UMTRI-96-18

A Bridging Analysis for Estimating the Benefits of Active Safety Technologies

Task One under Contract No. DTNH22-93-D-07000
Crash Avoidance Research Technology Support--Simulation Models

Final Report

by

Kenneth L. Campbell
Hans C. Joksch
Paul E. Green

April 1996

The University of Michigan
Transportation Research Institute
2901 Baxter Road
Ann Arbor, Michigan 48109-2150

This research was supported by the National Highway Traffic Safety Administration (NHTSA), U.S. Department of Transportation, under Contract No. DTNH22-93-D-07000. The opinions, findings, and recommendations contained herein are those of the authors, and do not necessarily represent those of the NHTSA.

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A Bridging Analysis for Estimating the Benefits of Active Safety Technologies		5. Report Date April 1996
		6. Performing Organization Code
7. Authors Kenneth L. Campbell, Hans C. Joksch, and Paul E. Green		8. Performing Organization Report No. UMTRI-96-18
9. Performing Organization Name and Address The University of Michigan Transportation Research Institute 2901 Baxter Road Ann Arbor, Michigan 48109-2150		10. Work Unit No.
		11. Contract or Grant No. DTNH22-93-D-07000
12. Sponsoring Agency Name and Address Office of Crash Avoidance Research National Highway Traffic Safety Administration 400 Seventh Street, S.W. Washington, D.C. 20590		13. Type of Report and Period Covered Task One Final Report 9/1/94 - 10/31/95
		14. Sponsoring Agency Code NHTSA/NRD-51
15. Supplementary Notes		

16. Abstract

The objective of this project was to develop a methodology for estimating the safety benefits of active safety technology. Accident data, the traditional safety measure, are not sufficient because too great an exposure is required to achieve adequate sample sizes and the events leading to the collision often cannot be determined. Consequently, a method relating pre-collision events to safety is needed. The Traffic Conflicts Technique developed in the United States, Europe, and Canada in the mid-70's to address intersection safety is such a method. Traffic conflicts are near-collision events that are recorded by trained observers.

This report describes an extension of the traffic conflicts technique to incorporate continuous measurement of crash margin measures, such as time to collision. This approach asserts that the probability of a collision can be estimated from the frequency of small crash margins. Extreme value theory is offered as a robust statistical method to compare probability levels in the tails of observed distributions of crash margins. The approach is illustrated using headway data from the FOCAS project. A final section describes future work that must be done in order to support application of this methodology to the estimation of the safety benefits of advanced technology.

17. Key Words	18. Distribution Statement
Active safety technology, safety benefits, traffic conflicts, near-collisions, extreme value theory	

19. Security Classification (of this report)	20. Security Classification (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	32	

SI* (MODERN METRIC) CONVERSION FACTORS

APPROXIMATE CONVERSIONS TO SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
in	inches	25.4	millimeters	mm
ft	feet	0.305	meters	m
yd	yards	0.914	meters	m
mi	miles	1.61	kilometers	km
		AREA		
in^2	square inches	645.2	square millimeters	mm^2
ft^2	square feet	0.093	square meters	m^2
yd^2	square yards	0.836	square meters	m^2
ac	acres	0.405	hectares	ha
mi^2	square miles	2.59	square kilometers	km^2
		VOLUME		
fl oz	fluid ounces	29.57	milliliters	ml
gal	gallons	3.785	liters	l
ft^3	cubic feet	0.028	cubic meters	m^3
yd^3	cubic yards	0.765	cubic meters	m^3

NOTE: Volumes greater than 1000 l shall be shown in m^3.

Symbol	When You Know	Multiply By	To Find	Symbol
		MASS		
oz	ounces	28.35	grams	g
lb	pounds	0.454	kilograms	kg
T	short tons (2000 lb)	0.907	megagrams	Mg
		TEMPERATURE (exact)		
°F	Fahrenheit temperature	5(F-32)/9 or (F-32)/1.8	Celcius temperature	°C
		ILLUMINATION		
fc	foot-candles	10.76	lux	l
fl	foot-Lamberts	3.426	candela/m^2	cd/m^2
		FORCE and PRESSURE or STRESS		
lbf	poundforce	4.45	newtons	N
psi	poundforce per square inch	6.89	kilopascals	kPa

APPROXIMATE CONVERSIONS FROM SI UNITS

Symbol	When You Know	Multiply By	To Find	Symbol
		LENGTH		
mm	millimeters	0.039	inches	in
m	meters	3.28	feet	ft
m	meters	1.09	yards	yd
km	kilometers	0.621	miles	mi
		AREA		
mm^2	square millimeters	0.0016	square inches	in^2
m^2	square meters	10.764	square feet	ft^2
m^2	square meters	1.195	square yards	ac
ha	hectares	2.47	acres	mi^2
km^2	square kilometers	0.386	square miles	
		VOLUME		
ml	milliliters	0.034	fluid ounces	fl oz
l	liters	0.264	gallons	gal
m^3	cubic meters	35.71	cubic feet	ft^3
m^3	cubic meters	1.307	cubic yards	yd^3
		MASS		
g	grams	0.035	ounces	oz
kg	kilograms	2.202	pounds	lb
Mg	megagrams	1.103	short tons (2000 lb)	T
		TEMPERATURE (exact)		
°C	Celcius temperature	1.8C + 32	Fahrenheit temperature	°F
		ILLUMINATION		
lx	lux	0.0929	foot-candles	fc
cd/m^2	candela/m^2	0.2919	foot-Lamberts	fl
		FORCE and PRESSURE or STRESS		
N	newtons	0.225	poundforce	lbf
kPa	kilopascals	0.145	poundforce per square inch	psi

* SI is the symbol for the International System of Units. Appropriate rounding should be made to comply with Section 4 of ASTM E380.

(Revised August 1992)

TABLE OF CONTENTS

TABLES

FIGURES

Introduction

Collision avoidance work has traditionally been supported by accident data. The study of collision avoidance necessarily focuses on the pre-collision events. However, it has been observed for some time, that accident investigation is limited in its ability to provide information on events that occurred before the collision. [1, 2, 3, 4] These authors have argued for the need to identify a pre-collision event, such as a near-collision, to better study traffic events that precipitate collisions. A similar approach in the human factors area is the study of critical incidents.

This view has been repeated more recently in the context of advanced collision avoidance technology. [1, 5, 6] The development of collision avoidance technology requires more detailed information on pre-collision situations and events. Just as important is the need to evaluate alternative technologies before implementation. The objective of this work is to develop an analytic framework for linking non-collision traffic/driving information with accidents. How can one extrapolate to the probability of collision from measurements, or distributions, collected over the range of "normal" (accident free) travel? Such a framework could support estimation of the change in crash frequency (and severity) that would result from the use of active safety technology.

The traffic conflicts technique (TCT) is a method first proposed in 1968 by Harris and Perkins of General Motors[2] as a way to study traffic risks. Over the past two decades, the technique has been developed in the United States, Canada, and Europe as a way to study traffic flow problems at individual intersections. Originally, traffic conflicts were observed collision avoidance maneuvers. This paper describes an extension of the traffic conflicts technique for estimation of the safety benefits of collision avoidance technology.

The approach relies on continuous measurement of crash margins that will characterize the entire distribution. The method assumes that the frequency of small crash margins (conflicts) can be related to the probability of an accident. This approach follows from the literature on traffic conflicts. This application requires analytic techniques for extrapolating beyond the observed crash margins in the tails of the distribution and estimating the proportion exceeding the minimum threshold defining a conflict. This problem is addressed using extreme value theory.

An overview of the approach is provided in the next section. This is followed by a brief description and history of the traffic conflicts technique. Extreme value theory is presented next as a method to address some of the analytic issues raised. The method is illustrated using data from the FOCAS project.[7] A final section discusses future work.

Background

A premise for this development is that continuous measurement of pertinent variables describing driver actions, vehicle motion, and the roadway is available. Several approaches to the collection of such information are under development. For example, in the NHTSA forward crash avoidance systems project (FOCAS)[7] at UMTRI, an instrumented vehicle records extensive information on the driver, including steering, throttle, and braking control, as well as information from the headway sensor. Also under development by NHTSA is a data acquisition system for collision avoidance research (DASCAR). This recording system would be appropriate for installation in larger numbers of vehicles, and would be much less intrusive. A comprehensive set of variables would be monitored, and a time history retained in the event of a collision, or other critical event.

At UMTRI, an instrumentation package to be deployed at fixed roadway sites is under development, called the Vehicle Motion Environment (VME). An overhead sensor would record a complete history of the motions of individual vehicles in the traffic stream as they moved through the field. The advantage of the vehicle-based instrumentation packages are that more complete driver information can be obtained. Information on surrounding vehicles is limited to those in the range of the available vehicle sensors. The VME will provide a complete record of the relative position and motion of all vehicles while they are in the field covered by the sensors, but will only be able to infer driver actions from the path of the vehicle. In combination, these new data systems offer the potential to study vehicles and drivers in actual traffic situations at a level detail and quantity not previously available. Thus, a premise for this work is the presumption that continuously recorded data on vehicle motions, or other pertinent variables, will be available.

The challenge is to develop an analytic approach to link this information to safety, where safety is ultimately reflected in the accident experience. Since we wish to carry out the evaluation before sufficient accident experience has accumulated, one issue is to identify performance measures that will be available and are related to the risk of accidents. Such performance measures will provide the basis for the safety evaluation, and the bridge to any accident projections. Dingus[5] discusses the problems in selecting such performance measures, and the difficulty in establishing the relationship to safety. Two broad classes of measures are described, human performance measures and vehicle performance measures. Dingus characterizes the vehicle performances measures as the relative positions and motions as determined from "first principles of physics." Measures based on the relative motions and positions of the vehicles, call them crash margin measures, are preferable because they can be shown to be related to the risk, or probability, of a collision.

This development only applies to two-vehicle collisions, for now. Possible extensions, for example, to pedestrian or bicycle collisions on the roadway would require the ability to identify such targets. The same problem is presented for any other object in the roadway (that is not a moving vehicle). Single-vehicle loss of control on the roadway accidents would pose special problems related to the lateral acceleration and stability of the vehicle. Any collision off the roadway could only be treated as a lane departure accident. Previous human factors work has looked at time to lane departure in relation to the driving task, so this might provide a suitable crash margin measure. Any unintended lane departure might be considered an "accident" regardless of whether damage or injury resulted. An overview of the approach for two-vehicle collisions follows.

Overview

Figure 1 shows a block diagram of the basic elements of the approach. The three blocks are Events (opportunities), Conflicts (the subset of events with the crash margin less than some threshold), and Accidents (only the events where the crash margin goes to zero). Traffic conflicts are shown as the link between normal driving at the bottom and accidents at the top of the diagram. The crash margin measure is a continuos variable that can be related to the probability of collision, and a threshold value of the measure is the definition of a conflict. Physical measures of the relative kinematic motions of the vehicles are preferred as the crash margin measure because collisions are explicitly identified (as crash margin equal zero).

Each block is partitioned into the different collision types. Each collision type must be related to corresponding Conflict and Event types. The relationship of traffic events, conflicts, and accidents may be different for each type, and even the crash margin may be different. The accidents and conflicts both are graded from mild to severe, reflecting a continuum of severity measured by the crash margin.

Strictly speaking, at any non-zero crash margin, the probability of collision *at that instant* is zero, and goes to 1 when the crash margin gets to zero. We already collect crash margin zero data; it's called accidents. The objective is to find traffic events that are not accidents, but can be used to predict changes in the probability of an accident.

To do this, an intermediate unit will be defined as a traffic conflict. If the collision type of interest is striking the lead vehicle in the rear, then the relevant traffic events are the periods of time (or travel) when there is a vehicle in front. When no vehicle is in front, the probability of a striking a vehicle in the rear is zero, and that travel poses *no* risk of rearend accidents. To make the definition operational, assume some arbitrary termination of the "event" when the crash margin exceeds some value, perhaps the range of the headway sensor. Rather than waiting for only those events where the crash margin goes to zero (accident), we choose to count events when the crash margin is less than some non-zero, but small level, such as 0.5 sec. or 1.0 sec. These traffic events with small crash margins are traffic conflicts. The assertion is that the proportion of such traffic events that include a zero crash margin (accidents) will be larger as the threshold is lowered. The number of collisions stays the same, but we count fewer and fewer of the non-collision traffic events as the crash margin is lowered.

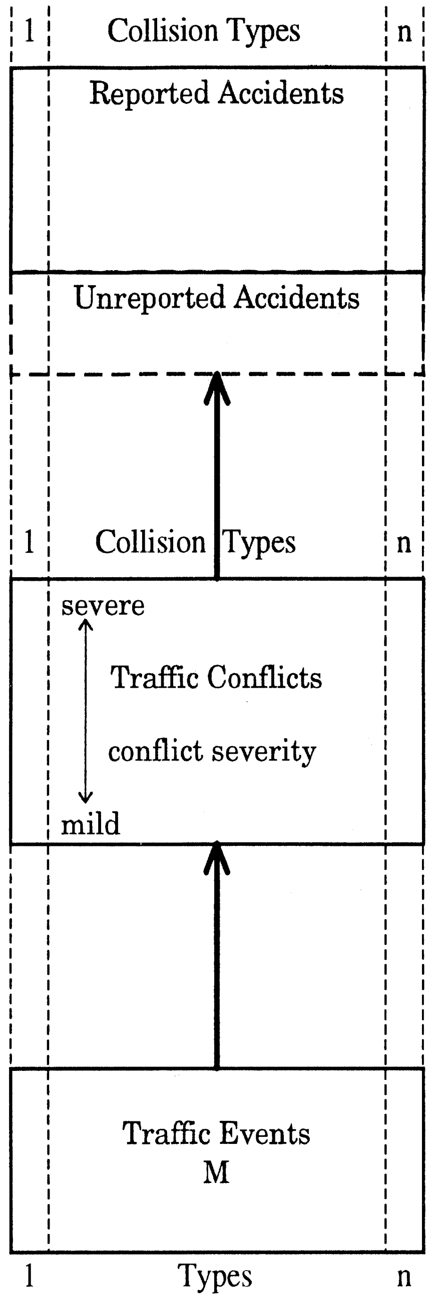


Figure 1. Block Diagram

Basically, the probability of a collision given a traffic event (the presence of another vehicle) is split into two conditional probabilities: the probability of a conflict given an event, and the probability of a collision given a conflict. The definition of a conflict is a traffic event in which the crash margin is less than some threshold value. The "opportunities" for a collision are all those events where a lead vehicle was present (or the time or mileage of such events).

The transition to zero crash margin may not be a continuous distribution. The distributions from normal driving may not include zero when extrapolated. This would imply that intervening events must act to precipitate the collision, loss of crash margin. However, probabilistic relationships can still be developed for the occurrence of accidents.

Burgett[8] describes a formulation for calculating the safety impact of collision avoidance technology as follows:

$$\mathbf{E = L(v - w) \bullet M / N}$$

where:

- E estimated effectiveness as the fraction of collisions remaining after implementation of the ITS system
- **L** relates the probability of near-collisions (conflicts) given the exposure vector M, to the probability of accidents
- v is the probability of conflicts for each collision type without the ITS system
- w is the probability of conflicts for each collision type with the ITS system
- **M** is the vector of exposure opportunities for each collision type
- N is the baseline number of collisions occurring without the system in place

The exposure vector, **M**, for each collision type is shown in the bottom box of Figure 1 as the traffic events corresponding to each collision (conflict) type. The vectors **v** and **w** would be estimated from the distribution of the relevant crash margin for each collision type, observed with and without the ITS system. L is the relationship between the distribution of the relevant crash margin measure and the probability of collision. The functional form and independent variables in this relationship may be different for different collision types.

Traffic Conflicts Technique

Background. In an effort to find a paradigm for the bridging analysis in the literature, a review of the traffic conflicts technique was carried out. The traffic conflicts technique (TCT) was developed in the late sixties by two engineers at the General Motor Research Laboratories, Joe Harris and Stuart Perkins.[2] A traffic conflict was defined as a potential accident situation. Harris and Perkins identified 24 specific conflict situations, each related to a particular collision type. TCT has primarily been used as a civil engineering tool to evaluate individual intersections. The primary advantages of TCT cited in the early literature were that conflicts occurred much more frequently than accidents, so a complete evaluation of a site could be conducted in a few days, as compared to waiting a few years for a sufficient accident record to accumulate, and the conflict method provided a clearer picture of the initial causes of the incidents, something often lacking from accident reports.

TCT was pursued in the United States through the 70's with support from FHWA. The primary issues then are pertinent to the problem at hand. What is the definition of a traffic conflict, and can the relationship between conflicts and accidents be demonstrated? Evaluations of the method published in the mid-70's by Hauer[9] and Glennon[10] described these problems. The available data linking conflicts with accidents was not too convincing. The ratio of conflicts to accidents varied widely, anywhere from a few thousand to several hundred thousand conflicts per accident. The definition of a traffic conflict was not producing consistent results from one application to the next.

In the late-70's, TCT was picked up by the Europeans. Starting in 1977, international conferences were hosted in Europe. At the first meeting in Oslo, a common definition of a traffic conflict was adopted:

> A traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is risk of collision if their movements remain unchanged.

These conferences frequently focused on a comparison of findings when teams from as many as 10 different counties would study the same intersection and compare the conflicts identified. The number of conflicts counted varied by as much as four to one across the teams.[11] One focus of the European work was to distinguish the severity of the conflict. When conflicts were classified by severity, much of the team to team differences were shown to be a consequence of the threshold level at which a conflict was recorded. The recording of "severe" conflicts was much more consistent. In 1976, Older and Spicer[12] developed a grading of five levels of conflict severity, ranging from precautionary braking to an emergency action followed by a collision.

The resolution of the conflict severity issue offers promise for the proposed bridging analysis. Video recording was used as a way to make more objective comparisons of conflicts recorded by different teams. Subsequent film analysis provided estimates of various objective measures such as the time to collision. In 1972, Hayward at Pennsylvania State University suggested 1.0 second as a good threshold for car to car conflicts.[13] A landmark study of this issue in Europe was a ten team comparison, complete with video recording and film analysis, conducted in Malmo, Sweden.[14] Based on a quantitative analysis of various objective measures from the video data, van der Horst reported in 1984[15] that time to collision was the major objective measure explaining the severity dimension. The more extensive use of video analysis has produced data on the number of conflicts in relation to the time to collision for specific conflict types and traffic situations. The European literature includes conflicts with times to collision (TTC) as short as 0.5 second. Van der Horst[16] observed a critical TTC of 1.6s and a minimum TTC of 1.1s in a closed course study of braking.

This finding supports the use of kinematic measures such as time to collision as the metric for a continuum that extends from normal driving to accidents. Near-collision events can be located on that continuum in terms of the time to collision when they occur. In this way, kinematic measures such as time to collision provide a way to relate quantitative data from disparate sources such as the VME and DASCAR.

Much of this material is taken from an excellent summary of traffic conflicts by van der Horst in 1990.[17] The Europeans recognized that focusing on severe conflicts produced a more consistently defined measure, and a measure that was more readily linked to accidents. However, they also concluded that video recording and time consuming film analysis was necessary to accurately identify severe conflicts. In the United States, these developments were seen as serious disadvantages for the evaluation of individual intersections. The requirement for video analysis made the process much more expensive and time consuming, and the limitation to severe conflicts substantially increased the observation period necessary for adequate sample sizes.

Work in the U.S. continued with the broader definitions of a traffic conflict.[18] A recognition that as the time to collision is longer, there are more opportunities for intervening events to cloud the relationship with accidents, led to more sophisticated analyses of this issue. In the mid-80's, the Midwest Research Institute published results from an FHWA-sponsored effort that provided more convincing evidence of the relationship of even minor conflicts to accidents.[19, 20] Example values taken from Glauz, et al, 1985 are presented in Table 1 to illustrate typical values of the ratio of conflicts to accidents, based on the work in the U.S.

Table 1
Conflict/Accident Ratio Statistics

Type of Conflict and Intersection Class	Number of Intersections	Mean Conflict/Accident
Left turn same direction		
Unsignalized medium volume	10	66,560
All same direction		
Signalized high volume	12	700,280
Signalized medium volume	14	375,516
Opposing left turn		
Signalized high volume	12	1,490
Signalized medium volume	14	5,408
Unsignalized medium volume	10	4,706
Through cross traffic		
Unsignalized medium vol.	10	1,359
Unsignalized low volume	9	2,044

Glauz, et al, TRB 1026, 1985

Discussion of Issues. TCT appears to be a successful example in which a pre-collision measure (a conflict) has been related to the probability of collision for specific situations occurring at intersections. If one can hypothesize a relationship between such a crash margin measure and the probability of collision, then the quantification of this relationship provides a common link between the normal traffic environment, near collisions, and actual collisions. Among other things, such a link would provide a basis to relate changes in the frequency distribution of the crash margin to changes in the collision experience. With TCT as a model, the approach would be to develop a definition for a pre-collision "critical incident" that is based on objective, measurable parameters. Then the more traditional approaches to human factors based on critical incidents could be extended to collision avoidance.

The first issue is the question of what the most appropriate measure is. Dingus[5] discusses the problem of converting practical performance measures to measures of the safety impact. The general problem is that safety cannot be practically and directly measured by empirical means, and performance measures are not directly related to safety. He observes that there is no single safety measure, and that the selection of measures generally will vary with the problem. In this case, each conflict type is potentially a separate problem that may require a unique safety measure. Conflict types correspond, in turn, to specific accident types. Dingus considers two broad categories of performance measures as potential safety measures: human performance measures, and measures of the kinematic motion of the vehicle. Human performance measures have the advantage that they quantify the changes in behavior of the driver that are responsible for the safety impact. However, Dingus observes that the human performance measures are " several theoretical steps and a complex relationship away from a direct safety relationship."[5] Also, human measures would not encompass automated countermeasures such as throttle changes or braking initiated by an automatic control system.

Consequently, kinematic measures of the vehicle motion, such as time to collision, seem preferable. One advantage is that collisions are explicitly included. When the range, or time to collision, reaches zero, a collision results by definition. While this formulation does not provide a guarantee of continuity, it does provide a basis for probabilistic statements about the relationship of the performance measure and accidents. Such a measure incorporates the "conflict severity" dimension, and the definition of a "conflict" is simply the specification of a threshold value. Some desirable attributes of these measures include:

- objective
- continuous variable
- kinematic validity
- applicable to different collision situations if possible
- applicable to different measurement methods
    - VME--conflict filter
    - DASCAR
    - Instrumented vehicle
    - Driving simulators
    - GES/CDS (link to collision situations)
- human factors validity

TCT applications in the literature are focused on conflicts between two or more vehicles at intersections. The approach has also been applied to pedestrian and bicycle conflicts. The 12 basic conflict types[21] (listed in Table 2) extend to most non-intersection conflicts between road users, but an extension to certain single-vehicle accidents, such as ran off the road, rollover, and impacts with fixed or non-fixed objects, may not be straight forward. The time to lane departure has been used in some human factors studies. While time to collision is the leading candidate from the literature, others have been suggested. Possible kinematic measures are:

range, R
available reaction time, R/V, (V=velocity)
headway time, R/Vp, (Vp=lead vehicle velocity)
time to collision, Range/(Range Rate)
minimum deceleration, $(\text{Range Rate})^2/ 2*\text{Range}$
and for crossing paths, post encroachment time [Allen, 1977]

Table 2
Basic Intersection conflicts

1. Left turn, same direction
2. Right turn, same direction
3. Slow vehicle, same direction
4. Lane change
5. Opposing left turn
6. Right turn cross traffic, from right
7. Left turn cross traffic, from right
8. Thru cross traffic, from right
9. Right turn cross traffic, from left
10. Left turn cross traffic, from left
11. Thru cross traffic, from left
12. Opposing right turn on red (protected left phase)
13. Pedestrian

Another important consideration is what measures the human uses for driving? Van der Horst[16] argues that such time-based measures are also relevant to the human control process. While the results are mixed, several studies ascribe some role to time to collision in the human control process in specific situations such as braking and lane-keeping (time to lane departure).

This issue has important implications. TCT has traditionally been used to study intersection design, not human factors. However, a need has been expressed[22] for identification of a "critical incident" for human factors studies of traffic accidents. An objective definition of a "critical incident," or near-collision, could provide a basis to relate information from various methods such as DASCAR, VME, instrumented vehicles, or driving simulators.

Another issue from the literature was the relationship of conflicts with accidents. From our review, this issue has pretty well been settled in both the U.S. and Europe. The key seems to be the recognition of an underlying time-based parameter, such as time to collision, that corresponds with conflict "severity." When only severe conflicts are considered, those with relatively short times to collision, then the relationship with accidents is more evident because there is less opportunity for interviewing variables to obscure the link. When less severe conflicts are studied, multivariate approaches are necessary in order separate out the effects of other factors from the relationship with collisions.

Based on our review of the literature and this preliminary analysis, it appears the log-linear models will be well-suited to quantifying the relationship between conflicts and accidents. The major hurdle to establishing conflict/accident relationships has been the limited number of accidents. Multivariate models would incorporate factors such as:

conflict type
collision situation
roadway characteristics
traffic characteristics
vehicle type
other factors

The collection of sufficient data and development of these models will be a major undertaking. A successful development of a traffic conflicts, or critical incident, method for the study of collision avoidance could be a very important tool.

Application of TCT to Collision Avoidance. The historical development of traffic conflicts was as a surrogate for accidents. Since a sufficient number of conflicts could be observed in a matter of days, instead of years for a limited accident count, the fundamental issue was whether the variance in the conflicts was improved. This variance depended both on how consistently conflicts could be identified and on the variability of the relationship between conflicts and accidents. However, once the advantages of conflicts were established, the general idea was to forget about accidents and use conflicts instead. That approach might be viable as a bridging analysis as well.

Two situations can be envisioned. Suppose that a measure like time-to-collision (TTC) can be related to the probability of a collision, and the VME provides distributions of TTC that can be extrapolated sufficiently to estimate a cumulative probability of collision. Now we want to calculate the effect of some collision avoidance technology. In one case, the effect of the technology might be to eliminate some of the very low TTC values in the distribution. Such a technology would reduce collisions by reducing the number of "near-collisions" (low TTCs), and the approach described would provide an estimate of the magnitude of the potential benefit.

In the second case, imagine that the effect of the technology is to allow traffic at smaller TTCs, but without increasing the probability of collision. Here the intended effect is to *change* the relationship between TTC and the probability of collision. In this case, one would have to quantify this new relationship before the benefit could be estimated. This case may be more likely for technology directed at improving traffic flow, and estimation of the safety impact would be much more difficult.

Extreme Value Theory

Information on traffic conflicts, or near-collisions, is to be gathered by observing normal (non-accident) driving/traffic with devices such as DASCAR or the VME. Such events are expected to be well out in the tails of the observed data on, for example, a crash margin measure. Most statistical methods are focused on estimating the mean or variance of a distribution. The material in this section addresses the question of how much confidence one can put in the extremes of the observed values. Typically, these values are suspected of being outliers, observations outside the intended distribution due to experimental or instrumentation errors. The issue here is to look into the reliability of the information in the tails of the observed values.

Method. A review of the literature on order statistics identified the topic of extreme values. Gumbel laid the foundation for the study of extreme values with the publication of *Statistics of Extremes* in 1958[23]. Gumbel describes two applications: (1) to determine if observed extreme values are outliers, and (2) to forecast further extremes (outside the available observations). Since then, extreme value theory has received much attention and undergone many changes. In particular, more recent developments have led to a unified approach to the theory. Gumbel showed that for large samples, depending on the parent distribution, distributions of extremes can follow one of three asymptotic distributions. The three asymptotic distributions that Gumbel referred to as the first, second, and third asymptotes are now commonly called the Gumbel, Frechet, and Weibull distributions, respectively.[24,25]

For these methods, the problem is formulated as follows. Start with several samples, each of size n. For example, each sample might be the level of a river measured daily during one year. One then selects only the largest (or smallest) observation from each sample (for example, the one day during each year when the river was highest). When this procedure is repeated over several years, this is the sample of extreme values for analysis. The basic questions addressed are: what is the distribution of the extremes, and how is this distribution related to the distribution of the original samples (all 365 levels observed in a year)?

The sampling procedure addresses an important issue; that is the independence of the observations. Consecutive observations of the level of a river, or the headway between two vehicles will not be independent since the process under observation is a continuous one. Consequently, one cannot apply standard distribution theory to the parent distribution. However, if the sample from which the extreme is selected is sufficiently large, the sample of extreme values will be independent.

The result is quite robust. No matter what distribution the parent observations follow, the distribution of the extreme values is one of only three asymptotic solutions. The extreme values from most common distributions (normal, lognormal, exponential, gamma, logistic) follow the first solution. The distribution of extreme values follows the second distribution when the original distribution is of the Cauchy or Pareto type. These distributions of the second type are characterized as having longer tails, so the asymptotic solution is somewhat different. Gumbel has also shown that distributions of the second type may be converted to the first type by a logarithmic transformation. Of course, if the log transformation is required, the implication is that the distribution of extreme values does not intersect zero. For these cases, the distribution of extreme values is the familiar Weibull distribution used in quality control applications.

Application of the method is straightforward. Given a sample of extreme values, one can fit the appropriate distribution. Inferences can then be made on the basis of the fitted distribution of extreme values. The most direct approach is to make extremal probability paper based on the first asymptotic solution. When the distribution of the observed extreme values are plotted on this paper, they should appear to follow a straight line. Outliers are recognized as not following the line of the rest of the observations. If the underlying distribution is of the second type, the log transformation can be used to plot it on the same probability paper. The examples reviewed clearly support extrapolation beyond the observed data. This might be regarded as the most straightforward approach to the bridging analysis. For example, one could take a distribution of the crash margin, all greater than zero since no collisions were actually observed, and simply extrapolate back to zero to estimate the proportion of collisions if this distribution were observed long enough.

Statistical methods for this are readily available and fairly robust. Extreme values are apparently quite predictable, even far beyond the actual observations. However, one must assume that there are no discontinuities in the distribution as the crash margin approaches zero. In the event of such discontinuities, the extreme value distribution is a technique to estimate the proportion of events that exceed some threshold (greater than zero) in the crash margin measure that corresponds to the definition of a conflict, or near-collision, event.

It will be interesting to see if collisions seem to arise as simply a rare event in the normal distribution of interactions among vehicles (the first solution), or if some discontinuity that is not reflected in accident-free travel seems necessary to precipitate collisions. Extreme value theory may be too simplistic as a bridging analysis. However, the original focus was to address the reliability of the information in the tails of the observed distribution (for example, a conflict may be defined as a time to collision of 1.0s). To that point, Gumbel states that "Forecasts of the extremes are more reliable than the median." On this later issue, we seem to be on firm ground. The technique can also be used to identify outlying observations. Additional information on the statistics of extreme values is included in an Appendix.

The FOCAS Data. The first year of the FOCAS project[7] produced a set of baseline data for 36 naive drivers. Each subject drove a specially equipped Saab 9000 over an approximately 50 minute freeway course on local limited access roads. Each subject drove the same course three times: once using manual headway control, once using conventional cruise control, and once using adaptive cruise control. Files for the 36 subjects manually controlling headway (no cruise control or headway control) were used for this analysis. Flags were added to identify the five separate sections of limited access road. Variables in the files are time, range, range rate, speed, acceleration, a valid target flag. From these variables, time available and time to collision can be calculated. Since the extreme value approach could be sensitive to erroneous data, the data were reviewed carefully.

A sample plot of range data from one subject is shown in Figure 2. Range (m) is shown versus time (sec). It took a subject about 50 min. (3,000 sec) to drive the 50 mile route. Due to the lengths of the five highway segments and differing traffic volumes, the character of the data are somewhat different for each highway segment. Only the segment on I94 East of Ann Arbor is shown here. Driving on entrance/exit ramps connecting the highways has been omitted.

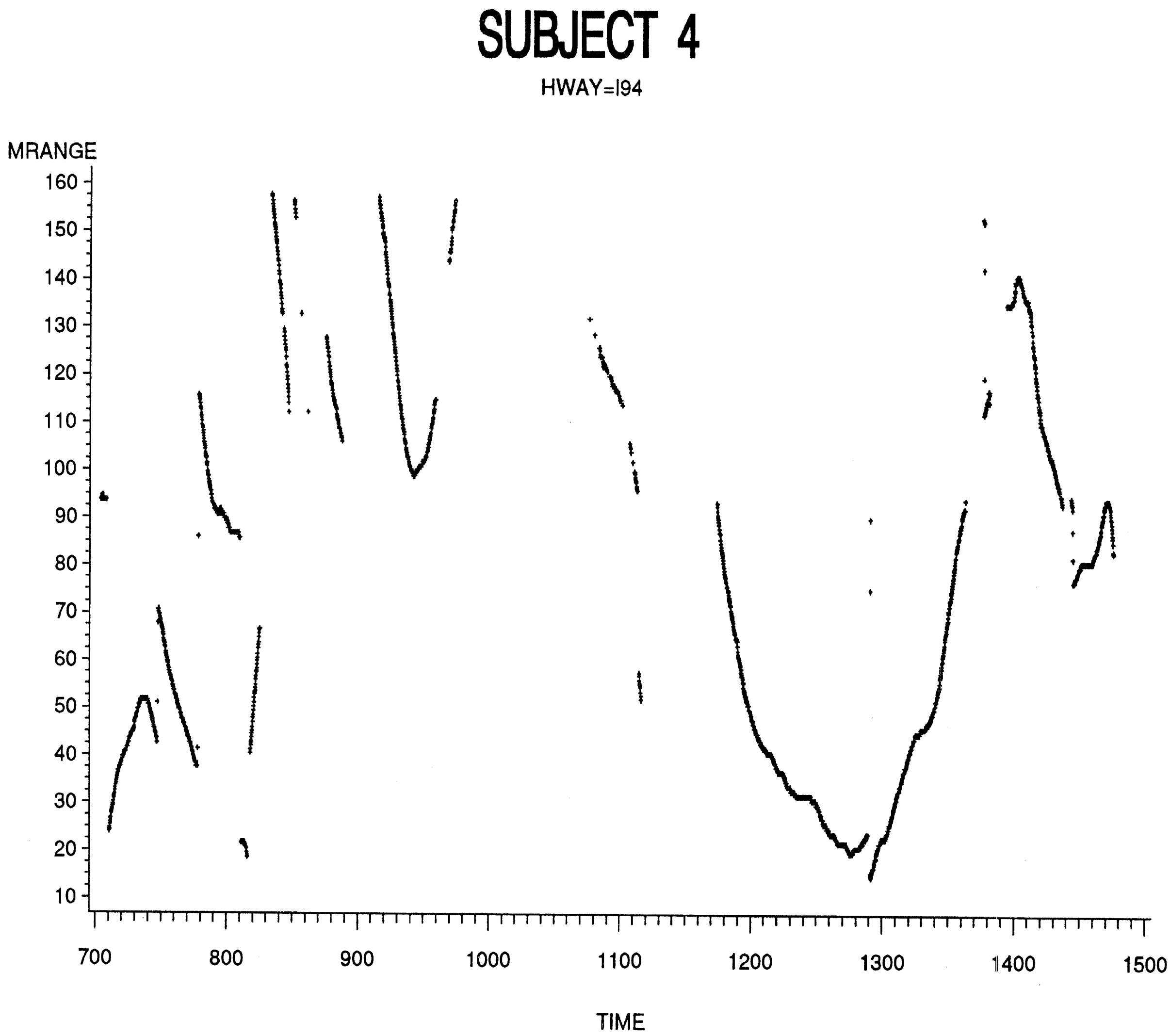


Figure 2. Range data from the FOCAS project

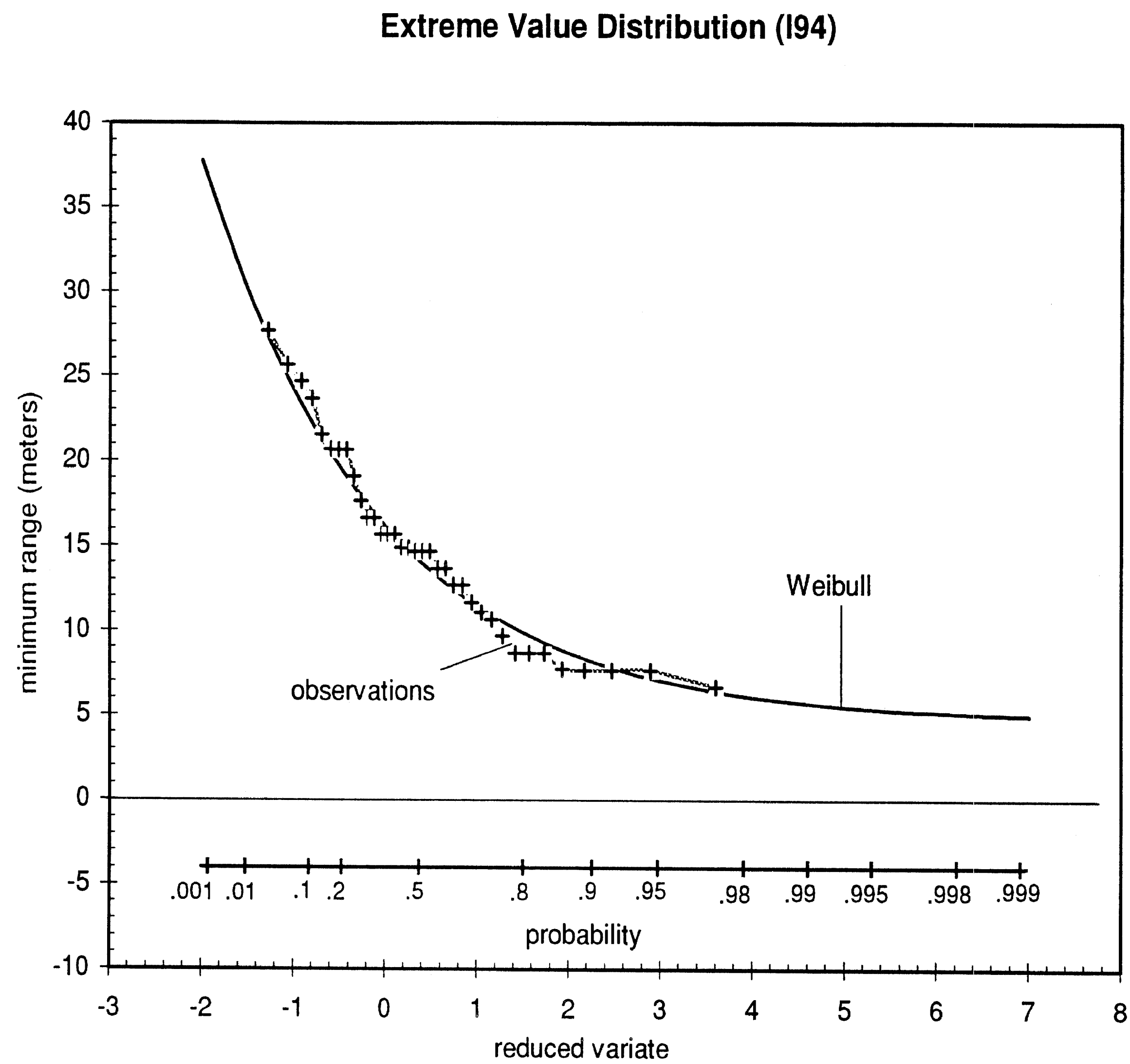
Gaps indicate that the range sensor did not return a valid value, either because there was no car within its range (500 feet) or the signal was erroneous (from a roadside object, for example). For the most part, the range values are continuous, apparently indicating measurements to a single lead vehicle. Occasionally, there are brief interruptions, perhaps due to brief loss of the target due to curves or grades or bumps.

Of more interest is the general character of the traces. Three distinct events are evident: (1) a lead vehicle is picked up at a distant range, and the range decreases rather sharply (subject driver closing), (2) the range fluctuates about a more or less stable value (subject driver following), and (3) a lead vehicle target is picked up at close range, and the range steadily increases (subject driver being passed). Occasionally, a closing event will continuously transition into a following event, but more often the closing event is broken off, presumably by a lane change, without any attempt by the subject driver to moderate the closing rate. Traffic conditions were relatively light for all subjects. Of course, these three events are an oversimplification in that there are other combinations of maneuvers and gaps in the headway signal that would also produce similar range-time traces, but they seem useful for discussion at this point.

As an example, in Figure 2, a target appears at about 1180 sec. at a range of about 90m. The subject drivers closes to about 20m, and then the trace is interrupted briefly at about 1290 sec and a target reappears closer, at about 12m. Over the next 60 sec or so, the range increases back to about 90m. The subject driver may have been closing on a vehicle while waiting for an opportunity to pass. As a faster vehicle passed, the driver changes lanes, and the headway sensor picks up the passing vehicle as it pulls away. Of course, other traffic events could produce similar data. This is a common pattern. When the subject driver appears to close on a target, the minimum headway is greater than when the lead vehicle has a positive range rate so the range immediately begins increasing. The proportion of events where the subject driver is closing, following, and being passed seems to change from subject to subject.

Thinking of the extreme value analysis, if all events are combined, the minimum range is likely to occur when the lead vehicle has a positive range rate. A different result would be obtained if only following, or only closing events were used. At this point, it seems that these events should be kept separate, since the nature of the conflict is different. In the closing and following situation, the subject driver is making decisions about controlling headway and changing lanes. However, when the subject driver is passed, the passing driver controls the headway decision when he/she moves into the subject driver's lane. These situations may distinguish conflicts that have different probabilities of collisions. These considerations are also relevant to the selection of a crash margin measure. Range and time headway do not consider the sign of the range rate, whereas the time to collision (range/range rate) does, and thus, can discriminate situations where the vehicles are separating rather than closing.

Extreme Value Plots. The theory is particularly useful for extrapolating outside the range of observed data to predict the occurrence of as yet unobserved events. In this regard, it is of interest to record the minimum crash margin observed with the 36 subjects of the FOCAS project. If the observed data follows the Gumbel distribution, then the data should plot as a straight line on Gumbel probability paper. If, however, the data follows the Frechet or Weibull distribution, the points will plot as a curve. The intent is to fit the most appropriate of the three asymptotic distributions to the observed data, and extrapolate from positive minimum headway where crashes have not occurred, to a headway of zero where crashes do occur. Examining the percentile of the fitted distribution where zero occurs can provide an estimate of the number of subjects required to participate in the study before an accident is observed. In extreme value theory this number is referred to as the *return period.*



Figures 3. Extreme value distribution for Range

After cleaning up the data for each of thirty-six subjects and deleting some points that appear as outliers, extreme value plots are shown in Figures 3 to 5 for minimum range (m), minimum time headway (sec), and minimum time to collision (sec), respectively. The results shown are limited to only one of the five highway segments, I94 East of Ann

Arbor. The observed data plots as curves on Gumbel probability paper, which is an indication that the Weibull distribution is a good candidate (rather than a straight line indicating the Gumbel distribution). To reduce subjectivity, the three parameters required to fit a Weibull distribution were estimated by methods proposed by Gumbel. Since the plots are generated on Gumbel paper for maximum values, larger probabilities correspond to smaller percentiles of the left tail. This has the effect of magnifying the region of interest, namely the left tail near zero. For example, a probability of .95 corresponds to the 5th percentile of the distribution.

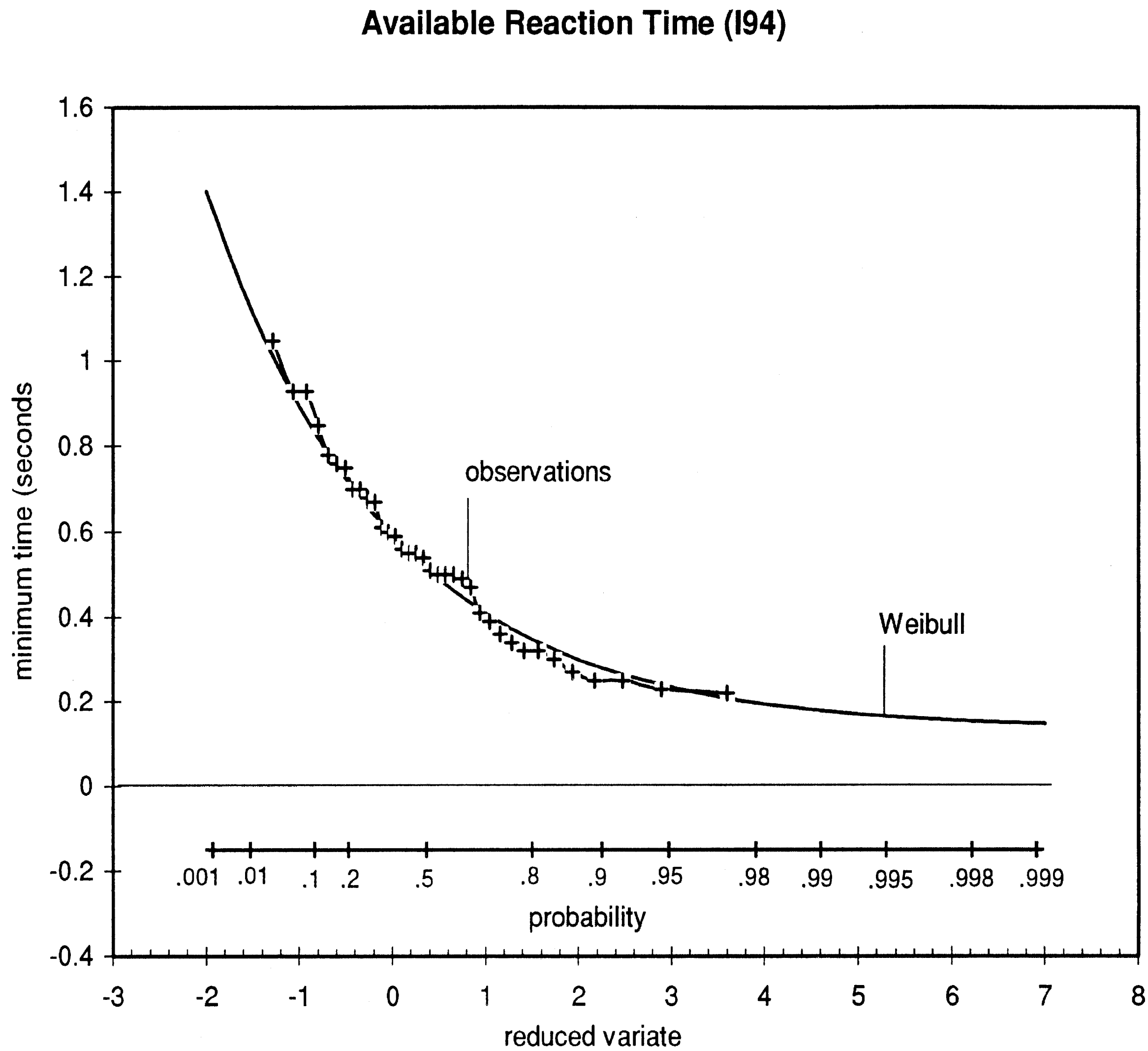


Figure 4. Extreme value distribution for available reaction time

Each plot consists of thirty-six data points, one per subject. The value plotted is the minimum crash margin observed for that subject on the I94 highway segment. Weibull distributions fit the observed data very well. The fitted distributions do not intersect zero, suggesting that under the experimental conditions of the FOCAS study, subjects do not exhibit driving behavior that permits extrapolation to zero. The 5th percentile of the fitted distribution of the minimum range shown in Figure 3 is about 7 m, and the fitted Weibull distribution is asymptotic at about 5 m. The corresponding available reaction time is about 0.2 sec, as shown in Figure 4. These observations are approximately equivalent except for differences in the traveling speeds, which were always over 55 mph. The minimum time to collision, shown in figure 5, also follows the Weibull distribution. Whereas the minimum range and available reaction time, 5m and 0.2 sec, seem quite short, the time to collision is asymptotic at about 3 sec, considerably longer. This result is a reflection of the very low range rate at the time the minimum value occurs. Given the carefully controlled experimental conditions under which these observations were made, the drivers were apparently quite cautious. It is perhaps not surprising that the asymptotic solutions resulted. The result might be different in a more realistic setting.

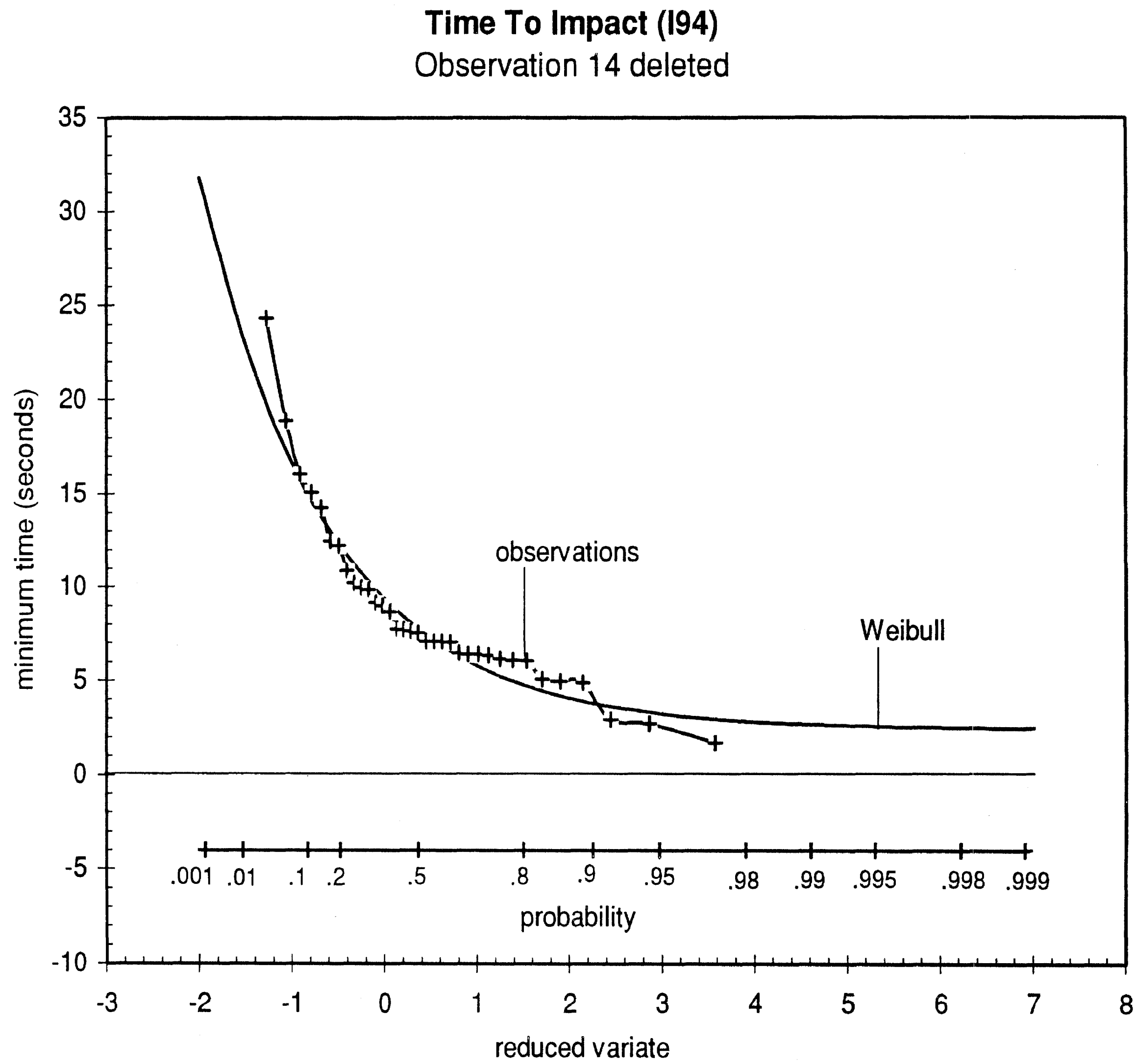


Figure 5. Extreme value distribution for time to impact

The difference between Figures 4 and 5 illustrates the influence of the choice of crash margin measure. It was noted that the minimum range often occurs when a passing vehicle cuts in front of the subject driver, usually with an increasing range rate. The first two plots, range and available reaction time include these events. However, time to collision (Figure 5) omits these events because time to collision is undefined if the vehicles are separating even though the initial range may be small. Thus, this choice of crash margin measure omits minimum headways that result from passing maneuvers if the vehicles are separating.

Future Work

The objective of this project was to develop a methodology for estimating the safety benefits of active safety technology. Accident data, the traditional safety measure, are not sufficient. Consequently, a method relating pre-collision events to safety is needed. The Traffic Conflicts Technique offers a paradigm for relating a pre-collision event to accidents. A conflict is a traffic event ("near-collision") in which a minimum threshold of the appropriate crash margin measure is exceeded. Currently, the identification of conflicts is done by trained observers at the site, or analysis of video tapes. The application of TCT has largely been limited to intersections. A great deal of work is necessary to support the extension of this technique to the broader problem of estimating the safety benefits of active safety technology. This last section describes some of this work.

The premise for this work was that technology (e.g.VME, DASCAR) was under development that would be capable of continuous measurement of the necessary kinematic variables (e.g. range, range rate) that form the basis for the appropriate crash margin measures (e.g. time available or time to collision). Thus, the necessary extension of the traffic conflicts technique is to reformulate the method to incorporate continuous measurement of the relevant parameters. This fundamental premise of the technique is that the probability of a collision can be estimated from the frequency of small crash margins. Extreme value theory is offered as a robust, statistical method to compare probability levels in the tails of observed distributions of crash margins.

The overview of this formulation has already been described in Figure 1, but many details remain to be filled in. Each collision type must be related to the conflict types and traffic events that precede the collision. Since application of TCT has been limited to primarily urban intersections, the major expansion will be to the connecting sections of highway and rural areas. However, the 12 basic conflict types would appear sufficient for most two-vehicle collisions, whether at an intersection or not.

An essential element is the identification of the appropriate crash margin measure. Different measures may be needed for different types of conflicts, such as the post encroachment time for crossing paths. Collection of data on the appropriate crash margin will be an essential specification for new technology under development to study collision avoidance such as DASCAR and VME.

A major activity will be research to quantify the relationship between small crash margins and accidents. In the traditional TCT formulation, the relationship is a simple ratio, and the magnitude varies with the conflict type, threshold value of conflict severity used, and other factors such as intersection geometry and control. TCT has also been a fair weather technique. Adverse weather or nighttime conditions were not addressed for obvious reasons. Previous studies to relate conflicts to accidents covered perhaps 20-30 intersections split among a few cities that were felt to represent a large number of similar urban intersections.

A much greater scope will be required for such studies to support benefit estimation for active safety technology over a wide range of highway conditions. These are likely to be large and costly studies since the scope must be sufficient to include enough accident data, and it will also be necessary for the crash margin by conflict type data to be sufficiently comprehensive to be representative of the underlying distributions of traffic events and crash margin distributions. Perhaps the most important single factor driving the frequency of traffic events (vehicle interactions) that, in turn, generate conflicts is traffic density. The ability to incorporate traffic density directly in the model would make it much easier to generalize. Since the entire crash margin distribution is known, more information is available for the model. For example, the frequency of each measured crash margin might be weighted to reflect the increased risk associated with smaller crash margins, instead of simply counting the number below a given threshold. Consequently, it is likely that more complex, multivariate models will be better suited for this research.

In the traditional application of TCT, once the ratio of conflicts to accidents was established, conflicts became a surrogate for accidents, and the need for accident data was diminished. That is not likely for the broader application to collision avoidance. It is more likely that the need for accident data will be even greater as research expands to improve our understanding of the relationship of pre-collision events and collisions. For example, van der Horst[6] proposes that the study of conflicts per se to determine why some result in collisions, when most do not, would be a fruitful approach.

Once the relationships between crash margin measures for specific conflict types and the probability of accidents is sufficiently quantified, how will the safety benefits of proposed active safety technology be estimated? One will still have to determine the effect of the technology on the incidence of small crash margins, or the probability of a conflict. This information may come from simulations of traffic events that incorporate the operating characteristics of the technology, driving simulators, or measurements in actual traffic with prototype technology in place. However, the conflict methodology provides a basis for relating these various approaches to safety by structuring the events studied to correspond to the appropriate conflict type and crash margin measure. Once the change in the incidence of small crash margins has been determined, the statistical models relating crash margin and the probability of accidents is used to estimate the safety benefits of the technology in question.

The approach described in the previous paragraph assumes that the relationship between the crash margin (conflicts) and accidents is not affected by the technology. In other words, the effect of the technology is to reduce the incidence of small crash margins. However, a technology might also alter the probability of collision given a small margin. This might be the case for a technology intended to enable shorter headways, for example. In this case, the previous relationships between crash margins and accidents would no longer be appropriate for estimating the safety benefit. Benefit estimates could only be based on studies including accident data, as were conducted to establish the relationship of conflicts and accidents in the first place.

Further work is hampered by the lack of technology to provide the automatic and continuous measurement of crash margin measures to support development of the methodology described here. However, a feasible intermediate step would be to use existing simulations, such as the Monte Carlo simulation of rearend collisions described by Farber,[26] or the ASCOM simulation developed at Michigan.[27] Shorter simulation runs would be set up as samples from which to select the extreme value (minimum). Sufficient numbers of these short runs would be made to provide an adequate sample to apply the extreme value theory to estimate the incidence of collisions. Running the simulation to conclusion also estimates the frequency of collisions, and comparison of the two estimates would illustrate the utility of the extreme value theory without waiting for real data.

References

[1] Brown, G.R. 1994. "Traffic conflicts for road user safety studies." *Canadian Journal of Civil Engineering*. Vol 21. pp 1-15.
[2]Harris, J. I.; Perkins, S. R. 1968. Traffic conflict characteristics. *General Motors Corporation Automotive Safety Seminar Proceedings.* General Motors Corporation, Safety Research and Development Laboratory. pp. 1-7.
[3] Allen, B.L., Shin, B.T., Cooper, P.J. and Parker, M.R. 1978. "Analysis of traffic conflicts and collisions." *Transportation Research Record 667.* Washington, D.C.: Transportation Research board. pp67-74.
[4] Williams, M.J. 1981. "Validity of the traffic conflicts technique." *Accident Analysis and Prevention.* Vol. 13, pp133-145.
[5] Dingus, T.A. 1995. Moving from measures of performance to measures of effectiveness in the safety evaluation of ITS products or demonstrations. Paper presented at the Workshop on ITS Safety Evaluation. May 1-2, 1995. Reston, VA.
[6] van der Horst, A.R.A. and Kraay, J.H. 1986. "The Dutch conflict observation technique DOCTOR." *Proceedings workshop on traffic conflicts and other intermediate measures in safety evaluation.* Budapest.
[7] Fancher, P.S., Bareket, Z., Sayer, J.R., Johnson, G.E., Ervin, R.D. and Medford, M.L. *Fostering Development, Evaluation, and Deployment of Forward Crash Avoidance Systems (FOCAS).* Annual Research Report. Ann Arbor : University of Michigan Transportation Research Institute. Report No. UMTRI-95-31. 1995
[8] Burgett, A. 1995. *DOT's approach to ITS safety evaluations.* Paper presented at the ITS America Workshop on ITS Safety Evaluations.
[9]Hauer, E. 1975. *The traffic conflicts technique - fundamental issues.* Toronto University, Department of Civil Engineering, Canada. 28 p. Sponsored by National Research Council, Ottawa, Canada. Report No. 75-01.
[10]Glennon, J. C.; Glauz, W. D.; Sharp, M. C.; Thorson, B. A. 1977. *Critique of the traffic-conflict technique.* Midwest Research Institute, Kansas City, Mo. 7 p. Transportation Research Record, No. 630, 1977, pp. 32-38.

[11]Kraay, J. H. 1985. *An international study on the calibration of traffic conflict techniques*. Institute for Road Safety Research SWOV, Leidschendam, Netherlands. 34 p. Report No. R-85-38.
[12]Older, S. J.; Spicer, B. R. 1976. "Traffic conflicts--a development in accident research." Transport and Road Research Laboratory, Crowthorne, England. 15 p. *Human Factors,* Vol. 18, No. 4, Aug 1976, pp. 335-349. UMTRI-53044
[13]Hayward, J. Ch. *Near miss determination through use of a scale of danger.* Report No. TTSC 7115, The Pennsylvania State University . 1972.
[14]Grayson, G.B. (ed.) *The malmo study: a calibration of traffic conflict techniques.* Report R-84-12, Institute for Road Safety Research SWOV, Leidschendam.
[15]Horst, A.R.A. van der. *The ICTCT calibration study at Malmo: a quantitative analysis of video recordings.* Report IZF 1984-37, TNO Institute for Perception, Soesterberg.
[16]Horst, A. R. A. van der; Brown, G.R. 1989. *Time to collision and driver decision making in braking.* Report IZF, 1989 C-23. Institute for Perception TNO, Soesterberg, Netherlands.
[17]Horst, A. R. A. van der. 1990. *A time-based analysis of road user behaviour in normal and critical encounters.* Institute for Perception TNO, Soesterberg, Netherlands. 180 p. UMTRI-79938
[18]Glauz, W.D. and Migletz, D.J. *Application of traffic conflict analysis a intersections.* NCHRP Report 219. TRB, National Research Council, Washington, D.C. 1980.
[19]Migletz, D.J.; Glauz, W.D.; and Bauer, K.M. *Relationships between traffic conflicts and accidents.* Report FHWA/RD-84/043, Vol. 2. 1984.
[20]Glauz, W. D.; Bauer, K. M.; Migletz, D. J. 1985. *Expected traffic conflict rates and their use in predicting accidents.* Midwest Research Institute, Kansas City, Mo./ Graham-Migletz Enterprises, Inc., Raytown, Mo. 12 p. Transportation Research Record, No. 1026, 1985, pp. 1-12. UMTRI-56287 A01
[21] Glauz, W.D. and Migletz, D.J. 1980. *Application of traffic conflict analysis at intersections.* NCHRP Report No. 219. TRB, National Research Council, Washington, D.C.
[22] *Human Factors Research in Highway Safety.* Transportation Research Circular 414. Transportation Research Board. September 1993. pp. 34-38.
[23] Gumbel, E.J. (1958). Statistics of Extremes. Columbia University Press, New York.
[24] Castillo, E. (1988). Extreme Value Theory in Engineering, Academic Press, New York.
[25] Resnick, S. (1987) Extreme Values, Regular Variation, and Point Processes, Springer Verlag, New York.
[26] Farber, E. 1994. "Using the REAMACS model to compare the effectiveness of alternative rear-end collision warning algorithms." *Proceedings of the Fourteenth International Conference on Enhanced Safety of Vehicles.* Volume 1. Washington, D.C.: National Highway Traffic Safety Administration. pp417-427.
[27] Ervin, R.D., Stein, J.L., Zachariou, N., Bogard, S.E. and Kleinsorge, K. 1996. *Predicting probability of encroachment conflicts posed by adaptive cruise control.* Paper submitted for the 1996 ASME Symposium on Transportation Systems.

Appendix
A Glance at the Statistics of Extreme Values

Let f(x) be the probability density of a random variable, F(x) its cumulative distribution. A sample of n values is taken, and the maximum value in this sample is y. What is the distribution of this maximum value?

Define u so that $F(u) = 1 - \frac{1}{n}$. That means that the expected number of observations among the n, which are above u is 1. u is the "location parameter" for the distribution of extreme values.

The distribution of the maximum value, y, can immediately be written down

$$\Phi_n(y) = (F(y))^n \qquad (1)$$

and its density is

$$\phi_n(y) = nf(y)(F(y))^{n-1} \qquad (2)$$

Though nice and simple, the expressions are practically useless for all but the smallest values of n.

For larger sample sizes, three asymptotic solutions have been developed. Distributions, f(x), fall into three classes, corresponding to the three solutions.

If

$$\frac{f(x)}{-f'(x)} \cdot \frac{f(x)}{1 - F(x)} \to 1 \qquad (3)$$

with $x \to \infty$, then f(x) belongs to the first class.

Examples are the normal, lognormal, exponential, gamma, logistic, and most of the commonly occurring statistical distributions. All distributions of this class have the same asymptotic-- for $n \to \infty$--distribution of extreme values y

$$\Phi(y) = \exp(-e^{-a(x-u)}) \qquad (4)$$

and the density

$$\phi(y) = a\exp(-a(x-u)\exp(-e^{-a(x-u)}) \qquad (5)$$

where u is defined as above, and a is a --sometimes complicated--function of the parameters of F(x). This distribution is called the Gumbel distribution.

If

$$x - \frac{f(x)}{1-F(x)} \to k > 0 \qquad (6)$$

for $x \to \infty$ the distribution is the second type (also called Cauchy type). Examples are the Cauchy distribution:

$$f(x) = \frac{1}{\pi(1+x^2)}$$

and

$$F(x) = \frac{1}{2} + \frac{\arctan g(x)}{\pi}$$

and the Pareto distribution:

$$F(x) = 1-(x-\varepsilon)^{-k}$$

which has the density, $f(x) = \dfrac{k}{(x-\varepsilon)^{k+1}}$

All have the same asymptotic distribution of the largest value

$$\Phi(y) = \exp\left[-\left(\frac{y-\varepsilon}{u-\varepsilon}\right)^{-k}\right] \qquad (7)$$

and the density

$$\phi(y) = \frac{k}{(u-\varepsilon)}\left(\frac{y-\varepsilon}{u-\varepsilon}\right)^{-k-1} \exp\left[-\left(\frac{y-\varepsilon}{u-\varepsilon}\right)^{-k}\right] \qquad (8)$$

where ε is a location parameter which is sometimes the lower limit of the underlying distribution. This distribution can be transformed into one of the first type, (4), by a logarithmic transformation. This distribution is called the Frechet distribution.

If a distribution has an upper limit w, and the condition

$$\frac{1}{w-x}\cdot\frac{f(x)}{1-F(x)} \to k > 0 \qquad (9) \qquad \text{for } w-x \to 0$$

is satisfied, it has a cumulative distribution of the third type,

$$\Phi(y) = \exp\left[-\left(\frac{w-y}{w-u}\right)^{k}\right] \qquad (10)$$

with the density

$$\phi(y) = \frac{k}{w-y}\left(\frac{w-y}{w-u}\right)^{k-1} \exp\left[-\left(\frac{w-y}{w-u}\right)^{k}\right] \qquad (11)$$

Except for different definition of the parameters, this is the same distribution as that of the second type, (7) and (8). It is the Weibull distribution.

If one studies smallest values, the results are analogous. One has just to note that if the original distribution is symmetric, one can use the mirror image of the distribution of the largest value; in other cases, if the original distribution is limited to the left, or has a left tail very different from the right one, then the distribution of the smallest value may be of a different type.

One can apply this theory in two ways. If one knows the underlying distribution, one can derive how the coefficients of the distribution of the largest value depend on the parameters of the original distribution. If the underlying distribution is not known, but extreme values are available, then one can fit the appropriate one of the asymptotic distributions to these values and thus estimate the parameters of the distribution of the largest values. In the case of a distribution of the first type, this can be done very simply by using a probability paper based on (4). In the other cases, one may still use the probability paper, but has to estimate the logarithmic transformation needed to result in a straight line for the cumulative distribution. This approach dates to times when the necessary exact computations were so time consuming as to be avoided wherever possible. Today, the problem can be efficiently solved by computer programs.

The theory can be applied to headway and rear-end collisions in two ways. The first approach starts with a theoretical distribution of headway. This distribution must allow a positive probability for a headway of 0, or even negative headway, because otherwise the distribution of the smallest value would never allow a positive probability for a headway of 0, or a negative headway, which would characterize a rear end collision. From this theoretical distribution, one can derive the asymptotic distribution of the shortest headway, and thus the probability of zero or negative headway, or the corresponding return periods for rear end accidents.

The second approach is based on empirical data. Actual headways--the analysis could similarly be performed with time-to-collision, or another suitable measure--would be collected. However, they would not all be used to derive a distribution of headway, and to fit a theoretical distribution of headway. Even if one could fit a simple analytical distribution to the bulk of the data, the question remains how well this function represents the more sparsely covered lower tail of the actual distribution. Rather, one studies the shortest headway in specified samples, e.g., during one hour at one location, or during one hour of driving by one driver. It is important that the periods are so long that the number of headway is large enough to make the asymptotic models valid. By using only the shortest headway, one avoids questions concerning what the actual analytical form of the underlying headway distribution is, because the analytic form of the distribution of the shortest headway is known--up to a logarithmic transformation.