

FINAL
CONTRACT REPORT
VTRC 08-CR2

**LIVE STOP-CONTROLLED INTERSECTION
DATA COLLECTION**

ZACHARY R. DOERZAPH
Research Associate

VICKI L. NEALE
Center Director

JODI R. BOWMAN
Senior Research Specialist

KENDRA I. WIEGAND
Project Associate

Center for Vehicle Infrastructure Safety
Virginia Tech Transportation Institute



Standard Title Page – Report on State Project

Report No.: VTRC 08-CR2	Report Date: December 2007	No. of Pages: 80	Type of Report: Final Contract	Project No.: 80637
			Period Covered: March 2006–December 2007	Contract No.:
Title: Live Stop-Controlled Intersection Data Collection			Key Words: Intersection, safety, collision avoidance, ICAS, CICAS, stop-controlled, unsignalized	
Author(s): Zachary R. Doerzaph, Vicki L. Neale, Jodi Bowman, and Kendra Wiegand				
Performing Organization Name and Address: Virginia Tech Transportation Institute 3500 Transportation Research Plaza (0536) Blacksburg, VA 24061				
Sponsoring Agencies' Name and Address: Virginia Department of Transportation 1401 E. Broad Street Richmond, VA 23219				
Supplementary Notes:				
<p>Abstract:</p> <p>This report describes an experimental investigation performed at live intersections to gather infrastructure-based naturalistic driver approach behavior data. This data was collected and analyzed with the goal of understanding how drivers approach intersections under various speeds and environmental conditions. Six stop-controlled intersection approaches across five intersections in the New River Valley, Virginia area were selected for data collection. The sites were selected based on the intersection characteristics and crash statistics. Data was collected from each site for at least two months resulting in over sixteen total months of data.</p> <p>A data acquisition system was devised and implemented to obtain the first intersection data set with fidelity sufficient for developing intersection collision avoidance threat assessment algorithms. An explorative analysis of driver stopping behavior and vehicle trajectories was also performed. Results indicate that an intersection collision system for stop-controlled intersections is feasible. Avenues for future research and potential uses of this new database are highlighted.</p>				

FINAL CONTRACT REPORT
LIVE STOP-CONTROLLED INTERSECTION DATA COLLECTION

Zachary R. Doerzaph
Research Associate
Center for Vehicle Infrastructure Safety
Virginia Tech Transportation Institute

Vicki L. Neale
Center Director
Center for Vehicle Infrastructure Safety
Virginia Tech Transportation Institute

Jodi R. Bowman
Senior Research Specialist
Center for Vehicle Infrastructure Safety
Virginia Tech Transportation Institute

Kendra I. Wiegand
Project Associate
Center for Vehicle Infrastructure Safety
Virginia Tech Transportation Institute

Project Manager
Catherine C. McGhee, Virginia Transportation Research Council

Contract Research Sponsored by
Virginia Transportation Research Council

Virginia Transportation Research Council
(A partnership of the Virginia Department of Transportation
and the University of Virginia since 1948)

Charlottesville, Virginia

December 2007
VTRC 08-CR2

NOTICE

The project that is the subject of this report was done under contract for the Virginia Department of Transportation, Virginia Transportation Research Council. The contents of this report reflect the views of the authors, who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Virginia Department of Transportation, the Commonwealth Transportation Board, or the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

Each contract report is peer reviewed and accepted for publication by Research Council staff with expertise in related technical areas. Final editing and proofreading of the report are performed by the contractor.

Copyright 2007 by the Commonwealth of Virginia.
All rights reserved.

ABSTRACT

This report describes an experimental investigation performed at live intersections to gather infrastructure-based naturalistic driver approach behavior data. Data were collected and analyzed with the goal of understanding how drivers approach intersections under various speeds and environmental conditions. Six stop-controlled intersection approaches across five intersections in the New River Valley, Virginia area were selected for data collection. The sites were selected based on the intersection characteristics and crash statistics. Data were collected from each site for at least two months resulting in over sixteen total months of data.

A data acquisition system was devised and implemented to obtain the first intersection dataset with fidelity sufficient for developing intersection collision avoidance threat assessment algorithms. An explorative analysis of driver stopping behavior and vehicle trajectories was also performed. Results indicate that an intersection collision system for stop-controlled intersections is feasible. The infrastructure required to implement a collision avoidance system will have a significant cost associated with it. Additional research will be required to determine if that cost can be outweighed by the safety benefits provided by such a system. Avenues for future research and potential uses of this new database are highlighted.

INTRODUCTION

Intersection crashes constitute over 35% of the nation's traffic-related fatalities (NHTSA, 2005). Currently, there is a major federal/industry initiative underway to determine the efficacy of Cooperative Intersection Collision Avoidance Systems (CICAS). One objective of a CICAS warning system is to mitigate crashes by providing imminent warnings to drivers who might otherwise violate a signalized or stop-controlled intersection control device. This report presents an initial step in an effort to determine driver approach behavior at intersections for the purpose of supporting development of a warning algorithm component of a stop-controlled intersection collision avoidance system (ICAS) to prevent violations.

To provide an effective violation warning to a driver, the ICAS must discriminate between compliant and non-compliant stopping behavior. This discrimination must be performed significantly upstream from the intersection such that drivers have sufficient time to perceive the warning, react, and stop their vehicles. Drivers with varying levels of experience, judgment, risk tolerance, and mood/emotions will respond differently as they approach an intersection. As a result, the behavior of a compliant driver may overlap with that of a violating driver at locations where a countermeasure should be deployed. This could lead to cases in which a driver that did not need a warning receives one, or vice versa. Thus, engineers and designers must work to construct a threat assessment algorithm that enhances safety while avoiding unnecessary alarms.

Over the past three years, a significant volume of test-track research has been completed on ICAS (e.g., Lee et al., 2005; Neale, Perez, Doerzaph, & Stone, 2005). From this research, a substantial amount of knowledge has been acquired about driver response to various ICAS countermeasure types. Countermeasures tested include in-vehicle visual, auditory, and haptic displays as well as visual alerts located on the roadway infrastructure. The test track environment has worked well for making relative comparisons between countermeasure options and determining at what point an inattentive driver must receive a warning in order to stop. Nonetheless, test track data are limited when applied to algorithm design. This is due to the complex and dynamic nature of the real-world event, experimental effects, and the interaction with human behavior (e.g., distractions, motivations, etc.) that cannot be re-created on the test track. This deficiency resulted in a need to gather observational data at actual intersections across a variety of experimental conditions. From this observational data, an algorithm can be created to better represent actual driver behavior as an intersection is approached. The algorithm can then be combined with the previous and ongoing countermeasure work to determine the ICAS performance (e.g., will a timely warning also result in an unacceptable number of nuisance alarms).

Warning timing will ultimately determine the effectiveness of an ICAS. Alarms that are too early will likely deflate the safety benefits of collision avoidance systems because of annoyance and loss of user trust in the system (Dingus, Jahns, Horowitz, & Knipling, 1998). Warnings that are too late will fail to prevent an intersection collision. Thus, it is imperative that researchers and engineers thoroughly investigate driver behavior at intersections. In particular, the frame-by-frame trajectories of the intersection approaches must be analyzed to understand how and when to warn appropriately.

This report describes an experimental investigation performed at live stop-controlled intersections to gather naturalistic driver approach behavior. Data were collected and analyzed with the goal of understanding how drivers approach intersections under various speeds and environmental conditions. The data from this collection were used to explore the feasibility of providing violation collision warnings to drivers who would otherwise violate the stop sign. Six stop-controlled intersection approaches across five intersections in the New River Valley, Virginia area were selected for data collection. The sites were selected based on the intersection characteristics and crash statistics. Data were collected from each site for at least two months resulting in over sixteen total months of data.

The database described and the analyses performed herein will feed into a growing body of knowledge regarding intersection collision avoidance. Intersection collision avoidance has been identified by the Federal Highway Administration (FHWA) as a key initiative and will thus receive a significant level of attention and funding over the next several years. Prior to a field operation test (FOT), the threat algorithm to warn drivers at stop-controlled intersections must be identified and optimized. It is the purpose of this study to contribute to this effort by developing the research protocols and data collection system for obtaining data for stop-controlled algorithm development. Furthermore, an explorative analysis was performed to understand driver stopping behavior and to demonstrate a proof of concept for a stop-controlled ICAS algorithm.

Previous ICAS and general intersection research provided a substantial amount of preliminary knowledge regarding the appropriate warning timing. The following sections introduce the work completed to date. Numerous studies, consisting primarily of database analyses, have been conducted in order to examine factors that may be involved in intersection control device violations and crashes. A review of these studies provides numerous scenarios, environmental factors, and driver characteristics that were found to contribute to the large number of intersection crashes each year. This ICAS background is used as a foundation for the experimental design and site selection presented in the methods section.

Purpose and Scope

Relative to other roadway segments, intersections occupy a small portion of the overall infrastructure; however, they represent the location for a large portion of the annual automotive crashes in the United States. Thus, intersections are an inherently dangerous roadway element and a prime location for vehicle conflicts. Traditional safety treatments are effective at addressing certain types of intersection safety deficiencies; however, cumulative traffic data suggests these treatments do not address a large portion of the crashes that occur each year.

ICAS represent a new breed of countermeasures that focus on the types of crashes that have not been reduced with the application of traditional methods. One objective of an ICAS is to mitigate crashes by providing imminent warnings to drivers who might otherwise violate a signalized or stop-controlled intersection control device. The present study is the first investigation of continuous driver behavior at stop-controlled intersections aimed at investigating ICAS. This report provides evidence for the efficacy of ICAS and roadside data collection for the purpose of ICAS algorithm development.

Overview of the Intersection Crash Problem

Crashes at intersections represent about 35% of all annual crashes (NHTSA, 2005). Most intersection crashes can be categorized as either rear-end or crossing path (CP) conflicts. Other crash types include single vehicle, pedestrian/cyclist, head-on and sideswipe. CP crashes cause an estimated 26.7% of all crash-caused delays totaling approximately 120.3 million vehicle hours annually (Wang & Knipling, 1994). Violation-related CP maneuvers account for 393,000 annual crashes at a cost of 39 billion dollars (Lee et al., 2005).

The investigation discussed herein addresses CP crashes at stop-controlled intersections resulting from a violation of the traffic control device (TCD). Overall, stop-controlled CP crashes account for 374,000 (38%) of the annual intersection crashes, resulting in 3,994 fatalities (Lee et al., 2005). CP crashes in which a violation was cited in the police report occur 184,000 times each year.

The purpose of the next sections is to decompose the intersection CP crash problem. It will begin by introducing terminology followed by a review of the factors that influence intersection crashes. The typical vehicle intersection approach will then be dissected by considering environmental factors and required driver decisions with respect to human capabilities and limitations. The focus of this discussion is to understand and attempt to predict the factors leading to intersection CP crashes.

Intersection and Crash Taxonomy

The incidence and severity of intersection-related crashes vary with, among other factors, the relative positions and travel directions of the vehicles involved. For this reason, various researchers have created different taxonomies of intersections and intersection crash types. These taxonomies are based on the various combinations of vehicle conflicts that can occur at a typical intersection. Use of these classification systems allows researchers to gather statistics, re-create, analyze, experiment, and communicate particular intersection-crash situations in a repeatable manner. For this study, these taxonomies were used to select intersections, identify maneuvers for study, and communicate the results throughout this report.

Crash taxonomies consider the two different vehicles that must be present for an intersection crash to occur: the subject vehicle (SV) and the principal other vehicle (POV). The vehicle of interest is the SV; its travel path is intersected by the POV. The actions of the SV initiate the conflict sequence since this is a violating vehicle. The POV has the right of way in these crash sequences. Combining taxonomies from several sources (Ferlis, 2001b; Wang and Knipling, 1994; Najm et al., 1995; Pierowicz et al., 2000), three primary taxonomies are presented in a graphical format (Figure 1, Figure 2, Figure 3).

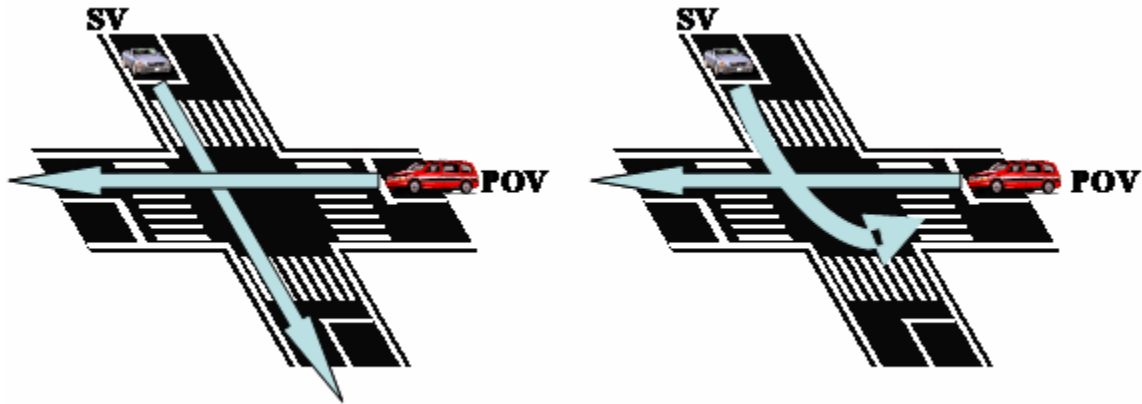


Figure 1. Intersection straight crossing path (SCP).



Figure 2. Left turn across path (LTAP).

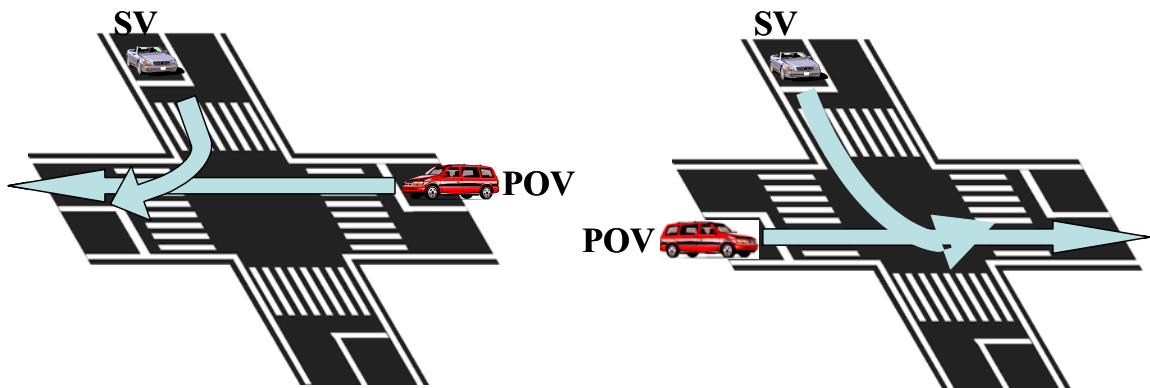


Figure 3. Turn into path (TIP) – Merge conflict: a) Right – RTIP, b) Left – LTIP.

A majority of intersection crashes result from a vehicle traveling straight through an intersection (over 50%), followed next by drivers making a left turn (Wang and Knipling, 1994). Left-turn-across-path lateral-direction (LTAP/LD) crashes are the most frequent at signalized

intersections, while SCP crashes are the most frequent at stop-controlled intersections (Bellomo-McGee, Incorporated [BMI], 2001a).

Another crash type related to intersections is rear-end crashes, which occur most frequently at or near signalized intersections (National Transportation Safety Board, 2001; Yan, Radwan, & Abdel-Aty, 2005). A majority of these collisions are due to an abruptly stopping lead vehicle, in which the following vehicle's driver is labeled by a reporting officer as either inattentive or following too closely (Yan et al., 2005).

Intersection Crash Causation

Several researchers have analyzed national databases, collected data at intersections, or collected in-vehicle driver data to understand the occurrence of particular intersection crash types and causation. Crash causation theories provide the reader with background in order to explain the intersection selection criteria used for the present research as well as general background of the primary ICAS objective: reducing intersection crashes. The following paragraphs summarize the crash causation findings of previous research efforts.

Intersection Geometry

Intersection geometry consists of multiple sub-variables including the angle of approaches, approach curvature, presence of pedestrian crossings, and dimensions. Intersection dimensions include the stopping sight distances, width between legs and the distance from the stopbar to the adjacent traffic lanes (i.e., the distance over the stopbar a vehicle can traverse before entering the adjacent traffic stream).

Wierwille et al. (2000) found a relationship between intersection complexity and crash rate; intersections with simpler geometric designs had fewer crashes. Research by Hendricks et al. (1999) analyzed a sample of 723 driver-at-fault crashes from 1996 to 1997. The most frequent fundamental factors cited for SCP crashes were "looked, did not see," "driver inattention/traffic control device violation," and "crossed intersection with obstructed view." The researchers provided typical infrastructure characteristics for each of these causal factors:

Looked, did not see

- All cases occurred at stop-controlled 90-degree intersections.
- In 71% of the cases, the victim vehicle (POV) was struck on the passenger side.

Driver inattention/traffic control device violation

- All crashes occurred at 90-degree intersections that typically use traffic signals.

Crossed intersection with obstructed view

- All cases occurred at stop-controlled 90-degree intersections.
- In 57% of the cases, the victim vehicle (POV) was struck on the passenger side.

Retting et al. (2003) completed research of stop-controlled intersection violations through reviewing police reports of 1,788 crashes from selected areas. A majority of the drivers (i.e., two-thirds) reportedly came to a stop before proceeding into the intersection, while 17% of drivers admitted to neglecting to stop before entering the intersection. Failure to see another

vehicle (44%) and obstructed vision (16%) were reportedly the most frequent reasons for a vehicle to proceed into an intersection despite the presence of another vehicle. It was also reported that approximately 50% of the stop-controlled collisions occurred at “T-intersections” (i.e., “locations where the stop sign-controlled approach road terminates at the intersection”) (Retting et al., 2003, p. 488).

Time of Day

As reported by Wang and Knipling (1994), a majority of intersection crashes occur during the peak hours of the day, with twice as many CP crashes occurring in the mornings (6:31 to 9:30) than during afternoon traffic hours (15:31 to 18:30). Additionally, the influence of the day of the week and day of the month were used to determine the effect on intersection violations. Ambient lighting conditions present during the times evaluated have not been considered directly.

Driving Conditions

Numerous studies, based on data collected from the General Estimates System (GES) databases, have reported that a majority of intersection crashes occur in clear weather and on dry pavement. SCP crashes at both signalized (Tijerina et al., 1994) and unsignalized (Chovan et al., 1994) intersections occur mainly on dry pavement, in good weather, and during daylight conditions. More specifically, 74% of unsignalized crashes take place on dry pavement, while only 24% occurred when the pavement was wet or snowy (Chovan et al., 1994).

Speed

Liu (2007) completed a study in which he examined two separate signalized intersections in order to determine contributing violation factors. Numerous decisions must be made as a driver approaches an intersection. The speed at which a driver is traveling when arriving at the intersection is a large determinant as to how the driver reacts.

Through a database analysis, Wang and Knipling (1994) reported that most intersection-related crashes occur when the posted speed limit is 15 m/s (35 mph) or less. Yang and Najm (2006) reported from their research that a majority of crashes occurred at around 8 m/s (18 mph) with an average speed of 14 m/s (32 mph). Similarly, Chovan et al. (1994) reported that a number of the crashes occurred on urban roads with lower speed limits. Conversely, Brewer et al. (2002) reported that the number of drivers who committed red-light violations increased as their approach speeds increased, with the largest number of violations occurring at travel speeds of 55 mph.

Definition and Prevalence of Violations

Not every violation of the TCD results in a crash. Wierwille et al. (2001; 2000) showed that TCD violations for both signalized and stop-controlled intersections are not uncommon (up to 15 violations per hour [ITE, 2003] have been reported). However, crashes result when the timing of the violation aligns with the presence of conflicting vehicle(s). Identifying when a crash will occur based on the interactions of multiple vehicles is a difficult problem. An easier

and equally effective method for preventing intersection crashes due to violations is to mitigate the violations themselves. As such, characteristics of violation approaches contribute to the design of the threat assessment algorithm. The purpose of this section is to determine factors that influence a driver's decision to violate. The following includes research on signalized intersections as complimentary studies relating to stop-controlled intersections that have not been completed to date.

The goal of an ICAS system is to mitigate intersection crashes by preventing violations. Thus, it is productive to review the types and causes of intersection violations. Fakhry and Salaita (2002) have provided some definitions of violations for stop-controlled intersections. These classifications, provided below in Table 1, categorize the vehicle as either "at speed" or "rolling stop". This particular classification does not provide a breakdown of the speed at which the driver is moving through the intersection under each classification.

Table 1. Summary of taxonomies of stop-controlled intersection violations.

Intersection Type	Violation	Definition
Stop-controlled	"At Speed"	Vehicle does not slow down and proceeds through the intersection. Also included are vehicles that "barely slowed" as they approached the intersection.
Stop-controlled	"Rolling Stop"	Wheels of a vehicle are slowing but not stopping and the vehicle continues through the intersection.

*As cited in Fakhry and Salaita, 2002.

Direct observations of violations provide some insight into the extensive violation problem. Fakhry and Salaita (2002) reported that "at speed" violations at stop-controlled intersections were recorded approximately 20 times more often than at signalized intersections. They also reported that out of 1,000 vehicles, 17.5 vehicles committed stop-controlled TCD violations and 1.5 committed signalized TCD violations. In addition, "rolling stop" violations greatly exceeded "at speed" violations at both signalized and unsignalized intersections.

Sudweeks, Neale, Wiegand, Bowman & Perez (Unpublished data) conducted an analysis of stop-controlled TCD violations and near violations. This study identified 77 drivers and 174 stop-controlled intersections from the data collected in the 100-Car Naturalistic Driving Study (Dingus et al., 2006), in which 100 cars were instrumented to collect behavioral and environmental data in a naturalistic driving setting. The study found that 61% of stop-controlled TCD violations occurred at speeds greater than 10 mph, while 39% occurred between 6 and 10 mph. Regarding driving maneuvers, 50% of violating drivers crossed straight through the intersection, 27% of violators performed left turn maneuvers, and 23% performed right turn maneuvers.

Driver error is frequently regarded as the primary casual factor in intersection CP crashes through either an intentional or unintentional disregard of the TCD, as reported by ITE (2003). For instance, 40% of red-light violators claimed that they "looked but did not see" the TCD. Another 12% mistook the signal color, claiming they had a green indication. Four percent blindly followed another vehicle into the intersection and did not look at the signal. Finally, 3% of the drivers were confused about which signal indication was theirs. In a study of driver error at intersections, Wierwille et al. (2000, 2001) found that 3.3% of all drivers entering a complex

signalized intersection committed some sort of error. Driver distraction may have contributed to many of these errors, particularly those that were made unintentionally. This group of inattentive drivers may represent the population which can most effectively be addressed by an ICAS.

Both Pierowicz et al. (2000) and Tijerina et al. (1994) have presented causal factors that imply a deliberate disobedience of TCD. In a survey of drivers, 8% admitted to intentionally running the red light (ITE, 2003). Drivers in the survey were provided with a GO/NO-GO scenario in which the GO decision could be interpreted as aggressive. Twenty-nine percent of the drivers opted to take the aggressive action. Of those drivers, 69% indicated that their motivation was due to being in a rush/save time and 12% reported doing it out of frustration (ITE, 2003). The decision to run or attempt to beat a traffic signal is due to a belief that a collision can be avoided. This belief could be based upon the failure to see cross traffic, misjudgment of the velocity, distance or direction associated with perceived traffic, or the assumption that other vehicles will yield to the violating vehicle. Although 99% of surveyed drivers acknowledged the dangers of red light running, they also perceived a low likelihood of receiving a ticket for the infraction (ITE, 2003).

Algorithm Development

Most human factors research is concerned with how a typical person responds under a given set of circumstances. Often, human factors engineers design a system to be used by a majority of people by considering a certain percentile of the population (Sanders & McCormick, 1993). In contrast, a collision avoidance system is specifically designed to address uncommon behaviors. For ICAS, these behaviors consist of two similar groups: 1) drivers that will violate the TCD and 2) somewhat aggressive drivers who appear to be violators even though they are complying with traffic regulations. To avoid false alarms, the ICAS algorithm must distinguish between aggressive drivers and violators such that a countermeasure is only delivered to those drivers for which it is a necessity.

An ICAS algorithm should, at a minimum, predict four possible scenarios for a driver approaching a stop-controlled intersection: 1) the approaching driver is aware of the TCD and stops appropriately, 2) the approaching driver is aware of the TCD, initiates stopping but does not stop (e.g. “rolling stop”), 3) the approaching driver is aware of the TCD but stops aggressively, and, 4) the approaching driver is either aware or not aware of the TCD and does not stop. “Rolling stop” cases are of particular interest because they have been found to be common and relatively safe vehicle maneuvers (Fakhry and Salaita, 2002). As such, in an effort to reduce nuisance alarms, perhaps these drivers should not receive a warning. Aggressive compliant drivers should also be identified in order to prevent unnecessary warnings. Distinguishing between attentive and inattentive violations may not be feasible through the methods proposed here, nor can it currently be integrated into an ICAS. Existing data supporting the three remaining scenarios will be reviewed for a variety of potential algorithm inputs. As with the previous sections, much of the literature reviewed in this chapter is from signalized intersections due to its availability, while there is a general lack of research concerning stop-controlled intersections.

Support for the feasibility of an ICAS algorithm at live intersections was provided by a naturalistic observation study of human intersection approach behavior (White & Ferlis, 2004). A hand-held radar gun was used to collect the intersection approach behavior of drivers at four intersections. The researchers collected data for two groups of drivers: those that were stopping for a red light and those that were going through a green light. It was presumed that drivers who went through a green signal would act identically to an inattentive violator. Drivers that received the phase change were not included in the sample so the scenario is similar to a stop-controlled approach (approaching a red light should be similar to approaching a stop sign).

The dataset consisted of 270 samples taken at four separate intersection sites. The speed and acceleration profiles of vehicles were then used to distinguish into which group the given driver fell in advance of the intersection. The initial speeds of both groups would overlap considerably. As the intersection was approached, drivers that intended to proceed through would exhibit a relatively constant speed profile, whereas the stopping group would exhibit decreasing speed. At some point the two curves would completely diverge, which was thought to indicate a location at which the threat assessment could be made. The authors were hesitant to suggest that this separation was at a sufficient distance to initiate a warning. They did suggest that performing the assessment on acceleration rather than speed could increase the assessment distance.

The ICAS work conducted by the Virginia Tech Transportation Institute (VTTI) under the Intersection Decision Support (IDS) contract performed some preliminary algorithm development (Neale et al., 2005). Three general types of algorithms were constructed based on three types of potential sensing equipment used in the ICAS: single point (i.e., vehicle data were measured at a single point on the approach), multi-point (i.e., vehicle data were measured at several discrete points on the approach), and continuous (i.e., vehicle data were measured at all points on the approach). From a comparison of the schemes it was clear that an algorithm based on continuous detection performed better than the other two options. In addition, the ongoing ICAS projects indicate that a continuous detection scheme is the most likely deployment architecture. Thus, it was necessary to collect continuous intersection approach behavior for the present study rather than measure the vehicle attributes at specified locations.

The IDS algorithm investigation was intended to be exploratory in nature; this was primarily due to the data source. In particular, the data were based on a small sample of drivers that approached a signalized intersection on a test track. The small sample of drivers did not capture the range of possible intersection approaches. Furthermore, the study did not fully account for the variety of factors that can affect the intersection approach (e.g., distractions, motivations, weather). Finally, the study considers the behavior of drivers at stop-controlled intersections. These factors highlight the need for the data collection described and executed during the present research project.

METHODS AND MATERIALS

Six stop-controlled intersection approaches across five intersections in the New River Valley area of southwestern Virginia were selected for data collection. While considering the literature described previously, these sites were selected based on intersection characteristics,

crash statistics, and Virginia Department of Transportation (VDOT) recommendations. Data were collected from the sites for approximately two months, resulting in sixteen months of intersection approach data.

To gather this data, an array of sensing and collection equipment was installed at the test sites. These systems included radar, video cameras, data pre-processing, and data storage. The parametric data and digital video were retrieved through a manual hard drive swap at the test site every five to seven days. The following sections describe the test sites, Data Acquisition System (DAS) installations, data post-processing, data validation, data reduction, and data analysis methods.

Site Selection

In preparation for the present research, potential test sites were identified based on geographic region, DAS compatibility, geometric attributes, design speed, traffic volume, and crash statistics. Detailed information about the intersections was gathered with the assistance of representatives from the Salem District VDOT, the Blacksburg Police Department, and the Christiansburg Police Department. The information provided anecdotal data as well as crash and violation data on intersections in the region. From these data, several intersections were chosen for site visits wherein further information regarding geometry, signage, locality, and intersection type was gathered.

Information collected during the site visits was added to a database and compared to a list of selection criterion. Sites meeting the selection criteria were kept for further consideration. The selection criteria, listed below, were defined based on the available literature, the requirements of the study, collection apparatus, and feasibility of implementation.

- Requirements

- Representative of intersections across the U.S.
- Contains balanced set of posted approach speeds
- Free from obstructions to radar sensor
- Stop sign located on the approach of interest
- Contains a suitable location to mount DAS
- Sufficient shoulders to allow for safe DAS access
- Located reasonably close to VTTI for data retrieval
- Operated by the Southwest Region VDOT for installation assistance and approval
- In range of differential global positioning system (GPS) corrections for calibration/validation

- Specifications

- At least 150m radar sight distance
- Contains speeds of 25 mph, 35 mph, or 45 mph
- A two hour period each day in which at least five satellites are available for calibration/validation

The final selection was based on the probability of obtaining relevant data, including crash rates and average daily traffic (ADT) statistics, and creating a balanced set of approach speeds. Six approaches at five intersections were selected representing the most common speed limits: 25 mph, 35 mph, and 45 mph.

Table 2 below provides a list of all the selected stop-controlled intersections and the corresponding speed limits. The following figures provide images depicting a map detailing measurements of each of the selected stop-controlled intersections, an aerial view, and ground images of each site.

Table 2. Stop-controlled intersections.

Intersection	Posted Speed Limit
Clubhouse & Lusters Gate	25 mph
Plank & Lusters Gate	25 mph
Nellies Cave & Woodland Hills	35 mph
Fairview Church & HW8	35 mph
Meadow Creek & Childress Eastbound	45 mph
Meadow Creek & Childress Westbound	45 mph

The intersection of Clubhouse Road and Lusters Gate Road is a three-way intersection with a single stop sign for vehicles traveling west on Clubhouse (Figure 4, Figure 5, Figure 6). No stop-sign ahead signs were present. A driveway opposite to the clubhouse approach may appear to some drivers as a fourth leg of the intersection. The posted speed limit on Clubhouse is 25 mph and the posted speed limit for traffic traveling on Lusters Gate is 45 mph. The entering ADT for this intersection is 2,084 (VDOT, 2005) and less than one crash typically occurs each year.

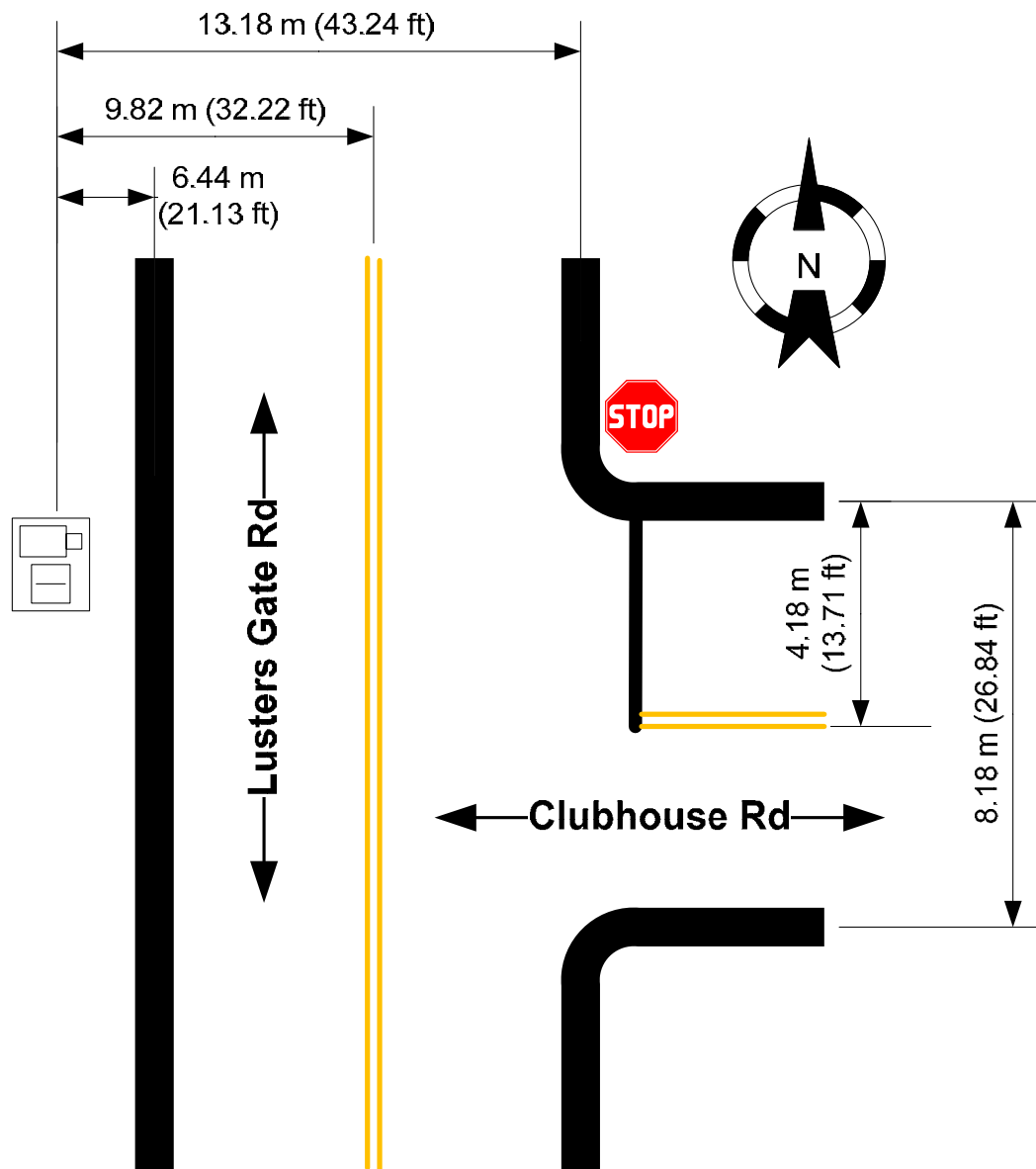


Figure 4. Diagram of Clubhouse & Lusters Gate intersection.



Figure 5. Aerial view of Lusters Gate & Clubhouse intersection.



Figure 6. Ground images of Clubhouse & Lusters Gate intersection; Top left: Southbound on Lusters Gate; Top right: Southbound on Lusters Gate; Bottom: Westbound on Clubhouse.

The intersection of Plank Drive and Lusters Gate Road is a three-way stop-controlled intersection with a single stop sign for vehicles traveling west on Plank (Figure 7, Figure 8, Figure 9). No stop sign ahead signs were present. The speed limit approaching the stop sign is 25 mph and traffic traveling on Lusters Gate has a posted speed limit of 45 mph. VDOT records show the entering ADT at 1,794 and less than one crash occurs in a typical year.

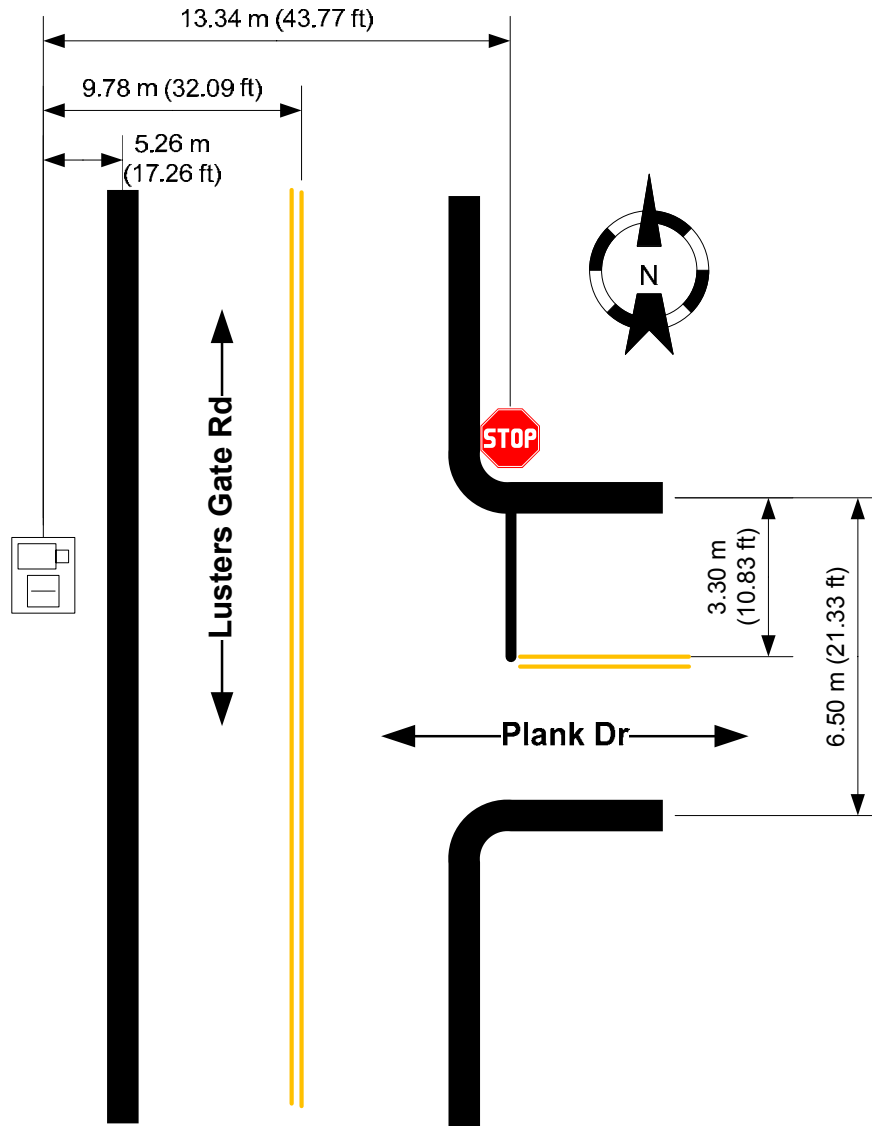


Figure 7. Diagram of Plank & Lusters Gate intersection.



Figure 8. Aerial view of Plank X Lusters Gate intersection.



Figure 9. Ground images of Plank & Lusters Gate intersection; Top left: Southbound on Lusters Gate; Top right: Westbound on Plank; Bottom: Northbound on Lusters Gate.

The intersection at Nellies Cave Road and Woodland Hills is a three-way intersection with a single stop sign for vehicles traveling north on Nellies Cave (Figure 10, Figure 11, Figure 12). There was a stop sign ahead sign on the Nellies Cave approach. Vehicles traveling through the Nellies Cave and Woodland Hills intersection have a posted speed limit of 35 mph. The entering ADT for this intersection is 1,674 and there was an average of one accident per year reported at this intersection.

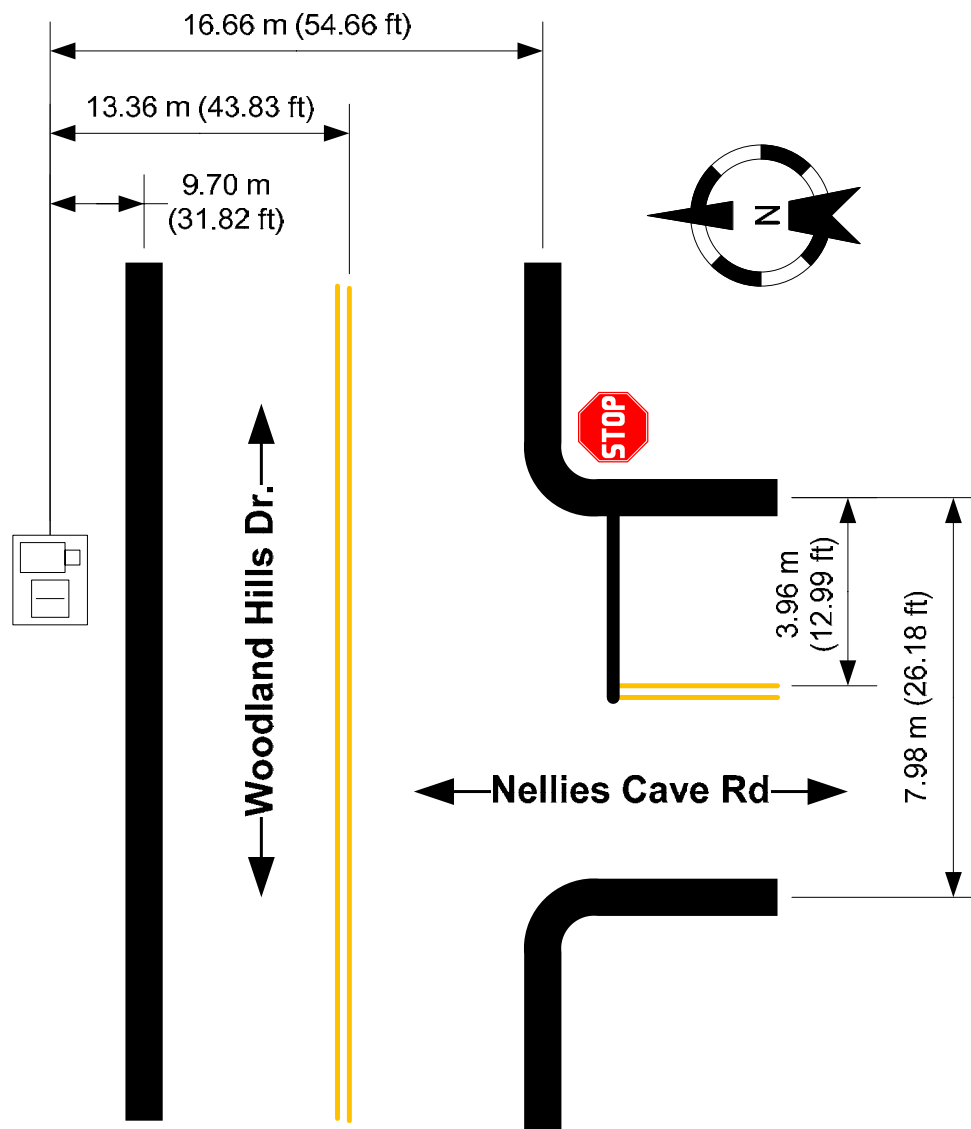


Figure 10. Diagram of Nellies Cave & Woodland Hills intersection.



Figure 11. Aerial view of Nellies Cave & Woodland Hills intersection.



Figure 12. Ground images of Nellies Cave & Woodland Hills intersection; Top left: Northbound on Nellies Cave; Top right: Northbound Nellies Cave; Bottom left: Westbound on Woodland Hills; Bottom right: Eastbound on Woodland Hills.

The intersection of Fairview Church Road and Highway 8 is a four-way intersection with stop signs presented to traffic traveling both directions on Fairview Church Road (Figure 13, Figure 14, Figure 15). There were stop sign ahead signs on both Fairview approaches. The posted speed limit for traffic traveling through this intersection is 35 mph. The entering ADT for this intersection is 8,110 and there is an average of 4 annual crashes at this intersection.

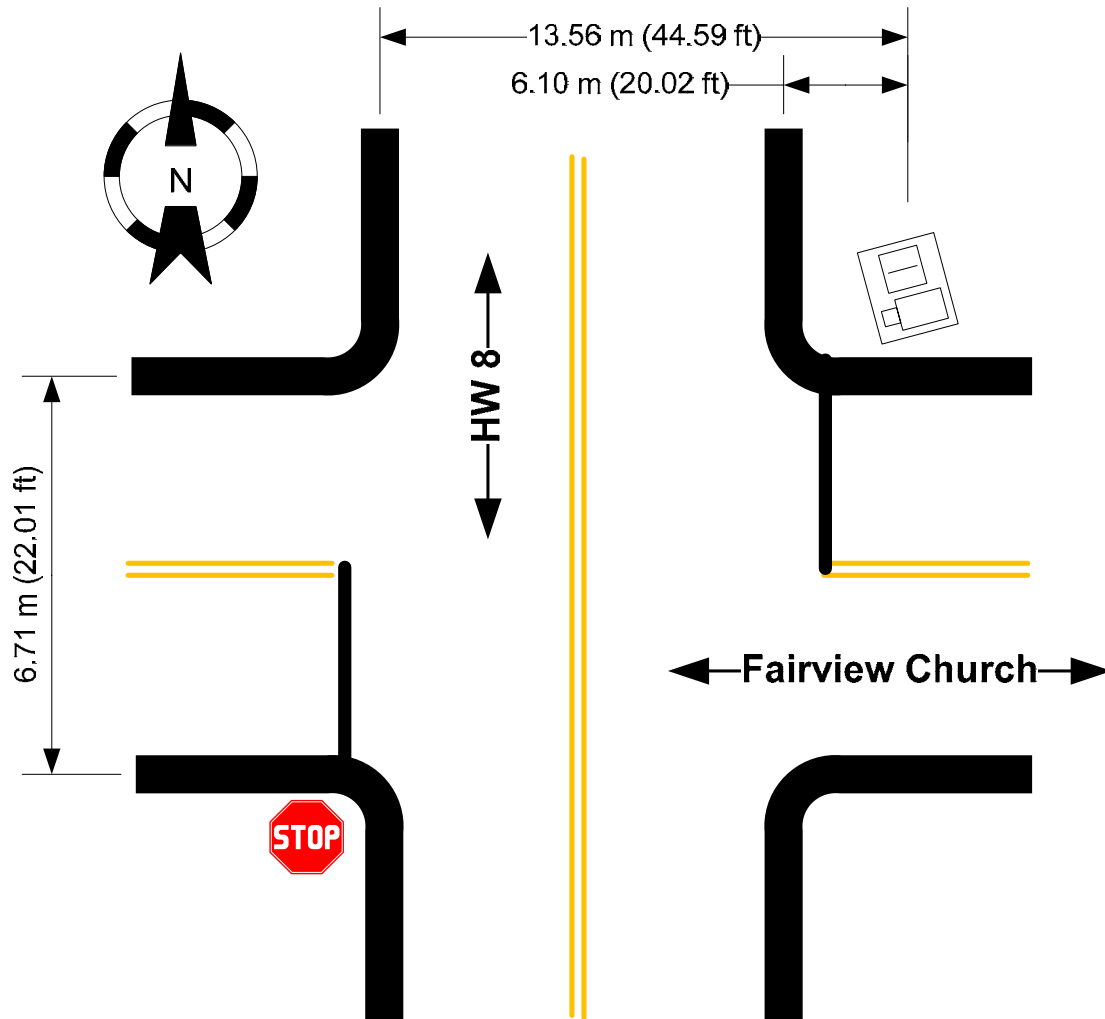


Figure 13. Diagram of Fairview Church & HW8 intersection.

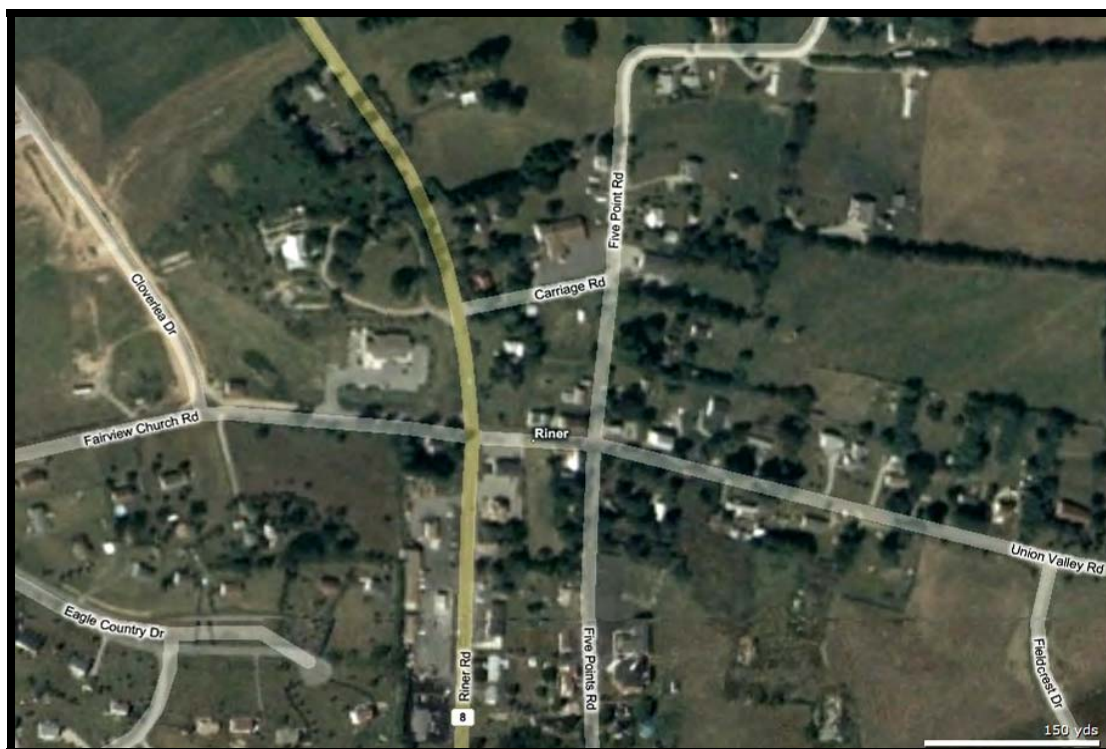


Figure 14. Aerial view of the Fairview Church & HW8 intersection.



Figure 15. Ground images of the Fairview Church & HW8 intersection; Top left: Eastbound on HW8; Top right: Southbound on Fairview Church; Bottom left: Northbound on Fairview Church; Bottom right: Westbound on HW8.

The intersection of Meadow Creek and Childress is a four-way intersection with multiple stop signs presented to traffic traveling in each direction on Meadow Creek (Figure 16, Figure 17, Figure 18). There were stop sign ahead signs on both the Meadow Creek approaches as well as rumble strips. There were also intersection-ahead signs on both directions of Childress. The posted speed limit for traffic on both Meadow Creek and Childress is 45 mph. The entering ADT for this intersection is 3,010 and there is an average of 5 annual crashes at this intersection.

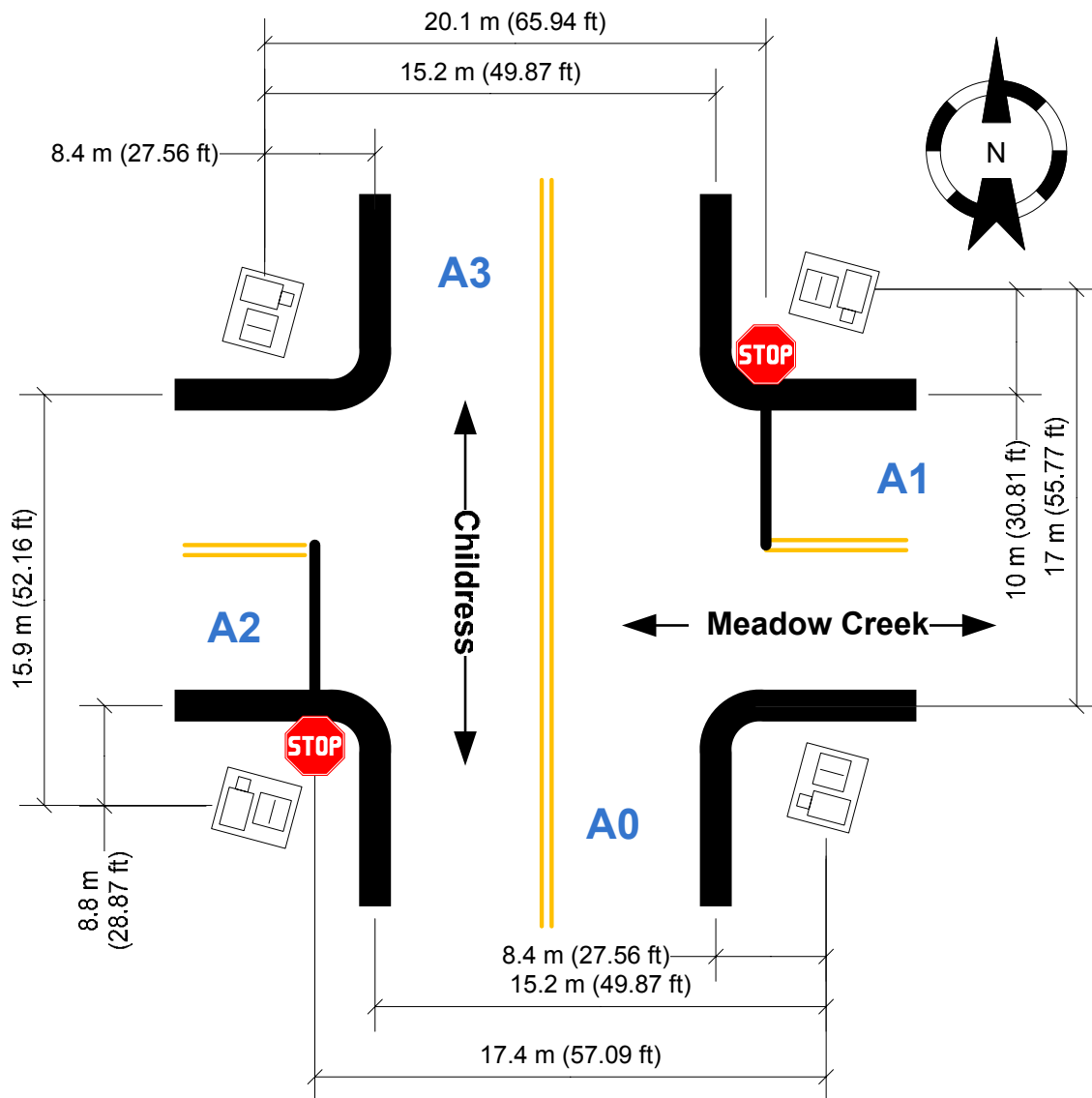


Figure 16. Diagram of Meadow Creek & Childress intersection.



Figure 17. Aerial view of the Meadow Creek & Childress intersection.



Figure 18. Ground images from Meadow Creek & Childress intersection; Top left: Eastbound on Meadow Creek; Top right: Northbound on Childress; Bottom left: Westbound on Meadow Creek; Bottom right: Southbound on Childress.

Data Acquisition System

Data acquisition at the stop-controlled intersections was accomplished through a custom non-obtrusive DAS installed at the selected intersection approaches. The characteristics of the stop-controlled intersections created some unusual requirements for the DAS. In particular, the rural nature of some sites meant the DAS would not have access to power or communication lines. Thus, the DAS was designed to be completely self-contained and self-powered. Furthermore, the relatively short data collection period suggested that tunneling under the roadway should be avoided. To instrument multiple approaches at the Meadow Creek intersection, four independent DAS had to be installed. The data from these four DAS were synchronized post-hoc based on the GPS time. The resulting DAS was designed from the ground up specifically for this collection effort.

The apparatus consisted of a sensing network, a custom digital signal processor (DSP) circuit board, a digital video recorder (DVR), and an enclosure with power source. The sensing network measured raw inputs and provided the measures to the DSP at 20Hz. The DSP pre-processed the inputs and assembled the dataset while archiving digital data files on the DVR. This system was completely contained at the intersection site and virtually invisible to drivers. The DAS time stamped the vehicle data obtained from automotive radar with millisecond time provided by the GPS. The parametric data were accompanied by a MPEG 4 video stream obtained from a charge coupled device (CCD) camera focused in the same orientation as the radar. To avoid tunneling below the roadway, each approach of the intersection was monitored by an independent DAS. The DAS is illustrated in Figure 19 with pictures provided in Figure 20. The subsequent sections contain detailed descriptions of each DAS component.

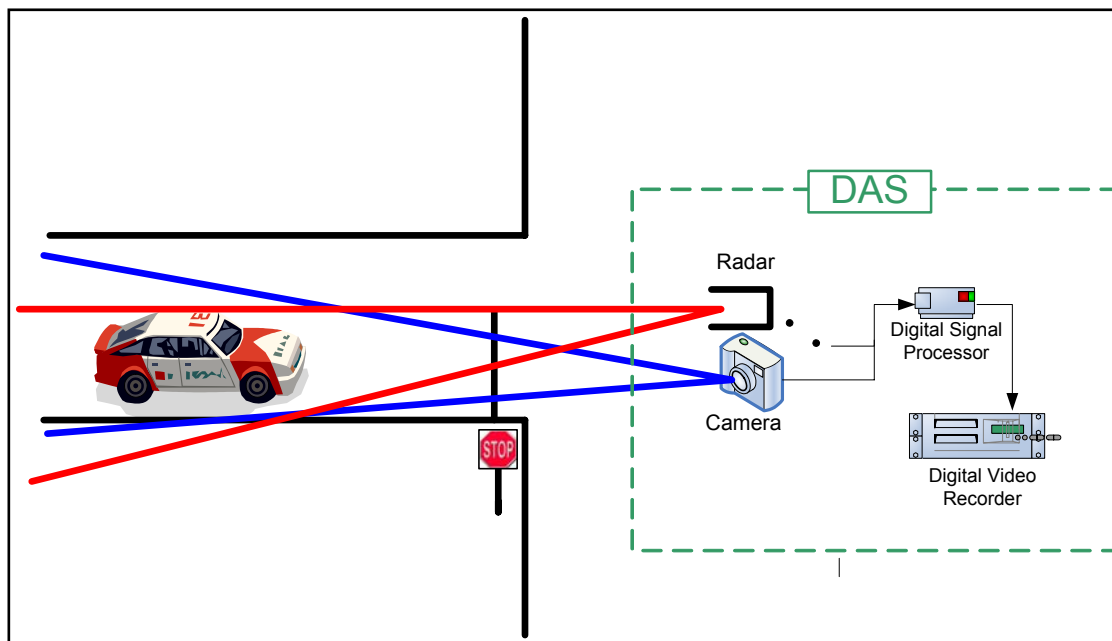


Figure 19. Diagram of stop-controlled data acquisition system including the radar, camera, digital signal processor, and digital video recorder

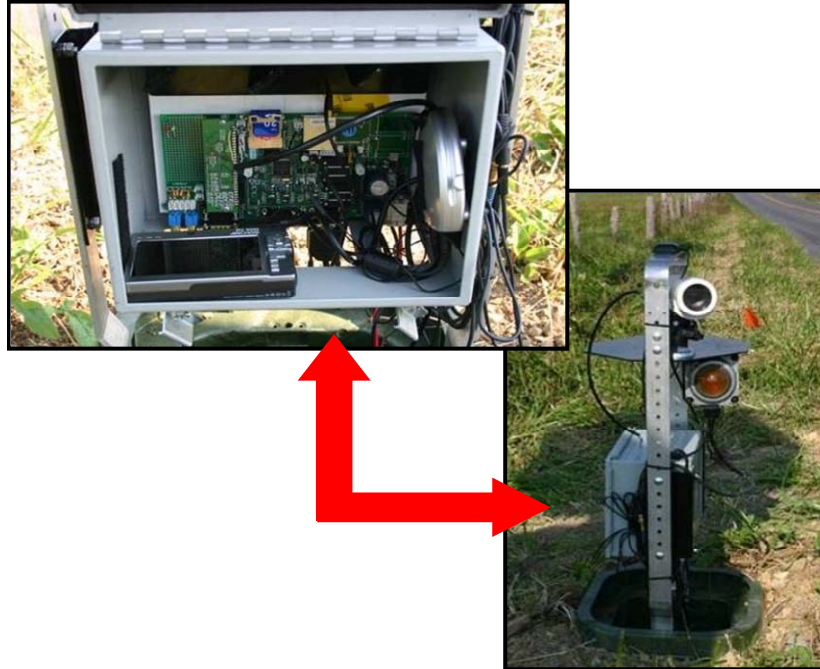


Figure 20. Stop-controlled radar and camera unit.

Sensing Network

The sensing network consisted of two major components: vehicle sensing and uncompressed video. A thorough investigation of current sensing technology identified radar as the most promising technology for obtaining vehicle kinematic information from the roadside (Neale et al., 2005). Previous research indicates that a minimum range of 150m with an accuracy of plus/minus 3 m and a range-rate accuracy of plus/minus 1 km/h is sufficient for ICAS operation (Neale et al., 2005). The radar must also be relatively inexpensive, weatherproof, and have Federal Communications Commission (FCC) approval. The AC20 Autocruise radar from TRW is adaptive cruise control (ACC) radar that exceeded the requirements of this study. The AC20 has the following specifications:

- Dimensions: 3.74 in x 3.74 in x 2.48 in
- Waveform: 76 GHz
- Interface: CAN
- Distance Measurement
 - Range: 3.28 ft – 656.17 ft
 - Accuracy: $\pm 5\%$ or 3.28 ft
- Speed Measurement
 - Range: ± 250 kph
 - Accuracy ± 0.1 kph
- Lateral Position Measurement
 - Range $\pm 6^\circ$
 - Accuracy $\pm 0.3^\circ$

The second raw data input provided by the sensing network are uncompressed video. This video was later used to derive measures via manual reduction techniques. The video was recorded using a robust all-weather NTSC video camera mounted in an inconspicuous location, providing a view of the desired approach. The camera selected was a SuperCircuits PC219ZWPH with the following specifications:

Horizontal Resolution: 480 Lines

- Illumination: 2.5 Lux/ F1.8
- Image Sensor: 1/3 in CCD Sensor Interline
- Power Requirements: 320 mA at 12V DC
- Video Format: NTSC
- Pixels: 492 (V) X 771 (H)
- Video Connection: BNC Female S/N Ratio 48 dB
- Lens Type: .20-1.97 in
- Zoom Lens Control: Auto Iris DC Driven
- Backlight: Built-in Backlight Compensation
- Weight: 27.87 oz (790 grams)
- Dimensions: 7.64 in X 3.50 in X 7.52 in

The NTSC signal provided by the camera was attached to the DVR for compression as choreographed by the DSP. The compressed video along with the other raw data from the sensing network were transmitted to the data DSP for pre-processing.

Digital Signal Processor Circuit

The DSP, housed on a proprietary circuit board with hardware and firmware, was designed specifically for this study. The DSP detects inputs from the sensing network, and pre-processes, aligns in time, integrates, and transfers them to the DVR and solid state memory for storage.

In addition to the data collection tasks, the DSP board housed the power management system and sampling scheme. The power management system controlled the on/off state of the sensing network and DVR. Battery voltage was monitored and if it dropped below a specified threshold, the entire DAS would systematically shut down to prevent data loss. Furthermore, to maximize battery life, the DSP would switch the sensing components and the DVR off when vehicles were not present at the approach.

Data from the sensing network was sent to two separate locations by the DSP. Parametric data were processed by the DSP, time stamped, and sent to a 2 GB solid state memory card. Video data were time stamped and sent to the 100 GB DVR for compression and storage.

Digital Video Recorder

An Archos AV 500 was selected for compressing and storing the video collected by the DAS. This highly portable DVR uses an extremely efficient MPEG4 compression algorithm for reducing video file size. The hardware based compression system is configurable such that the balance between file size and quality can be manipulated. For the purpose of this study, a high compression was selected to minimize file size and reduce the possibility of attaining personal information (e.g., license plate number – discussed further in the Institutional Review Board [IRB] section).

The compressed video was sent a 100 GB hard drive housed within the DVR. The storage space needed to be sufficient for collecting over a week of uninterrupted video. This hard drive was retrieved as needed and transported to VTTI for storage. The AV 500 has the following specifications:

- Capacity: 100 GB Hard drive
- Display: 4 in LCD 480x272 pixels, 262 000 colors
- Video recording: MPEG-4 SP up to 640x480 @ 30 f/s, in AVI format.
- Video playback: MPEG-4 SP up to 720x480 @ 30 f/s
- AV connections: Audio & Video line out. IR emitter
- Interfaces: USB 2.0 high-speed device
- Power source: Rechargeable Lithium-Ion Battery
- Battery life: Up to 15 hours
- Dimensions: 2.99 in x 4.88 in x .94 in
- Weight: 315 g - 11.11 oz

Enclosure and Power Source

To minimize behavioral adaptation by the driver, it was essential for the DAS to be as unobtrusive as possible. The sensing network was mounted inside a standard telecommunications box frequently seen on roadsides. These boxes were buried approximately 1/2 m underground with roughly 1 m protruding above the surface. The above-ground portion included a lid that was field-removable and was secured with a security screw. The telecommunications box was constructed from thin uniform plastic which was easily penetrated by the radar without any cutting, making it completely invisible to drivers. A small hole was drilled for the camera which provided a clear image of the intersection. Finally, a small box was mounted inside the enclosure that contained the DSP board and the DVR. The telecommunications box located the sensing equipment at the recommended heights while protecting them from direct moisture and ultraviolet rays (See Figure 21).

In addition to the above-ground sensing enclosure, a second enclosure resided underground near the telecommunications box. The second enclosure was a Pelican® waterproof high-impact case with security lock. The enclosure provided storage for the power source. The enclosure was buried in order to be inconspicuous and to provide temperature stability for the power source. The enclosure had the following specifications:

- Temperature Rating: Minimum -23° C, maximum +99° C

- Inside dimensions: 21.73 in x 16.81 in x 7.87 in
- Outside dimensions: 24.25 in x 19.45 in x 8.66 in
- Water proof
- Weight: 6 kg

The DAS was designed to operate at low voltage for at least six days, at which time the batteries (two large capacity 12v gel-cell built by MK part number VRLA-Gel 8G31) would need to be exchanged for a freshly charged cell. At the time of the battery exchange, the DVR was also swapped for empty units. Data from the test site were transported back to VTTI and uploaded to a secure fiber channel server for long term storage. The battery has the following specifications:

- Cranking amps: 550 amps@ - 17° C, 780 amps @ 0° C
- Discharge time:
 - 100 hrs at 1.08 amp to 1.75VPC @ 27° C
 - 48 hrs at 2.15 amp to 1.75VPC @ 27° C
- Weight: 32.2 Kg
- Dimensions: 12.95 in X 6.73 in X 9.37 in



Figure 21. Stop-controlled DAS.

Test Site Installation

Test sites were outfitted with the DAS sequentially, with each site taking approximately one day to install, calibrate, and validate. VDOT assisted in the installation by providing the necessary signage and equipment support. The enclosures were mounted first, followed by the installation of the DAS hardware and cabling. The battery pack, used to provide power at each site, was buried in order to remain as unobtrusive to traffic as possible.

A calibration procedure was initiated once the hardware was installed and powered up. During calibration, the camera focal length and zoom was set to obtain the desired image. This image captured the entire vehicle approach into the intersection. The radar was aimed to capture the vehicle from at least 150 m continuously through the stopbar. Calibration values, such as the distance from the radar to the stopbar, were also set to correct sensor mounting configurations.

Data Retrieval and Management

As with most stop-controlled intersections, traffic volume at the test sites would decrease substantially during non-peak hours. Collecting data when vehicles were not present on the intersection approach would be inefficient. The power supply would drain faster and the storage devices would reach capacity sooner. Thus, a triggering scheme was used to determine when to collect data. In particular, data were only written to the storage devices when a vehicle was present on the intersection approach. When no vehicles were sensed, the DAS entered a low-power mode in which the camera and assorted other components were powered off or put into a standby mode.

The frequency with which data were collected from the stop-controlled intersections was dependent on the traffic volume at a given site. Sites with higher traffic volumes placed higher demands on the DAS, consuming battery life at a higher rate. On average, data retrieval occurred every five (Meadow Creek intersection) to seven (Nellies Cave intersection) days.

Trained data retrievers maintained each of the sites in order to eliminate down time in which data were not collected. The system was temporarily shut down for approximately two minutes in order to replace the DVR, memory card, and batteries.

Following data retrieval, the files were transported to VTTI and uploaded onto the VTTI database. Each of the files was named based on the intersection, day, and time at which they were retrieved. This process was completed automatically using a custom software program.

Data Overview and Post-Processing

To overcome limitations present in the raw dataset, an extensive post-processing effort was undertaken. Post-processing applied a series of data cleansing, extrapolation, and smoothing techniques to prepare the dataset for analysis. The following sections first describe the raw dataset that resulted from the data collection effort. Next, the validation procedure and subsequent analysis to validate the DAS and post-processing method are presented. Finally, the post-processed data are described and summary statistics for the final dataset are provided. This final dataset was used for the remaining analyses described in the subsequent sections.

Raw Dataset

Data were continuously collected at six stop-controlled intersections in Montgomery County, Virginia. The data were natively recorded by the DAS in two file formats. The first file type stored the parametric data in a compact binary format. In conjunction with the binary files an MPEG4 video file was written. Each of these files was written for every hour of clock time

(e.g., 48 files per day). The total number and size of files collected for each intersection are provided in Table 3.

Table 3. Raw data collection file information.

INTERSECTION	NUMBER OF BINARY FILES	SIZE (binary)	NUMBER OF VIDEO FILES	SIZE (video)
Meadow Creek & Childress (2 approaches equipped)	3,270	2.59 GB	2,732	574GB
Clubhouse & Lusters Gate (1 approach equipped)	1,501	568 MB	1,205	169 GB
Fairview Church & HW8 (1 approach equipped)	1,331	622 MB	1,331	119 GB
Nellies Cave & Woodland Hills (1 approach equipped)	1,192	483 MB	1,027	142 GB
Plank & Lusters Gate (1 approach equipped)	1,659	384 MB	1,147	94.8 GB
TOTAL	8,953	2.06 GB	6,416	1.10 TB

The raw dataset was converted from binary to SQL database format to allow rapid testing of algorithms. The raw SQL dataset is essentially a copy of the binary files stored in a single clustered index table with null values removed. This resulted in a table containing 135,063,981 rows and 38 columns.

The raw data were limited in utility. The best radar sensor available at the time of the stop-controlled data collection was the AC20. While the AC20 likely works well in its intended application, performance was disappointing when it was placed at the intersection. Although the radar did return accurate measurements of range and velocity, it did not reliably return these measures for a given vehicle. It was common for the sensor to provide sparse data such that a measurement was returned at rates well under the data collection rate. Data dropouts were also common in which a vehicle was only returned for a portion of the overall approach.

In addition, the radar sensor did not provide a unique vehicle ID. The track ID variable used by the sensor would repeat as soon as the vehicle with a given ID was not visible. For example, three vehicles approaching the intersection may be assigned IDs 1 through 3. If a fourth vehicle were to enter the field of view of the sensor just after vehicle two exited the view, the new vehicle would be designated as vehicle two as well. For the analysis it was imperative that each vehicle be indexed individually so the algorithm could be executed on each vehicle trajectory. Data provided by each vehicle needed to be continuous in order for the algorithm to be functional on each data frame.

Post-Processing

Post-processing began once the raw binary files were uploaded to the server at VTTI. The first step was to populate a structured query language (SQL) database from the binary data files. This was accomplished by using a file conversion tool programmed specifically for this purpose. The resulting database stores the raw data and does not perform any computations.

The erroneous radar returns (e.g., image shadows, trees, deer, etc.) were then cleaned from the database. This process was performed primarily through a filter that passed data depending on a combination of values for several metrics calculated by the sensor. These metrics are measures of signal quality and magnitude. Queries were then constructed to filter out the erroneous data.

Once the first-pass of filtering was complete, a more sophisticated set of Matlab programs was developed and executed. This set of programs performed three primary functions. First a program was developed to assign a unique identity to each vehicle that approached the radar. The radar reported the location of 4 vehicles at a time but could track up to 24 vehicles simultaneously (i.e., six frames would be needed to report all 24 vehicles). The radar would assign an ID to each vehicle on a temporary basis. Each time a vehicle was out of view, the radar would recycle that ID. This method used by the radar made the derivation of a unique ID difficult. In addition, IDs were frequently assigned to the inappropriate vehicle (e.g., track switching).

To develop the unique ID and fix the track switching problem, a post-hoc tracking algorithm was developed to enhance the radar tracking. This was accomplished by developing software that crawled through the data frame by frame. For each frame, the associated radar returns were compared to the returns in the previous points in time at which the subject vehicle was believed to exist in the data. Thus, even if a vehicle dropped out for some time, the program would hold information about that vehicle's prior dynamic state while waiting for it to reappear.

The time-based comparison was made by the program for up to 16 simultaneous vehicles. For a point to be assigned the current vehicle ID, the tracking algorithm required a series of criteria to be met. The criteria were based on a dynamic prediction of the vehicle's location at the current frame based on the dynamic state of the vehicle during its previous incarnation in the data. This prediction was made using the fundamental kinematic equations of motion. In addition, the time between returns, overall distance traveled for the vehicle, and overall number of points returned was also part of the comparison. For the vehicle ID, the assigned track could not last more than 15 s, had to contain at least two data points, have reported ranges greater than zero, and travel at least 1 m. The resulting vehicle ID was a database-wide unique ID associated with each vehicle that approached the intersection. Thus, all of the data points returned from a single vehicle were assigned a single ID.

In conjunction with the vehicle ID, a "fit ID" was also created. The fit ID was essentially a more stringent version of vehicle ID. Where vehicle ID contained all the data points from one vehicle, the fit ID contained only the series of points that include sufficient information for repairing broken tracks and deriving the acceleration measure. For a set of points to be assigned a fit ID, they were required to contain at least 10 data points, cover at least 2 m, and drop out for no more than four seconds. It is possible for a single vehicle ID to have multiple associated fit IDs. For instance, consider a vehicle track that was returned by the radar at a long range, then disappeared for some time, and then reappeared later. If that dropout was less than 4 s, a single fit ID would have been created; however, if that dropout lasted more than 4 s, a second fit ID would have been assigned to the second collection of points within that vehicle ID. The resulting fit ID grouped sets of points together in which it was feasible to perform additional post-processing to improve the data.

The dataset contained dropouts in which the radar would stop providing updates for a period of time. In addition, the radar sensor used in this project performed some real-time smoothing using a Kalman filter; however, based on the data, it was apparent that additional smoothing would reduce data dither. This dither is inherent in all radar systems as a result of the changing scattering center of the returned reflections over time. Thus, a non-parametric smoothing spline was fit to each collection of data points within a single fit ID. The smoothing spline was a knotted piecewise polynomial that responds very quickly to changes in the underlying form of the data. No latency was introduced through the use of this fitting technique.

The polynomial produced by the smoothing spline was used to repair vehicle ID tracks with short dropouts. Thus, the fit ID provided the means to reproduce a complete vehicle approach despite missing data points. In addition, the derivative of the polynomial function created by the spline was used to calculate the vehicle acceleration. Thus, derived acceleration is available within each fit ID. In general, the fit ID will be used for all analysis as it contains data of the fidelity required for assessment.

In addition to acceleration, four other continuous measures were also calculated. These included the time to intersection (TTI), required deceleration parameter (RDP), traffic volume, and brake status. TTI is a calculated value based on the current distance to the intersection and speed of the vehicle. TTI is believed to be a good approximation of the measure used by the human visual system to judge whether or not to stop (Horst, 1990). It combines the effects of speed and distance into a single measure (Equation 1).

Equation 1

$$TTI = \frac{D_s}{V}$$

Where:

V = Vehicle speed at the point when driver initiated braking (ft/s)

D_s = Distance to the trailing edge of the stopbar or start of perpendicular roadway when a stopbar was not present (ft)

RDP represents the calculated average acceleration required to stop at the stopbar based on the vehicle's current speed and distance from the intersection. RDP is an easily interpreted variable representing the required braking effort to stop at any point during the intersection approach. This equation is provided below.

Equation 2

$$RDP = \frac{V^2}{2 \cdot D_s \cdot g}$$

Where:

V = Vehicle speed at the point when driver initiated braking (ft/s)

D_s = Distance to stopbar when driver initiated braking (ft)

g = gravitational acceleration constant (32.2 ft/s²)

Brake status was approximated based on the acceleration data. Brake status is a binary indication of whether or not the driver is pressing the brake. While this variable cannot be directly measured with the radar, it can be accurately approximated. To approximate brake status an algorithm was devised to monitor the acceleration rate throughout each vehicle's intersection approach. The algorithm first used a ten point zero-phase-shift moving-average-filter to smooth the acceleration. The smoothed acceleration was monitored for a change point that dropped below -0.075 g. Searching for this change point allowed vehicles to coast, slowing through engine braking, without actively pressing the brake. This threshold was selected through investigation of VTTI databases as discussed below and in Appendix A.

First, the data used for the validation procedure described previously were used to determine the threshold at which the radar would reliably show braking when compared to the actual brake status information provided by the vehicle data. This study indicated that a -0.05 g rate was a reliable indicator of brake activation. Next, stop-controlled approaches in the 100-car database were analyzed to determine the minimum acceleration rate present before drivers initiated braking (i.e., beyond slowing provided by engine braking). This threshold was found to be approximately -0.06 g. To ensure that drivers were indeed actively braking, the threshold was set at -0.075 g for identifying the initial brake activation. Data from the 100-car analysis were then used to correct the brake point by the average reaction time such that the brake status flag was activated when the driver was predicted to have actually pressed the brake. After initial brake activation, the brake flag remained on until the smoothed acceleration variable exceeded -0.05 g. After initial braking, -0.05 g was used as the threshold for determining ongoing brake status throughout the remainder of the approach. Additional details of the 100-car analysis to determine the brake activation status are provided in Appendix A. Each vehicle tracking variable and its associated operational definition are provided in Table 4.

Table 4. Variables and operational definitions.

Variable	Operational Definition
Vehicle_ID	Groups a series of radar measurements into a single vehicle observation. This ID is unique across the entire database
Frame	Incrementing counter for the current frame. Started at one for each file and incremented at 20 Hz
File_ID	DAS assigned unique file ID. One file per hour of clock time was written to the hard drive
Fit_ID	Groups a series of radar measurements into a single observation. This group of data contained a series of high quality densely populated vehicle measures. Only point series contained in a Fit_ID were used for deriving measures such as acceleration. This ID is unique across the entire database
Range	Range as measured by the radar but cleaned and smoothed during post processing. Also, the range was offset such that stopbar or adjacent traffic lane is used as the origin (ie. Range=0 is the stopbar)
Velocity	Velocity as measured by the radar but cleaned and smoothed during post processing
Accel	Acceleration as measured by the radar but cleaned and smoothed during post processing.
RDP	The kinematically calculated average deceleration required to stop at the stopbar given the present vehicle speed and distance to the intersection
Brake_Status	A on/off flag that approximates when the driver is actively braking the vehicle
TTI	The time that will elapse until the vehicle reaches the stopbar based on speed and intersection distance assuming there is no change in speed
Angle	Angle as measured by the radar but cleaned and smoothed during post processing
GPS_Time	Common time base that permits synchronizing all four intersection approaches.

System and Post-Process Validation

A system accuracy validation was performed to ensure accuracy of the stop-controlled DAS and post-processing methods. A small experiment was devised and executed on the Smart Road test track in Blacksburg, Virginia. The DAS described above was installed at the Smart Road intersection and set to collect data in the same manner as the live intersection sites. Next, a series of intersection passes were performed using an instrumented vehicle with an independent DAS.

To evaluate the DAS accuracy over a range of realistic scenarios, a total of 36 runs were performed. These runs included three replicates at 25 mph, 35 mph, and 45 mph stop approaches with soft (~0.2 g), medium (~0.4 g), and hard (~0.6 g) average braking rates. In addition, one violation at each speed was also performed.

The vehicle used for this experiment was a highly-capable VTTI test vehicle. The instrumentation included a high-accuracy differential GPS system that has been validated in past studies and measures position to within a few centimeters and velocity within one-fourth of a meter per second. The data collected by the roadside and the in-vehicle DAS were compared post-hoc. Data alignment was performed using the millisecond GPS time recorded in each of the two systems. Each of the reported measures were compared using the vehicle DAS as the ‘true’ measure. Any deviation of the infrastructure DAS from the vehicle DAS was considered error. These errors were evaluated to develop an overall estimate of system accuracy.

Measurements from both the infrastructure and vehicle DAS were recorded and subsequently compared post-hoc (Figure 22). The figure below depicts an exemplar vehicle approach trajectory as measured by the vehicle and infrastructure DAS. Each subplot contains a line series for the post-processed radar data, a scatter plot of the raw radar data, and a line series for the vehicle data, as appropriate. Note that an approach with sparse raw radar data were intentionally selected to demonstrate the effectiveness of the post-processing method (only 13 points were returned by the radar during this approach). The top subplot depicts the range as reported from the three data sources. On this particular approach, the vehicle was initially detected by the radar at 178 m and was tracked up to the stopbar. The second subplot shows the vehicle's velocity during the approach initially traveling at 11.5 m/s (3.36 mph). The third plot depicts the acceleration as measured by an accelerometer inside the vehicle and the acceleration that was computed as the first derivative of the post-processed radar velocity. Finally, the bottom plot depicts the difference between the radar-derived acceleration and directly measured acceleration. For this example the maximum error in the computed acceleration occurred near the stopbar and resulted in a 0.03 g difference.

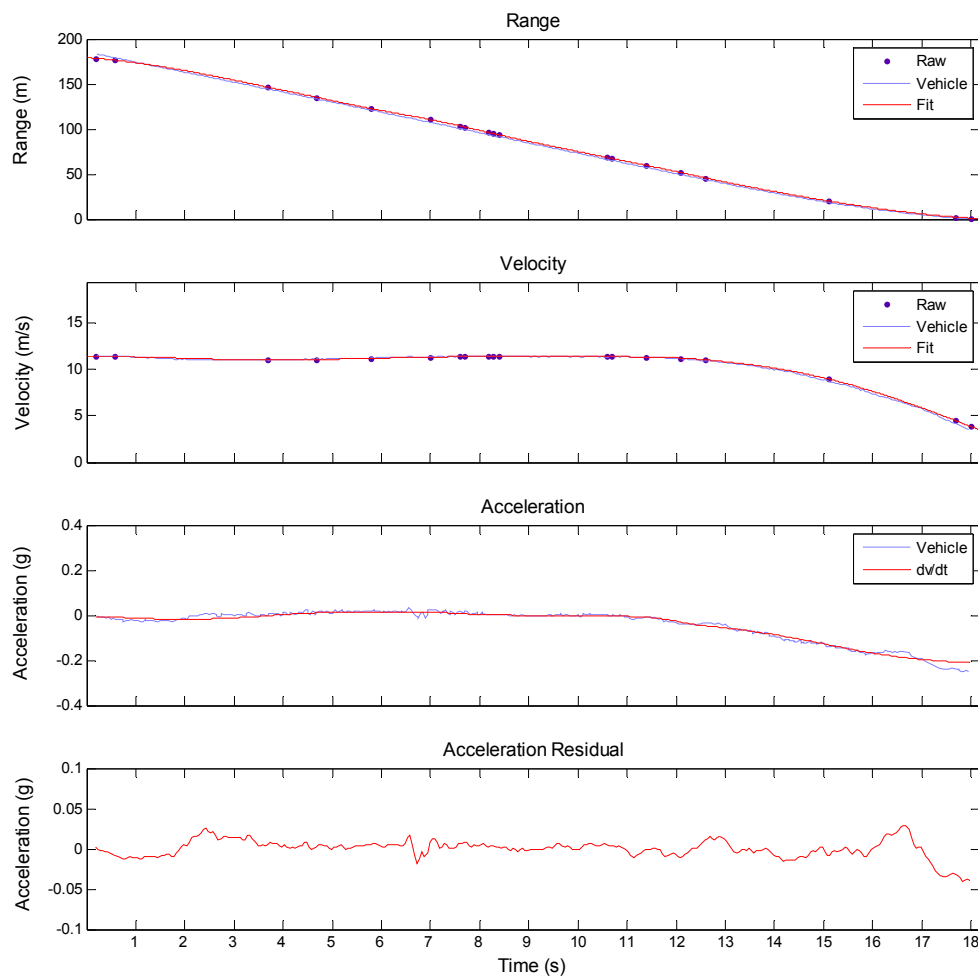


Figure 22. Infrastructure and vehicle DAS measurements.

Based on the results (Table 5), the final dataset was deemed sufficient for the purposes of this analysis. The overall accuracy levels demonstrated through the validation met or exceed the accuracy guidelines presented in the ICAV Task 5 report (Lee et al., 2005). No speed- or approach-type bias was observed in the error data (Appendix C). Thus, the radar error should be consistent across a variety of intersection approach types. There are some occasional large errors produced by the radar; however, these events are rare and will be mitigated through the statistical techniques used in the subsequent analysis. Given the moderate quality of the initial data, the relatively high accuracy of the final dataset demonstrates a successful post-processing method. However, if feasible, future DAS should make use of alternative radar that provides a more reliable data source that will not require such an involved post-processing.

Table 5. Infrastructure DAS system error relative to the vehicle high-accuracy DAS.

	Range (m)	Range (ft)	Speed (m/s)	Speed (mph)	Acceleration (g)
Mean	0.828549	2.71834	0.062148	0.139021	-0.00252
Standard Deviation	2.40099	7.87726	0.151491	0.338876	0.026217

Post-Processed Dataset

During post-processing, the data structure was re-organized for ease of analysis and efficiency. This resulted in a database one-third the size of the raw dataset despite containing significantly more useable data. The final database that was used for all analysis contained 40,996,295 rows. In this database, there were a total of 311,753 vehicles as indicated by the number of Unique Vehicle IDs generated. The breakdown of traffic at each site is provided in Table 6 below. In addition to the vehicle ID, a fit ID (as discussed in the methods section) is also provided in the table. Recall from the methods section that the fit ID contains the level of fidelity needed for assessing frame by frame vehicle trajectories.

Table 6. Unique vehicle IDs generated for each site.

INTERSECTION APPROACH	UNIQUE VEHICLE IDs	UNIQUE FIT IDs
Meadow Creek & Childress (westbound)	73,328	14,714
Meadow Creek & Childress (eastbound)	79,510	7,759
Plank & Lusters Gate	14,152	877
Clubhouse & Lusters Gate	25,348	1,592
Nellies Cave & Woodland Hills	22,240	5,073
Fairview Church & HW8	10,141	610
TOTAL	224,719	30,625

Event Validation and Video Reduction Process

The data collected for this study were obtained primarily from an infrastructure-mounted radar sensor. Radar has some additional limitations relative to in-vehicle sensors. While the measurements of speed and range are accurate, it is the association of those measures with a particular vehicle that is prone to error. This means that a vehicle reported by the radar is not necessarily a valid vehicle. The quintessential example of this behavior occurs with large

vehicles and trailers. The radar used for this study would frequently treat a large vehicle as two separate targets, particularly if a trailer was in tow. As a result the subject vehicle may have completed a stop, however, the secondary false-target located on the rear of the vehicle or trailer would appear to violate the TCD as it was pulled through the intersection at-speed.

The primary purpose of the video reduction was to validate the events of interest. It was important to ensure that false targets were not inadvertently being included in the sample of violating and aggressive driver approaches. In addition to false triggers, other invalid events also needed to be removed from the dataset. These included violations that were a result of atypical scenarios such as a funeral procession and the crossing of in-service emergency vehicles. All of the invalid events were marked and were removed from the analysis.

The dataset contains over 300,000 unique vehicle trajectories. A typical validation average is five minutes per event, suggesting that well over 600 person-hours would be required to reduce all events. While reducing all of the events would provide an excellent dataset for investigation, it was simply not possible to perform within the time and budget constraints of this study.

Therefore, a strategy was devised to methodically select approaches of interest for an ICAS investigation. The selection process occurred through the development of a trigger that swept through the parametric data and flagged events requiring attention. The flagged events were automatically collected for easy retrieval and accessible to the data reductionists. A detailed discussion of the method used to identify validation events is available in Appendix B.

The validation process required a reductionist to view each event. Once the event was opened and viewed, it was logical to collect a few additional measures that were not available natively in the data collection system. These measures included environmental factors such as weather, lead and following vehicle presence, violation type, vehicle type, and maneuver. However, while these measures were recorded, they were not analyzed as part of this research effort.

RESULTS AND DISCUSSION

One of the main objectives of this project was to demonstrate that a data collection system and data management scheme could be constructed to collect the data required for the development of an ICAS threat assessment algorithm. The second objective was to perform an exploratory analysis to demonstrate the efficacy of an ICAS threat assessment. The following results and discussion provide evidence to support the feasibility of an ICAS system. An explorative analysis of the intersection approaches is performed; it includes stopbar behavior, stopbar speed, brake onset, and overall vehicle trajectories. These analyses are performed with the goal of working toward the development of an ICAS algorithm.

Driver Behavior Analyses

An ICAS algorithm must be capable of predicting the driver's stopping behavior at a distance that provides sufficient time for the driver to stop prior to entering conflicting traffic

paths. The analysis reported in this section has two purposes. The first is to investigate the different types of stopping maneuvers that are performed by drivers as they approach stop-controlled intersections. The second purpose is to investigate driver behaviors that are likely to be assessed by an ICAS algorithm.

Stopping Behavior

In theory, drivers are expected to perform a complete stop each time they approach a stop-controlled intersection regardless of the environmental state (e.g., traffic, sight distance, etc.). In practice, however, drivers frequently cross stop-controlled intersections without placing their vehicle in a stationary condition. For this report, these slow-moving violations are referred to as ‘rolling violations’.

Consider a driver approaching a stop-controlled intersection at the suggested speed limit. This driver may slow his/her vehicle at a sufficient rate to stop at the stopbar, if required. However, if visibility of the adjacent traffic lanes is good near the intersection box, the driver may cease braking and perform a rolling violation. While a rolling violation is technically illegal, if the driver is cognizant and prepared to stop it is unlikely that this behavior contributes to the target population (i.e., behaviors leading to crashes).

From the stop-controlled data collected it is clear that rolling violation behavior is very common. If a warning is provided to all drivers performing a rolling violation, a nuisance alarm problem is likely to result. Thus, a binary discrimination based on traffic laws is insufficient for determining the algorithm effectiveness. Rather, a stopping behavior classification scheme is required to discriminate the infrequent unsafe behaviors from the numerous safe behaviors. The purpose of this section is to determine the classification system based on driver behavior at the collection sites. This classification system will later be used to determine the algorithm’s performance.

Objective clustering of driver stopping behavior

Although stopping behavior was classified subjectively as part of the data reduction, there are reasons to develop a corresponding objective partitioning scheme. In particular, manual reduction was only performed on a subset of the data and therefore a majority of the intersection approaches were not classified in the reduction. During algorithm development it is desirable to use a large dataset. An objective measure of stopping behavior will provide a performance metric that is applicable to intersection approaches that were not considered during the reduction. In addition, objective measures are generally preferred over subjective measures due to higher repeatability and accuracy.

A cluster analysis was performed to objectively explore and categorize stopping behavior. Cluster analysis is a statistical method designed to partition data into subsets such that data within each subset shares some common trait. In the present context the clusters represent different driver behavior groups.

There are several potential measures by which the driver’s stopping behavior may be classified. For instance, the speed at the stopbar may be indicative of the driver’s stopping

behavior (i.e., a violator would have a higher speed than a compliant driver). Other possible measures include minimum TTI or average deceleration.

The variable selected for classifying driver stopping behaviors was the maximum RDP. This measure was selected because the value has a direct relationship to the ability of the vehicle to stop and is therefore easily interpreted. For instance, if the maximum RDP is 0.3g, the driver could have easily stopped by the stopbar. On the other hand, a driver with a maximum RDP of 4g committed to the violation and would not have been able to stop by the stopbar. The maximum RDP was only calculated for ranges greater than 2 m from the stopbar.

A 2 m cutoff was selected because of the behavior of RDP near the stopbar. As a vehicle draws near to the stopbar, RDP will tend toward infinity regardless of travel speed. By only considering RDP values that occur at a range larger than 2 m, this portion of the RDP trajectory is not considered, making the measure sensitive to differences in approach behavior.

The non-parametric kmeans (MacQueen, 1967) method of clustering was selected for this analysis. Kmeans is a partitional unsupervised learning algorithm that minimizes a measure of distance between each data point and the center of the corresponding cluster. The optimization occurs in an iterative process that moves the cluster center and recalculates the distances until a minimum is obtained. Kmeans was selected because it is a common and well understood clustering method (Davidson, 2002) that is computationally efficient (Matlab, 2007) and sensitive to rare observations (Hauskrecht, 2003) making it a logical choice for this large dataset.

The cluster analysis was performed on a subset of the data that contained a sample of complete vehicle approaches. As discussed in the post-processing section, it was not unusual for a new vehicle track to be generated when vehicles deployed from a standing queue. Thus, only vehicle approaches in which the vehicle was present from 100 m to 2 m were included in the analysis. The 100 m cutoff also ensures that the vehicle was present for the entire warning region which will be important when the clusters generated during this analysis are used for algorithm development. This resulted in a sample size of 30,623 observations.

The kmeans procedure requires prior selection of the distance measure used for optimization, initial locations for the clusters' centers (seeds), and the number of clusters to partition the data. The squared Euclidian distance was used as the distance metric over which the optimization was performed. Euclidian distance was preferred because of the low dimensionality of the clustering measure and ease of interpretation.

To mitigate the chance of a local optimization, the seed locations were determined using three methods. First, the seeds were allowed to be selected at random. Next, the seeds were selected based on the expected cluster groupings. These seed values were selected based on expert knowledge of vehicle deceleration kinematics. The last method selected seeds by evenly distributing them across the range of the cluster variable. For each of the three seed selection methods the kmeans optimization was allowed 100 iterations for convergence and was replicated 10 times. For all seeds and replicates, convergence was obtained and resulted at the same total sum of distances. Therefore, local minimums did not appear to exist in the clustering variable; thus, seed selection strategy did not influence the resulting clusters.

Selection of the number of clusters was based on the stopping behavior groups described in the literature review. In general, the previous research and experimenter experience suggested a range of potential groupings from a simple two level to a more comprehensive five level grouping scheme. The potential groups are provided below (Table 7).

Table 7. Potential cluster groupings.

2 Level Clustering	3 Level Clustering	4 Level Clustering	5 Level Clustering
Violation	Stop	Stop	Conservative stop
No violation	Moderate violation	Rolling violation	Normal stop
	Severe violation	Moderate violation	Rolling violation
		Severe violation	Moderate violation
			Severe violation

To evaluate the appropriate number of clusters, the kmeans analysis was repeated for two, three, four, and five groups. The quality of the clusters was evaluated through 1) the overall and within cluster silhouette widths (Rousseeuw, 1987) and 2) the functional implications of the cluster thresholds. The silhouette width is based on the proximity of an observation to other observations in its cluster versus the proximity to observations assigned to its brother cluster; the width ranges from -1 to 1. A positive silhouette width indicates an observation that belongs in the assigned cluster; the higher the value, the stronger the association. Negative values represent likely misclassifications.

Considering the analysis results (Appendix D.), the ‘four group’ classification scheme was selected. Although the two group scheme resulted in the highest overall silhouette width (0.9924), the groupings were too coarse to segregate the desired behaviors. For instance, the first group contained drivers who demonstrated a maximum RDP of up to 2.23 g. It is desirable to break this group down further as a driver with a very low RDP should not be grouped together with a driver with a 2g+ RDP. On the other hand, the five cluster model contained too much resolution with partitions in the data that were not necessary for the algorithm to discriminate. In addition, the five level models failed to converge in 100 iterations for a few of the replicates. This may indicate over-partitioning of the data which could lead to lower repeatability in future investigations.

The three and four cluster models performed nearly identically when the overall silhouette widths were compared. However, the within-cluster silhouette widths tended to be better for the four cluster model. In addition, the clusters created (Figure 23) generate logical partitions of RDP. Cluster 1 consists of the normal driver approach containing nearly 25,000 approaches with an average RDP of 0.25 g. Cluster 2 contained 5573 drivers with an average RDP of 0.75 g and likely contains most safe rolling violations. Cluster 3 contained 51 drivers with an average RDP of 3.15 g representing violations. Cluster 4 contained the most severe violations with only four drivers and an average RDP of 9.8 g. The error bars displayed on the figures represent the 5% and 95% of the population determined by a distribution fit, discussed below.

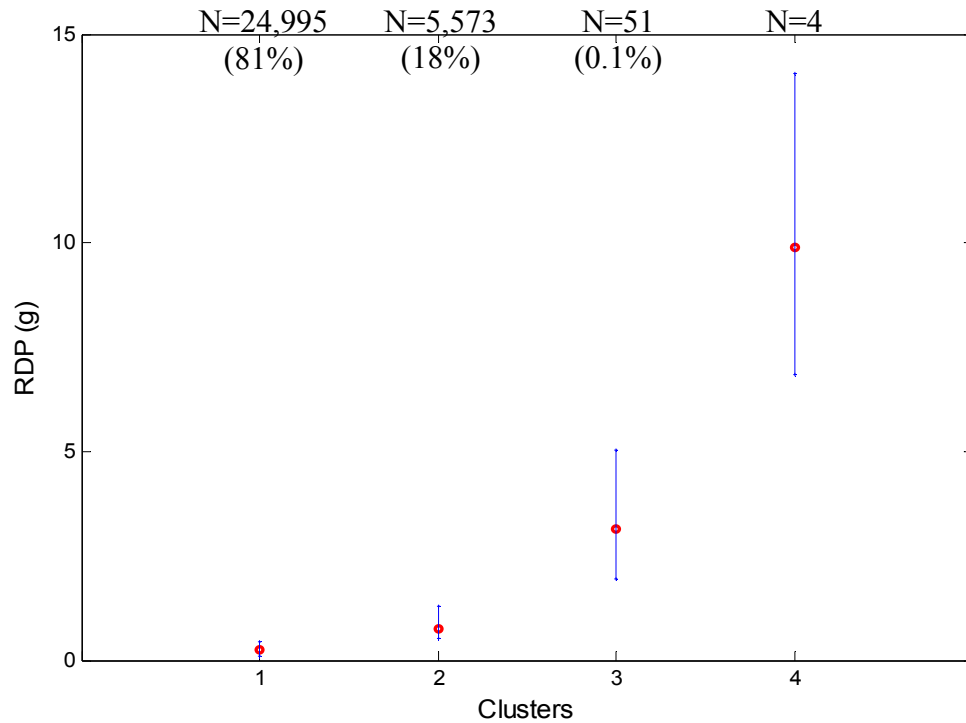


Figure 23. Average RDP and the 5% and 95% population estimates show as error bars for each of the four assigned clusters describing different types of stop-controlled approach behavior.

The cluster analysis provided a partitioning scheme to describe stopping behavior within the sample provided. However, clustering in itself does not describe how the population is distributed into the partitions. To obtain this information, a distribution fit was performed to the overall max RDP measure and within cluster groups. A variety of distributions were considered including normal, exponential, inverse Gaussian, Generalized Pareto, Rayleigh, Birnbaum-Sanders, and Gamma. However, the best performing and most applicable distribution for the measure was the Generalized Extreme Value (GEV) distribution. The GEV distribution is a family of continuous probability distributions that are well suited for modeling the tails of other distributions (Kotz and Nadarajah, 2001). As such, it is particularly well suited for the max RDP measure that was used as the clustering measure. In addition, the GEV is good for capturing skewed data. This data includes unusual events which are of particular interest in this application (e.g., serious violations which are uncommon). Finally, when graphically comparing the GEV to the other potential distributions listed above, the GEV was superior at capturing the empirical distribution of the data.

To ensure that the GEV was appropriate, a one sample Kolmogorov-Smirnov (KS) goodness-of-fit test was performed. The KS test is a statistical comparison between the empirical distribution function of the data and the theoretical underlying population distribution (Chakravart, Laha, and Roy, 1967). The KS test can be sensitive to large data samples, thus the datasets were ordered and down-sampled to 50 equally distributed observations. These 50 observations were used to construct the cumulative distribution function which was then

compared to the GEV using the KS test. The KS test was significant at a 0.05 alpha for the overall max RDP ($p=0.195$) as well as for each of the within-cluster groups (Appendix D.). Thus, the GEV distribution may be used to model driver stopping behavior at stop-controlled intersections. The resulting cumulative distributions are provided below (Figure 24) with the parameter estimates provided in Appendix E.

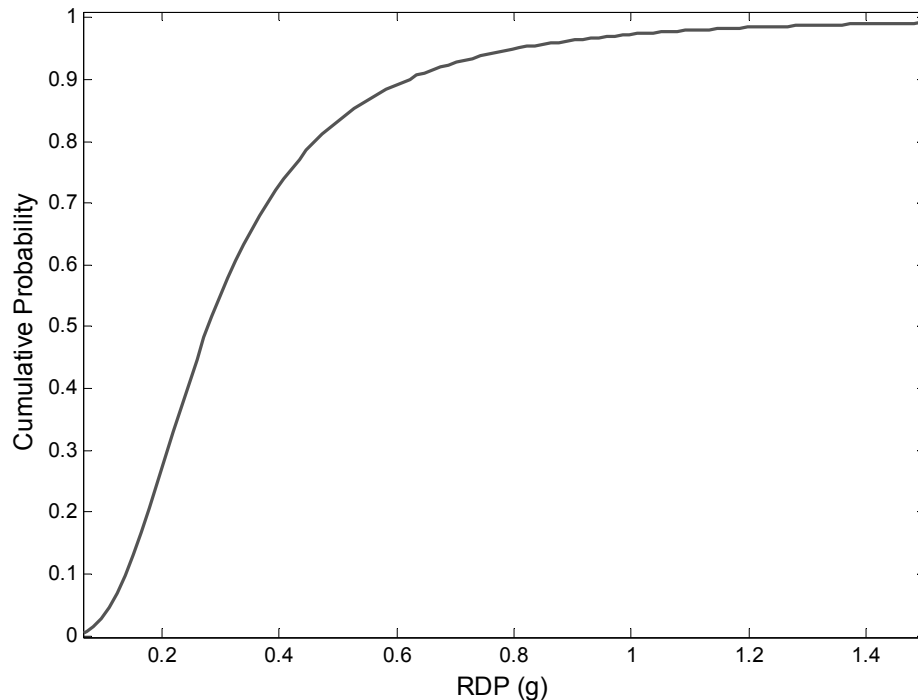


Figure 24. Cumulative distribution for the overall max RDP exhibited by drivers as they approached the stopbar.

For the overall distribution, the average driver RDP was 0.35 g which is consistent with typical driver braking levels. Ninety-five percent of the population will approach the stopbar with an RDP of 0.8 g or less. This suggests that most drivers approach the stopbar with an RDP that allows them to come to a complete stop if required. However, as will be demonstrated in the subsequent section, most of these drivers will roll through the intersection without bringing their vehicle to a complete resting state. These rolling violations result in the seemingly high RDP values which represent deceleration rates at which most drivers would not intentionally brake.

The GEV fit also works well for describing the distribution of drivers within each of the cluster groups created (Figure 25). These GEV fits were used to calculate the error bars displayed in Figure 23. While the first cluster contains considerably more data than any of the other three clusters, it is also substantially more compact. As the clusters' average RDP increases, the number of observations decreases and the variance increases. The fourth-cluster fit is based on only a few data points and is explorative in nature. Considering the four cluster scheme and the data reduction results described in a previous section, it appears that the target groups are primarily contained in the third and fourth clusters. These two clusters represent

drivers that violate at speeds that do not permit them to come to a complete stop prior to entering adjacent traffic.

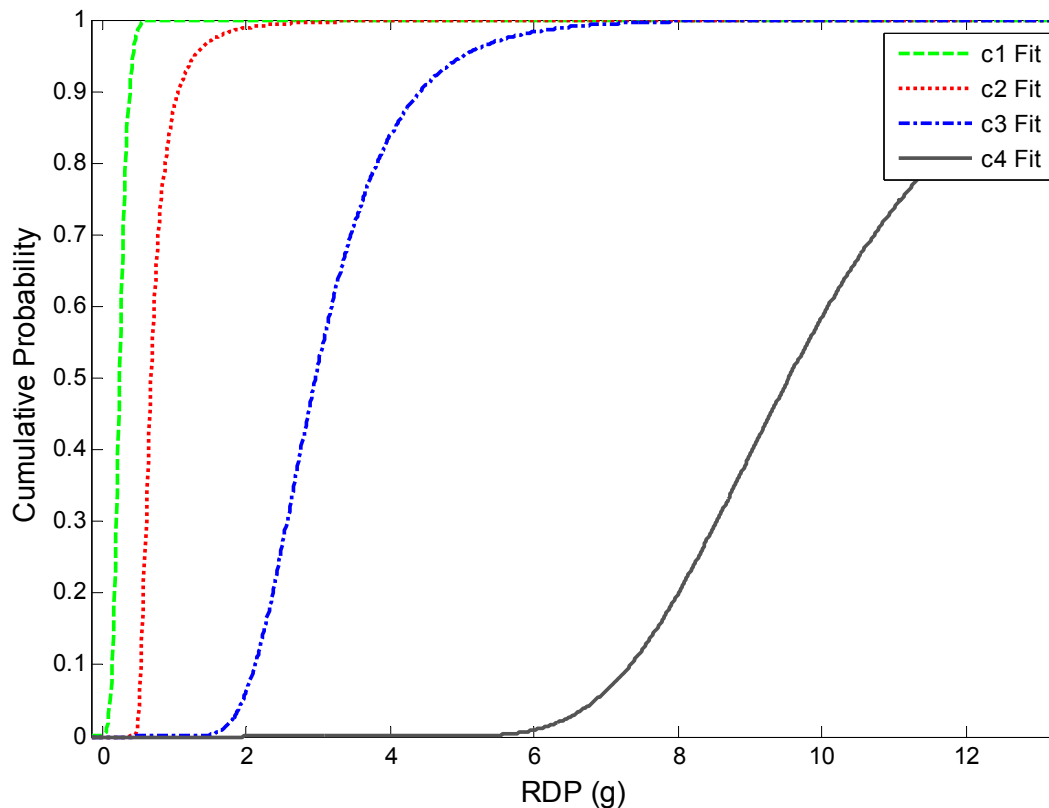


Figure 25. Cumulative distribution for RDP within each of the clusters

Stopbar Speed Analysis

The average RDPs found for the cluster analysis suggested that a large portion of drivers do not completely stop at the stopbar. These drivers are performing a “rolling stop” (i.e., a low speed violation performed by an attentive driver). This type of stopping behavior is performed intentionally by a driver that does not want to come to a complete resting state. Therefore, if a warning is issued to this driver it will likely be considered a nuisance. Such nuisance alarms will have a negative impact on driver acceptance which may reduce the effectiveness of a necessary alert.

To investigate the stopbar speed of drivers at stop-controlled intersections, the distribution of minimum speed from 2 m prior to the stopbar up to 1 m over the stopbar was evaluated (Figure 26). This region was selected because it allows for variations in the location where drivers stop. Only vehicle approaches in which the vehicle was reported traveling through the stopbar region were included in the analysis for a total sample of 28,880 observations. As expected, a significant number of drivers did not fully stop their vehicle. Half of the drivers

maintained a speed of 2 m/s (4 mph) or more as they crossed the intersection. However, 90% of the drivers exhibited a minimum stopbar speed of less than 4 m/s (9 mph).

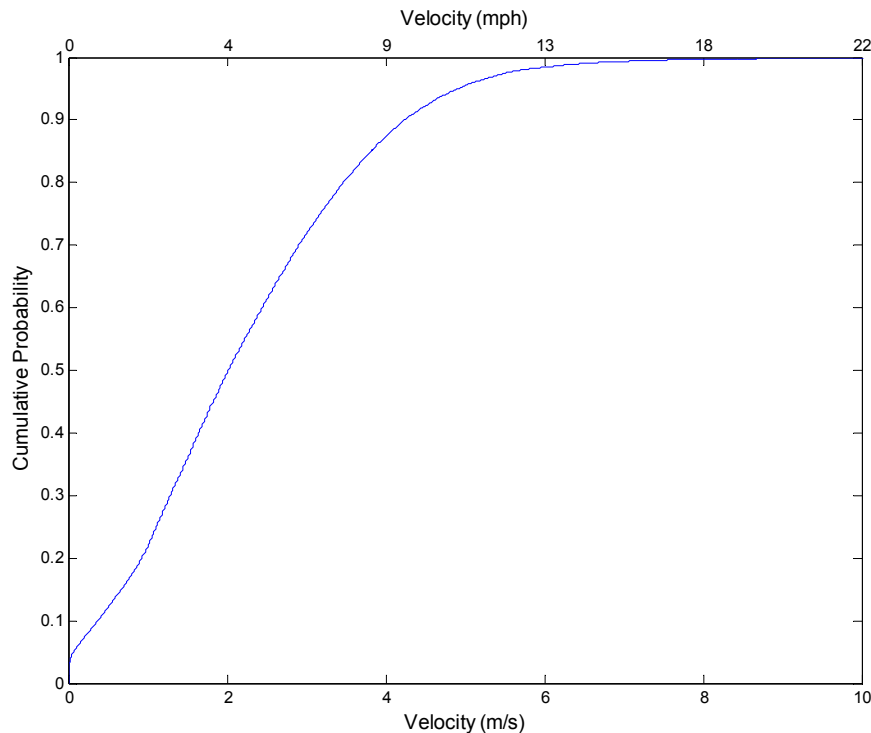


Figure 26. Empirical cumulative probability of stopbar speed at stop-controlled intersections.

This overall distribution may be partitioned into the driver stopping groups defined by the cluster analysis. The behavior grouping provides insight into the relationship between the RDP measure of stopping behavior and the driver's corresponding stopbar speed. Figure 27 depicts the distribution for stopbar velocity for each of the clusters. A generalized extreme value distribution was fit to cluster 1 while normal distributions were fit to clusters 2, 3 and 4 (parameters are provided in (Appendix F.)). However, the distribution fits for clusters 1 and 4 are for explorative purposes only. Cluster 4 only contains four observations which are not sufficient for modeling the population. Cluster 1 exhibited a mixed distribution with a substantial number of drivers completing a full stop; only the drivers that did not stop are modeled by the depicted fit.

As expected, stopbar speed tends to increase with the stopping behavior group. The cluster analysis suggested that the goal should be to warn all drivers in clusters 3 and 4. It also suggested that warning drivers in the tail of the cluster 2 distributions would not be considered a false alarm. This logic suggests that a minimum speed threshold of 4.4 m/s (10 mph) may be appropriate as a warning criterion. Any driver that is traveling at less than this speed threshold would not be warned. As illustrated in Figure 27, a speed threshold of 5 m/s (11 mph) is predicted to warn over 99.9% of the drivers in the third and fourth cluster while only providing alerts to 20% of the cluster 2 drivers.

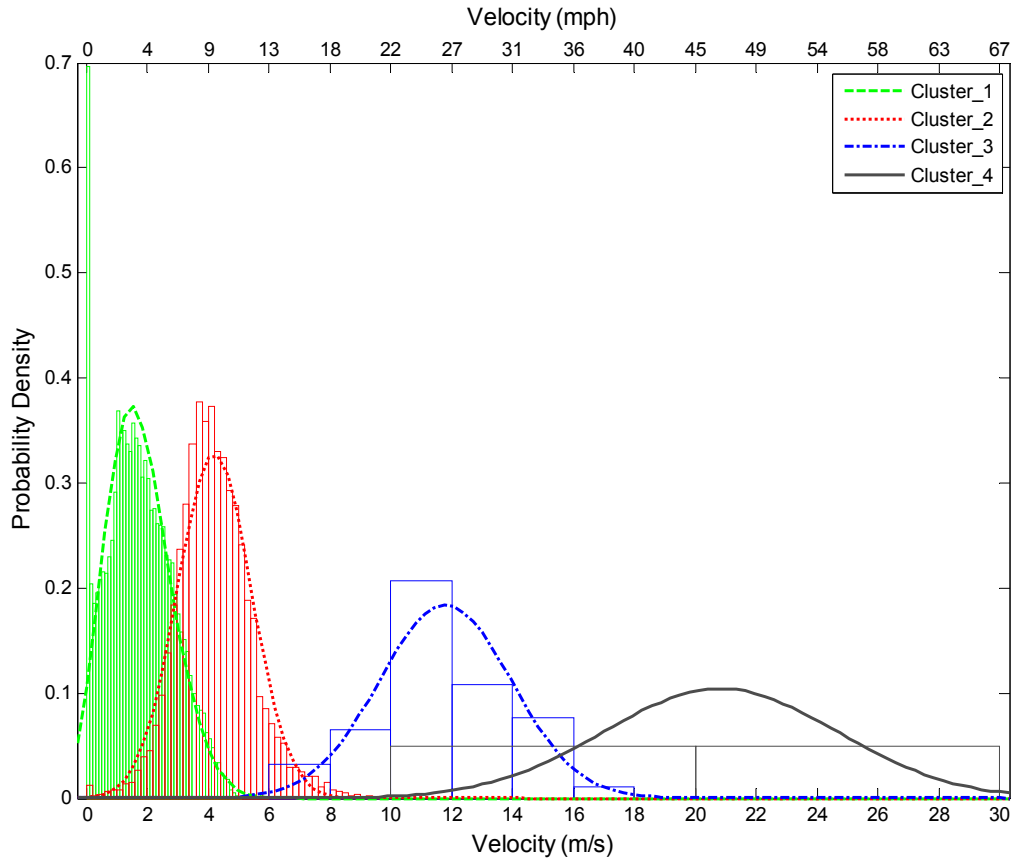


Figure 27. Distribution of stopbar speed partitioned by cluster. Curve fits for clusters 1 and 4 are for explorative purposes only.

Brake Onset Analysis

The literature reviewed in the introduction suggested that nuisance alarms should be minimized to ensure warning acceptance and effectiveness. One potential method to avoid alerting an attentive driver is to monitor the brake status. If the driver is braking, it may be reasonable to assume that he/she is aware of the intersection and does not require an alert. Thus, when the driver is actively braking, an ICAS algorithm would suppress the warning regardless of the vehicle's other kinematic measures. To investigate the use of brake status as a component of the algorithm, a box plot of brake onset for clusters 1 through 3 was drawn (Figure 28). A box plot for the fourth cluster is not included as only one driver in this group applied the brakes during the approach. Cluster 3 only includes ten observations as the rest of this group also did not press the brake during their approach. Thus, the box plot is explorative.

A box plot is a statistical method that visually shows the empirical distributions of different populations without any assumptions of the underlying distribution. Most of the data within a group is contained inside the box which is bounded by the first and third quartile. The line within the box represents the median. The whiskers depict the regions that lie within 1.5 times the corresponding quartile. Values outside the whisker are unusual observations and may be treated as outliers.

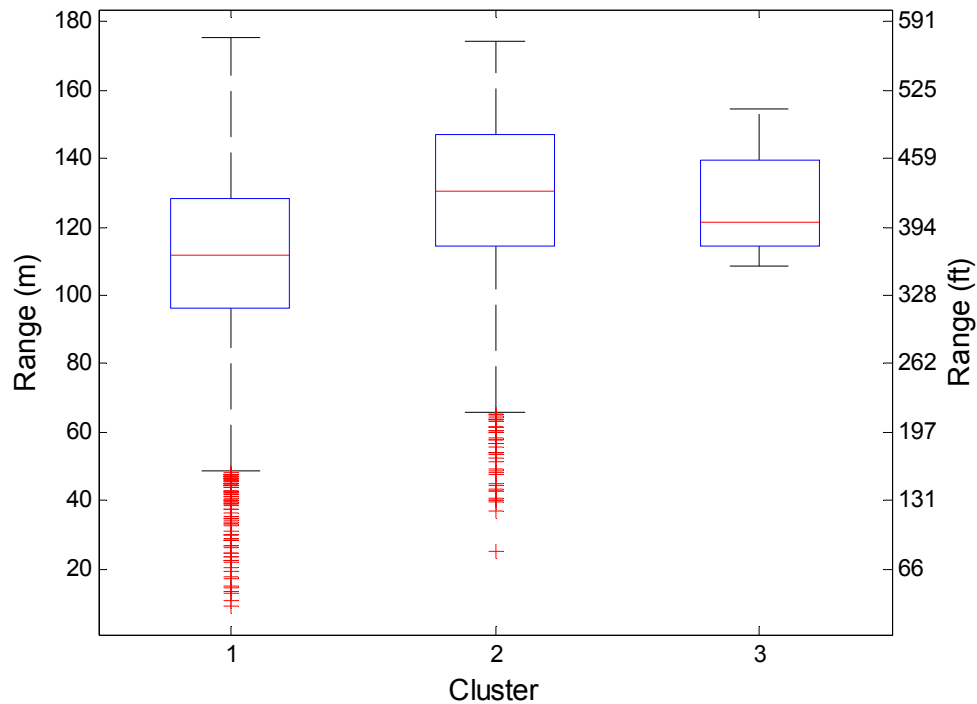


Figure 28. Box plot for brake onset.

The overlapping boxes indicate that a statistical difference for the onset of braking is unlikely. It is interesting to note that the conservative drivers of cluster 1 appear to brake closer to the intersection. This result was initially unexpected but may be explained by the approach style of the conservative drivers. Discussed further in the subsequent section, drivers in cluster 1 tended to initiate slowing earlier by coasting; thus, taking advantage of engine braking without actually applying the brake. At some point closer to the intersection an increase in slowing is required such that the brake is eventually applied.

The primary lesson that should be extracted from this plot is that conservative drivers do not necessarily brake further from the intersection than do aggressive drivers. Thus, algorithm designers should not assume that a braking criterion in the algorithm will preclude compliant drivers from receiving a warning. It may, however, help to suppress the warning for aggressive drivers that will perform a compliant stop.

Given that group differences for drivers that apply the brake do not appear to exist, the groups were collapsed while the distribution of brake onset was evaluated. Brake onset followed a normal distribution (Figure 29 & Figure 30, parameters available in Appendix G.). The distribution of braking was calculated using both distance and TTI as the dependent measures. There are indications that TTI is a better measure for representing variables such as brake onset because it is a construct thought to be used by the human visual system to judge when slowing should occur.

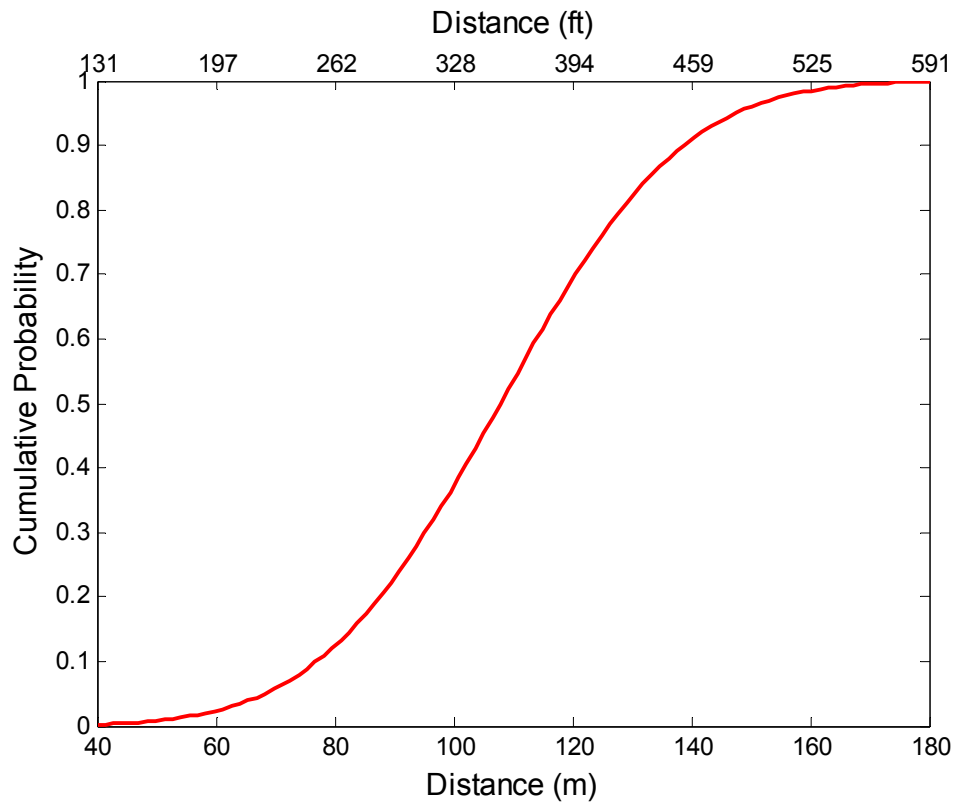


Figure 29. Cumulative distribution of brake onset as a function of distance to the intersection.

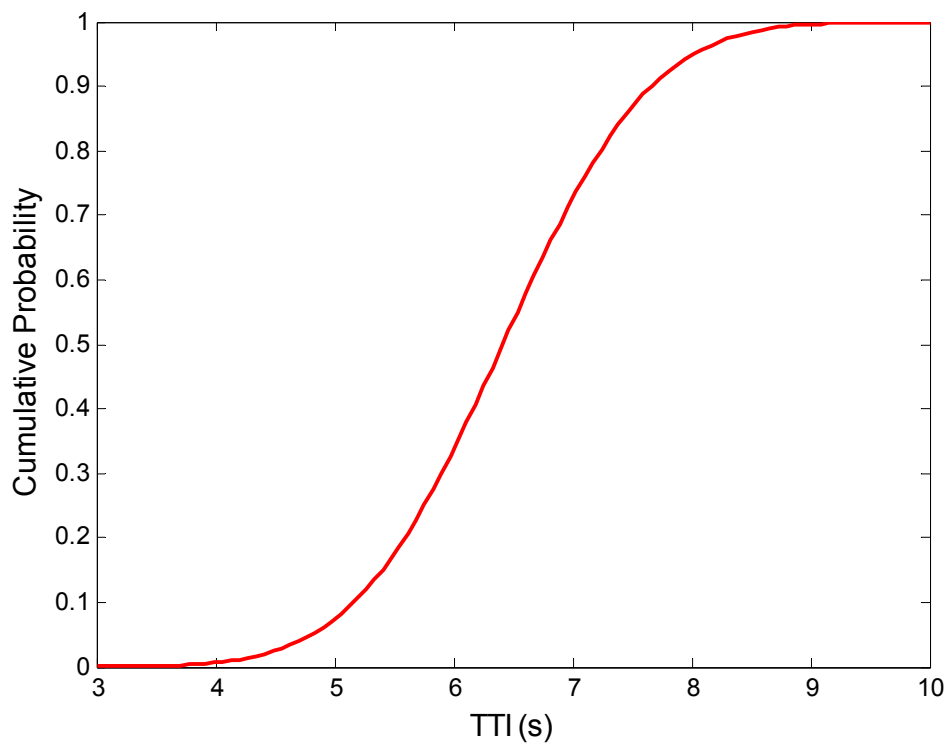


Figure 30. Cumulative distribution of brake onset as a function of TTI.

In general, 99% of the drivers appear to have applied the brake by 51 m or a TTI of 4.3 s or greater. This brake onset does provide support for the use of a brake onset component in an ICAS algorithm. Smart Road experiments testing an ICAS warning interface have suggested a timing of 2.44 s TTI for issuing a warning and obtaining a high compliance rate. Based on the data above, most drivers that will brake would have applied the brake prior to this warning range.

Investigation of Driver Trajectories

Moving beyond the investigation of driver behaviors at a specified point, it is important for an ICAS algorithm to consider measures across the entire intersection approach. The algorithm will need to operate on a continuous basis in order to provide a timely warning to the violating driver (Neale et al., 2005). Considering the exploratory nature of this project the vehicle trajectories will be graphically analyzed for trends that may be exploited during future statistical analysis. It is beyond the scope of the present work to develop statistical methods to fully characterize the trajectory data.

Plots of the vehicle kinematic measures provide insight into the differences between driver groups that may be used for threat assessment. A figure for each kinematic variable used in this study is provided below along with a discussion of the characteristics of the plot. Each plot was created by calculating the average of the depicted kinematic measure every 2 m from the stopbar to 100 m from the intersection. The observations used to produce the figures are the same sample that was described in the driver stopping behavior section above. For the sake of clarity, the figures in this section only depict the mean of each measure. When applicable, the confidence intervals for the data will be described.

The velocity trajectory plot (Figure 31) demonstrates the large separation between the cluster 4 stopping behaviors. While caution should be exercised not to over-interpret the small sample in this group, drivers did tend to exhibit higher velocities that demonstrate a minimal decrease over the entire intersection approach. Cluster 1 also tends to exhibit a different approach than the other clusters. As described above in the brake onset results, drivers in cluster 1 brake earlier resulting in a lower velocity at the maximum range reliably measured by the radar (100 m).

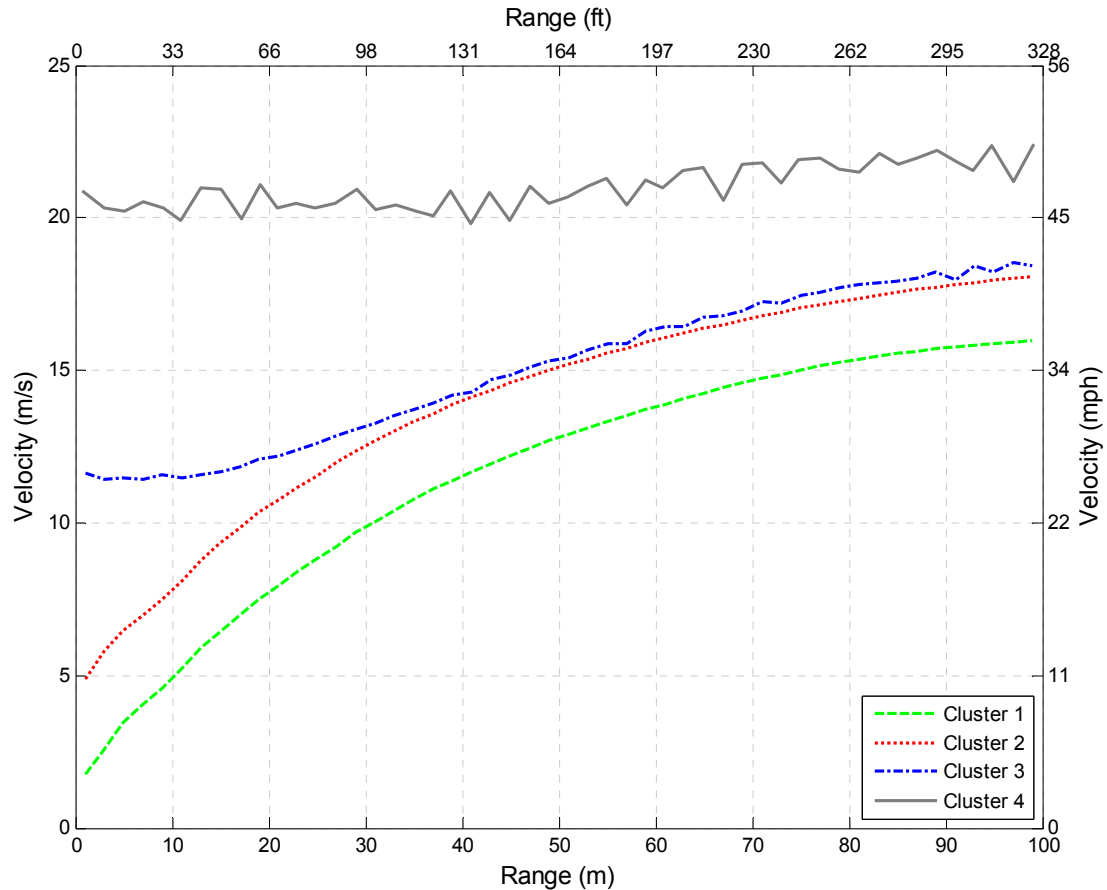


Figure 31. Mean velocity trajectory of vehicles partitioned by the stopping behavior cluster.

The second cluster parallels the velocity trajectory exhibited by the first cluster. It appears that drivers in cluster 1 and 2 are similar except in their desired stopbar speed. Functionally, their velocity patterns are identical, except cluster 2 drivers appear to brake later and thus carry more velocity into the intersection creating the “rolling stop” behavior.

Clusters 3 and 2 are extremely similar at longer distances from the stopbar. As the intersection is approached, the two curves deviate with cluster 3 drivers carrying much higher speeds into the intersection than cluster 2. Partitioning drivers from clusters 2 and 3 will be the challenge for an ICAS algorithm. For the mean velocity depicted, the curves for these two groups diverge around 33 m. This distance from the intersection is likely sufficient for providing a warning to drivers. However, when the confidence intervals are considered, the point of separation (between the 15% and 85% confidence limits) is reduced to just 10 m. This distance may not provide sufficient warning time for higher speed approaches.

Like the velocity trajectory plots, the acceleration plots also provide substantial insight into the differences among the driver stopping behavior clusters (Figure 32). Drivers in the fourth cluster exhibit a minimal amount of braking early in the approach which tends to drop off to nearly no braking after that point. It is doubtful that this mild acceleration is indicative of an

intentional violation as none of the intersections have sight distances that would permit drivers to see the adjacent traffic at the speeds exhibited by this group. It is possible that this early acceleration was the result of some other intersection characteristics (e.g., the incline present at the Nellies Cave intersection).

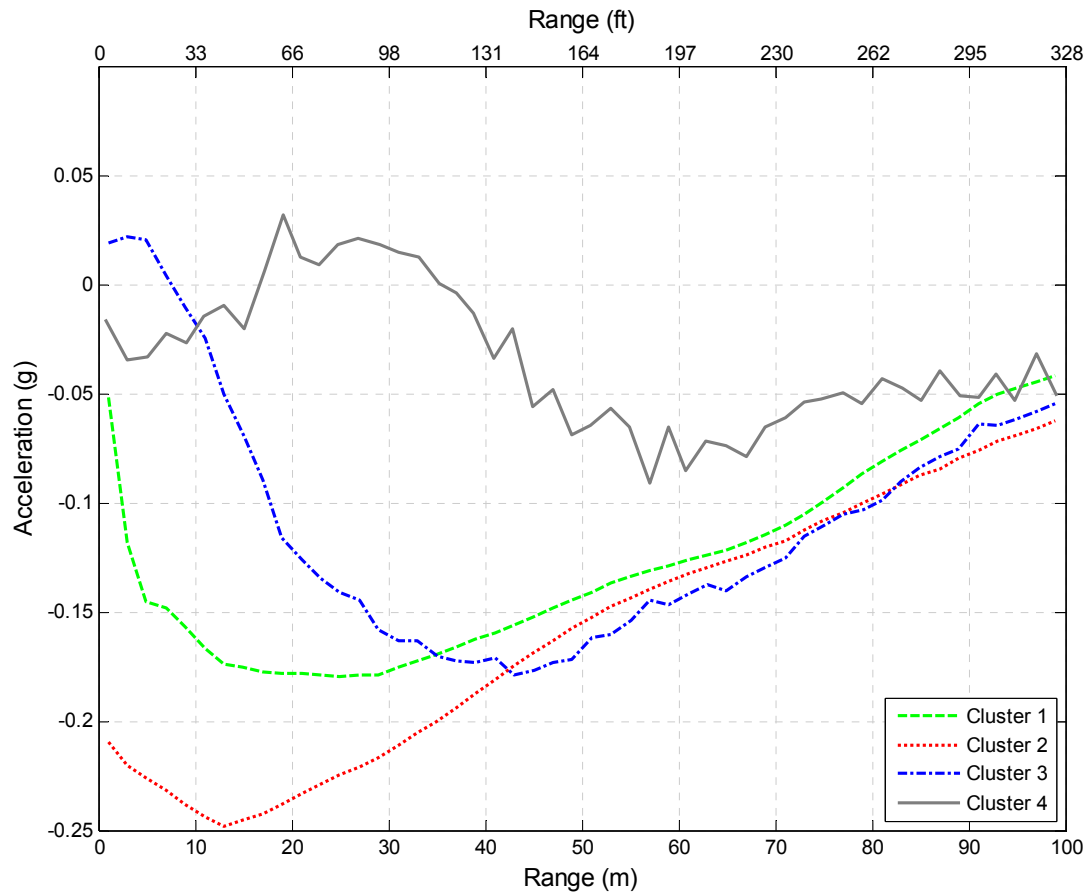


Figure 32. Mean acceleration trajectory of vehicles partitioned by the stopping behavior cluster.

Clusters 1, 2, and 3 all exhibited similar decelerations for the first 50 m of the approach. Cluster 1 drivers demonstrated the lowest acceleration initially, likely because they pressed the brake furthest from the intersection, providing more time over which their speed was able to decrease. Interestingly, the cluster 2 drivers appear to brake harder toward the end of the intersection approach. This may indicate that this group of drivers is more aggressive with their approach behavior. These drivers are aware of the intersection and initiate braking fairly early, but carry more speed further into the intersection and brake harder and later than cluster 1 drivers. Drivers in cluster 3 may initiate braking such that they can stop if cross traffic is present. However, it appears that at around 40 m, these drivers may feel confident enough that cross traffic is not present such that they cease braking behavior.

The velocity plots indicated that clusters 2 and 3 would be difficult to discriminate. This discrimination may actually be somewhat more sensitive for acceleration. In particular, the cluster 3 group appears to stop braking earlier in the approach than does cluster 2. This behavior

may indicate an intentional violation of the stop-controlled intersection. However, clusters 1 and 2 are actually more similar with regard to their acceleration patterns. Perhaps the algorithm will be most effective if it considers a combination of speed and acceleration to take advantage of the strengths of each measure.

In terms of algorithm assessment, RDP appears to have some significant advantages over the other kinematic measures (Figure 33). In particular, the confidence intervals for clusters 1 and 4 never overlap at any point during the entire intersection approach. Again, care should be taken not to over interpret cluster four; however, this suggests that discrimination between these groups will be more effective than for any other kinematic measure. The cluster 1 and 4 groups, in general, are substantially more divergent from every other group demonstrated by the additional kinematic measures. The difficulty in discrimination is once again apparent when clusters 2 and 3 are compared. The confidence limits of these two groups do not completely diverge until the intersection stopbar is within 10 m. It is likely that some drivers in cluster 2 will be warned in order to catch most of the drivers in cluster 3. However, the drivers on the tail of cluster 2 will have relatively aggressive approaches and will contribute to roughly less than 5% of the overall population based on the models presented in the cluster analysis.

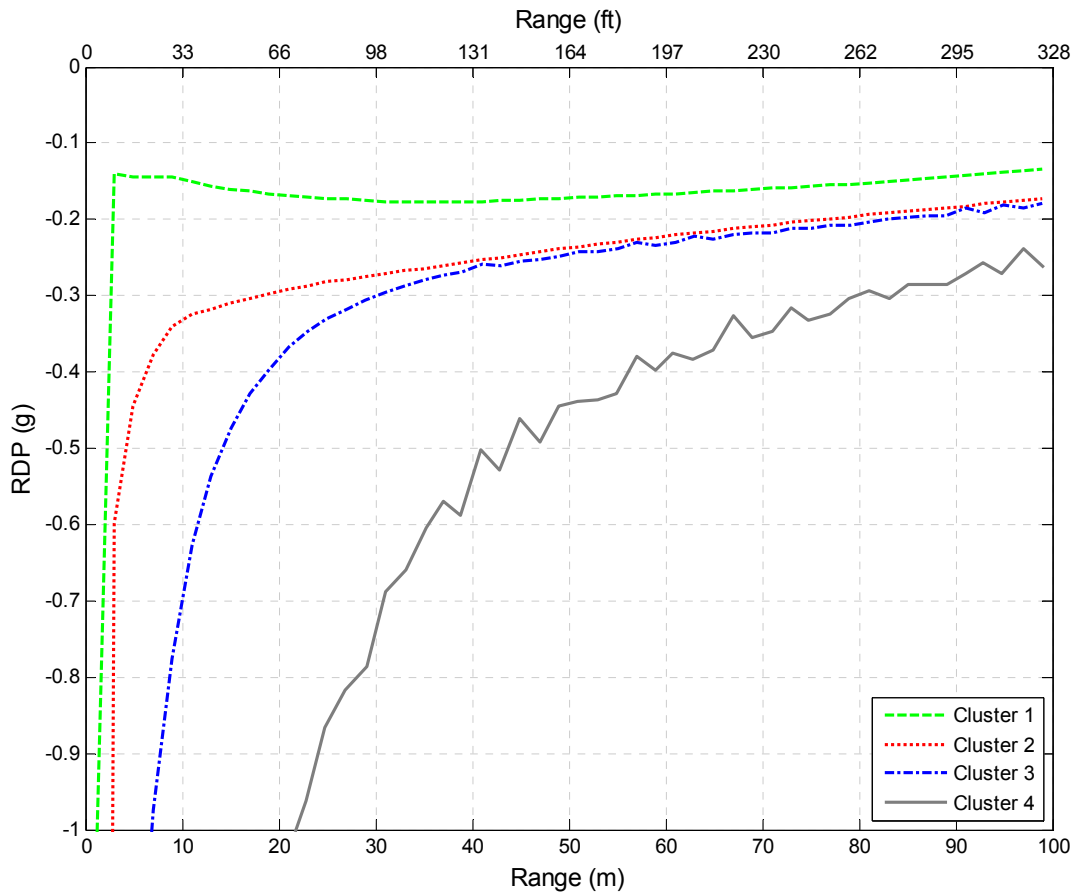


Figure 33. Mean RDP trajectory of vehicles partitioned by the stopping behavior cluster.

The TTI trajectories contain the largest variability of any kinematic measure (Figure 34). The spikes present in the data represent locations in which vehicles began to enter queues at the intersection. When a vehicle slows, the velocity approaches zero which in turn causes TTI to approach infinity. Thus, the TTI for any given vehicle may tend toward infinity at a variety of locations depending on the queue length at a specific stop-controlled intersection. However, this behavior is one-sided such that TTI could make an effective algorithm component if the algorithm looks for values below a set value. This value, however, must be dependent on the distance to the intersection as TTI will always approach zero as the intersection stopbar draws nearer.

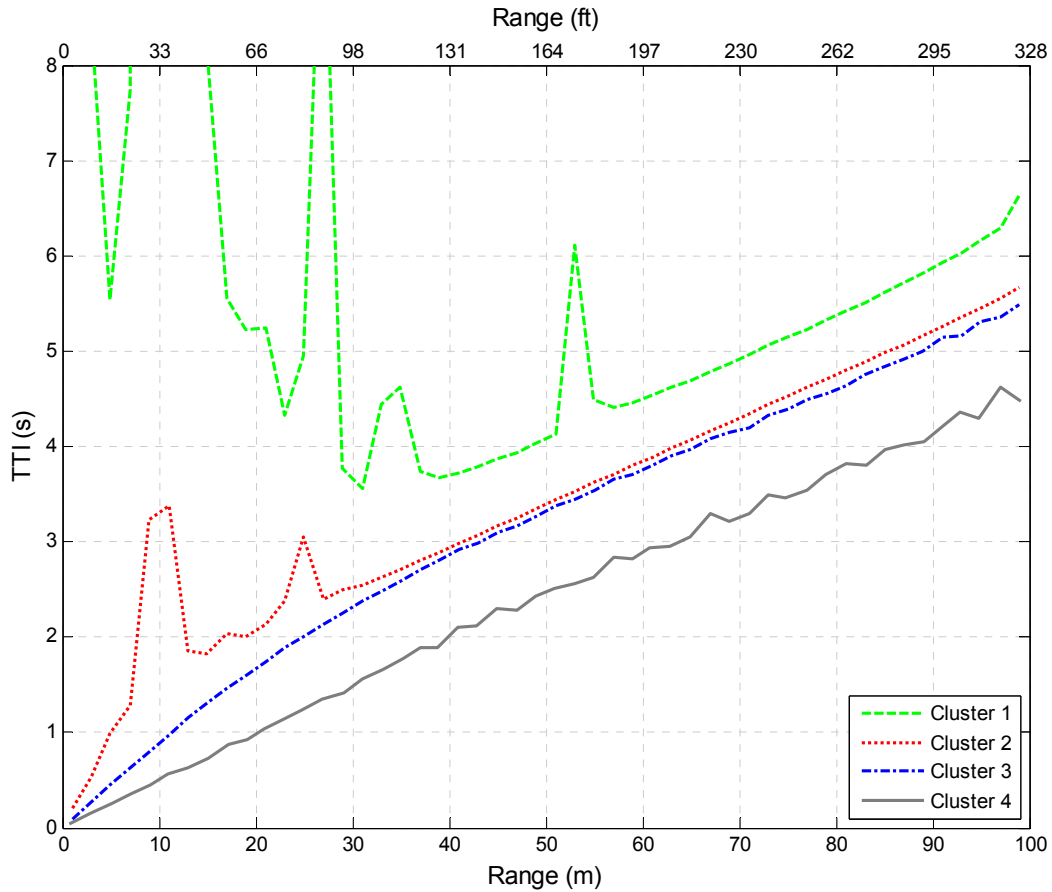


Figure 34. Mean TTI trajectory of vehicles partitioned by the stopping behavior cluster.

From the figure above it appears that only clusters 1 and 2 regularly approached the intersection with a queue. Drivers in cluster 1 tended to encounter queues more frequently and at greater distances than cluster 2. If the cluster 2 TTI trajectory is projected for cases in which the queue was not present, it does not appear substantially different than cluster 3. This may indicate that TTI is not the optimal measure for determining when the warning should be initiated.

LIMITATIONS OF THE PRESENT STUDY

As with most research projects, there are certain limitations that need to be considered when interpreting the results. First, the geographic region was limited to southwest Virginia. It is possible that intersection approach behavior depends on regional differences which could affect the some of the presented results. Furthermore, the data collection took place over two consecutive months at each site. Thus, the database collected may not reflect seasonal differences in driving behavior.

Placing the DAS at the intersection was required to obtain the volume of data necessary to construct a valid ICAS algorithm. However, placing the data at the intersection rather than in the vehicle has certain limitations. This study did not provide information about the driver actions that led up to the violation. For instance, it is not be possible to know for certain whether a violation was intentional or unintentional.

Finally, the radar used for this research did not perform at the level anticipated. The radar tended to provide sparse data where continuous data should have been provided. This resulted in a significant post-processing effort to improve the data. During this effort only vehicle tracks that contained sufficient fidelity were carried through to the analysis portion of the study. While there was no direct evidence to suggest that this systematic selection resulted in confounding the data, it is possible that certain types of vehicles or vehicle approach characteristics were prone to degraded radar performance. Thus, a certain type of vehicle or vehicle approach may be unknowingly underrepresented in the dataset.

FINDINGS AND CONCLUSIONS

The primary objective of this research project was to collect a database that could be used to develop ICAS algorithms and to provide a proof-of-concept for the efficacy of an ICAS threat assessment. An explorative analysis of driver stopping behavior including the vehicle trajectories was performed. The results presented indicate that an intersection collision system for stop-controlled intersections is feasible. The following primary conclusions are provided:

- It is feasible to develop and install a non-obtrusive self-contained DAS for collecting continuous vehicle data from stop-controlled intersections. While this data collection was occurring, higher quality radar has become available that will provide more accurate data that will require less post-processing and alleviate the issues with dropped vehicle tracks. Future studies should incorporate this radar into their data collection efforts.
- Data from the DAS can be successfully mined to explain driver behavior through monitoring vehicle kinematics. In particular, speed, distance, and acceleration trajectories for drivers that stop and violate can be compared. This DAS cannot capture in-vehicle driver characteristics, such as distraction. Driver-level measures such as this must be captured using in-vehicle methods and may be best evaluated during a FOT with an algorithm based on data such as the data recorded herein.

- Driver stopping behavior can be grouped by differing styles. Most drivers exhibit compliant and non-aggressive intersection approaches. The sample of intersections studied suggests that approximately 18% of drivers will perform a more aggressive approach and tend not to slow their vehicle below 4 m/s (10 mph) while crossing the stopbar. The cluster analysis suggested that this may be a good starting point for a low-speed threshold in the algorithm.
- A step in ICAS development will be defining the target population. This determination should be made cooperatively by state and federal DOTs, OEMs, policy makers, and other stakeholders. To assist in this decision, Table 8 provides a summary of the target population as a function of the stopbar speed and the minimum RDP. The final target population should be based on integration of the data provided in this report with additional research projects. In particular, future research should work toward determining a relationship between the driver's behavior at the stopbar and crash risk.

Table 8: The four selected violation threshold criteria and the resulting samples.

Stopbar Velocity	Minimum RDP	Percent of Sample	Percent of Population	Percent of Cluster 1	Percent of Cluster 2	Percent of Cluster 3	Percent of Cluster 4
2.2 m/s (5mph)	0.34g	35.5%	36.5%	17.6%	100%	100%	100%
4.52 m/s (10mph)	0.80g	5.5%	5.0%	0%	27.7%	100%	100%
6.72 m/s (15mph)	1.36g	0.6%	1.1%	0%	4.2%	99.9%	100%
8.92 m/s (20mph)	2.00g	0.2%	0.3%	0%	1.2%	93.8%	100%

- For drivers that brake during the intersection approach, there do not appear to be any significant differences in brake onset for different stopping types. This implies that brake status early in the approach does not indicate that a driver is attentive to the stop sign. Brake status may be used to suppress warnings for aggressive drivers who will comply with the stop sign and do not require an alert.
- The trajectories of the vehicles approaching a stop-controlled intersection tend to diverge for the different stopping clusters upstream of the stopbar. Some measures such as RDP appear to be better predictors of stopping clusters than other measures. In addition, it appears that increases in algorithm performance may be obtained by combining several of the measures during threat assessment. Overall, the exploratory trajectory analysis provides evidence for the feasibility for the success of an ICAS threat assessment at stop-controlled intersections.

OBSERVATIONS AND FUTURE DIRECTIONS

The research discussed herein is a key step in the development of an ICAS. Nonetheless, future research will be required to bring the ICAS concepts into fruition.

1. VTRC and other state and federal DOTs must work with industry to apply the methods devised in this research to signalized intersections. An expanded database of both signalized and stop-controlled intersections should be obtained in future analysis to fully develop driver approach models across a wide range of intersection geometries, localities, and control types.
2. An algorithm must be devised and tested using the collected data. By overlaying an algorithm on the collected data the actual performance of an algorithm can be measured. A variety of algorithm types should be devised and tested using a systematic optimization procedure. The end result of this research should be an optimal ICAS threat assessment algorithm. Several decisions will need to be made during this process. Many of the decision regarding the algorithm formulation can be made objectively by the researchers. Other decision, such as who to warn should be jointly determined by state and federal DOTs, OEMs, and other stake holders.
3. The warning interface must be optimized by researchers. Studies should be completed to determine the appropriate modalities and physical locations for the driver vehicle interface (DVI). There is an interaction between the algorithm timing and the effectiveness of a DVI. For example, a more effective DVI will allow the driver to be warned later in the approach. DVI tests are planned as part of the CICAS -V project and should eventually be integrated with the results of this research to verify the algorithm feasibility.
4. A full cost benefit analysis should be performed using information from the VTRC and other state DOTs, OEMs. Warning effectiveness should first be evaluated based on the results from the algorithm development and DVI experiments. The effectiveness data should next be combined with crash statistics to determine the national benefit in terms of reduction in injuries and loss of life. Finally, state DOTs and OEMs can provide cost estimates for installing and maintaining the roadside and in-vehicle equipment required for the ICAS system.
5. In addition to the human factors work, a significant volume of engineering development must be completed for an ICAS to become reality. Engineers need to develop the on-board equipment, the roadside equipment, and the communications backbone necessary to support such a system. These systems will need to meet specifications necessary to compute and deliver the warning in a timely and appropriate manner.
6. After integrating all the components of the threat assessment algorithms, the warning interface, and the hardware/software, the entire system must be validated. This validation should include a large scale FOT to demonstrate the system's effectiveness on the open roadway under natural conditions. The VTRC is a stakeholder and should be involved in the planning and execution of the FOT. Furthermore, if the FOT is performed in Virginia the local VDOT region should be used for planning, engineering, and installation assistance.

COSTS AND BENEFITS ASSESSMENT

The data collected through this study provides the first naturalistic dataset on stop-controlled intersection approach behavior and represents the first step towards the development of an algorithm for collision avoidance at stop-controlled intersections. The data indicate that those drivers most likely to violate the stop-controlled intersection at a high rate of speed can be identified at a distance sufficient to provide a warning and hopefully prevent a collision with a crossing vehicle. Without a doubt, the infrastructure required to implement a collision avoidance system will have a significant cost associated with it. Additional research will be required to determine if that cost can be outweighed by the safety benefits provided by such a system.

ACKNOWLEDGMENTS

The authors would like to thank the project contract monitor Catherine McGhee for her support of this project. Additionally, we appreciated the assistance of Chris McDonald and John Holst of the Southwestern Region of the Virginia Department of Transportation in the site selection process and for acquiring the necessary permissions for installation of the data acquisition equipment.

REFERENCES

- BMI. *Infrastructure-based intersection collision avoidance concept study: Technical memorandum 2 – Task 2 – Analysis of crashes at intersections selected from cooperative states* (Contract No. DTFH61-96-C-00077): Bellomo-McGee Incorporated, 2001a.
- Brewer, M.A., J. Bonneson, & K. Zimmerman. *Engineering countermeasures to red-light-running*. Paper presented at the Proceedings of the ITE 2002 Spring Conference and Exhibit (CD-ROM), Washington, D.C., 2002.
- Chakravart, Laha, & Roy. *Handbook of Methods of Applied Statistics: Volume I*. (pp. 392-394). John Wiley and Sons, New York, 1967.
- Chovan, J., L. Tijerina, J. Pierowicz, & D. Hendricks. *Examination of unsignalized intersection, straight crossing path crashes, and potential IVHS countermeasures* (Report No. DOT-HS-808 152). Washington, D.C.: National Highway Traffic Safety Administration, 1994.
- Davidson, I. *Understanding K-Means Non-hierarchical Clustering*. Albany, New York: State University of New York, 2002.
- Dingus, T. A., S. K. Jahns, A. D. Horowitz, and R. Knipling. Human factors design issues for crash avoidance systems. In W. Barfield and T. A. Dingus (Eds.), *Human Factors In Intelligent Transportation Systems* (pp. 55-94). Mahwah, NJ: Lawrence Erlbaum Associates, 1998.
- Dingus, T.A., S.G. Klauer, V.L. Neale, A. Peterson, S.E. Lee, J. Sudweeks, M.A. Perez, J. Hankey, D. Ramsey, S. Gupta, C. Bucher, Z.R. Doerzaph, J. Jermeland, & R.R. Knipling. *The 100-car naturalistic driving study, Phase II – Results of the 100-car field experiment*. (Contract No. DTNH22-00-C-07007). Washington, D.C.: National Highway Traffic Safety Administration, 2006.
- Fakhry, S. M. & K. Salaita. Aggressive Driving: A Preliminary Analysis of a Serious Threat to Motorists in a Large Metropolitan Area. *Journal of Trauma*, 52(2), 2002, pp. 217-224.
- Ferlis, R. *Infrastructure systems for intersection collision avoidance*. Paper presented at the Proceedings of the First International Driving Symposium on Human Factors in Driver Assessment, Training, and Vehicle Design, Aspen, CO, 2001.
- Hauskrecht, M. Clustering. Presentation for CS2750 Machine Learning. Pittsburgh, Pennsylvania: University of Pittsburgh, 2003.
- Hendricks, D.L., J.C. Fell, and M. Freedman. *The relative frequency of unsafe*

- driving acts in serious traffic crashes*. (Contract No. DTNH22-94-C-05020). Washington, DC: National Highway Traffic Safety Administration, 1999.
- Horst, A.R.A. van der. *A time-based analysis of road user behaviour in normal and critical encounters*. Doctoral Thesis, Delft University of Technology, Delft, 1990.
- ITE. Making intersections safer: A toolbox of engineering countermeasures to reduce red-light running. Washington, D.C., Institute of Transportation Engineers, 2003.
- Kotz, S. and S. Nadarajah. *Extreme value distributions: Theory and applications*. River Edge, NJ: World Scientific Publishing, 2001.
- Lee, S. E., M.A. Perez, Z.R. Doerzaph, S.R. Stone, S.B. Brown, V.L. Neale, R.R. Knipling, & G.T. Holbrook. *Task 5 of the Intersection Collision Avoidance - Violation Project: Final Project Report*. Blacksburg: Virginia Tech Transportation Institute, 2005.
- Liu, B.S. Association of intersection approach speed with driver characteristics, vehicle type and traffic conditions comparing urban and suburban areas. *Accident Analysis and Prevention*, Vol. 39, 2007, pp. 216-223.
- MacQueen, J.B. (1967). Some methods for classification and analysis of multivariate observations. *Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability*. Berkeley: University of California Press, 1, 281-297.
- Matlab. (2007). *Statistics Toolbox 6: Users Guide*. Natick, MA: The Mathworks, Inc.
- Najm, W., M. Mironer, J. Koziol, J.S. Wang, & R.R. Knipling. *Synthesis report: Examination of target vehicular crashes and potential ITS countermeasures* (Report No. DOT HIS 808 263). Washington, DC: National Highway Traffic Safety Administration, 1995.
- National Transportation Safety Board (2001). *Special Investigation Report—Vehicle- and Infrastructure-Based Technology for the Prevention of Rear-End Collisions*. (NTSB/SIR-01/01, PB2001-917003).
- Neale, V. L., Perez, M. A., Doerzaph, Z. R., & Stone, S. R. *Intersection Decision Support Final Report*. Blacksburg: Virginia Tech Transportation Institute, 2005.
- NHTSA. *Traffic Safety Facts 2005*. (Report No. DOT HS 810 631). Washington, D.C.: National Highway Traffic Safety Administration, 2005.
- Pierowicz, J., E. Jocoy, M. Lloyd, A. Bittner, & B. Pirson. *Intersection Collision Avoidance Using ITS Countermeasures: Task 9 Final Report – Intersection Collision Avoidance System Performance Guidelines*. (Contract No. DTNH22-93-C- 07024). Seattle, WA: Battelle Human Factors Transportation, 2000.
- Retting, R.A., H.B. Weinstein, & M.G. Solomon. Analysis of motor-vehicle

- crashes at stop signs in four U.S. cities. *Journal of Safety Research*, Vol. 34, 2003.
- Rousseeuw, P.J. Silhouette: A graphical aid to the interpretation and validation of cluster analysis. *J. Computational Applied. Mathematics*, Vol. 20, 1987, pp. 53–65.
- Sanders, M. S., & E.J. McCormick. *Human Factors in Engineering and Design*. New York: McGraw-Hill, 1993.
- Tijerina, L., J. Chovan, J. Pierowicz, & D. Hendericks. *Examination of signalized intersection, straight crossing path crashes, and potential IVHS countermeasures* (Report No. DOT-HS-808 143). Washington, DC: National Highway Traffic Safety Administration, 1994.
- VDOT. *Average daily traffic volumes with vehicle classification data of interstate, arterial, and primary routes*. Washington, DC: Federal Highway Administration, 2005.
- Wang, J.S., & R.R. Knipling. *Intersection crossing path crashes: Problem size assessment and statistical description* (Report No. DOT HS 808 190). Washington, DC: National Highway Traffic Safety Administration, 1994.
- White, B., & B. Ferlis. *Algorithm for determining inattentive signal violators to be used in infrastructure-based intelligent system for signal violation prevention*. Paper presented at the 83rd annual meeting of the Transportation Research Board, Washington, D.C., 2004.
- Wierwille, W.W., D.A. Keiliszewski, R.J. Hanowski, A.S. Keisler, & E.C.B. Olsen. *Identification and evaluation of driver errors: Task E report – Investigation of critical incidents* (Interim Report for Federal Highway Administration Contract No. DTHF-61-97-C-00051). Washington, DC: Federal Highway Administration, 2000.
- Wierwille, W.W., J.M. Hankey, C.A. Kieliszewski, R.J. Hanowski, A. Medina, S.E. Lee, A.S. Keisler & T.A. Dingus. *Identification and Evaluation of Driver Errors, Task H Report: Final Project Report* (Contract No. DTFH61-97-C-00051). Blacksburg, VA: Virginia Tech Transportation Institute, Blacksburg, 2001.
- Yan, X., E. Radwan, & M. Abdel-Aty. Characteristics of rear-end accidents at signalized intersections using multiple logistic regression models. *Accident Analysis and Prevention*, Vol. 37, 2005, pp. 983-995.
- Yang, David C.Y. & W.G. Najm. *Analysis of Red Light Violation Data Collected from Intersections Equipped with Red Light Photo Enforcement Cameras* (Report No. DOT HS 810 580). Washington, DC: National Highway Traffic Safety Administration, 2006.
- Zabyshny, A.A. & D.R. Ragland. *False Alarms and Human-Machine Warning Systems*. Berkeley, California: University of California, Berkeley, U.C. Berkeley Traffic Safety Center, 2003.

APPENDIX A.

DECELERATION THRESHOLD AND DRIVER REACTION TIME

Introduction

The following study was performed to determine the relationship between brake pedal status and acceleration for drivers approaching a stop-controlled intersection. The goal of the analysis was to predict driver-induced brake status (on/off) based on the vehicle's instantaneous rate of deceleration. This relationship was used to infer brake status using radar data which does not natively contain brake information. The primary measures evaluated in this study included the deceleration level at which the brake was initially pressed and the time from brake press to various pre-determined deceleration levels. The outputs of this task included a threshold below which the brakes will be considered active and the corresponding delay to determine the initial point at which the brakes were applied. This information was used to determine the point of brake activation.

Methods

The 100-car database was mined for the desired braking information (Dingus et al., 2006). This database includes naturalistic continuous in-vehicle data for over 100 participants who drove a personal or leased vehicle for one full year. The parametric data included a variety of kinematic and environmental variables collected at 10 Hz. This parametric data were accompanied by a digital video feed containing images of the driver and vehicle environment collected at 30 Hz. The 100-car database was accessed and analyzed using the VTTI Data Analysis and Reduction Tool.

To extract the relevant data samples, a query was created to identify regions in which drivers would approach stop-controlled intersections. From the 1,400 resulting observations, 10 approaches were selected at random for each of five different vehicle models providing a total of 60 observations. Only intersection approaches that contained straight and flat geometry were considered in the evaluation. The task was to determine the threshold values and time offsets for each of the approaches. The parameters that were collected from the database were:

- Trigger ID for each vehicle
- Vehicle type (e.g. Ford Explorer, Ford Taurus, Chevy Malibu, etc.)
- Time sync and deceleration values at which time the driver initiated braking; similarly, the sync and deceleration values when the car reached -0.05g, -0.075g, -0.10g, -0.12g.
- Time elapsed before abovementioned deceleration values were calculated. This was obtained by subtracting the sync numbers from the start of braking to the sync number when the deceleration reached the specified levels.

Results

Based on the method described above, Table A-1 provides the average deceleration thresholds determined for each vehicle. The table also contains the time it took each vehicle to reach pre-determined deceleration values.

Table A-1. Deceleration thresholds and pre-determined deceleration values from 100-car data.

Threshold deceleration when braking starts(g)	CAR TYPE							
		Chevy Malibu	Toyota Corolla	Toyota Camry	Leased Cavalier	Ford Taurus	Ford Explorer	
	Average	-0.042	-0.040	-0.012	-0.025	-0.052	-0.007	
	Std Dev	0.034	0.038	0.022	0.034	0.022	0.017	
	Min	-0.11	-0.05	-0.04	-0.07	-0.04	-0.04	
	Max	0.01	0.07	0.03	0.04	0.03	0.02	
Time Taken to reach pre-determined deceleration value (seconds)	-0.05g	Average	0.257	0.600	1.450	0.778	1.778	0.840
		Std dev	0.257	0.644	2.514	1.180	1.300	0.609
		Min	0	0	0.1	0	0.6	0.1
		Max	0.6	2.1	8.5	2.9	4	1.7
	-0.075g	Average	1.600	1.230	2.180	1.200	2.289	1.120
		Std dev	3.090	0.979	2.785	1.217	1.274	0.890
		Min	0.1	0.5	0.2	0.1	0.8	0.3
		Max	9.8	3.4	9.3	3.4	4.2	3.1
	-0.10g	Average	1.840	1.730	2.520	1.290	2.844	1.3
		Std dev	3.769	1.636	3.134	1.251	1.455	0.885
		Min	0	0.6	0.3	0.1	1.3	0.5
		Max	12.5	6	10.7	3.7	4.9	3.1
	-0.12g	Average	2.660	2.440	2.810	1.640	3.712	1.460
		Std dev	4.884	2.872	3.252	1.510	1.726	0.916
		Min	0.1	0.7	0.4	0.2	1.7	0.6
		Max	16.2	10.3	11.3	4.1	6.2	3.2

Conclusion

The purpose of this study was to determine 1) the deceleration threshold at which it may be stated that a vehicle is actively braking and 2) the corresponding time from initiation of braking to the selected threshold deceleration value. Rather than selecting an average threshold deceleration value, the 95% value was selected for three primary reasons:

1. The vehicle data collected for this study were measured by radar. As radar does not directly measure acceleration, the acceleration had to be derived from the velocity. As a derivative, any noise inherent in the velocity data were amplified in the acceleration data. Thus, a more strident threshold would minimize early identification of braking.
2. Drivers initiate braking at a variety of initial levels of acceleration. It was not unusual for the acceleration at the brake onset to be positive (i.e., driver accelerating just prior to pressing the brakes). This explains a number of the low average threshold values displayed in the table above. Given the safety application of an ICAS system, it is desirable to select a threshold in which a majority of drivers will have applied the brake.

3. The braking points were determined by identifying the threshold deceleration level and subsequently subtracting the average time elapsed from the onset of braking until the threshold is reached. Thus, the time measure will take into account the differences in the actual brake onset described in the table above.

The overall average threshold value for the 60 observations was -0.015 g with a standard deviation of -0.031 g. Thus, the 95th percentile driver has initiated braking at a threshold of approximately -0.077 g. From the table above, this corresponds closely to the -0.075 g threshold suggesting an average response time value of 1.60 s with a standard deviation of 1.89 s. Thus, the brake onset will be identified at 1.60 s before the vehicle acceleration reaches -0.075 g.

APPENDIX B.

VIOLATION TRIGGER PROCESS

The data needs, identified in the methods section, require the validation of aggressive intersection approaches. In particular, it was desired to identify approaches in which drivers either committed a violation or performed a stop or near stop in such a way that their behavior would be difficult for the algorithm to discriminate. Thus, the trigger development focused on finding a metric that would identify aggressive stopping behaviors. In addition, a second important criterion was to identify a trigger that functioned on sparsely populated data. As discussed in the methods and results section, the radar used for this study did not reliably return a measurement for every collection frame. Thus, the trigger could not operate at a single location as it would miss vehicles that were not reported by the radar in that location.

Several possible trigger strategies were considered with the goal of identifying violations. In particular, triggers evaluating the stopbar speed, average deceleration, peak deceleration, and minimum speed were assessed. However, a measure of the average deceleration required to stop at the stopbar was identified as the appropriate triggering variable.

The required deceleration parameter (RDP) is a calculated value computed at each frame of data. RDP is the kinematic relationship between instantaneous velocity and distance to the stopbar as described by Equation B-1.

Equation B-1. Required Deceleration Parameter.

$$RDP = \frac{V^2}{2 * R * g}$$

Where: V = Instantaneous velocity
R = Instantaneous range from stopbar
G = Gravitational constant

RDP has several advantageous characteristics that make it a particularly good metric for triggering the data reduction. First, RDP uses velocity and range which are frequently measured by the radar. This has the advantage of not relying on a derived measure such as acceleration which is prone to amplified noise. Second, RDP is easily interpreted in the context of stopping behavior. A driver that performs an aggressive intersection approach would have to brake hard to stop before the stopbar. Thus, this driver would also exhibit a high RDP. If that driver did not stop, the RDP would likely exceed the capabilities of the vehicle as the stopbar neared. On the other hand, a conservative driver would exhibit a low required deceleration. Furthermore, unlike evaluating velocity at a particular point, RDP can be evaluated at any point along the intersection approach and remain valid. For instance, a driver who initially approached the intersection at a high speed followed by a hard brake will be missed by a stopbar speed trigger. A RDP trigger, on the other hand, will catch the high deceleration that was required to slow that vehicle upstream of the stopbar.

To develop an effective trigger, there are some additional criteria that must be considered during the RDP computation. First, RDP tends towards infinity as the vehicle nears the stopbar. This tendency will have a negative impact on the sensitivity of the measure to segregate stopping behaviors. Investigation of RDP indicated that the tendency does not exist for typical approaches until the vehicle was within 1 m from the stopbar. Thus, RDP was only evaluated at distances greater than 1 m. In addition, RDP is a continuous measure existing over the entire time period in which the vehicle was tracked. Thus, to enable a simple trigger comparison the maximum RDP was extrapolated and compared to a trigger threshold. The maximum RDP over the entire vehicle approach represents the highest rate at which the driver would have needed to stop. For a violating driver the maximum RDP will exist near the stopbar. However, for an aggressive driver that stops rapidly, the maximum RDP may exist at some point upstream; prior to their high deceleration stop. This makes the trigger particularly advantageous as it is sensitive to violators and aggressive drivers as these are the groups of primary interest for the algorithm.

With the triggering metric identified, the next step was to set a threshold for separating the reduction events from the non-reduction events. To identify the trigger threshold, the maximum RDP was calculated for each vehicle approach in the dataset. The distribution of maximum RDP was then analyzed to determine an appropriate threshold. This analysis will be described through the subsequent figures.

First, in light of time and budget constraints it was important to select a threshold that provided a reasonable number of reducible events. The number of events was evaluated as a function of the threshold selected (Figure). As the RDP threshold is lowered, the number of events rises sharply. Based on the constraints, a cap was set at 10,000 events, with a desire to reduce the number of events to a lower number if possible. This criterion suggested a threshold in the region of 0.8 g to 1.2 g. While initially these values may appear high, they are actually well within the region that provides the data of interest. Additional details will be provided with the figures below as well as during the stopping behavior analysis discussed in the results section.

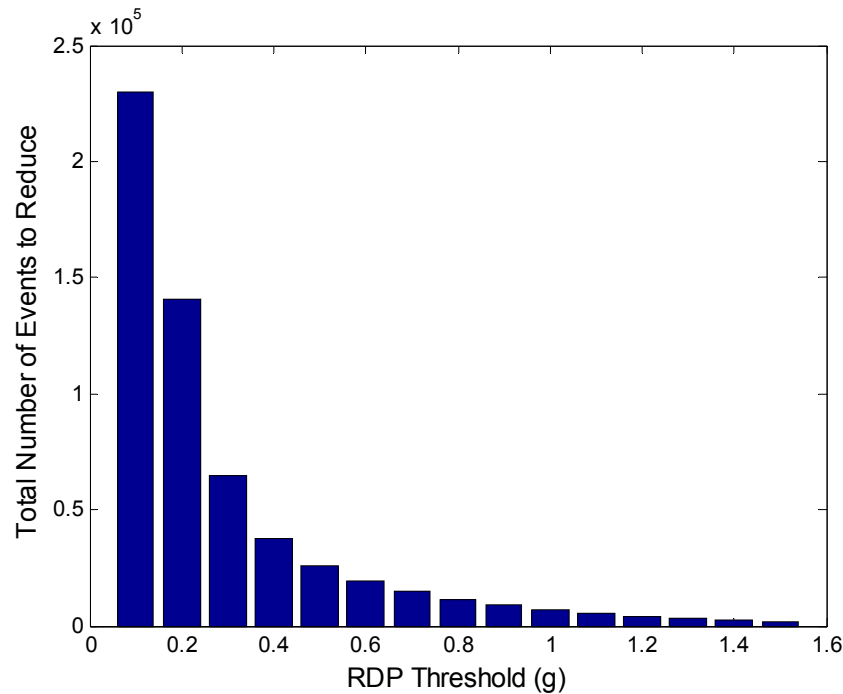


Figure B-1. Number of events that would be reduced as a function of the RDP threshold selected.

In conjunction with the results above, the average stopbar speed as a function of the RDP threshold was also considered (Figure B-2). Stopbar speed provides an indication of the severity of a violation that is readily understood. Furthermore, results discussed during the introduction indicated that a stopbar speed of 4.47 m/s (10 mph) appeared to separate intentional “rolling stops” from a violation resulting from some form of inattention. Considering the stopbar speed, it appeared that a RDP threshold of approximately 1.2 g was appropriate as it corresponded. To ensure that these approaches were obtained a cutoff closer to 1 g looked attractive.

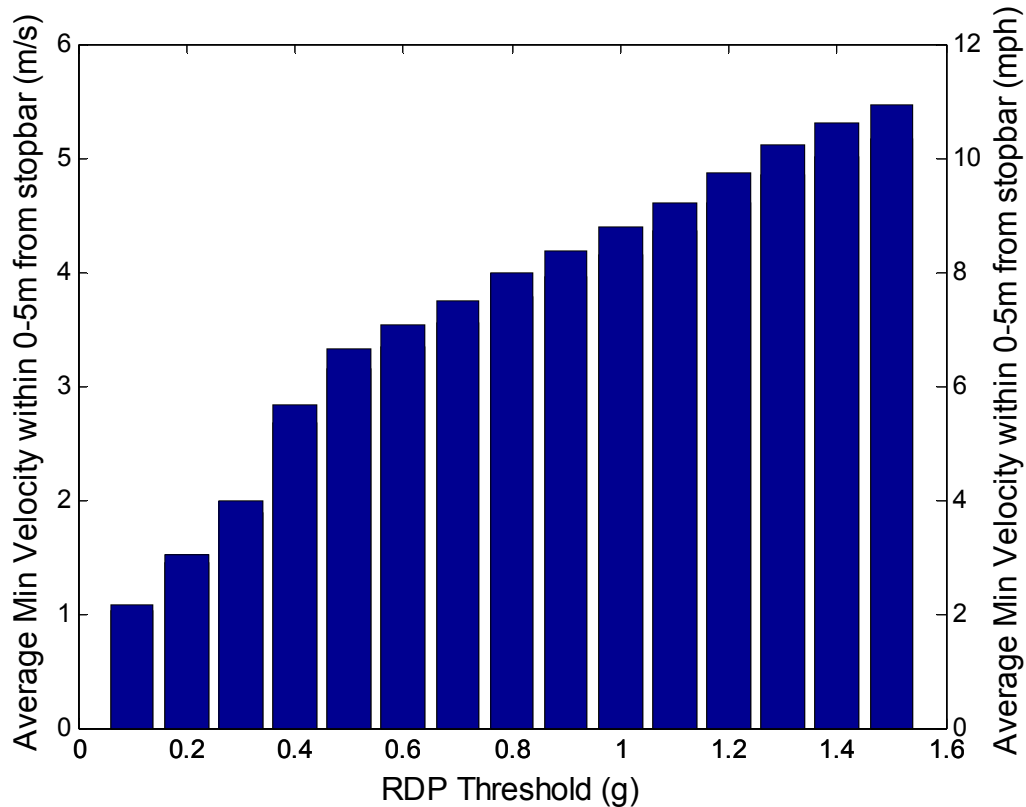


Figure B-2. Average minimum velocity for vehicles near the stopbar.

Finally, the empirical distribution of RDP across the entire sample of drivers was evaluated (Figure B-3). With warning systems, such as the ICAS, researchers are interested primarily in the tail of a distribution which represents uncommon behavior; in this context, these behaviors are dangerous violators. Considering the distribution, a 1 g RDP threshold addresses 5% of the sample. Focusing on the top 5% of the population provides a convenient and logical cutoff for investigation. Furthermore, this cutoff should be conservative as ICAS developers are aiming for warning rates significantly lower than 5% to avoid excessive nuisance alarms. Setting the reduction threshold for RDP at 1 g provides a conservative cutoff that will include nearly all of the severe violations and a sample of the more aggressive intersection approaches.

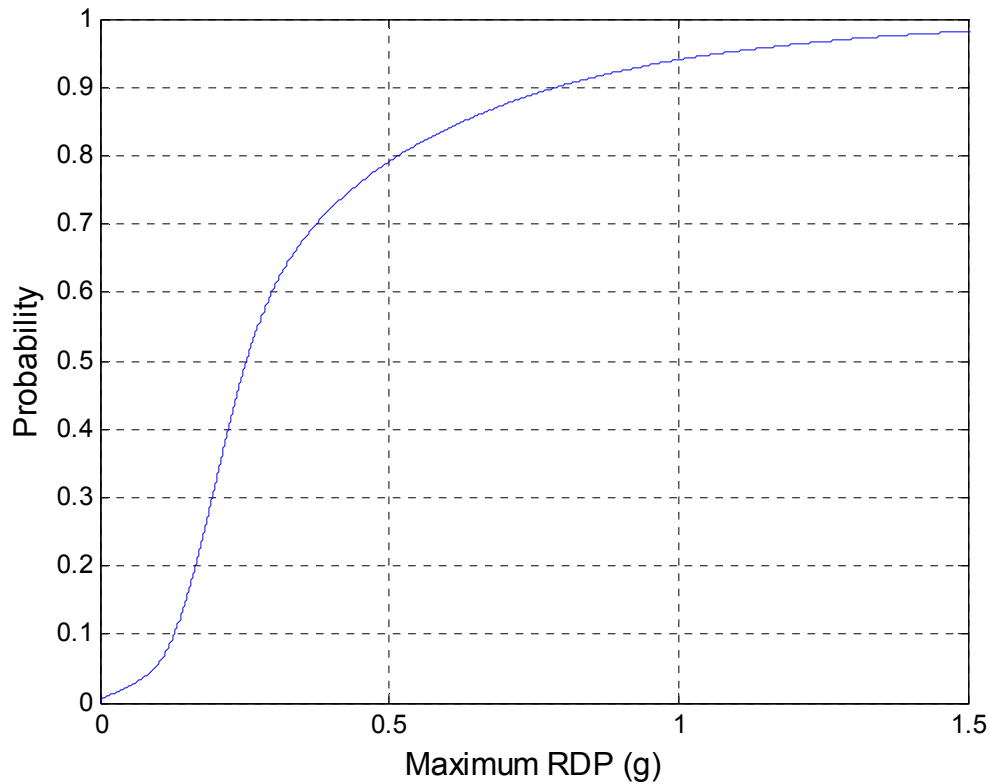


Figure B-3. Empirical cumulative probability distribution for RDP.

Based on the data presented above, a trigger threshold of 1 g was selected for identifying the approaches for reduction. This resulted in the reduction of 6,171 events. Considering that a typical braking maneuver occurs around 0.3 g, the 1g threshold may appear larger than one might expect. There are a few reasons for this. First, as discussed in the results section, most drivers do not come to a complete stop. Thus, even the drivers performing a slow rolling stop will exhibit an elevated RDP; however, these slow rolling cases occur with such frequency that it would be a mistake to issue a warning. Such a low threshold warning system would have an impact on driver acceptance and would not be implemented by the automotive manufacturers. Furthermore, the purpose of the ICAS system is to mitigate crashes through a violation warning. In general, drivers that perform a slow rolling stop are as attentive as a driver that completely stops, suggesting no increase in crash likelihood.

APPENDIX C. STOP SIGN POST-PROCESSING AND RADAR-INDUCED ERROR

Table C-1. Stop sign post processing and radar-induced error.

		Range (m)	Range (ft)	Speed (m/s)	Speed (mph)	Acceleration (g)
45 mph	Mean	0.506	1.660	0.164	0.367	-0.013
	Standard Deviation	2.154	7.067	0.245	0.548	0.054
	Min	-6.579	-21.585	-0.095	-0.213	-0.318
	Max	2.446	8.025	1.520	3.400	0.076
35 mph	Mean	1.061	3.481	0.076	0.170	-0.003
	Standard Deviation	2.200	7.218	0.159	0.356	0.023
	Min	-6.207	-20.364	-0.085	-0.190	-0.140
	Max	4.247	13.934	0.934	2.089	0.047
25 mph	Mean	0.833	2.733	0.034	0.076	0.000
	Standard Deviation	2.505	8.219	0.102	0.228	0.014
	Min	-6.292	-20.643	-0.278	-0.662	-0.115
	Max	5.562	18.248	0.690	1.543	0.029
Overall	Mean	0.829	2.720	0.062	0.139	-0.003
	Standard Deviation	2.401	7.877	0.151	0.338	0.026
	Min	-6.579	-21.585	-0.278	-0.622	-0.318
	Max	5.562	18.248	1.520	3.400	0.076

APPENDIX D.
SUMMARY OF RESULTS FOR THE CLUSTER ANALYSIS PERFORMED TO
DETERMINE THE APPROPRIATE PARTITIONS FOR UNSIGNALIZED
INTERSECTION DRIVER APPROACH BEHAVIOR

Table D-1. Summary results from the performed cluster analysis

		Within Cluster Information					Overall Cluster Information		
		Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Clusters	Silhouette Width	Sum of Differences
Cluster	Count	30581	42				2	0.9924	1896.82
	Centroid	0.34418	4.1308						
	Standard Deviation	0.23626	2.152						
	Min	0.000339	2.4224						
	Max	2.2254	13.1572						
	Silhouette Width	0.99328	0.35142						
Cluster	Count	25083	5499	41			3	0.8453	740.833
	Centroid	0.25316	0.75972	4.1724					
	Standard Deviation	0.098582	0.23753	2.1615					
	Min	0.000339	0.50647	2.4885					
	Max	0.506337	2.42239	13.1572					
	Silhouette Width	0.9221	0.49996	0.1854					
Cluster	Count	24995	5573	51	4		4	0.84656	590.728
	Centroid	0.25228	0.75228	3.1587	9.8971				
	Standard Deviation	0.09762	0.22803	0.93244	2.5392				
	Min	0.000339	0.5023	1.9724	7.2506				
	Max	0.502221	1.93996	5.15076	13.1572				
	Silhouette Width	0.92056	0.51739	0.56914	0.62755				
Cluster	Count	21494	6784	2303	38	4	5	0.79096	576.093
	Centroid	0.22356	0.51428	0.96885	3.5238	9.8971			
	Standard Deviation	0.070347	0.10409	0.22779	0.79739	2.5392			
	Min	0.000339	0.36893	0.74161	2.4224	7.2506			
	Max	0.368885	0.741524	2.2254	5.15076	13.1572			
	Silhouette Width	0.8716	0.64893	0.45761	0.76373	0.56416			

APPENDIX E.
SUMMARY TABLE OF RESULTS FOR THE GENERALIZED EXTREME VALUE
DISTRIBUTION FITS AND THE KOLMOGOROV-SMIRNOV GOODNESS OF FIT
TEST

Table E-1. GEV summary

Generalized Extremely Value Fit Information							
Overall Max RDP GEV Fit	Results		Centroid	0.3494			
			95%	0.8017			
			5%	0.1142			
	Quality		H	0			
			p	0.1946			
			KSTAT	0.1497			
			CV	0.1884			
	Parameters		k	0.2787			
			k_95% LC	0.2701			
			k_95% UC	0.2874			
			Sigma	0.1235			
			Sigma_95% LC	0.1223			
			Sigma_95% UC	0.1248			
			Mu	0.2309			
			Mu_95% LC	0.2294			
			Mu_95% UC	0.2325			
Four Partition Within Cluster GEV Fit	Results		Centroid	0.25228	0.75228	3.1587	9.8971
			95%	0.426	1.2885	5.0285	14.0513
			5%	0.11	0.5187	1.9494	6.8464
	Quality		H	0	0	0	0
			p	0.4678	0.7098	0.1571	0.9994
			KSTAT	0.1177	0.0973	0.2152	0.1665
			CV	0.1884	0.1884	0.2591	0.6239
	Parameters		k	-0.1249	0.3677	0.0674	-0.0279
			k_95% LC	-0.1343	0.3339	-0.3048	-1.6251
			k_95% UC	-0.1155	0.4016	0.4396	1.5693
			Sigma	0.0864	0.1224	0.7085	1.8173
			Sigma_95% LC	0.0855	0.1189	0.5351	0.6093
			Sigma_95% UC	0.0872	0.1259	0.9382	5.4205
			Mu	0.2116	0.6292	2.6987	8.8712
			Mu_95% LC	0.2104	0.6252	2.4492	6.4333
			Mu_95% UC	0.2128	0.6332	2.9481	11.3091

APPENDIX F. STOPBAR VELOCITY

Table F-1. GEV fit for Cluster 1.

GEV Fit Information			
			Cluster 1
Fit	Quality	H	0
		p	0.2708
		KSTAT	0.0444
		CV	0.0604
	Parameters	k	-0.1392
		k_95% LC	-0.1506
		k_95% UC	-0.1279
		Sigma	0.9756
		Sigma_95% LC	0.9653
		Sigma_95% UC	0.986
		Mu	1.203
		Mu_95% LC	1.189
		Mu_95% UC	1.2171

Table F-2. Normal fit for Cluster 2 through Cluster 4.

Normal Fit Information					
			Cluster 2	Cluster 3	Cluster 4
Fit	Quality	H	0	0	0
		p	0.1894	0.796	0.9526
		KSTAT	0.05	0.089	0.2373
		CV	0.0626	0.1866	0.6239
	Parameters	Mu	3.8542	11.6836	20.7889
		Mu_95% LC	3.8168	11.0576	14.6622
		Mu_95% UC	3.8917	12.3096	26.9155
		Sigma	1.4254	2.2258	3.8503
		Sigma_95% LC	1.3994	1.8623	2.1811
		Sigma_95% UC	1.4523	2.7668	14.356

APPENDIX G. BRAKE ONSET

Table G-1. Brake onset.

Normal Fit Information				
			Range	TTI
Fit	Quality	H	0	0
		p	0.7886	0.0761
		KSTAT	0.0295	0.0569
		CV	0.0614	0.0604
	Parameters	Mu	115.6876	6.4111
		Mu_95% LC	115.1761	6.3985
		Mu_95% UC	116.1991	6.4237
		Sigma	26.2639	0.9767
		Sigma_95% LC	25.9071	0.9679
		Sigma_95% UC	26.6306	0.9858