median. The height of the LIDAR horizontal scan is set to approximate the height of the vehicle bumper. The LIDAR scanner sweep could have missed the bumper, occlusions from other vehicles, sensor alignment (as the tracked vehicle passes from one sensor data collection field to another), or other algorithm processing factors may also contribute to the error. Of course, such anecdotal evidence are not enough to pinpoint the source of the difficulty and certainly are not enough to allow the creation of automated correction methodologies, which, in any case, are outside the scope of this project. The aforementioned issues constrained our ability to process large numbers of possible near-crashes. Regardless, the rest of this chapter includes the analysis, lessons learned, and modeling results for one near-crash event at the North Carolina CICAS site.

5.2 CICAS NORTH CAROLINA US-74E NEAR CRASH CASE: 12:11PM, APRIL 25th, 2007

Lacking video for this selected event we can only deduce the sequence of events from the animated trajectory information for the vehicles involved. Figure 5.4 shows a frame captured from our visualization tool.

At the left-hand side there is schematic of the intersection with the lines representing lane and turning centerlines, while on the right-hand side the speed and acceleration graphs for the mainline and side road vehicles are provided. A solid red line indicates the side road vehicle while magenta points indicate mainline vehicle. On the graph side, the green vertical line indicates the values at the current frame displayed. In Figure 5.5, a zoomed view of the trajectories is presented.

In Figure 5.5, the red circle indicates the mainline vehicle involved in the near-crash. The vehicle moves toward the intersection in subsequent frames. The green (short) part of the side vehicle trajectory is the movement already performed and the blue (longer) part is the trajectory in future frames. The conflict point for these trajectories is indicated in the figure. For the purposes of developing a model describing the vehicle trajectories, we translated all coordinates to a system that has its center at the conflict point and the X axis being parallel to the mainline vehicle trajectory. This way we managed to simplify the kinematic equations of the vehicles by using only the distance to the conflict point. The speed and acceleration values provided in the CICAS database already represent the projected vectors on the trajectory of each vehicle. From the animation of the event as well as the graphs of speed and acceleration, we deduced a description of the event. Specifically, the side road vehicle is performing a north bound crossing of the intersection beginning from a standing position approximately 11 meters (36 feet) from the conflict point. Considering this trajectory and movement rate, the mainline vehicle, which initially is moving with an approximate speed of 26.8 meters/sec (60 mph), will eventually collide with the side vehicle unless a change of attitude is implemented. The main road driver, realizing this, performs a sequence of decelerations of up to more than -

DRAFT

2 meters/sec^2 (-6.6 ft/sec^2) in order to prevent the crash. The maneuver is performed successfully with room to spare. It is important to note that the described behavior does not take into account any lateral movements of the vehicles, signifying types of evasive action other than deceleration, since if such movements did happen they were smaller than the threshold used in the database for projecting the vehicle to the closest lane centerline. As pointed out earlier the animation of the side vehicle, which depended on the XY coordinates, indicated that the reported speeds and accelerations do not agree. As can be seen in Figure 5.5 the side vehicle apparently never stops, but appears moves with a constant speed of approximately 11 mph (4.9 meters/sec) even while the acceleration increases from zero until well after the vehicle is clearly moving. For this reason for the rest of the analysis we only discuss the XY coordinates of the side vehicle and try to infer speed and acceleration from them.

The analysis of the data followed two routes. An empirical review of the information was performed to determine the concurrence of the provided information, followed with a more in-depth analysis producing Bayes estimates for each vehicle's initial speed, the time points at which accelerations changed, and the accelerations in all stages. The reason for the empirical review was based on the quick realization of errors and discrepancies in the information provided, some of which was described earlier in this chapter and will be illustrated in the following sections specifically for this event. For the first part of the analysis, approximately 19 seconds of data were used for the mainline vehicle, 12 seconds of these were spent approaching the conflict point from a starting point approximately 276 meters away. Figure 5.6 presents a subset of these for clarity.

From this figure we can immediately observe certain problematic measurements both in speed and distance. Specifically, speed exhibits a single point drop of ~1.7 kph (1.0 mph), which apart from being physically unrealistic is not evident in the distance measurements. To explore the extent of this phenomenon, we proceeded to use the available data of speed and acceleration to estimate displacement per interval and compared this to the displacement produced by the XY coordinates. This comparison can be seen in Figure 5.7. For the purpose of the estimation, we assumed the starting values were known and progressed from there. The estimates of distance based on speed and acceleration, although not identical, show better agreement with each other than they do with the provided distance from the database. This can be seen as an indication that displacement uses the sensor range measurement while the other two utilize range-rate. Still it is evident that significant filtering has been introduced.

In Figure 5.8, a similar comparison is made on speed values, one being the provided and the other calculated based on acceleration. For the speed estimates calculated from acceleration, the abnormal speed drop discussed earlier is not observed, but this can be either due to the introduction of an error in post-processing or an artifact of the sampling methodology that will not allow it to propagate the acceleration measurements. That is, if accelerations were originally estimated as finite differences of speed, then a step change in speed will generally not be present in the acceleration estimates. Although in this case this single error can be easily corrected, further investigation of the sensor characteristics and data reduction methodology would be

DRAFT

required to facilitate better use of CICAS collected mainline data with our trajectory based methodology.

Figure 5.9 presents the best information regarding the side road vehicle, distance to conflict point calculated from XY coordinates. From the figure it is evident that the measurements have both noise/errors and are not produced at the 10 Hz interval. The profile shows an unnatural step-wise progress for the vehicle. In this figure, back-and-forth movements of the vehicle can be observed as discussed earlier in this chapter. This case of a near-crash was selected for analysis after a large number of events were reviewed to locate a relevant case with the fewest discrepancies, so that manual post-processing with a minimal amount of abductive inference was possible. Specifically, in this case the problems can be found mainly in time periods that are not important for the analysis (before the vehicle comes to rest at the stop line and after it has sufficiently cleared the conflict point. Hence further investigation utilized only information from time -6 second to time 1 second, approximately.

Exploratory modeling for both vehicles was carried out empirically in Excel to identify a plausible piece-wise acceleration model. For the main line vehicle, we fit a three-stage model, where a gentle deceleration lasted for about 3 seconds, followed by a stronger deceleration for 1 second, and an even stronger one until the vehicle passed the conflict point. For side road vehicle, we fit a 1-stage model, where a constant mild acceleration carries it over the first part of the intersection. To avoid some of the errors and data discrepancies discussed earlier, further analysis utilized a subset of the data starting at point 8.2 seconds prior to the arrival of the mainline vehicle at the conflict point and ending at that point.

Bayes estimates for each vehicle's initial speed, the time points at which accelerations changed, and the accelerations in all stages were computed using WinBUGS. For this analysis we felt it was necessary to remove some of the most obvious problems exhibited in the provided data. Specifically, in the case of the mainline vehicle, we corrected the single point speed drop by adding a fixed value to all subsequent data points, while for the side vehicle we removed some abnormal back-and-forth data points at the very beginning of the time period. The WinBUGS results are displayed in Table 5.1, while Figures 5.10 to 5.12 display the comparison between model estimates and provided measurements.

One interesting observation resulting from the Bayes estimates of the driver behavior parameters involves the hypothetical mainline vehicle driver reaction time. Although the data we have are not sufficient to draw precise conclusions on what were the actions of the mainline driver, we can reasonably assume that the reaction to the side road vehicles encroachment on the mainline must be somewhere between the first and second change time points, -5.24 to -4.323. If we ignore the Bayes estimate for the side vehicles starting time of encroachment and literally follow the data provided, we can say that encroachment starts at around time point -5.012. This indicates a possible reaction time zone of -0.224 to 0.689. This is a very tight reaction time for a driver noticing movement at a distance of more than 70 meters (~230 feet). Alternatively, though, if we accept the Bayes estimate for the side vehicle time of encroachment, then we have a reaction time zone of 0.9 to 1.8 seconds, which is a much more reasonable.

DRAFT

In conclusion, in Chapter 4 we described how our method of trajectory-based reconstruction could be applied to site-based video data in order to estimate values of physical and behavioral variables characterizing crash and near-crash events. Reduction of the video data was carried out manually; issues concerning automatic reduction of video data, although important, were treated as outside the scope of this project. The main objective of the work described in Chapter 5 was to assess the ability of an alternative technology, based on using Doppler shift methods, to support our analytic approach. One advantage of this alternative technology is that the very real problem of automatically extracting vehicle trajectories from video is avoided, and that a database of potentially useful events currently exists. Our experience indicated that our methods of structural modeling and counterfactual screening can be applied to these data, but that to support this research the existing data collection and storage methods would require some technical modifications.

DRAFT

DRAFT

CHAPTER 6

6. CONCLUSIONS AND RECOMMENDATIONS

6.1 SUMMARY

Before proceeding it may be helpful the recall the primary goal of this project: to develop and test analytic methods that could then be applied to data produced by the SHRP2 safety field studies. The general approach was to first identify interesting research problems and then attempt to solve these using data similar to what is expected from the field studies. In Chapters 3-5 illustrated how Bayesian statistical methods can be used to fit models to trajectory data, the parameters of these models being related to interesting event features. Although the empirical results presented in Chapters 3-5 offer some tantalizing suggestions regarding the relationships between crashes and conflicts, and the degree to which non-crash events provide information about crashes, our position is that the number of events analyzed is too small to justify drawing general conclusions. Such conclusions of course are what one hopes will result from the field studies. In this chapter then we will restrict our discussion to methodological issues.

In the work plan submitted at the end of Phase I, we identified three research problems on which progress was needed. The first of these was identification of an appropriate class of structural models describing how crash and near-crash events developed, together with analytic tools for fitting these models to data expected from the vehicle-based and site-based field studies. The second problem involved counterfactual screening of supposed near-crash events to determine their similarity to crashes, and in the Phase I report we developed a theoretical result, which indicated that those near-crash events that are most like crashes would have evasive actions more extreme than those used in crashes. The third problem involved developing plausible models of how drivers select evasive actions as functions of the situations in which they find themselves. In our Phase I report we indicated that given an adequately large sample of crash and near-crash events for which estimates of background conditions and driver actions were available it should be possible to conduct exploratory modeling of evasive action selection using regression-type models.

Solutions to our second and third research problems are contingent on solution of our first, so the bulk of our effort during Phase II has been devoted to structural modeling of crash and near-crash events. In Chapter 2 of this report, we described a class of models that characterized driver behavior as a sequence of discrete changes in acceleration, and illustrated how an ordinary differential equation taking this acceleration sequence as an input, together with initial speeds and positions, could be integrated to give a predicted trajectory for the vehicle's motion. We then illustrated how the parameters describing a driver's acceleration sequence and the initial conditions could be estimated from vehicle trajectory data. We also showed how the identified model and estimates could be used to address the second research problem by computing the probability a crash would have resulted, other things being equal, as a function of a range of counterfactual evasive

DRAFT

actions. Chapter 3 of this report described the application of our methods to seven rear-ending crash and near-crash events obtained from the 100-car vehicle-based field study. Chapter 4 described application to six rear-ending crash and near-crash events using site-based video data obtained from the MTO, while Chapter 5 described a pilot application to an intersection angle conflict using site-based Doppler shift data obtained from the CICAS system.

6.2 CONCLUSIONS

Statistical analyses of crash frequency data can identify reliable associations between crash experience and roadway or driver features, but are of limited value in discovering how crashes occur. Our first conclusion is that at least for situations where the direction of travel is roughly constant, trajectory-based reconstruction of crash-related events, where trajectory data are used to fit parsimonious models of driver behavior, is feasible using both vehicle-based and site-based data. The product of such a reconstruction is a set of estimates of when and how much drivers changed their acceleration, and the background conditions associated with these changes. These estimates can in turn be used to produce estimates of driver reaction times, following distances, and selected/rejected gaps. Bayes estimates, and especially estimates of posterior probability distributions, can be obtained using Markov Chain Monte Carlo simulation. This approach is especially helpful in studying crash-related events involving two or more vehicles, where information on the behavior of drivers in non-instrumented vehicles is required.

One goal of traffic safety research is identifying causes of traffic crashes. The notion of cause has historically been rather difficult to pin down, but a good case can be made that a common core exists in what is meant by cause in statistical estimation of crash-reduction factors, in reconstruction of individual crashes, and in simulation of crash events. This common core is a counterfactual definition of cause (*17*). As we indicated in Chapter 1, a counterfactual component is also present in working definitions of traffic conflict and near-crash, and that given a simple structural model of an event along with Bayes estimates of the posterior distribution for the model's parameters, it is possible to quantify the degree to which a near-crash could have been a crash, by perturbing a driver's evasive action and computing the probability a crash results. Our second conclusion is that it is possible to extend the methods of counterfactual analysis to more complicated structural models involving differential equations.

In our Phase I report, we presented a theoretical argument to the effect that for a near-crash event to be similar to a crash event, the near-crash would have an evasive action more extreme than that seen in the crash. If crashes tend to involve evasive actions that are extreme then this would imply that those near-crashes that are most similar to crashes would tend to be less frequent than crashes. Although a conclusive test of this prediction would require more data than we had available in this study, Table 6.1 at least hints that this might not be as big a problem as feared.

The maximum evasive deceleration observed for the crash events was about -12.9 feet/sec^2, while the maximum evasive deceleration observed for a near-crash event was about -21.8 feet/sec^2, with the second-most extreme successful evasive deceleration being

DRAFT

about -18.3 feet/sec^2. Our third conclusion is that, at least for rear-ending events, there is some limited evidence that the distributions of evasive actions for crashes and near-crashes share some overlap, so that it should be possible to find near-crash events that are similar in other respects to crashes.

Vehicle-based data configurations have definite limits as to information provided about non-instrumented vehicles involved in multi-vehicle crashes or near-crashes. In Chapter 4, we illustrated how site-based video with manual extraction of vehicle trajectory data can support structural modeling and counterfactual analysis of multi-vehicle events, but clearly this is not feasible for processing very large numbers of events that might be expected from a longitudinal study. The CICAS data collection system, based on radar and LIDAR units, collects and processes large amounts of vehicle trajectory data at intersections. The current CICAS configuration and architecture is adapted to providing available gap information to minor approach drivers rather than processing and analyzing data on crashes and conflicts. With some technical modifications, however, this system has the potential to provide data needed for structural modeling of crash and near-crash events, at least at lower-volume intersections.

Conclusions are only as reliable as the data upon which they are based, and as described in Chapter 3, in many cases the vehicle-based data provided to us from the 100-car study showed incompleteness, inconsistencies, or errors that limited our ability to use them. These included cases where the forward radar data were missing or corrupted, cases where the speedometer data were clearly in error, and cases where there were marked differences in the vehicle trajectory as implied by the accelerometer measures and as implied by the speedometer and heading measures. It was also true for all cases we studied that the GPS coordinates were not sufficiently refined to determine vehicle trajectories. We recognize that the 100-car study was a pilot effort and that detailed quantitative reconstruction of events was probably not one of the study's objectives. None the less, our final conclusion is that the usefulness of the data produced by the SHRP 2 vehicle-based field study will be strongly dependent on the ability to calibrate and maintain the data collection systems.

DRAFT

6.3 RECOMMENDATIONS FOR FUTURE WORK

All but one of the crash and near-crash events we have analyzed so far show essentially straight line trajectories for the involved vehicles. Although we have developed preliminary model code that allows for two-directional trajectories, it was not possible within the time and resource constraints to bring this to the degree of maturity we achieved for straight-line events. Since two-directional motion occurs when a driver uses swerving as an evasive action, and when a left-turning driver selects a gap, our first recommendation is for development and testing of trajectory estimation tools that handle two-directional trajectories.

In the reconstruction of crash and near-crash events, driver inputs such as acceleration rates, reaction times, and following distances can be treated as exogenous

quantities to be estimated, and the issue of how drivers select evasive actions does not arise. When attempting to include crash events in a microscopic traffic simulation, however, plausible models that close the feedback loop between existing conditions and driver actions will be necessary. Our second recommendation is for research on this issue, using the data from the SHRP 2 field studies.

In our Phase I report we pointed out that in some cases the residuals obtained after fitting a trajectory model showed serial correlation. When serial correlation is present but unaccounted for, the standard errors and confidence intervals associated with parameter estimates tend to suggest more precision for those estimates than is justified. That is, although in all our analyzed cases the trajectory models we fit appear to give reasonable descriptions of data, it may be that there is greater uncertainty in the parameter estimates than we have so far acknowledged. Our third recommendation is that our model estimation methods be enhanced to allow for possible serial correlation.

Our experience with vehicle-based data indicated that except for car-following events where both vehicles remain primarily in the same lane, the quantitative information available about driver behavior in non-instrumented vehicles is essentially non-existent. Compiling data on gap-selection and other intersection-related events will then require a different data setup. Our fourth recommendation is that the vehicle-based study be complemented with site-based research.

During the course of this project a recurring factor was the unknown or uncertain influence of the measurement method/hardware, as well as the post-collection filtering, on the available data. We believe that in all future studies consistency and transparency in data collection and processing methodologies is paramount. Past efforts naturally were affected by the objectives and priorities of the projects funding, so that future dissemination of data may have received less emphasis. Considering that the primary objective of SHRP2 S01 was to examine the usability of existing data, it is clear that the broader utility of data in a project be recognized. For these reasons our final recommendation is that, beginning with SHRP2 and similar federally-funded projects, a clause be added to each RFP, making sure that the data collection setup, post-collection processing, and storage/availability information is clearly described in the final report.

DRAFT

REFERENCES

1. Hauer E., "The engineering of safety and the safety of engineering", In: Rothe, J. (Ed.) *Challenging the Old Order: Towards New Directions in Traffic Safety Theory*, Transaction Publishers, London, UK (1990), pp. 29-71.

2. Freedman, D., "From association to causation via regression", in V. McKim and S. Turner (eds.) *Causality in Crisis?* University of Notre Dame Press, Notre Dame, IN (1997), pp. 113-162.

3. Bonneson, J., and Lord, D., "Theory, explanation, and prediction in road safety", presentation at *85th Annual Transportation Research Board meeting*, Washington, DC. (2006).

4. Transportation Research Board, SHRP 2 Request for Proposals, Development of Analysis Methods Using Recent Data, Washington DC, September 11, 2006.

5. Pearl, J., *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, UK (2000), 384 pp.

6. Brill, E., "A car-following model relating reaction times and temporal headways to accident frequency," *Transportation Science*, 6 (1972), 343-353.

7. Davis, G. and Swenson, T., "Collective responsibility for freeway rear-ending accidents? An application of probabilistic causal models", *Accident Analysis and Prevention*, 38 (2006), pp. 728-736.

8. Guttinger, V., "Conflict observation in theory and practice", in E. Asmussen (ed.) *International Study of Traffic Conflict Techniques*, Springer-Verlag, Berlin (1984), pp. 17-24.

9. Dingus, T., Klauer, S., Neale, V., Petersen, A., Lee, S., Sudweeks, J., Perez, M., Hankey, J., Ramsey, D., Gupta, S., Bucher, C., Doerzaph, Z, Jermeland, J., and Knipling, R. *The 100-car Naturalistic Driving Study Phase II-Results of the 100-car Field Experiment*, Report DOT HS 810 593, National Highway and Transportation Safety Administration, Washington, DC (2006).

10. Fambro, D., Fitzpatrick, K., and Koppa, R., *Determination of Stopping Sight Distances*, NCHRP Report 400, National Academy Press, Washington DC (1997).

DRAFT

11. Bayarri, M., Berger, J., Higdon, D., Kennedy, M., Kottas, A., Paulo, R., Sacks, J., Cafeo, J., Cavendish, J., Lin, C. H, and Tu, J. *A Framework for Validation of Computer Models*, Technical Report 128, National Institute for Statistical Sciences, NC (2002).

12. Seber, G. A. F. and Wild, C. J., *Nonlinear Regression*, Wiley, New York, NY (1989), 792 pp.

13. Lunn, D., A. Thomas, Best, N., and Spiegelhalter, D., "WinBUGS- A Bayesian modeling framework: concepts, structure, and extensibility," *Statistics and Computing*, 10 (2000), 325-337.

14. *Using MATLAB (Version 6)*. The Math Works, Inc., MA (2002).

15. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org, (2008).

16. Balke, A. and Pearl, J., "Probabilistic evaluation of counterfactual queries", *Twelfth National Conference on Artificial Intelligence* (AAAI-94), Seattle, WA (1994).

17. Davis, G., "Towards a unified approach to causal analysis in traffic safety using structural causal models", *Transportation and Traffic Theory in the 21st Century* (2002), 247-266.

DRAFT