Speed metrics and crash risks – statistical assessment and implications for highway safety policy

by

Thomas Brooks Stout

A dissertation submitted to the graduate faculty
in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY

Major: Civil Engineering (Transportation Engineering)

Program of Study Committee:
Reginald R. Souleyrette, Major Professor
Alicia Carriquiry
Shauna Hallmark
Edward Kannel
Thomas Maze
David Plazak

Iowa State University
Ames, Iowa
2005

Copyright © Thomas Brooks Stout, 2005. All rights reserved
This is to certify that the doctoral dissertation of

Thomas Brooks Stout

has met the dissertation requirements of Iowa State University

Signature was redacted for privacy.

Major Professor

Signature was redacted for privacy.

For the Major Program
# TABLE OF CONTENTS

1.0 GENERAL INTRODUCTION ........................................................................................................ 1  
1.1 Introduction .......................................................................................................................... 1  
1.2 Dissertation Organization ..................................................................................................... 4  
1.3 Statement of purpose and objectives .................................................................................... 4  
2.0 LITERATURE REVIEW ......................................................................................................... 7  
3 STUDY METHODOLOGY ........................................................................................................... 26  
3.1 ATR Site Identification and Road Segmentation .................................................................... 27  
3.2 Crash Data Collection .......................................................................................................... 28  
3.3 Speed Data Collection ......................................................................................................... 28  
3.4 Descriptive Statistics .......................................................................................................... 33  
3.5 F-Test of Two Distributions .................................................................................................. 34  
3.6 Correction for False Discovery ............................................................................................. 34  
3.7 Crash Models ..................................................................................................................... 35  
4 RESULTS .................................................................................................................................. 40  
4.1 Speed as a function of time-of-day by roadway type ............................................................ 40  
4.2 Speed as a function of day of year by roadway type ............................................................. 44  
4.3 Correlation between dispersion and departure ...................................................................... 47  
4.4 Evaluation of standard deviation of speed distributions ....................................................... 50  
4.5 Crash risk vs. speed and volume measures ......................................................................... 53  
4.6 Crash Models ..................................................................................................................... 66  
4.6.2 Non-weather crashes ....................................................................................................... 72  
4.7 Weather Effects .................................................................................................................. 74  
5 CONCLUSIONS AND RECOMMENDATIONS ........................................................................ 76  
5.1 Discussion of results ............................................................................................................ 76  
5.2 Policy Implications .............................................................................................................. 78  
5.3 Limitations of Current Research .......................................................................................... 81  
5.4 Suggestions for Further Research ...................................................................................... 82  
REFERENCES ............................................................................................................................. 84  
APPENDIX A -- VISUAL BASIC CODE FOR DATA PROCESSING ........................................ 89  
APPENDIX B – SUMMARY TABLES OF SPEED METRICS BY TYPE OF FACILITY ................ 115  
APPENDIX C – SAS INPUT & OUTPUT PRINTOUTS ............................................................. 140  
APPENDIX D – ATR SITE DESCRIPTIONS .............................................................................. 167
1.0 GENERAL INTRODUCTION

1.1 Introduction

The safety of the motoring public has long been an important consideration of professionals in a wide range of disciplines. For the purposes of this document safety is defined as the relative freedom from crash risk, so that an increase in safety is equivalent to a decrease in crashes or their risk.

Each of these disciplines has its own area (or areas) of expertise and concern. For example, the medical profession is concerned about the amount and types of injuries, the speedy transport of the injured to care centers, and providing appropriate medical care. Mechanical engineers are interested in improving the safety of the vehicles, to reduce the potential for crashes and to reduce the risk of injury for those vehicles that are involved in crashes. For that portion of the civil engineering profession involved in traffic safety, the interest, concern, and focus have been on finding ways to improve the safety of the roadway environment.

There has long been an interest within this portion of the profession in achieving a balance between transportation efficiency and safety. It is a broad but important generalization that efficiency (characterized as the speed of movement) comes at the expense of safety. The goal of providing the most efficient mobility without unduly sacrificing safety equates to setting speed limits as high as possible without significantly raising crash risk. There is an issue of public policy with regard to speed limits and their acceptance and/or observation; anecdotal evidence is that increased speed limits are no more observed than were lower limits. This was recently the subject of discussion in the Iowa legislature during debate on changing the speed limit on rural interstate highways to 70 mph (Des Moines Register, April 20, 2005).

The role of technological innovations in the evaluation of roadway safety has been a peripheral focus in a number of studies (see, for example, the use of speed radar in Wiley et al (1949)), although it does not appear to have been the specific object of a study. There is a variety of new technologies, as well as new applications of old technologies, that appear to
have the potential to provide valuable information to improve the quality of roadway safety research. One technology, currently used by such agencies as the Iowa DOT to evaluate mean speed on the roadway system is the automated traffic recorder or ATR. It is the purpose of this document to develop a method for and to report on the use of automated traffic recorders to evaluate the risk of crash involvement on a variety of highways in Iowa.

In 1964 the Federal Highway Administration (FHWA) published a report by David Solomon in which he reported an exhaustive study of the relationship between crashes on 2-lane and 4-lane roadways and a number of factors (Solomon, 1964). The factor which generated the greatest interest and has had a major influence on highway operations and traffic safety was that of the speed of the vehicles involved in the crashes and how this speed compared to the speeds of non-crash vehicles in the same sections. Solomon’s work appears to be seminal in that it is often cited as the source of the 85th percentile speed rule for setting speeds (see for example Cirillo (1968) and Kloeden et al (2002)). Since this time there have been numerous changes in the roadway system, most of which have made driving a safer endeavor. The current research makes a broadly based evaluation of the relationships between speed, the variation in the speed distribution (measured by the standard deviation and by the difference between the mean speed and the 85th percentile speed), the difference between the 85th percentile speed and the speed limit, and the risk of crash involvement. The goal of the research is to evaluate various metrics of speed on highways, determining their relationship to safety and thus providing a uniform basis for setting reasonable and enforceable speed limits that will strike a balance between the somewhat competing issues of mobility and safety.

Solomon’s study was based on data from the mid-1950s. Since that time there have been a number of changes that have at least the potential to impact on highway safety and thus on the relationship between speed and crash risk. Some of these are as follows:

- The mix of vehicles has changed substantially. At the time the study data were collected trucks were less prevalent and not heavily used for the long-distance movement of goods; automobiles constituted a much higher percentage of the total vehicles. According to Solomon’s data, approximately 22 percent of the vehicles
were commercial; this increased to 30 percent at night. Based on 2002 data from the Iowa DOT, trucks averaged approximately 37 percent of the vehicles on I-80. According to preliminary 202 data from the Bureau of Transportation Statistics, trucks hauled about 58 percent of the tonnage of freight in the U.S.

- The vehicles themselves have changed substantially. Safety belts were just becoming available on cars in at the time of the study; today we have three-point seat belts at most or all locations within our vehicles and we have multiple airbags in all automobiles and their cousins (SUV and mini-vans). The litany of vehicle changes goes on to include better tires, energy-absorbing bumpers, disk and anti-lock brakes, more powerful headlights, side impact protection, and energy-absorbing vehicle structure. All of these have an impact on the safety of highway travel.

- The roadway environment has changed substantially as well. Speed limits are set somewhat more rationally, based on research such as Solomon’s and others. Signing, striping, lighting, and traffic signals have been changed to improve visibility. Clear zones, as well as better protection of fixed objects in the clear zone, now provide a more forgiving environment for errant vehicles. Much has been done to make the roadway environment more understandable to drivers. We are well along in the process of modifying freeways and other limited access roadways to have all entrances and exits on the outside, providing a uniform expectation to drivers.

- The driving population has changed. Based on Census Bureau data, the 50-54 age group experienced the highest rate of growth, increasing by 55 percent over the last decade. The Bureau attributed this to baby-boomers entering this age group. The next largest growth for an age group was the 45-49, also part of the baby boom group. With the aging of the “baby-boomer” generation, a greater proportion of drivers is reaching an age at which they are losing visual acuity and cognitive functions, putting them at greater risk.

- Travel patterns have changed. According to Census Bureau data (ACS 2000), the average drive time has increased by nearly 20 percent since 1980, from 21.7 minutes to 25.5 minutes (in 2000). There has been about a 30 percent increase in the number
of drivers traveling over 45 minutes to work, from 11.6 percent in 1980 to 15 percent in 2000. People travel farther for employment, recreation, and education today than in the mid-1950’s. In addition to creating a greater degree of exposure for the average driver, the tendency to travel farther also has fostered a tendency toward multi-tasking (eating, cell phone usage, and shaving for three examples) in the moving vehicle, reducing the driver’s ability to respond properly to roadway events by diverting his/her attention. The longer travel distances seem to go hand-in-hand with increased speeds, a not unreasonable coincidence given the reduced tolerance for delay.

1.2 Dissertation Organization

This dissertation is separated into two parts. The first part presents the development of a methodology to utilize automated traffic recorder and GIS-based crash data to evaluate the relationship between roadway speed, the variance in that speed, and crashes in various types of roadways. The second part applies the methodology to data from the State of Iowa, to present a comprehensive evaluation of these above-mentioned relationships. The dissertation is organized as follows:

- Introduction
- Literature review
- Study methodology
- Results
- Conclusions and recommendations

1.3 Statement of purpose and objectives

Solomon’s study used data from 1955-56, covering about 600 miles of 2-lane and 4-lane rural highways (Solomon, 1964). Crash speeds were estimated by investigating police or were reported by the drivers involved. Interviews were conducted with approximately 290,000 drivers passing through these 600 miles, to determine a variety of personal socio-economic factors; their travel speeds through the roadway section of interest were monitored before they were stopped for interviewing. Crash risks were evaluated, based on these data, and were presented in what has come to be known as “Solomon’s Curves.” At the time of
the report, travel on these two types of highways constituted about one-third of total travel. Today travel on 2-lane and 4-lane rural highways is less than 15 percent (2003 data from FHWA). Yet the results of his study are frequently cited (see for example Cirillo (1968) or Kloeden et al (2002) in Chapter 2) and have been incorporated into policies for setting speed limits on all types of roadways.

There are several ways new technologies can be used to provide an updated and improved evaluation of the relationship between speed and crash risk. In Solomon (1964), speed data were obtained for selected drivers by concealed speed measuring devices (no additional information is provided on the specifics of these devices and more importantly, no information is provided on their accuracy). Speed profiles were determined for each study section using (apparently) the floating car technique in which a test car is driven through the test section several times. Currently, each state’s network of automatic traffic recorders (ATR) can provide speed data for a wide variety of roadway sections. These records permit an accurate assessment of the speed distribution and the variation in speeds in sections of interest. Crash records are now stored in data bases that are accessible to researchers, permitting more in-depth analysis of the records and a better correlation between crashes and their locations. Many new automobiles with air bags include what is called an event data recorder (EDR) that is similar to aviation’s “black box” data recorders. This device provides a record of various inputs from the vehicle (speed, braking, throttle setting, air bag deployment, and change in speed) for the five seconds preceding a crash severe enough to deploy the air bag (NHTSA EDR Working Group summary report, 2001). Geographic data bases are available that include aerial photography and these can be used to characterize access density and a variety of potentially contributory parameters in a corridor and their relationship to crash risk. In today’s political and social environment, as well as a purely practical matter, it would be impossible to stop 290,000 drivers to solicit their personal information.

Another factor that was not considered in Solomon was weather. Most states have automated weather recording systems that may permit an estimation of and possibly allow the (experimental) control of the effects of adverse weather conditions on crashes.
Other factors that have the potential to affect the relationship between speed and crash risk include the following:

- Sight distance and design speed. According to AASHTO (2001), “The available sight distance on a roadway should be sufficient to enable a vehicle traveling at or near the design speed to stop before reaching a stationary object in its path.” Vehicles exceeding the design speed (irrespective of the speed limit), are subject to a reduction in the amount of time available to perceive and react to objects in the roadway.

- Stopping ability. As discussed above, there have been a number of changes in vehicles since Solomon’s data were collected. Improvements in durability and traction capability of tires have generally improved vehicles’ ability to make full use of the improvements in braking that have come from the almost universal use of disk brakes and the considerable use of anti-lock brakes. Improved smoothness and rideability of pavements also contribute to improved stopping ability. This improved stopping ability will have an impact on the overall crash risk and thus will modify the relationship between speed and crash risk.

- Vehicle energy. The increased number of heavy commercial vehicles and large sport utility vehicles in the highway mix means that there is a potential for more kinetic energy in a crash, thus potentially affecting the severity of crashes.

It is hypothesized that the one or more parameters related to the speed and variation of the traffic stream, such as the mean speed or the difference between the 85th percentile speed and the speed limit from the time immediately preceding a crash, will demonstrate a greater variance of distribution than the same parameters of the similar period one week earlier (in the absence of a crash). This document describes this test of this hypothesis, which utilizes speed data from Iowa’s network of automated traffic recorders and crash data from the Iowa DOT’s crash record system.
2.0 LITERATURE REVIEW

The history of research into the relationship between speed and crashes is long and varied. Early speed limits were set somewhat at the whim of elected officials, who were concerned with the then new automobile and its impacts on a largely rural society. Later limits were set based on a consensus of reasonable speed for certain environments. Speed limits have been set as an element of national policy, for example as an energy conservation measure as during the Second World War or the “energy crisis” of the 1970’s. Safety concerns prompt speed limits such as in school zones. A synthesis of research over the past 55 years into this relationship is presented in this section.

2.1 Illinois urban experience (1949)

An interesting study was conducted by Wiley et al (1949) on the effects of speed limit signage on vehicular speeds in Champaign and Urbana, Illinois. Their objective was to determine the actual travel speeds on various sections of roadway, while varying the posted speed limits. They note that speed is not well understood and that speed limits depend more on (page 1), “…custom, assumption, and tradition in establishing speed limits. The result is that most present-day speed limits have been handed down from ‘Model-T days’ and officials still seem loath to change them. Furthermore, it is evident that a strong, popular belief still exists that alleged high speeds, per se, are fruitful causes of traffic accidents and that lower speeds can be obtained by posting low speed limits.” They go on to note (page 1), “Traffic engineers, however, have long observed that modern traffic tends to run at speeds which the motorists themselves consider to be reasonable and safe, irrespective of any posted limits.”

They note the difficulty of enforcing an artificially low speed limit and that the practice was to set speed limits low and for the police to then tolerate a large deviation from that limit on the faster side. As they observe (pages 1 & 2), “In effect, this establishes an actual speed limit, which in itself may be reasonable, but which is unknown because of the
false value on the speed limits signs and the unknown and often variable tolerance. Such practice is not only unfair and confusing to the motorist and the police but also extremely misleading to the public."

Their methodology was straight-forward. They selected four roadway sections on state highways passing through Champaign. Intersections were stop controlled. Speeds were measured using Enoscopes and stop watches; they were verified to be accurate within 5 percent by the use of an Illinois DOT radar speed meter. Initially the streets were posted at 25-mph. They reposted the limits to 30-mph or 35-mph, then measured speeds 1 and 4 days after the change. They then removed the signs, leaving no posted limits, and measured speeds at 1 and 5 days. Next the speed limits were set at 20-mph, with measurements at 1 and 5 days afterwards. Finally the limits were reset to 25-mph and the speeds were checked to confirm the first measurements. In all cases counts were taken over a period of about ½ hour and included no more than 100 cars. The duration of the study lasted about five months, beginning in January of 1948 and concluding on the first of June. They do not discuss weather conditions that might have differing impacts on the vehicle speeds.

They conclude that (page 4), “…speed ranges and variations are so nearly alike in all cases for the same street as to demonstrate clearly that the posted speed limits are entirely ignored and that traffic ran at ‘natural’ speeds irrespective of whether the signs read 20, 25, 30, or 35 mph or whether there were no signs at all.” They recommend a speed limit for each roadway segment studied that would be consistent with the 90th percentile speed and that would coincide with the tolerance in local enforcement of the posted limits. They present graphs showing (page 5), “…the relation between the average speed and the 85- and 90-percentile speeds at the sign-speed stations for each speed limit. They indicate that these percentile speeds are not excessive…” The issue of importance to the current study is that even 50-plus years ago it was known that speed limits, if they are to be observed, must reflect the composite judgment of drivers as to the appropriate speed for the road in question.
2.2 Study of rural highways across the country

David Solomon’s research (1964) was an extensive evaluation of the relationship between speed and crash risk (among other factors). In his introduction (page 1) he notes, “Many of these relationships (between speed, vehicle and driver characteristics, and accidents) have not been clearly understood in the past.” This research formed the basis of many policies for setting speed limits, on all types of highways. This is clearly beyond his intent as he states in the preface, “The study was confined to 2- and 4-lane main rural highways of the nonfreeway (sic) type, and the findings are limited to these types of main rural highways.” He observes that at the time (page 6) “…the study sections are representative of highways that accommodate more than one-third of all the vehicle-miles of highway travel in the United States.” His study covered 600 miles of these highways, on 35 sections in 11 states, and included crash records of approximately 10,000 drivers, speed profile determination, spot speed studies, and interviews with 290,000 drivers.

His approach was first to determine the average speeds within each roadway section by the use of a study vehicle and team moving with the normal flow of traffic. Speed profiles of the sections were plotted and were evaluated by highway department staff to select a specific site within each section that the staff members felt was representative of the overall section. Next the teams measured the spot speeds of the 290,000 drivers, with (page 7) “Concealed, speed measuring devices were used to record the speed of individual drivers…” At a point farther down the road the drivers were stopped and were interviewed to gather data for characterizing them on the basis of sex, age, etc. Crash data were obtained from state records and covered 10,000 drivers that had crashed in the study sections over a 3 to 4-year period ending in 1958. The speeds of the vehicles involved in the crashes were taken from the crash records. Speed distributions, involvement rate plots, and speed – cumulative percentage plots were presented.

The speed related findings of particular interest to the present study showed that the accident involvement rate (number of vehicles involved in accidents per 100 million vehicle-miles of travel) reached a low point (at a rate of about 90 crashes per 100 MVMT) at a travel speed of about 65-mph in the daytime (see Figure 1). The low point for night-time travel
was at a rate of about 230 per 100 MVMT, at a speed of about 57-mph. Travel at low speeds (20 mph or less) presents a substantial additional risk of crash involvement.

Two of his conclusions are of continuing interest and relevance, that the lowest risk of being involved in an accident comes at a speed close to the average travel speed in a corridor and that the risk of being involved in an accident is reduced if the variation between drivers’ speeds is reduced (see Figure 2, below). Figure 2 differs from Figure 1 in that it presents the involvement rate in terms of what Solomon referred to as “variation” from average speed, rather than the travel speed he used in Figure 1. The use of the variation from average speed has the effect of normalizing data from a variety of speed limits. It is not explained why the minimum crash involvement came at a speed variation that is about 5 mph above the average speed rather than at the average speed as might be expected. Although he is often credited with recommending the use of the 85th percentile speed as the speed limit, there is no mention of the 85th percentile speed in his report. It is important to remember his observation that “nearly half of all the accident involvements were either rear-end collisions or same direction sideswipes.” His roadway sections did not have controlled access and although there were no major intersections there is no information as to other points of access such as driveways. He notes the possibility that as many as half of the low-speed crashes
could have occurred at intersections but does not discuss the fact that for each low-speed rear end collision there must also be a vehicle that was going faster. Crashes due to drivers that were inattentive, speeding, or following too close may have been included in these low speed crashes.

Figure 2 Variation from Average Speed vs. Involvement Rate

2.3 The interstate system

In 1968 Cirillo presented a brief report covering speed aspects of a larger study, the “Interstate System Accident Research Study II.” Speed data were provided by 20 state highway agencies, reporting only day-time speeds (between 9:00 A.M. and 4:00 P.M.). The crash data used were limited to the same time period. Her analysis methods and results were similar to those of Solomon, although she noted that the average speed on the interstate system is about 7 mph higher than on main rural highways (if referring to Solomon’s findings this would put the speed on the interstate highways at about 72 mph. She presented a curve showing accident risk (rate per 100 MVMT) versus speed variation (in mph, the difference from the mean speed). The lowest risk was at a speed 10 mph greater than the mean speed, at a rate of 25 vehicle involvements (as opposed to crashes) per 100 MVMT). She also plotted Solomon’s data for conventional rural highways; the lowest risk on his curve was at about 5 mph above the mean speed at a rate just over 100 per MVMT. The
importance of her research to the present study is that her findings paralleled Solomon’s, albeit for the interstate highway system.

2.4 State highway in Indiana

In this small study on a single highway in Indiana, the researchers (West and Dunn), utilized data from magnetic loop detectors in state highway 37 near Bloomington, Indiana. Crash speeds were estimated by an accident investigation team from the Institute for Research in Public Safety at Indiana University; their findings were correlated with the speed data from the detectors. They studied 36 cases, with good or better correlation in speed for 1 or more of the vehicles in 32 of the 36 cases.

Their results, similar to those of Solomon, were that the risk of crash involvement was highest at very low speeds and that the lowest risk occurred at a speed around 10 mph above the mean speed. They then went on to remove crashes at intersections, reasoning that turning traffic would be forced to reduce speed and thus was not representative of free-flowing traffic, and found that (when turning related accidents were not included) the rates for speed deviations above and below the mean speed were essentially the same.

They conclude (page 55), “This also indicates what these limits should be for safe travel on the highways. Since the standard deviation of speed is approximately 6 to 8 mph for these roads in Indiana, if the standard 85\textsuperscript{th} percentile speed were used and enforcement were provided at the 95\textsuperscript{th} percentile for the upper limits, this would provide speed enforcement for the high-speed drivers in the high accident involvement region.” The main importance of this study is of the need to control for atypical traffic situations such as turning movements in determining the operating speeds of roadway sections (unlike Solomon).

2.5 Socially optimal speed limit

In a report that does not directly relate to highway safety but is nonetheless of interest due to its discussion of the bases for drivers’ personal speed decisions, Crouch (1976) developed a philosophical argument considering the socially optimal speed limit. The study, conducted shortly after the institution of the 55-mph national speed limit in the United States
in 1974, suggested that if the police act as society’s agents they will not attempt to enforce a speed limit that is set below the optimum limit; rather their enforcement will be limited to the level needed to bring speeds down to the optimum level. Crouch concludes by noting (page 198), “Somewhat ironically, however, if the police do behave as postulated here not much harm is generated by a speed limit which is too low.” This conclusion should be compared with the comments of Wiley et al as cited above regarding the setting of artificially low speed limits and relaxed enforcement and that this practice is unfair to both the public and the police.

2.6 Legal approach to speed management

Another approach is taken by Ruschmann, et al. (1981) in their evaluation of how to manage the risk of unsafe speed, that is, the risk of increased crashes associated with higher speeds. This report reviews and evaluates other studies, summarizing a number of earlier reports (including Solomon’s 1964 study). They cite a 1977 study by Treat and associates from Indiana, in which the researchers found (page 9). “...excessive speed was a definite causal factor in seven to eight percent and at least a probable causal factor in 16 to 19 percent of the crashes...” Ruschmann, et al. further referred to other studies that confirmed that this was an appropriate estimate of the range. They point out that clinical studies provided some support for Solomon’s finding that excessive negative speed deviation (driving too slowly) presented a high crash risk; they further note, “Many instances of slow driving may be due to conditions over which the driver has little or no control, such as slowing to turn, or slowing on account of pedestrians or other vehicles, rather than a discretionary and inadvisable choice of a slow speed.” This is consistent with the findings of West and Dunn as were noted above but actually seems to conflict with Solomon’s discussion of the higher crash risk associated with low speeds and his inclusion of this in his crash risk assessment.

Ruschmann, et al. conclude with several recommendations, one of which is germane to the present discussion. Speed limits should be reasonable, that is they should not prohibit non-risky speeds. Punishment for “those speeding offenses that create relatively low risk compared to criminal conduct and that are not accompanied by criminal intent should be
handled through noncriminal (sic) procedures." (page 42) Minimum speeds as well as maximum speeds should be included. They also note that using our present system of criminal enforcement of speed laws will not provide an effective deterrent to speeding and that it would require an order of magnitude increase in police officers to control speeding using our current procedures. They discuss some interesting concepts regarding incentives and disincentives for speeding, pointing out, for example, that trucking companies that pay drivers by the mile are encouraging their drivers to speed in order to maximize their income. They further discuss the use of incentives to reward non-speeders, while noting that it might be difficult to identify them, as well as the possibility of disincentives levied on speeders by insurance companies.

2.7 Swedish experience

In a study reported in the proceedings of an OECD (Organisation for Economic Cooperation and Development) symposium, Nilsson (1981) reported on some overall crash statistics and their relationship to changes in speed limits on national highways in Sweden. Speed limits were reduced in response to an energy crisis; speeds measured on the highways were lower by 7 to 11 percent following the change in limits. Crashes were significantly reduced in most cases, with reductions ranging from 5 to 25 percent of personal injury accidents and up to 52 percent for fatal accidents. The traffic flows were reported to be unchanged over the two periods (before and after). Comparisons were made between those roadways whose speed limits were changed and other, similar roadways that remained unchanged; the author noted that there was a crash reduction over the study period on the unchanged roadways but that the reduction was much less than on those roads where the speed limits were reduced.

2.8 German experience

This brief report by Lenz in the OECD symposium (1981) evaluated the impacts on crashes on the autobahn of posting a speed limit of 130 kph (equivalent to about 80 mph), versus posting a speed advisory of the same speed. Two sets of sites were considered, in the
first set the speed limit was in place for one year; it was changed to a recommended (advisory) speed for the second year. At the second set of sites the order of signing was reversed. The study found that the number of crashes was reduced for all cases where speed limits were set. The reduction in crashes resulting from implementing a speed limit of 130 kph was the same for one year as it would have been for two years of the long-term temporal reduction in autobahn crashes that began in 1968. They note that the speed limit provided “a reduction in accident figures of 11 percent and in serious injury and fatality figures of 23 percent” when compared to the speed recommendation. The greatest impact of a speed limit was found on the rate of night-time crashes, where the reduction in crash rate was twice as great as during the day.

2.9 Finnish experience

In a study similar in concept to Lenz’s above, Salusjarvi (proceedings of the OECD symposium in 1981) found that recommending the appropriate speed on rural main roads reduced speed by 2 kph (not an operationally significant figure). The study involved changing the speed limits on these roads to one of three cases; above, at, and below the 85th percentile speed. In the first case the standard deviation of the speed distribution decreased while the mean speed and total number of crashes increased. In the second case the mean speed did not change while the standard deviation decreased, the total number of crashes remained the same, and the number of injury crashes decreased. In the third case the mean speed decreased, there was no change in the number of crashes, and the author did not report any change in standard deviation. Overall, the average effect of the imposition of speed limits on crash results had a more significant impact on crashes, resulting in a 10 percent decrease in total crashes, a 25 percent increase in property damage only crashes, and a 46 percent decrease in crashes involving fatalities or injuries. The author notes a significant limitation in that the roadway sections chosen had a high crash rate, which would be “...a major limitation on generalizing the effect.” The author did not discuss the likelihood of the regression to the mean phenomenon as a possible explanation for the reduction of crashes following the change in speed limits, although that may have been his intention. The importance of this to the current research is to emphasize the possible contribution of the
regression to the mean to crash reductions if sites are not randomly selected. The current research is based on using the ATR sites without selecting for high crash rates and, therefore, regression to the mean should not be an issue.

2.10 Irish experience

The Irish experience is reported by Hearne in the proceedings of the OECD symposium (1981). On one hand the Irish experience somewhat confirmed the concept that speed limits should be set at the 85th percentile speed, while on the other hand they found that the average speed of trucks is a more important concern in terms of crash risk. The speed limit for trucks was set at 15 mph less than that for cars. As Hearne notes a speed differential for trucks was implemented in spite of the conflict with Solomon’s findings regarding speed deviation. Their finding that the average speed of trucks is a more important concern in crash risk suggests that the impact of truck traffic needs to be considered in speed studies.

2.11 External cost approach to speed limits

In this interesting abridged report of a study by Jondrow, et al. reported, in the Transportation Research Record 887 (1982), a rationale for calculating speed limits based on the external (i.e., societal) costs of individual driver’s decisions regarding appropriate speed. The authors first present an equation for the optimal speed for an individual driver, bringing in values of the driver’s time value, the cost of gasoline, the increase in usage of gasoline from an increase in speed, the (negative) value of an increased fatal crash risk, and the increase in probability of having a fatal crash. They then modify this equation to include the impact the speeding driver has on the crash risk of other drivers. Finally they apply this equation to the (at that time current) national 55-mph speed limit to determine if it is optimal. They conclude that the optimal speed (considering various values of human life) is above 55 mph and that the national speed limit was an expensive way to save lives.
2.12 Other factors, the Michigan experience

In this 1992 report Parker conducted a survey of speed zoning procedures throughout the United States, and compared how these procedures would impact the speed limits within Michigan. The study employed a before and after design with a comparison group. He found that (page 85) “The current Michigan practice of posting speed limits within 5 mi/h of the 85th percentile speed has a beneficial effect, although small, on reducing total accidents, but has a major beneficial effect on providing improved driver compliance.” He also found that there was no benefit to setting limits lower than 5 mph below the 85th percentile speed. Other speed zoning methods would not improve traffic safety in Michigan. He recommended that speed studies be conducted over a 24-hour period, to reduce the impacts of short-term variations in the rate. He recommended against the use of radar methods, finding that they underestimated speed by 3 mph. This study provides support for the 85th percentile value as the most reasonable speed limit and the effect of doing so on crashes.

2.13 Urban low-speed, Australian experience

Kloeden et al. (1997) conducted a case-control study of injury crashes on urban roadway sections in Adelaide that had speed limits of 60 kph. They used a criterion for injuries in that one or more people in the crash had to have been transported to a hospital, as well as other control criteria. In this study they also conducted an extensive review of the literature of speed versus crashes, with some interesting observations. They take exception to the findings of Solomon (and others) concerning the risks related to very low speeds in a high (roadway) speed environment. They note that Solomon's data could well have been influenced by the drivers executing turns off of or onto the highway and thus would not reflect drivers' free-moving speed choice. The issue of speed variance is also criticized (Page 21),

…the speed variation idea gained weight, more through successive restatements than through good research, it would seem.

Conceptually it is possible to separate speed variance from mean speed, but practical demonstrations of separate effects are difficult. This is because, in reality, both factors are
strongly tied to the characteristics of the road, which are fundamental determinants of the local accident rate. (In theory, the role of speed variation would best be addressed by examining accident rates for a set of roads that were matched for geometry and other characteristics, but which had a different degree of speed variation for the same mean speed).

In their study they evaluated the relative crash risk of traveling at a speed other than 60 kph in a 60 kph zone; they found that the risk at 70 kph (10 kph over) was more than 4 times the risk at 60 kph. The risk of traveling slower than 60 kph was lower until a speed of 40 kph was reached, when it was 1.4 times as risky. They do note that the confidence intervals show that this increase in risk at lower speed could be due to "random variation."

An interesting additional discussion relates to the possible causes for the increased crash risk with increased speed. The factors the authors discussed include the reduction in the reaction and braking distances due to excess speed, the relationship between the crash energy and speed, the relationship between vehicle speed and loss of control, and a combination of these factors. Efforts were made to control for other, driver-related risk factors, especially that of alcohol impairment. The value of this study to the present research is in identifying contributory factors to be evaluated and in casting some doubt on the issue of speed variance as the causative factor in roadway crashes.

2.14 Non-controlled access highways

This study by Parker (1997) covered the effects of raising and lowering speed limits on 98 sites in 22 states, during the period October 1985 to September 1992. He concluded that there was no practically significant change in the 85th percentile speed due to changes in the posted speed limit, even for speed limit reductions of up to 15 mph. For sites where the speed limits were lowered by 15 mph the change in average speed was less than 1 mph. Using several statistical methods to analyze the before and after conditions, he determined (Page 71) "...that there is not sufficient evidence in this dataset to reject the hypothesis that total crashes or fatal and injury crashes changed when posted speed limits were either raised or lowered." He concluded (Page 82), "Based on the best information available to date, there
is no evidence to suggest that lowering or raising posted speed limits on nonlimited access roadways has an effect on crashes. Reducing the posted speed limit without utilizing other enforcement, educational, and engineering measures does not appear to be an effective safety treatment.” These comments address the social issues of reducing speed limits in the expectation that this will in and of itself make the roadway safer. The basis of the current research is to determine the relationship between speed (and variance) and crash risk.

2.15 Australian experience part 2

In a follow-up to the Australian study discussed above, Kloeden et al. (2002) applied a logistic regression model to the data used in the earlier study with the goal of determining a curve showing the relationship between speed and crash risk. Their results were somewhat paralleled and confirmed the earlier results, in that each 5 kph increase in speed (above the 60 kph reference speed) resulted in a doubling of the crash risk. A secondary element of the study was “…to examine the effect of hypothetical speed reductions on this set of crashes and urban crashes in general…” This portion of the study found that the result of eliminating all speeds above 60 kph should be a reduction of casualty (injury) crashes of 25 percent. They also determined, under some hypothetical scenarios that reducing the general urban area speed limit from 60-kph to 50-kph would result in a large reduction in casualty crashes. This assumed that compliance with the new speed limit would be in proportion to the compliance under the old speed limit. Based on other research (see for example Parker 1997) evaluated above this is a less than sanguine assumption in the absence of special speed enforcement measures.

2.16 Early effects of return to the 65 mile per hour speed limit

In a report based on slightly more than one year’s data from 32 states that restored the 65 mph speed limit, Chang and Paniati (1990) found no statistically significant difference in fatalities on rural interstate highways that could be attributed to the higher limit. They point out several possible factors that may influence such findings including the small amount of
speed data after the change in limit, changed observance of the new speed limits, and changed seat belt use.

2.17 Saving lives with higher speeds

In a study funded by the AAA Foundation for Traffic Safety, Lave and Elias (1992) evaluated the impact of the end of the 55 mile per hour “National Speed Limit” on highway safety. They concluded that other researchers, in reporting just the increases in crashes on the interstate system that followed the speed limit change, were ignoring reduced crashes on other roadways. As postulated by the current researchers, drivers that were seeking to avoid the lower speed limits on the interstates by traveling on less rigorously patrolled highways may have had more crashes on these secondary highways. When the 55-mph limit was rescinded the drivers would then return to the interstates. Therefore it would be appropriate to consider the changes in crashes on a statewide basis when evaluating the impact of the speed limit change. Lave and Elias found that there was an overall 3.4 percent reduction in the fatality rates following the repeal of the 55-mph speed limit, in those states that adopted the higher speed limit. They attributed this to the diversion of traffic from the less safe rural highways to the safer interstate system.

2.18 Speed limit impacts in Iowa.

In an unpublished creative component for Iowa State University, Munandy cited earlier work by Garber and Gadiraju (1991) and Brown, et al. (1990) in discussing speed dispersion as an explanatory variable. Speed dispersion was defined as the 85th percentile speed minus the mean speed.

\[
\text{Speed dispersion} = (85^{\text{th}} \text{ percentile speed}) - (\text{mean speed})
\] (2.1)

Using a group of time-series plots of dispersion for various types of roadways, he concluded that the speed dispersion increased on rural interstate highways following the change to a 65-mph speed limit; the speed dispersion decreased for rural primary and secondary roads.
In modeling the speed-related parameters he found an increase of 28 injury crashes per year for each increase of one mph in the speed dispersion on urban interstate highways. For rural primary highways the result was even more dramatic, with an increase of 60 injury crashes per year for each one mph increase in speed dispersion (p.49).

2.19 Bayesian analysis of speed data in Iowa

In this study on rural interstates in Iowa, Raju et al (1998) applied Bayesian techniques to evaluating changes in fatal crashes following the increase to a 65-mph speed limit in 1987. Bayesian techniques make use of prior information to predict future behavior and are especially useful in dealing with phenomena that are relatively rare and that may be subject to the regression to the mean phenomenon. The researchers evaluated both quarterly and annual crash rates and found that there were changes in these rates attributable to the increased speed limit; the quarterly frequency increased by 4 fatal crashes while the annual frequency increased by 9 fatal crashes. The report does not discuss the apparent difference between these estimates. There also is no discussion of traffic volumes or of other factors that may contribute to changes in fatal crashes. The authors state (p. 53) “It is important to note, however, that overall traffic safety may be improved by the increase, if sufficient to attract a large enough number of users from less-safe facilities.” It appears from the context that this “increase” is meant to be the increase in fatal crashes and that the intended meaning is that there would be a net reduction in fatal crashes statewide. The authors conclude (p. 54-55) “However, regardless of the specific number of fatalities resulting from increased speed limits, it is clear that increased speed limits have serious safety implications.”

2.20 Minimum Speed Research

In a recently completed study of traffic on rural interstate highways in Florida, Muchuruza and Mussa (2005) found that 9 percent of the crash-involved vehicles were traveling less than the 40 mph minimum speed; this in contrast to the 0.14 percent of the total number of vehicles that were traveling at that slow speed. They also determined that the minimum crash risk came at the 85th percentile speed. They found that the mean speed was
73 mph, or 3 mph above the speed limit, and that some 56 percent of drivers were exceeding the speed limit. It should be noted that their analysis of the crash speeds is derived from crash records, which speeds are lower than those measured at automated traffic monitoring stations. As they observe (page 5), “Examination of the two distributions show that the estimated pre-crash speeds are skewed to the left of the actual vehicle speeds collected at the sites.” The difference is significant, at least 5 mph from their plot of the two distributions, and it may be that this difference should be applied to the crash data. If this is the case it would be expected to reduce the percentage of crash involved vehicles at the low end of the speed distribution. It also calls into question the validity of crash speed estimates in crash records in general, at least those that are not derived from actual vehicle data. This possibility bears on the importance having good speed data on which to base decisions.

2.21 Urban Freeways in Los Angeles

Some interesting relationships were found between mean speed, lane-to-lane speed variation, and volume to types of collisions on on heavily traveled freeways in Southern California. Golob and Recker (2003) used 30-second slices of speed data from detectors in the vicinity of crashes to evaluate the impact of speed, weather, and lighting on crash types. Their evaluation covered three lanes of several freeways, having up to six lanes; these lanes were the innermost, the outermost, and one other lane in the group. Some of their findings are as follows:

- Crashes in the innermost lanes were related mostly to volume
- Crashes in the outermost lanes were related mostly to speed differences between adjacent lanes
- The severity of crashes was mostly related to volume, rather than speed.
- Only 13 percent of the crashes were related to wet weather conditions
• They were able to predict types of crashes in terms of traffic volume but they could not predict crash rates

Of particular interest to the current research is their use of the difference between the 90th and 50th percentile speeds as a measure of variation.

2.22 Speed Cameras in England

In a study of work zone traffic management, Freeman et al (2004) reported on an analysis of various control measures and their effectiveness at reducing the number of personal injury (and fatality) crashes in work zones. They reported that one of these control measures, the use of speed cameras, had a positive impact on work zone crashes but that these cameras were associated with a higher crash risk in areas outside the work zones. They suggest (p. 42), “However it should be noted that it appears in general sites with speed cameras were chosen as they were thought to have a high accident risk.” This may be a form of selection bias, in that due of the expense of installing speed cameras their use is reserved for areas that are experiencing above average crash risk. They did not report on the effectiveness of the cameras in these non-work areas, that is, if the addition of the cameras to high risk areas had had a positive impact on crash risk.

2.23 Conclusions

Over the course of the last 55 years there has been considerable effort expended in evaluating the relationship between speed and crash risk in a variety of roadway environments. Solomon concluded that it was the variation from the mean speed that presented greater crash risk, yet he did not control for factors such as turning vehicle crashes that would at least modify the shape of the curves he offered to demonstrate the risk. His estimates of the speed of crash involved vehicles was to some degree based on driver self-reporting (police reports provided other crash speeds); as noted by Muchuruza and Mussa even official crash reports appeared to be skewed to the left (toward lower speeds) when compared to the measured speed distributions in the areas of the crashes.
A number of researchers examined the change in crashes following the repeal of the national maximum speed limit of 55 mph, with varied conclusions. Chang and Paniati found no statistically significant change attributable to the speed limit change. Lave and Elias, while noting an increase in crashes on interstate highways found that for states that raised the speed limits on interstate highways there was a decrease in crashes overall (considering all highways in the states). Raju et al concluded that an increase in fatal crashes in Iowa was due to the change to a speed limit of 65-mph on rural interstate highways. Munandi found a significant increase in crash risk with increases in speed dispersion on rural interstate highways in Iowa, following the institution of a 65-mph speed limit on those highways. An examination of fatal crash rate data (taken from the Iowa Safety Management System report for 2002) for the period (Figure 3) shows that this finding may be due to variations in the data. It appears that there was a general long-term downward trend in the fatal crash rates for the rural interstates as well as for the state as a whole. For the three years preceding the change in speed limit the fatal rate on the interstates was low and this may have impacted their findings. It does appear that the rate of decline in the statewide fatal crashes rates was less following the change in speed limits and that the crash rates on the rural interstates flattened out. These data do seem to indicate that speed is related to crash risk, although as discussed by Kloeden et al (1997) it would be difficult to separate speed per se from variations in the speed distribution.

![Iowa Fatal Crash Rates Over Time](image)
There is a broad although not necessarily universal agreement among the researchers cited that there is a relationship between some element of speed and crash risk. Some of the researchers have tied crash risk to a broadly defined variation from mean speed (Solomon, Cirillo, e.g.) while others (Kloedien et al) have tied crash risk to absolute speed. There have been efforts to tie the change in the interstate speed limit from 55 mph to 65 mph to crash risk (Muniandi, Raju et al, Lave & Elias, e.g.) with somewhat varying results. In none of these cases was there an analysis of vehicle speeds and how they varied with changes in the speed limits. It may well be that the increased crash risk could be related to an increase in the variation of the speeds, as some drivers increase their speed more than other drivers after the speed limit is raised. The methods used have ranged from extensive interviews with drivers to statistical analyses based on aggregated data over wide geographic areas. Less common was any discussion of possible factors that would influence risk, only Kloedien et al (2002) discussed possible factors that could influence crash risk and they appear to have been the only ones that controlled for such factors as alcohol impairment. Some cases cited were narrowly based (West and Dunn 1971, for example) or used broad extrapolations of speed measures (Solomon 1964, for example). The need exists to examine a broad range of locations and roadway types, using speed data specific to each type and limited to representative sections. There is also a need to examine a wide variety of speed metrics (statistical and other measures of the variations in the speed distributions), in order to determine what if any of these metrics is related in a statistically significant manner to crash risk.
3 STUDY METHODOLOGY

The purpose of this study was to develop and evaluate a methodology for quantifying the relationship between vehicular speed and crash risk, utilizing available data sources. The process has involved a number of tasks, described in more detail in the balance of this chapter. The first task was to identify and locate the sites of the automatic traffic recorders (ATR) that capture speed data. The next task was to identify roadway sections, develop criteria for selecting study segments, and select appropriate study sections. Following a meeting with staff from the Iowa DOT's Office of Transportation Data, the next task involved the collection of six years of data (1998-2003) on six compact disks. Other data collection involved the acquisition of FHWA data regarding Vehicle Miles of Travel for 1950 and 2000; census data from the Census Bureau and the American Community Surveys of 1980, 1990, and 2000; and crash rate summaries from the Iowa DOT. The next major task was the processing and analysis of the speed data, utilizing Visual Basic code to assemble data and calculate descriptive statistics such as the mean speed, 85th percentile speed, and standard deviation and variance of the speed distributions. The processing and analysis of crash data, obtained from the Iowa DOT crash data bases, and selected using ArcGIS. The first step in the process was to develop an ArcGIS model with the Iowa DOT statewide road data base, a database containing the identification and geographic coordinates of the ATR's, and that portion of the DOT crash records containing time and date information covering the same time frame as the speed data. Roadway sections of interest were identified using the criteria described above to select homogeneous sections. Finally, statistical analyses were conducted on the case hour vs. control hour data. The case hour is the hour immediately prior to the hour in which the crash occurred and the control hour is the same hour but one week earlier. Using SAS, statistical evaluations were made of the relationships between the speed profiles in the roadway sections and the risk of crash involvement.
3.1 ATR Site Identification and Road Segmentation

The process of selecting candidate roadway segments was relatively straightforward. The latitude and longitude data for each ATR, provided by the Iowa DOT, were transformed to state plane coordinates within the Arc Map software. These transformed data were utilized to locate the ATR stations on the state road data base. One criticism of Solomon’s work, raised by several of the later researchers, is exemplified by the comment of Kloedén et al (p. 10), “It is difficult to comprehend how speeds measured at one location can be considered to be adequately representative of speeds on road sections up to 91 miles in length.” In order to minimize potential speed differences that might arise from using roadway segments that were too long, a restrictive procedure was used to select the candidate segments. Using the Iowa state road and aerial photographic bases, adjacent sections were evaluated visually to determine their homogeneity with the ATR sections in terms of the following criteria:

- Uniformity of access
- Type of development
- The presence of major alignment changes
- The presence of rail or highway crossings.

Uniformity of the speed limits was also a factor in selecting adjacent sections. On interstate highways and other freeways, homogeneous sections were established between the immediately adjacent interchanges on either side of the ATR. In other cases, the homogeneous sections were taken as bounded by a change in speed limit, by a change in the through traffic volume on the main line roadway as reported in the statewide road data base, by the location of a crossroad with traffic volume (as determined by the DOT listing of segment AADT) greater than ten percent of the traffic volume on the study section, or by a crossroad with more than a single crash recorded during the study period. Figure 4 presents the ATR locations in the context of Iowa’s major state highways; appendix D contains a summary table of all the speed-related ATR’s, with capsule descriptions.
3.2 Crash Data Collection

Crashes within each selected segment were selected utilizing the Iowa DOT’s crash data base and the State Road data base with the ESRI ArcMap software. Selected crashes for each location were exported as data base files for the analysis of the speed data. These files (generally one per year per ATR) were combined into a single data base with the roadway information from the DOT’s State Road Database, the several crash data files, and a summary sheet combining all of the crash data.

3.3 Speed Data Collection

Raw speed data were provided by the Iowa Department of Transportation Office of Transportation Data, covering the years 1998 through 2003 with minor gaps. These data were provided on compact disks, with one disk per year. Each disk contains one folder per month with each folder containing from 4,000 to 7,000 files. In a typical month more than 2,000 of these files contain speed data (the balance are index files or vehicle classification
The overall number of files is about 420,000; Figure 5 is a representation of the file structure and size.

For Each CD, One Folder Per Month

For Each Folder, One Index File per Day

For Each Day, ~120 Classification Data Files

For Each Day, ~80 Speed Data Files

Total ~420,000 Files!!!

Figure 5 The file management problem

The initial processing of the desired speed data from these files began with opening the space-delimited index file for the appropriate date and then visually identifying the appropriate data file. In the sample shown (Figure 6) ATR 100 is identified as being on files 3 and 4 for March 1, 2000. The first (D0301003.PRN) is the vehicle classification file and the second (D0301004.PRN) is the speed data file.
Each speed data file contains two, three or four rows of data per hour, consisting of one row for each lane of through traffic (roads with more than four lanes use one ATR for each direction of travel). Figure 7 presents a sample speed data file, again for March 1, 2000. The first three rows contain identification data including the location, time, and date. The fourth line contains the column headings for the speed bins with each bin being five miles per hour wide. The speed for each bin as listed is the maximum speed assigned to the bin; for the purposes of this paper the mid-point of the speed range within each bin is used. The first bin represents speeds 40 mph and lower; the “mid-point” of this bin was taken as 37.5 mph. Similarly, the last bin represents speeds from 85 to 147 mph; the “mid-point” of this bin was taken as 87.5 mph. These are a simplification but due to the generally small proportion of vehicles in these bins should not significantly impact the mean speed. The desired data file was then opened as a space-delimited text file and then the descriptive statistics for the desired hour were computed. These statistics are the mean speed, the 85th percentile speed, the dispersion (difference between the 85th percentile speed and the mean speed), the standard deviation of the speed distribution, the departure, the variance, and the volume. For
The standard deviation of the speed distribution is calculated as follows:

1. Calculate the mean speed for any given period.
   - The mean speed for any given period is calculated as follows:
   - Multiply each of these squares by the count in the respective speed bin; sum these products.
   - Calculate the difference between the mean speed and the "mid-point" speed of each bin.
   - Multiply the "mid-point" speed of each bin by the number of vehicles in that bin.

2. Immediately before the hour in which the crash occurred, and the control hour, the same hour of the day one week earlier. The case hour was used because it was thought that speed during the hour in which the crash happened would reflect the presence of the crash and not the immediate pre-crash conditions. The control hour one week earlier was chosen to reflect conditions that were representative of a normal, non-crash flow regime.

The mean speed for any given period is calculated as follows:

- Multiply the “mid-point” speed of each bin by the number of vehicles in that bin.
- Sum the products of all the bin multiplications; then divide this sum by the total volume of traffic during the stated period to calculate the mean speed.

The standard deviation of the speed distribution is calculated as follows:

- Calculate the difference between the mean speed and the “mid-point” speed of each of the speed bins; square each of these differences.
- Multiply each of these squares by the count in the respective speed bin; sum these products.

Figure 7 Sample Speed Data File
• Divide the sum of the products by the total vehicle count during the period to get the variance of the speed distribution; take the square root of the variance to yield the standard deviation.

Because of the magnitude of the data processing required, a Visual Basic program was developed to calculate the mean speeds, volumes, and variances for the case hour and control hours. The program is set up to access the data base for each ATR that was created as described above. It utilizes the date and time information to find the appropriate data in the ATR data files, calculates the basic statistics for the study and control hours of each crash, and adds the results to the data base for the ATR. It also performs several error checks on the data, checking for anomalies such as missing data or problems with data collection. The data for each crash or control hour were combined into data bases by type of facility for further processing including analysis using the Statistical Analysis Software (SAS) package. Other Visual Basic programs were developed to compute the hour-by-hour and the day-by-day distributions of speeds at each ATR. Code for these programs is included in Appendix A. Typical problems identified include the following:

• All traffic in a single speed bin, indicating a malfunctioning loop
• A lane volume that exceeds the single lane capacity, indicating another type of loop malfunction

The 85th percentile speed was calculated using a simple routine as follows:

• Calculate the cumulative traffic volume for each bin, including all the traffic in preceding bins
• Multiply the total volume for the hour by .85
• Using simple logical functions, determine the two bins whose cumulative volumes bracket the volume calculated in the second step
• By simple proportions calculate the 85th percentile speed

The 85th percentile speed for each case or control hour was entered into the data base for each ATR; the speed dispersion was then calculated by subtracting the mean speed from the 85th percentile speed. The departure is calculated by subtracting the speed limit from the
case hour 85\textsuperscript{th} percentile speed. The calculations were performed within the SAS data bases for use in the SAS modeling and within the categorical data bases (freeway, expressway, etc.) for the general statistics.

### 3.4 Descriptive Statistics

An example of the descriptive statistics for a typical crash is presented in Table 1. These statistics were used to develop “box and whisker” plots as an aid to visualizing the differences between the case and control hours. Figure 8 presents a sample plot for the data in Table 1. As may be seen, for this example the mean speed (the line across the middle of the box) in the immediate pre-crash hour is lower than for the week before. Also the minimum speed (the lower “whisker”) is somewhat lower and there is a greater spread between the 25 percent and 75 percent values (the lower and upper limits of the boxes, respectively).

<table>
<thead>
<tr>
<th></th>
<th>1 HOUR BEFORE</th>
<th>1 WEEK BEFORE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>66.31</td>
<td>68.06</td>
</tr>
<tr>
<td>Std Dev</td>
<td>5.18</td>
<td>3.75</td>
</tr>
<tr>
<td>Median</td>
<td>67.76</td>
<td>68.97</td>
</tr>
<tr>
<td>Q1</td>
<td>65.88</td>
<td>67.39</td>
</tr>
<tr>
<td>Q3</td>
<td>69.20</td>
<td>69.87</td>
</tr>
<tr>
<td>Min</td>
<td>48.50</td>
<td>53.33</td>
</tr>
<tr>
<td>Max</td>
<td>71.04</td>
<td>72.69</td>
</tr>
<tr>
<td>25th Pct</td>
<td>65.88</td>
<td>67.39</td>
</tr>
<tr>
<td>50th Pct</td>
<td>1.88</td>
<td>1.57</td>
</tr>
<tr>
<td>75th Pct</td>
<td>1.44</td>
<td>0.90</td>
</tr>
</tbody>
</table>
3.5 F-Test of Two Distributions

The next step in the process was to compare the variances of the speed distributions for the case and control hours, using the F-test to determine if the differences were significant at the 95 percent confidence level (one-tailed test at 0.025). Because most tables for the F-test have no detail beyond a degree of freedom of (typically) 120, an F-test calculator was used to obtain the “p” value for each pair of distributions. F-test comparisons were made on all crashes for which there were paired case-control hour speed data.

3.6 Correction for False Discovery

One of the problems identified in some statistical references is that of the false positive in the analysis of large data sets, sometimes referred to as the “data mining problem”. It relates to the fact that in a single test of significance there may be, for example, a 0.95 probability of correctly rejecting the null hypothesis. If a number of tests are conducted with the same probability criterion for each there is a significant chance that there will be a false positive. That is, some of the tests might lead to incorrectly rejecting the null
hypothesis. Using as an example a series of 20 tests, each with a 0.95 probability of supporting the null hypothesis, the possibility that none will be significant is $0.95^{20}$, or a value of 0.35849. The possibility of having at least one false positive is the difference, $1-0.35849 = 0.64151$.

In the current research the F-test has been applied to some 1,244 pairs of distributions. The one-tailed level-of-significance was taken as 0.025, equivalent to a 95 percent confidence interval. Applying the F-test to the data resulted in approximately 18 percent (218 cases) of the cases having a case hour distribution that was statistically different from the control hour distribution, i.e., rejecting the null hypothesis. The Benjamini-Hochberg False Discovery Rate (FDR) correction method was then applied to these data, resulting in a corrected rate of rejection of the null hypothesis of 13 percent (159 cases).

The application of the correction is relatively straightforward. When the analyses (including F-test) of all valid crash and speed data was complete, all of the “p” values were ranked in ascending order, and then the correction for each value was computed by dividing that value’s rank by the total number of values. The correction was then multiplied by the original “p” value. The result is that the largest value of “p” is not changed and that the smallest values of “p” are changed the most. In the cusp region around a value of 0.025 the correction increased the P values by about 3 times. Therefore the equivalent “raw” significance level to remain valid was about 0.008 rather than 0.025.

3.7 Crash Models

The statistical tools of SAS were used to model the relationship between the dependent variable crash risk and several explanatory variables including the following:

- Mean speed
- Standard deviation of the speed distribution
- Variance (of the speed distribution)
- Speed dispersion (the difference between the case hour 85th percentile and mean speeds)
• Speed departure (the difference between the case hour 85th percentile speed and the speed limit)
• Ratio of the traffic volume to the reported average daily traffic (ADT)

The type of model selected for these analyses was the logistic regression model. Logistic regression modeling was chosen because the response variable – whether a crash occurs or not – is categorical. Categorical variables are those that represent conditions or states and often take only the values of 0 and 1. The dependent variable evaluated is the crash occurrence; since it can only take two values, 0 and 1 (there was no crash or there was a crash) it is necessary to use logistic regression. The use of linear regression assumes continuous response variables, and that the explanatory variables are additive would not be appropriate because it also requires that the error in response variables are normally distributed and homoscedastic. The variables listed above are the continuous explanatory variables; the following variables are the categorical explanatory variables that were included in the data input to the modeling software:

• Time of day. “T” which takes any one of four states depending on the time of day. “T” is equal to 1 if the case hour is between 7:00 and 9:00 AM, to 2 if the case hour is between 4:00 and 6:00 PM, 3 if the case hour is between 11:00 PM and 5:00 AM, and 0 otherwise.
• Week day or week end

The input files for the SAS modeling were prepared using the following steps:
• The individual ATR data bases were combined into a single data base for each type of facility
• Categorical variables were set for crash status, time of day (morning rush hour, evening rush hour, or late night), type of roadway (freeway, expressway, or undivided), and weekend
• Formatting changes were made to make the files consistent with the requirements for SAS input files
The dispersion, departure, and hourly volume divided by the reported average daily traffic were entered.

In logistic regression the model takes the following form:

\[
\ln \left( \frac{P}{1-P} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3
\]  

(3.1)

In this case the dependent variable “P” is the probability of a crash and the explanatory variables are the dispersion, the departure, and the volume to ADT ratio, respectively. It should be emphasized that the logistic regression modeling was conducted on two cases, with approximately equal numbers of each, the immediate pre-crash hour and the control hour one week earlier. The probability “P” that results from this modeling reflects the methodology used to conduct the modeling and not the actual probability of a crash’s occurrence.

A sensitivity analysis was made of the logistic regression model results. This involved varying the value of one variable while holding the other variables at their respective average values (average for the type of facility).

An analysis was made of the relationship between the explanatory variables, to determine the extent of correlation. The multi-variate correlation method of the software package JMP was used for this evaluation. Table 2 presents the results of this analysis for the interstate highway segments.

<table>
<thead>
<tr>
<th></th>
<th>casemean</th>
<th>controlmean</th>
<th>case85</th>
<th>control85</th>
<th>vol</th>
<th>casedep</th>
<th>controldep</th>
<th>disp</th>
</tr>
</thead>
<tbody>
<tr>
<td>casemean</td>
<td>1.00</td>
<td>0.82</td>
<td>0.55</td>
<td>0.43</td>
<td>-0.25</td>
<td>0.55</td>
<td>0.43</td>
<td>-0.51</td>
</tr>
<tr>
<td>controlmean</td>
<td>0.82</td>
<td>1.00</td>
<td>0.34</td>
<td>0.57</td>
<td>0.00</td>
<td>0.34</td>
<td>0.57</td>
<td>-0.54</td>
</tr>
<tr>
<td>case85</td>
<td>0.55</td>
<td>0.34</td>
<td>1.00</td>
<td>0.79</td>
<td>-0.13</td>
<td>1.00</td>
<td>0.79</td>
<td>0.43</td>
</tr>
<tr>
<td>control85</td>
<td>0.43</td>
<td>0.57</td>
<td>0.79</td>
<td>1.00</td>
<td>0.10</td>
<td>0.79</td>
<td>1.00</td>
<td>0.35</td>
</tr>
<tr>
<td>vol</td>
<td>-0.25</td>
<td>0.00</td>
<td>-0.13</td>
<td>0.10</td>
<td>1.00</td>
<td>-0.13</td>
<td>0.10</td>
<td>0.13</td>
</tr>
<tr>
<td>casedep</td>
<td>0.55</td>
<td>0.34</td>
<td>1.00</td>
<td>0.79</td>
<td>-0.13</td>
<td>1.00</td>
<td>0.79</td>
<td>0.43</td>
</tr>
<tr>
<td>controldep</td>
<td>0.43</td>
<td>0.57</td>
<td>0.79</td>
<td>1.00</td>
<td>0.10</td>
<td>0.79</td>
<td>1.00</td>
<td>0.35</td>
</tr>
<tr>
<td>disp</td>
<td>-0.51</td>
<td>-0.54</td>
<td>0.43</td>
<td>0.35</td>
<td>0.13</td>
<td>0.43</td>
<td>0.35</td>
<td>1.00</td>
</tr>
</tbody>
</table>
The case and control departure are fully correlated to the case and control hour 85\textsuperscript{th} percentile speeds, as would be expected since they are derived therefrom. The dispersion is somewhat correlated to the case and control hour departures, although it is more highly correlated to the crash density.

Table 3 presents the correlation matrix for the metrics of the expressway speed data. Unlike the interstate case, the dispersion is highly, albeit negatively, correlated to most of the explanatory variables. There is a positive correlation between the dispersion and the case hour volume. Most importantly, there is not a great degree of correlation between the dispersion and the case or control hour departures. To reiterate, the following defines the explanatory variables in the correlation matrices:

- Case and control means – the mean speed during the case or control hours, respectively
- Case and control 85 – the 85\textsuperscript{th} percentile speed during the case or control hours, respectively
- Vol – the volume during the case hour
- Case and control dep – the difference between the 85\textsuperscript{th} percentile speed and the speed limit during the case or control hours (departure)
- Disp – the case hour dispersion (difference between the 85\textsuperscript{th} percentile and mean speeds)

<table>
<thead>
<tr>
<th></th>
<th>case mean</th>
<th>control mean</th>
<th>case 85</th>
<th>control 85</th>
<th>vol</th>
<th>case dep</th>
<th>control dep</th>
<th>disp</th>
</tr>
</thead>
<tbody>
<tr>
<td>case mean</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control mean</td>
<td>0.95</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case 85</td>
<td>0.99</td>
<td>0.91</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control 85</td>
<td>0.95</td>
<td>1.00</td>
<td>0.92</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vol</td>
<td>-0.70</td>
<td>-0.78</td>
<td>-0.68</td>
<td>-0.78</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case dep</td>
<td>0.42</td>
<td>0.12</td>
<td>0.49</td>
<td>0.13</td>
<td>-0.12</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>control dep</td>
<td>0.45</td>
<td>0.16</td>
<td>0.52</td>
<td>0.17</td>
<td>-0.21</td>
<td>1.00</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>disp</td>
<td>-0.96</td>
<td>-0.98</td>
<td>-0.93</td>
<td>-0.97</td>
<td>0.71</td>
<td>-0.20</td>
<td>-0.24</td>
<td>1.00</td>
</tr>
</tbody>
</table>
Table 4 presents the correlation matrix for the non-interstate freeways. For these facilities there is a high degree of correlation between the dispersion and the departures; there are also significant negative correlations between the dispersion and the other explanatory variables.

Table 4 Multivariate Correlation Matrix for Freeway Analysis Metrics

<table>
<thead>
<tr>
<th></th>
<th>Case mean</th>
<th>Control mean</th>
<th>Case 85</th>
<th>Control 85</th>
<th>vol</th>
<th>Case dep</th>
<th>Control dep</th>
<th>disp</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case mean</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control mean</td>
<td>0.90</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case 85</td>
<td>0.97</td>
<td>0.80</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control 85</td>
<td>0.91</td>
<td>0.96</td>
<td>0.86</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vol</td>
<td>-0.82</td>
<td>-0.88</td>
<td>-0.68</td>
<td>-0.75</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case dep</td>
<td>-0.40</td>
<td>-0.68</td>
<td>-0.18</td>
<td>-0.51</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>control dep</td>
<td>-0.66</td>
<td>-0.65</td>
<td>-0.52</td>
<td>-0.49</td>
<td>0.87</td>
<td>0.75</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>disp</td>
<td>-0.73</td>
<td>-0.87</td>
<td>-0.54</td>
<td>-0.73</td>
<td>0.92</td>
<td>0.86</td>
<td>0.83</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 5 presents the correlation matrix for the two-lane highway analysis variables.

Table 5 Multivariate Correlation Matrix for Two-Lane Highway Analysis Metrics

<table>
<thead>
<tr>
<th></th>
<th>Case mean</th>
<th>Control mean</th>
<th>Case 85</th>
<th>Control 85</th>
<th>vol</th>
<th>Case dep</th>
<th>Control dep</th>
<th>disp</th>
</tr>
</thead>
<tbody>
<tr>
<td>case mean</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control mean</td>
<td>0.89</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case 85</td>
<td>0.89</td>
<td>0.77</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>control 85</td>
<td>0.83</td>
<td>0.90</td>
<td>0.86</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>vol</td>
<td>-0.27</td>
<td>-0.14</td>
<td>-0.40</td>
<td>-0.34</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>case dep</td>
<td>0.89</td>
<td>0.77</td>
<td>1.00</td>
<td>0.86</td>
<td>-0.40</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
<tr>
<td>control dep</td>
<td>0.83</td>
<td>0.90</td>
<td>0.86</td>
<td>1.00</td>
<td>-0.34</td>
<td>0.86</td>
<td>1.00</td>
<td></td>
</tr>
<tr>
<td>disp</td>
<td>-0.05</td>
<td>-0.10</td>
<td>0.40</td>
<td>0.22</td>
<td>-0.34</td>
<td>0.40</td>
<td>0.22</td>
<td>1.00</td>
</tr>
</tbody>
</table>

In the case of the two-lane highways there is little correlation between the dispersion and the other explanatory variables. It is interesting to compare this matrix to the others; there is little consistency between the correlations on one type of facility to those on another.
4 RESULTS

As discussed in the methodology, analyses were made of the crash and speed data to include model evaluations of case vs. control hour speed metrics as well as sensitivity analyses of the resulting models, correlations between explanatory variables, and speed variations by time-of-day and day-of-year vs. crash count. The results of these analyses are presented in the following sections.

4.1 Speed as a function of time-of-day by roadway type

There is a difference of about 2.6 mph in the hourly mean speed for all interstate ATR’s over the course of an average day, ranging from a minimum of 68.1 mph at 5:00 AM to a high of 70.7 mph from 4:00 to 7:00 PM. As expected, there are also variations between the lanes; the inner lanes (those next to the median) on the typical, 4-lane interstate highway average from 3.0 to 3.7 mph faster than the outer lanes. This difference between lane groups also varies throughout the day, ranging from a low of 2.1 mph between midnight and 6:00 AM to a high of 5.0 mph from 4:00 to 5:00 PM. See Figure 9 for a summary of speed by lane.

![INTERSTATE HOURLY SPEED SUMMARY](image-url)

Figure 9 Interstate Highway Speed Summary by Time of Day
Figure 10 presents the crash frequency distribution for the interstate highways by time of day. Speed is found to explain 24 percent ($R^2 = 0.24$) of the variation in crash count when regressed by time of day.

On freeways (speed limits less than 65 mph) there is a variation of about 3.1 mph in the average speed by hour of the day; the overall average speed is 68.6 mph. Between the hours of 3:00 to 7:00 PM the average speed was nearly 69.5 mph; the hours from 2:00 to 4:00 AM had an average speed of approximately 66.5 mph. As with the interstate highways, there is a noticeable difference in the speed distributions by lane with the maximum difference of 3.3 mph occurring in the afternoon from 2:00 to 4:00. Figure 11 presents a summary of the distribution of the mean speed by lane and time of day. Figure 12 presents the distribution of freeway crashes by time of day. Speed is found to explain 24 percent ($R^2 = 0.24$) of the variation in crash count when regressed by time of day. The time of the lowest frequency of crashes corresponds to the lowest mean speed, although this is typically also the time of lowest volume traffic which may influence the results.
The results of the expressway analyses are somewhat limited by the small number of expressway ATR sites with speed data, which total only five. For some reason all of the sites had the highest speeds in lane number 3, with the mean speed in lane 3 at nearly 68 mph. Lane 2, the other inside lane, had a mean speed of just over 63 mph, lower by 1.5 mph than the mean speed in its companion outside lane (number 1). These variations may be due to commuting or travel patterns, or random errors in the data. Figure 13 presents a summary of the speed distribution on the expressways by lane and time of day.
Figure 13 Expressway Speed Summary by Time of Day

Figure 14 presents the crash count by hour of the day. Speed is found to explain 32 percent \((R^2 = 0.32)\) of the variation in crash count when regressed by time of day.

The hourly mean speed for the two-lane highways reaches a peak of 59.9 mph at 4:00 AM and a minimum of 58.6 mph at 10:00 PM. There is not an obvious relationship between the mean speed and the crash frequency distributions, as seen on Figures 15 and 16, except that the period of highest speed corresponds to the time of day of the lowest crash frequency.
This may be due to a low volume of traffic than to the speed. Speed is found to explain only two percent ($R^2 = 0.02$) of the variation in crash count when regressed by time of day.

![Two-Lane Highway Speed Summary by Time of Day](image1)

**Figure 15 Two-Lane Highway Speed Summary by Time of Day**

![Two-Lane Highway Crashes by Time of Day](image2)

**Figure 16 Two-Lane Highway Crashes by Time of Day**

### 4.2 Speed as a function of day of year by roadway type

As was discussed above, on the interstate highways the mean speed varies between lanes when evaluated over the course of a day. As Figure 17 shows this pattern is also present throughout the year. The mean speed for the outer lanes is 68.7 mph while the mean
speed for the inner lanes is 3.7 mph faster. There is also a seasonality to the daily mean speed, that is, the mean daily speeds tend to be slightly higher during the summer months. Possible explanations for this include the influence of weather (on winter travel) or vacation travel (related to higher volumes or to different travel purpose). There are two minima to the inner lane mean speeds that correspond roughly to the Labor Day and Thanksgiving Day holidays. At these times the lane differential nearly disappears.

Non-interstate freeways show a similar pattern to the interstates, with the inner lanes averaging about 3 mph faster than the 68 mph mean speed in the outer lanes. There appears to be more variation in the individual lane means, and a somewhat smaller difference between the inner and outer lane pairs. The daily variation of mean speed by lane is presented in Figure 18.
The expressways show a very different pattern, with Lane 3 having a mean speed of 68.4 mph, 3.7 mph higher than the other inner lane (Lane 2). Lanes 1 and 4 are very similar, with an identical mean speed of 61.9 mph. It should be noted that two of these five expressways have 55 mph speed limits (the other three are 65 mph). ATR 217, one of the two with a 55-mph speed limit and one with a very large difference between Lane 3 and the other lanes, has the highest crash rate and crash density observed on the expressways studied. Iowa DOT staff members have indicated that ATR 217 is not currently being used, because of problems with the ATR and its loops. However, none of the data for ATR 217 show any obvious problems that would dictate removing them from the study, although removing the ATR 217 data from the data base reduces the differential between lanes two and three. Figure 19 below presents the expressway mean speed by lane (data from ATR 217 included).
4.3 Correlation between dispersion and departure

Prior to developing the models, an evaluation was conducted to determine possible correlations between the candidate explanatory variables. It was expected that either dispersion or departure (or both) would be related to crash risk. If speed alone is the primary factor influencing crash risk then the expectation would be that the departure would be correlated to the crash risk. Conversely, if the variation in speeds between vehicles is the primary factor then the dispersion would be correlated to crash risk. Because the dispersion and departure are both speed related measures it was deemed important to evaluate the extent to which they are correlated. The use of correlated explanatory variables would violate the assumption of independence inherent in the regressions used. Using the hourly values of mean and 85th percentile speeds, which cover the entire six-year study period, average values of dispersion and departure were calculated for each ATR. The data were grouped by type and plotted, yielding Figures 20 through 23.
INTERSTATE HIGHWAYS DISPERSION VS. DEPARTURE

Figure 20 Dispersion vs. Departure Interstate Highways

y = -0.4633x + 7.3327
$R^2 = 0.7059$

FREEWAYS DISPERSION VS. DEPARTURE

Figure 21 Dispersion vs. Departure Freeways

y = -0.4937x + 7.2733
$R^2 = 0.355$
There is a significant level of correlation between the dispersion and the departure, with the degree of correlation being dependent on the type of facility. It is an interesting, if not unexpected, observation that the slope of the regression line for the two-lane highways is positive, versus the negative slope on all of the facilities with four lanes. This is likely an artifact of the nature of driving on two-lane roads, with their sometimes limited passing opportunities.
4.4 Evaluation of standard deviation of speed distributions

An analysis was conducted of the standard deviations of the case hour and control hour speed distributions for each category of ATR site. The case hour is the hour ending immediately prior to the time of the crash and the control hour is exactly one week earlier than the case hour.

4.4.1 Interstate Highways

The case hour speed distributions on the interstate highways (defined as rural, with 65mph speed limits) had a mean standard deviation of 6.12 mph versus 5.96 mph for the control hour speed distributions. The standard deviations of these standard deviations were 1.31 and 1.12, respectively. Figure 24 presents a box and whisker plot summarizing this comparison.

![INTERSTATE HIGHWAY STANDARD DEVIATIONS OF SPEED DISTRIBUTIONS](image)

Figure 24 Box and Whisker Plot of Interstate Highway Standard Deviations

4.4.2 Freeways

The case hour speed distributions on the freeways (non-interstates as well as interstates with speed limits less than 65 mph) had a mean standard deviation of 6.36 mph
versus 6.28 mph for the control hour speed distributions. The standard deviations of these standard deviations were 1.70 and 1.54, respectively. Figure 25 presents a box and whisker plot summarizing the analyses of these standard deviations.

![Box and Whisker Plot of Freeway Standard Deviations](image)

Figure 25 Box and Whisker Plot of Freeway Standard Deviations

### 4.4.3 Expressways

The case hour speed distributions on the expressways had a mean standard deviation of 6.93 mph versus 6.89 mph for the control hour speed distributions. The standard deviations of these standard deviations were 2.05 and 1.87, respectively. Figure 26 presents a box and whisker plot summarizing the analyses of these standard deviations.
4.4.4 Two-lane highways

The case hour speed distributions for the two-lane highways had a mean standard deviation of 6.14 mph \textit{versus} 6.21 mph for the control hour speed distributions. The standard deviations of these standard deviations were 1.26 and 1.43, respectively. Figure 27 presents a box and whisker plot summarizing the analyses of these standard deviations.
4.5 Crash risk vs. speed and volume measures

A number of possible correlations were evaluated for the different roadway types. Various explanatory variables, listed in the modeling subsection below, were compared to the crash rate (crashes per million vehicle miles) and crash density (crashes per mile per year). Models were examined for all of the comparison plots and the key findings are summarized in the following paragraphs.

4.5.1 Interstate highways

For the interstate highways the most significant model with regard to crash risk was the crash density versus overall average departure (the departure considering all of the data, not just the crash data). Figure 28 plots this relationship. Models for Figures 28 through 44 were developed using the statistical software package JMP.

The model and $R^2$ values for the data plotted in Figure 28 show a relatively strong relationship. The form of the model is similar to that found by Solomon, that is, concave upward. The polynomial (2nd order) model calculated for these data had an $R^2$ of 0.56. The model is strongly influenced by the data from ATR 117 (I-80 east of the east system...
interchange near Des Moines). With the values for ATR 117 included the mean of the crash density is 1.68, while with it removed the mean is 1.25. The standard deviation of the crash density is 1.68 with ATR 117 included and just 0.56 with it removed. The effect of the removal of ATR 117 from the analysis changes the $R^2$ of the model from 0.56 to 0.33. The crash density vs. the case hour departure showed no significant model with ATR 117. When the ATR 117 data are removed the model becomes somewhat significant with an $R^2 = 0.39$.

![Figure 28 Interstate Highway Crash Density vs. Overall Average Departure](image)

Figure 28 Interstate Highway Crash Density vs. Overall Average Departure

Figure 29 presents the scatter plot for the crash rate versus the case hour departure. The model had an $R^2$ of 0.22.
The model of crash density versus the case hour dispersion shows a significant relationship. The $R^2$ coefficient was 0.4852. Figure 30 presents the scatter plot of this comparison. Equation 4.1 presents the regression equation for the model, which reduces to equation 4.2.

$$498y = 1.1603409 (x - 3.25589)^2 + 0.511577 x - 0.802576 \quad (4.1)$$
\[ y = 1.1603 x^2 - 7.0443 x + 11. \]  \hspace{1cm} (4.2)

Figure 31 presents the scatter plot of the comparison of the crash rate to the case hour dispersion. The model had an \( R^2 \) of 0.388. Equation 4.3 is the model equation as returned by JMP, it reduces to equation 4.4.

\[ y = 0.0564 (x - 3.25571)^2 - 0.0006148 x + 0.14732 \]  \hspace{1cm} (4.3)

\[ y = 0.0569 x^2 - 0.3714 x + 0.7513 \]  \hspace{1cm} (4.4)

The analysis found little correlation between the crash density and the case hour speed departure. Figure 32 presents the scatter plot of these data.
INTERSTATE HIGHWAYS
CRASH DENSITY VS. CASE HOUR SPEED DEPARTURE

Figure 32 Interstate Highway Crash Density vs. Case Hour Departure

It should be noted that all of these study segments are in areas with a 65-mph speed limit; two interstate highways in areas with lower speed limits are evaluated in the freeways’ category because it is expected that these facilities have been designed using a lower design speed.

4.5.2 Freeways

Statistical analyses of the freeway data found significant relationships between the crash density and both the dispersion and the departure (both case hour). The crash density vs. departure, shown in Figure 33, showed an $R^2$ coefficient of 0.84. Equation 4.5 presents the model equation from the JMP regression which reduces to the form shown in Equation 4.6:

\[
y = 0.1654665 (x-6.24333)^2 + 0.973713 x -3.207252 \\
y = 0.1654 x^2 - 1.0926 x + 3.2451
\]  (4.5)  (4.6)
The results of the regression on the crash rate versus the case hour speed departure showed no correlation. The JMP analysis returned a 2\textsuperscript{nd} order polynomial regression model with an $R^2$ value = 0.049. Figure 34 presents the scatter plot of these data.

The crash density versus case hour dispersion analysis showed a strong correlation, with the 2\textsuperscript{nd} order polynomial model having an $R^2$ equal to 0.925. Figure 35 presents the
scatter plot of the analysis data. Equation 4.7 presents the model equation from the JMP regression which reduces to the form shown in equation 4.8.

\[ y = 0.4552337(x - 3.53)^2 + 2.0784969 x - 4.086517 \]  
\[ (4.7) \]

\[ y = 0.4516 x^2 - 1.1077 x + 1.5387 \]  
\[ (4.8) \]

Figure 35 Freeway Crash Density vs. Case Hour Dispersion

The linear regression of the crash rate versus the case hour dispersion showed no correlation, with an \( R^2 = 0.0004 \) on a linear model. Figure 36 presents the scatter plot of the crash rate vs. dispersion data.
4.5.3 Expressways

The results of the analysis of the crash density versus the case hour departure yielded a model with little correlation. The scatter plot of these data is presented as Figure 37. The \( R^2 \) value was equal to 0.0426. No model equations are presented.
The evaluation of the crash rate *versus* the *case hour* departure returned a linear model with an $R^2$ of 0.417. The scatter plot of these data is presented in Figure 38. Equation 4.7 presents the model equation.

$$y = 0.01204 x + 0.456509$$

(4.9)

The crash density *versus case hour* dispersion showed an $R^2$ of 0.857 on the polynomial model. Figure 39 presents the scatter plot of the data evaluated. Equation 4.10 presents the model equation which reduces to the form shown as equation 4.11.

$$y = 0.128411 (x - 4.61)^2 + 0.302259 x + 0.3314007$$

(4.10)

$$y = 0.1288 x^2 - 0.8851 x + 3.0679$$

(4.11)
The evaluation of the crash rate versus case hour dispersion showed an $R^2$ value of 0.426 on a 2nd order polynomial model. Figure 40 presents the scatter plot of these data. Equation 4.12 presents the model equation which reduces to the form shown in equation 4.13.

\[ y = 0.0228644 (x - 4.61)^2 + 0.0038321 x + 0.4607059 \]  

(4.12)

\[ y = 0.023 x^2 - 0.208 x + 0.949 \]  

(4.13)
It should be recognized that due to the small number of expressways sites, the statistical foundation is not particularly strong for arguing any conclusions from these results.

4.5.4 Two-Lane Highways

The analysis of the crash density versus case hour departure for the two-lane highways found a weak correlation, with an $R^2$ of about 0.223 for a polynomial model. Equation 4.14 presents the model from the JMP regression, which reduces to the form shown in equation 4.15. Figure 41 presents the scatter plot of the data.

$$y = -0.0110029 (x - 7.75909)^2 - 0.1371806 x + 1.8836334 \quad (4.14)$$

$$y = -0.0063 x^2 - 0.045 x + 1.5256 \quad (4.15)$$

![Two-Lane Highways Crash Density vs. Speed Departure](image)

Figure 41 Two-Lane Highways Crash Density vs. Case Hour Speed Departure

The analysis of the crash rate versus case hour speed departure found a weak correlation, with an $R^2$ of 0.1633 for a polynomial model. Equation 4.16 presents the model from the JMP regressions, which reduces to the form shown in equation 4.17 respectively. Figure 42 presents the scatter plot of the data.
\[ y = 0.0094525 (x - 7.75909)^2 - 0.0195665 x + 0.9896329 \]  
(4.16)

\[ y = 0.0094525 x^2 - 15.53775 x + 61.19311 \]  
(4.17)

The analysis of the crash density versus case hour dispersion found essentially no correlation; the $R^2$ values ranged from 0.03 to 0.05 for linear to second order polynomial models. No model equations were developed. Figure 43 shows the scatter plot for these data.
The analysis of the crash rate versus the case hour dispersion found no correlation, with the JMP linear model having an $R^2$ of 0.002. No equations are presented. Figure 44 presents the scatter plot of these data.

![TWO-LANE HIGHWAYS
CRASH RATE VS. DISPERSION](image)

Figure 44 Two-Lane Highway Crash Rate vs. Case Hour Dispersion

Table 6 presents a summary of the various average metrics utilized in these analyses, without weather related crashes.

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>INTERSTATE</th>
<th>FREEWAY</th>
<th>EXPRESSWAY</th>
<th>TWO-LANE</th>
</tr>
</thead>
<tbody>
<tr>
<td>CASE HOUR MEAN SPEED</td>
<td>69.2</td>
<td>64.9</td>
<td>62.7</td>
<td>58.7</td>
</tr>
<tr>
<td>CONTROL HOUR MEAN SPEED</td>
<td>70.0</td>
<td>66.3</td>
<td>63.5</td>
<td>58.8</td>
</tr>
<tr>
<td>CASE HOUR 85TH %ILE SPEED</td>
<td>72.4</td>
<td>68.5</td>
<td>67.3</td>
<td>62.2</td>
</tr>
<tr>
<td>CONTROL HOUR 85TH %ILE SPEED</td>
<td>73.5</td>
<td>69.9</td>
<td>67.5</td>
<td>62.7</td>
</tr>
<tr>
<td>DISPERSION (MPH)</td>
<td>3.3</td>
<td>3.5</td>
<td>4.6</td>
<td>3.8</td>
</tr>
<tr>
<td>DEPARTURE (MPH)</td>
<td>7.4</td>
<td>6.2</td>
<td>6.3</td>
<td>7.8</td>
</tr>
<tr>
<td>CRASH RATE (PER MVMT)</td>
<td>0.19</td>
<td>0.41</td>
<td>0.53</td>
<td>0.96</td>
</tr>
<tr>
<td>CRASH DENSITY (PER MI PER YR)</td>
<td>1.68</td>
<td>3.80</td>
<td>2.03</td>
<td>0.79</td>
</tr>
<tr>
<td>VOLUME</td>
<td>970</td>
<td>1231</td>
<td>535</td>
<td>117</td>
</tr>
<tr>
<td>VOL/ADT</td>
<td>0.044</td>
<td>0.043</td>
<td>0.052</td>
<td>0.052</td>
</tr>
</tbody>
</table>
Appendix B includes tables summarizing the speed metrics as well as plots of crash rate or crash density versus the various parameters evaluated, for each type of facility.

4.6 Crash Models

Logistic regression modeling was conducted on all crashes and then on only crashes that were not weather related. The following repeats from Chapter 2 the explanatory variables considered:

- Mean speed
- The variance of the case hour speed distribution
- Volume
- Dispersion (difference between the mean and 85th percentile speeds) – representative of the variation in the vehicle stream
- Departure (difference between the 85th percentile speed and the speed limit) – representative of the speed of the vehicle stream normalized for the speed limit
- Ratio of the hourly volume to the reported ADT for the study segment – a normalized measure of the vehicle volume
- Time of day
- Weekday or weekend

4.6.1 All crashes

The first model considered all of the crashes, including those that were related to weather or had weather as a contributing circumstance. The model for the freeways (combining interstate highways and other freeways) showed a strong correlation between the speed departure, speed dispersion and the ratio of the case hour volume to the segment ADT on one hand, and crash risk on the other hand. The mean speed did not enter the model as it is strongly correlated to the dispersion. Recall that the logistic regression model takes the following form:

\[
\ln \left( \frac{P}{1-P} \right) = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3
\]

(4.17)
In this case the dependent variable “P” is the probability of a crash and the explanatory variables are the dispersion, the departure, and the volume to ADT ratio, respectively. For the freeway case this equation becomes:

\[
\ln \left( \frac{P}{1-P} \right) = -0.0431 + 0.1523 \text{ (dispersion)} - 0.1141 \text{ (departure)} + 5.2652 \text{ (case hour volume/segment ADT)}
\]

(4.18)

The standard errors of the regression coefficients are (respectively) 0.16, 0.03, 0.01, and 1.85. The equation yields the following equation to solve for P:

\[
P = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3}}
\]

(4.19)

Table 7 presents a sample sensitivity analysis on P for values of departure on either side of the average. It is developed holding the dispersion and vol/ADT ratio constant at their respective average values while varying the departure. It should be emphasized that the probability “P” does not directly correlate to crash frequency or crash rate. The sensitivity analyses provide insight as to the impact of changes in the explanatory variables on this probability and thus an inference as to the impact of these changes on crash risk or crash frequency.

<table>
<thead>
<tr>
<th>Value of Explanatory Variable</th>
<th>intercept</th>
<th>dispersion</th>
<th>departure</th>
<th>vol/ADT</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-0.0431</td>
<td>0.1523</td>
<td>-0.1141</td>
<td>5.2652</td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>3</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>5</td>
<td>0.1</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>0.1</td>
<td>6</td>
<td>0.1</td>
<td></td>
</tr>
</tbody>
</table>

Average Freeway Values

<table>
<thead>
<tr>
<th>P/(1-P)</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.75</td>
<td>5.12</td>
</tr>
</tbody>
</table>

- \( P/(1-P) = 2.1177 \quad P = 0.67925 \)
- \( P/(1-P) = 1.8893 \quad P = 0.6539 \)
- \( P/(1-P) = 1.6856 \quad P = 0.62764 \)
- \( P/(1-P) = 1.5038 \quad P = 0.60061 \)
- \( P/(1-P) = 1.1548 \quad P = 0.53592 \) (Avg)
For the two-lane highways dispersion and departure entered the model. The model form is similar to that for the freeways; with the coefficients on the explanatory variables the equation becomes:

\[ \ln \left( \frac{P}{1-P} \right) = 0.3930 - 0.0457 \text{ (dispersion)} - 0.0573 \text{ (departure)} \] (4.20)

The standard errors of the coefficients are (respectively) 0.13, 0.02, and 0.02. The equation solving for \( P \) becomes:

\[ P = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2}} \] (4.21)

Table 8 presents a sensitivity analysis on "P" for values of departure on either side of the average while holding dispersion constant at its average.

<table>
<thead>
<tr>
<th>Value of Explanatory Variable</th>
<th>intercept</th>
<th>dispersion</th>
<th>departure</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>4</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>6</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Average Two-Lane Values:

\[ \frac{P}{1-P} = 3.91 \quad P = 0.50957 \]
\[ \frac{P}{1-P} = 6.98 \quad P = 0.49525 \]
\[ \frac{P}{1-P} = 0.98118 \quad P = 0.49525 \]
\[ \frac{P}{1-P} = 0.92654 \quad P = 0.48093 \]
\[ \frac{P}{1-P} = 0.87494 \quad P = 0.46665 \]
\[ \frac{P}{1-P} = 0.83061 \quad P = 0.45373 \quad \text{(Average)} \]

The logistic regression modeling of the expressway data found no significant results for any of the parameters considered; only the categorical variable \( W \) (weekday or weekend) entered the model and it was not statistically significant.
As noted above, sensitivity analyses were conducted of each of the models, to determine the impact of changing one of the explanatory variables on crash probability "P" while holding the other variables constant. Table 9 summarizes these analyses.

Table 9 Sensitivity Analysis of Model Results (For All Crashes)

<table>
<thead>
<tr>
<th>CHANGE IN VARIABLE</th>
<th>CRASH PROBABILITY CHANGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Each 1 mph change in dispersion yields a</td>
<td>+7.5%</td>
</tr>
<tr>
<td></td>
<td>-2.5%</td>
</tr>
<tr>
<td>Each 1 mph change in departure yields a</td>
<td>-5.1%</td>
</tr>
<tr>
<td></td>
<td>-2.9%</td>
</tr>
<tr>
<td>Each 1% increase in Vol/ADT ratio yields a</td>
<td>+2.3%</td>
</tr>
<tr>
<td></td>
<td>N/A</td>
</tr>
</tbody>
</table>

Average values of the explanatory variables were used in the models, as found during the data analysis process. Single variables were then changed in unit increments to determine the change in the dependent variable crash probability ("P"). In both models the results were nearly linear on the variables dispersion and departure within a range of 4 or 5 mph on either side of the average value. The volume/ADT ratio for the freeways' model was approximately linear from the mean value of about 0.04 to 0.11. In Table 9 the positive sign on the probability change indicates that the change in probability moves in the same direction as the change in the variable; the negative sign indicates that the change in probability has the opposite direction to the change in the variable. For example, a 1-mph increase in freeway dispersion yields a 7.5 percent increase in crash probability; a similar increase in departure yields a 5.1 percent reduction in crash probability. Figures 45 through 49 present plots of the models (those with the highest correlations).
Figure 45 Freeway Crash Probability vs. Dispersion

Figure 46 Freeway Crash Probability vs. Departure
Figure 47 Freeway Crash Probability vs. Volume/ADT

Figure 48 Two-Lane Highway Crash Probability vs. Dispersion
4.6.2 Non-weather crashes

When weather related crashes are removed there is a noticeable change in the results of the modeling. The modeling for the freeways (combining interstate highways and other freeways) shows a strong correlation between the speed departure, variance of the case-hour speed dispersion and the weekday status on one hand, and crash risk on the other hand. For the freeway case the logistic model equation becomes:

$$\ln \left( \frac{P}{1-P} \right) = 1.3272 + 0.00598 \text{ (variance)} - 1.8834 \text{ (status)}$$  \hspace{1cm} (4.23)

The standard errors of the coefficients are (respectively) 0.22, 0.003, and 0.19. Equation 4.23 yields the following equation to solve for $P$:

$$P = \frac{e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2}}{1 + e^{\beta_0 + \beta_1 x_1 + \beta_2 x_2}}$$  \hspace{1cm} (4.24)

Table 10 presents a sensitivity analysis of the freeways’ model.
Within this range of the variance the value of \( P \) increases by about 0.35 percent for each 1 mph\(^2\) increase in the variance of the case hour speed distribution. The change is the same for both the weekday and weekend cases. It is interesting to note that the weekend crash probability \( P \) is only about one-half of that for weekdays; however, if the exposure difference (5 days vs. 2 days) is taken into consideration there is a greater crash risk on the weekends than during the week. The weekend risk is about 24 percent greater than that during the week.

For the expressways the removal of the weather related crashes did not change the results of the modeling; in either case only the intercept entered the model.

When weather related crashes are removed from the input to the two-lane highways' model the dispersion entered the model. The model form is similar to that for the freeways; with the coefficients on the explanatory variables the equation becomes:

\[
\ln \frac{P}{1-P} = 0.2262 - 0.0483 \text{(dispersion)} \tag{4.25}
\]
The standard errors of the coefficients are (respectively) 0.12 and 0.02. The equation solving for probability “P” becomes:

\[ P = \left( e^{b_0 + \beta_1 x_1} \right) / \left( 1 + e^{b_0 + \beta_1 x_1} \right) \]  

Table 11 presents a sensitivity analysis of probability “P” for some sample values of the dispersion for the two-lane highways. The average value of the dispersion for all two-lane highways was 3.74 mph, which yielded a “P” value of 0.511. The sensitivity analysis indicates that a 1.0 mph increase in the case hour speed dispersion is associated with a 2.4 percent decrease in crash risk.

<table>
<thead>
<tr>
<th>DISPERSION</th>
<th>ODDS</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1.195</td>
<td>0.544</td>
</tr>
<tr>
<td>2</td>
<td>1.138</td>
<td>0.532</td>
</tr>
<tr>
<td>3</td>
<td>1.085</td>
<td>0.520</td>
</tr>
<tr>
<td>4</td>
<td>1.034</td>
<td>0.508</td>
</tr>
<tr>
<td>5</td>
<td>0.985</td>
<td>0.496</td>
</tr>
<tr>
<td>6</td>
<td>0.938</td>
<td>0.484</td>
</tr>
<tr>
<td>7</td>
<td>0.894</td>
<td>0.472</td>
</tr>
<tr>
<td>8</td>
<td>0.852</td>
<td>0.460</td>
</tr>
</tbody>
</table>

4.7 Weather Effects

Weather has a major impact on travel safety. Of the total 1,937 crashes within the ATR segments during the six years of data analyzed, 556 or 29 percent were weather related. On a percentage basis, weather-related crashes ranged from 0 percent (6 sites) to 63 percent (ATR 104 on I-35 in Hamilton County) of the total. If weather-related crashes are removed from the data base, the crash rates are affected as follows:

- Interstate highways – 0.32 crashes per Million Vehicle Miles of Travel (MVM) for all crashes and 0.19 without weather-related crashes.
- Freeways – 0.55 crashes per MVM for all crashes and 0.41 without weather-related crashes
- Expressways – 0.65 crashes per MVM for all crashes and 0.53 without weather-related crashes
• Two-lane highways – 1.26 crashes per MVM for all crashes and 0.96 without weather related crashes

With regard to the modeling of crash probability, the removal of weather related crashes had an important impact on the results. For the freeways’ model the impact was the removal of the dispersion and departure and the addition of the variance of the case hour speed distribution and the weekday/weekend status. For the two-lane highways the impact was to remove the speed departure from the model leaving only the dispersion.

The data in the current research show that the crash parameters during this 6-year range of data ranged as follows:

• Two-lane highways – crash rate of 1.26 per million vehicle miles of travel (MVMT); the crash density was 0.98 crashes per mile per year for all crashes. The crash rate for non-weather related crashes was 0.96; the density was 0.79.

• Expressways – crash rate of 0.65 per MVMT and crash density of 2.49 per mile per year for all crashes. For non-weather related crashes the rate was 0.53 and the density was 2.03.

• Freeways – crash rate of 0.55 per MVMT and crash density of 5.15 per mile per year for all crashes. For non-weather related crashes the rate was 0.41 and the density was 3.80.

• Interstate highways – crash rate of 0.32 per MVMT and crash density of 2.14 per mile per year for all crashes. For non-weather related crashes the rate was 0.19 and the density was 1.68.
5 CONCLUSIONS AND RECOMMENDATIONS

There are a number of issues that must be considered in evaluating the results of this study. These are discussed in the following sections.

5.1 Discussion of results

The modeling results and the statistics on the dispersion and departure suggest that speed is not the main issue, as these values on the freeways are very close to those for the interstate highways. However, the analyses do not show a single factor that can be consistently evaluated to suggest the crash risk on any type of roadway. This should not be unexpected, considering how differently the various categories of highways operate. For example, passing is clearly more problematic on two-lane highways than on multi-lane facilities. If variation from the mean speed is the causative factor, it would be reasonable to assume that dispersion, the difference between the 85th percentile and mean speeds, would be a good indicator of crash risk. While this is supported by the analyses for the divided facilities (interstates, freeways, and expressways), it is not supported for the two-lane roads based on the regression utilizing the overall data.

The logistic regression model for the two-lane highways shows that a 1-mph increase in dispersion results in a 2.5 percent reduction in crash risk. On freeways and interstates, that same 1 mph increase in dispersion is expected (based on the logistic regression model) to result in a 7.5 percent increase in crash risk. It should be emphasized that these conclusions come from the modeling, which is based on comparing the case hour to a control hour. One should exercise caution in extrapolating these results directly to crash rate or density. If only those crashes that are not related to weather are considered, the model results suggest that the variance of the speed distribution on the freeways has a positive correlation with the crash probability, such that a 1-mph$^2$ increase in the variance is associated with a 0.35 percent increase in probability.

The relation of departure to crash risk was found by the regression model to be more consistent between the freeways and two-lane roads, although the results are not consistent with the expectation that speed alone is the greatest risk factor on the highway. In both cases
(freeways and two-lanes) an increase in the departure is associated with a decrease in crash risk; for a 1 mph increase in the departure the freeway crash risk is expected to decrease 5.1 percent while for the two-lanes the reduction is 2.9 percent. It should be noted that this does not mean that speeding is safer. It is likely that speeding is acting as a surrogate for some other factor. The ability to drive at higher speeds is constrained as traffic volumes increase, which may indicate that volume is that factor. It may also be that speeds that are inappropriate for the conditions (such as traffic or weather) are still important.

On the freeways the model showed that an increase of 1 percent in the ratio of the case hour volume to the ADT is associated with a 2.3 percent increase in crash risk. This again would tend to suggest that the enforcement emphasis should be on the highest volume roadways, the urban freeways. Two of these, I-29/80 in Council Bluffs and I-74 in Bettendorf, have the highest crash rates and crash densities of the freeways segments in the study.

Do the results of the research reject the initial hypothesis, that there is a speed related parameter that will demonstrate a greater variation of distribution in the pre-crash hour than in the control hour one week earlier? Some of the speed metrics evaluated did demonstrate greater variation in the crash hour versus the control hour. The more important question is whether aggregated speed data can be used to determine if speed or variance contributes more to crash risk. Based on the current research it appears that the data from automatic traffic recorders do not provide sufficient detail to make that determination.

Based on the results of the analyses conducted in this study, there is a definite relationship between the dispersion of the speed distribution (the difference between the 85th percentile speed and the mean speed) and the probability of a crash when all crashes are considered. On freeways, each 1 mph increase in the speed dispersion is associated with a 7.5 percent increase in the crash probability; on two-lane highways the effect is a 2.5 percent decrease in the crash probability. There is also a relationship between the normalized speed of the traffic stream (as measured by the departure, the difference between the 85th percentile speed of the case hour and the speed limit) and the crash risk. On freeways each 1 mph increase in the departure is associated with a 5.1 percent reduction in the crash probability. On two-lane highways the effect is smaller, a 2.9 percent reduction in the crash probability.
One additional measure entered the model for the freeways; the ratio of the case hour volume to the reported ADT. In the freeways’ model a 1 percent increase in the volume/ADT ratio was associated with a 2.3 percent increase in the crash probability. There were more definitive results in some of the statistical analyses of the measures of departure and crash rate or density, for the interstate highways. The regression model produced a minimum crash density of 0.68 crashes per mile per year at a departure of 9.1 mph. This implies that a speed of 74 mph should be the safest speed. Solomon (see Figure 1) found the safest speed (on the rural highways he evaluated) to be about 65 mph. This speed was equal to or greater than the speed limits in approximately 75 percent of the roads he studied. It remains to be seen if the change to a 70 mph speed limit, established in Iowa in July of 2005, results in fewer crashes. Raju et al (1998) reported an increase in crashes on the rural interstates in Iowa when the speed limit was raised from 55-mph to 65-mph. It could be argued that as speed limits have increased, so have the speeds that people drive with the result that 85th percentile speeds are generally 5 to 10 mph over the speed limit.

If non-weather related crashes are considered alone, the results are quite different. For the freeways the logistic regression model showed that the variance of the case hour speed distribution was directly related to crash risk. It found that for each 1 mph increase in variance the crash risk is expected to increase by about 0.35 percent. The model also showed that for the freeways the weekend is about 29 percent more risky than a two-day period during the week. For the two-lane highways dispersion remained as a model variable but departure did not. According to the model each 1 mph increase in the dispersion is associated with a 2.3 percent reduction in crash risk.

5.2 Policy Implications

There are a number of policy implications that result from the analyses and conclusions of this study. They are discussed in the following subsections.

5.2.1 Speed Externality

As discussed in the introduction, an externality is defined as a cost (or benefit) that is imposed by one person’s actions on others. The dispersion of the speed distribution is somewhat linked to an increased crash risk; thus the driver whose driving behavior increases
the dispersion may be increasing the crash risk and thus imposing an external cost on the other drivers. Under this scenario, enforcement should be focused on minimizing the dispersion by enacting and enforcing proper speed limits, increasing minimum speed limits and enhanced enforcement of the speed limits.

5.2.2 Speed Enforcement

The results of the current research do not directly address the issue of speed enforcement, at least with regard to a direct relationship between speed and crash risk. Neither the mean speed nor the speed departure was found to be significantly correlated to crash risk on the interstate highways or freeways. The current research determined that the mean speed on Iowa's interstate system is over 70-mph, more than 5-mph over the speed limit (at the time of the data collection). Fewer than 16 percent of the drivers travel within the speed limit and the 85th percentile speed is nearly 75-mph.

Although it would be tempting to link the findings of this research with recommendations for setting speed limits or their enforcement, caution should be exercised in advocating such a linkage. All of these analyses have been based on aggregated data, those speed data from either the case hour or the control hour. These speed data are grouped into 5-mph bins and thus do not permit the determination of the actual speed of crash-involved vehicles. It is not necessarily intuitive to assert that increasing the speed of the vehicle stream will reduce the crash risk. It may be more reasonable to conclude that the results indicate that there is something happening some time before the crash that makes the dispersion increase and the 85th percentile speed decrease. It may be advisable to consider the development of a predictive model to identify these deviations in real-time, thus perhaps permitting the intervention of law enforcement prior to a crash in order to prevent the crash from occurring. With regard to two-lane highways, it may also be the case that the dispersion of the traffic stream must decrease with increasing volume. An analysis of this relationship found a weak correlation between the case hour volume and the dispersion; however, it should be noted that the highest value of the case hour volume was 364, not a significant percentage of a two-lane highway's capacity. It does appear that consideration
should be given to setting and enforcing more realistic speed limits for weather conditions that create greater crash risk.

5.2.3 Automated Enforcement

The results of the current research cannot be construed as supporting a specific conclusion with regard to automated enforcement. There are issues that would have to be addressed if automated speed enforcement is to be used. Privacy, accuracy, and spoofing are some of these issues. These were discussed in the introduction. Advanced technology radar systems could also be used, such as frequency-agile systems where the frequency is changed on a millisecond basis, to frustrate both radar detectors and spoofers. There is a new passive radar detection system that makes use of signals from cell-phone towers which works by detecting the Doppler shift in the background radiation. Called CELLDAR (for cell phone radar), it has been tested on a proof-of-concept basis and presumably could be used to detect speeding vehicles without emitting radiation.

5.2.4 Automated Speed Control

Because the current research indicates that dispersion may be related to crash risk, there may be a benefit from the control of vehicle speeds to minimize variation of vehicle speed. One method that could be used is the automated control of speeds of vehicles. There are policy issues that arise from the possible use of automated speed control, as follows:

- Whether the posted speed limit is always appropriate, such as for weather problems, construction zones, or high volume traffic (if not modified to reflect these differing conditions)
- Variation in speed determination of vehicles, which could be addressed by simple GPS systems to report the exact vehicle speed to the vehicle operator
- How to pay for a system
- Who should pay for the system
- How to deal with older vehicles
How to deal with attempts to defeat the speed control system, such as by spoofing devices that would provide an excessive speed limit to the speed control system

5.3 Limitations of Current Research

It is well known that crashes are rare events. The total count of crashes in the study segments is 1,937, over 6 years, or about 323 per year. Over the study period the average number of reported crashes (according to Iowa DOT data) was about 62,000. The study area crashes thus represent approximately \( \frac{1}{2} \) percent of the total. Expanding the number of sites for which speed data are available would help in improving the statistical base for such a study. As discussed above, emerging technology such as CELLDAR has the potential to provide speed data unobtrusively and relatively cheaply, so it should be possible to generate speed data at crash “hot spots” or at a wider variety of locations.

The current research relies on aggregated speed data, from the state’s network of ATR’s. While some conclusions have been reached about the statistics of crash risk from certain speed measures, the overarching question remains somewhat unanswered, “does speeding in and of itself creates a crash risk, or is it the variation of that speeding vehicle from the speeds of the other vehicles in the traffic stream that creates the risk?” It appears that answering this question will require the use of disaggregated data, that is, the speeds of crash-involved vehicles as well as the speed profile of the traffic stream. To obtain these data may require access to the event data recorders (EDR) that are present in most vehicles equipped with air bags. Current versions of these devices record the speed, brake switch status, throttle position, seat belt use, and air bag deployment for (typically) five seconds prior to the deployment. Each vehicle manufacturer has a proprietary code for their device; at the present time only Ford and GM have made this code available to the public. One technology firm has developed software to permit the recovery of these pre-crash data from EDR-equipped Ford and GM vehicles. At the present time only law enforcement agencies have access to these data and in Iowa they are generally only accessed in the case of fatal crashes. Access to these data could facilitate future extensions of the current research for all crashes in those areas where useful speed profile information can be obtained. There are also social barriers to the acquisition of EDR data, including privacy and 5\textsuperscript{th} Amendment
considerations. Access to these data, stripped of personal identifiers, would be very useful to future studies on the relations between speed and crash risk.

5.4 Suggestions for Further Research

At the present time in Iowa the DOT’s Office of Transportation Data recovers individual vehicle data from certain weigh-in-motion (WIM) ATR sites. These data are processed to reformat them into the 5-mph bin, aggregated data such as were used in this research. Retention of these data in their original format (or processed to reduce their size while retaining their information content) would be helpful for future research.

CELLDAR-like technologies would seem to promise the ability to accurately determine the speed of individual vehicles and the vehicle stream, without affecting drivers’ behavior. However, while proved in concept, the technology should be tested with regard to its capability to reliably and rapidly detect the speed of individual vehicles in a traffic stream under all sorts of environmental conditions and locations.

Over the time frame of the data used in the current research there have been changes in vehicle technology and driver distractions. Still unknown are potential changes in speed behavior of drivers using cell phones and there are devices available now that permit the identification of vehicles in which a cell phone is being used. Using the passive speed measuring technology discussed above, it should be possible to measure and compare the speed profiles of cell-phone users vs. non-users.

One avenue of research that would be expected to receive great interest from governmental safety agencies is into the identification in real time of potential traffic problems via monitoring and sensing algorithms. If a law enforcement agency could be alerted to the onset of conditions that have an increased crash risk, such as an increase in the dispersion, they might be able to take action to modulate the traffic to reduce the risk. This could also be used with a system of automated speed control as discussed above, to reduce the speed of vehicles approaching an anomalous condition and thus mitigate the impacts of that condition.
Weather-related crashes play a significant role in crash performance. Weather-related speed limits intended to reduce crash risk could be studied, especially with regard to driver awareness of the need for such limits (Comte et al 1997).

Finally, the State of Iowa recently (July 2005) raised the speed limit on rural interstate highways to 70 mph. A future study should evaluate the impact of this change on speed metrics such as the dispersion and departure, as well as to see what changes have occurred in crashes and in the volumes on other highways.
REFERENCES

Des Moines Register, April 20, 2005


Gonzalez, Javier S., Elio R. Espino, and Albert Gan. *Effects of Regression to the Mean in the High Crash Identification Process Using the Rate Quality Control Method.* Presented at the 2004 Annual meeting of the Transportation Research Board.


Box and whisker plot method for Excel, accessed at http://peltiertech.com/Excel/Charts/BoxWhiskerV.html

F-test calculator accessed at (Hyperstat Online at http://www.davidmlane.com/hyperstat/F_table.html)


**ACKNOWLEDGMENTS**

I would like to acknowledge the support of a number of individuals and agencies during the preparation of my research. The following lists each of them; the order in which they are listed has no bearing on their relative importance or their contribution:

- Bob Thompson of the Governor's Traffic Safety Bureau of the State of Iowa, whose agency provided much encouragement and funding for this research
- Tom Welch of the Iowa Department of Transportation's Office of Traffic and Safety, whose support for several research projects provided many opportunities for publication and whose crash data were pivotal in my research
- The Iowa Department of Transportation Office of Transportation Data, who provided speed data files and other important information to facilitate my research
- Dr. Reginald R. Souleyrette, my major professor, for his support and guidance in the completion of my research
• Ann Stout, my wife, whose constant support, encouragement, and tolerance kept me going throughout this long process
• Hank Zaletel, Iowa DOT librarian, whose ability to get copies of obscure reports made things a lot easier
• Xu-Dong Chai, fellow graduate student, whose Visual Basic programs made my data analysis manageable
Code for ATR Query Module

Private qYear, qMonth, qDay, qFileName As String
Private qATR, IntCrashTime As Integer

Private Sub Command1_Click()

    Dim qDate, NewDate, sNewDate, ExcelFileName As String
    Dim LaneN As Integer
    Dim MSpeed As Single, SDev As Single, Varia As Single, Vol As Single
    'Cannot write it like this:
    ' Dim MSpeed, SDev, Varia, Vol As Single
    ' Otherwise type missing
    Dim StartPos, tempMonth, tempDay As Integer
    Dim WeekBefore As Integer
    Dim CrashNum, k, qtime As Integer
    Dim CrashYear, CrashMonth, CrashDay, CrashTime As Integer
    Dim ErrorMsg As String

    CommonDialog1.ShowOpen
    ExcelFileName = CommonDialog1.FileName

    Dim myExl As Excel.Application
    Dim wb As Workbook
    Dim ws As Worksheet
    ' Dim var As Variant
    Set myExl = New Excel.Application
    Set wb = myExl.Workbooks.Open(ExcelFileName)
    Set ws = wb.Worksheets("Base")
    '-> Set ws = wb.Worksheets("sheet1")
    ' var = ws.Range("A1").Value
    ' or
    ' var = ws.Cells(1, 1).Value
    ' ws.Cells(1, 1).Font.Bold = True
    myExl.Visible = True
    ws.Cells(1, 9) = "Calculation:"
    ws.Cells(1, 10) = "MeanSpeed"
    ws.Cells(1, 11) = "StdDev"
    ws.Cells(1, 12) = "Variance"
    ws.Cells(1, 13) = "Volumn"
    ws.Cells(1, 15) = "bMeanSpeed"
    ws.Cells(1, 16) = "bStdDev"
    ws.Cells(1, 17) = "bVariance"
    ws.Cells(1, 18) = "bVolumn"
qATR = ws.Cells(2, 1)
CrashNum = 0
Do

CrashNum = CrashNum + 1

qATR = ws.Cells(CrashNum + 1, 1)
CrashMonth = ws.Cells(CrashNum + 1, 4)
CrashDay = ws.Cells(CrashNum + 1, 5)
CrashYear = ws.Cells(CrashNum + 1, 6)
CrashTime = ws.Cells(CrashNum + 1, 8)
WeekBefore = 0
qDate = CStr(CrashMonth) + "/" + CStr(CrashDay) + "/" + CStr(CrashYear)

NewDate = qDate 'if in the hour of 2400, NewDate will be changed.
qtime = Fix(Crashtime / 100)
't = Mid(qTime, 3, 2)
IntTime = CInt(Left(qtime, 2)) * 100
If qtime < 1 Then 'Use the hour of 2400, one day before the crash happened
NewDate = DateAdd("d", -1, qDate)
IntCrashTime = 2400
Else
IntCrashTime = qtime * 100
End If
sNewDate = NewDate
quit4: qYear = Right(NewDate, 2)
StartPos = InStr(1, NewDate, Chr(47), vbTextCompare)
tempMonth = CInt(Mid(NewDate, 1, StartPos - 1))
If tempMonth < 10 Then
qMonth = "0" + CStr(tempMonth)
Else
qMonth = CStr(tempMonth)
End If
NewDate = Right(NewDate, Len(NewDate) - StartPos)
StartPos = InStr(1, NewDate, Chr(47), vbTextCompare)
tempDay = CInt(Mid(NewDate, 1, StartPos - 1))
If tempDay < 10 Then
qDay = "0" + CStr(tempDay)
Else
qDay = CStr(tempDay)
End If
'time 2359 means that you need to use 2300 data. 2300 is the end of that time slot.
qFileName = "c:\ATR\" + qYear + "/" + qYear + qMonth + "ATR\" _
Call ATRProcess(ErrorMsg, MSpeed, SDev, Varia, Vol, LaneN)

If ErrorMsg <> "" Then
    ws.Cells(CrashNum + 1, 9) = " -->"
    ws.Cells(CrashNum + 1, 19) = ErrorMsg
Else
    If WeekBefore <> 111 Then
        If LaneN = 4 Then
            ws.Cells(CrashNum + 1, 9) = " -->"
        Else
            ws.Cells(CrashNum + 1, 9) = CStr(LaneN) + "Lane"
        End If
    Else
        ws.Cells(CrashNum + 1, 10) = MSpeed
        ws.Cells(CrashNum + 1, 11) = SDev
        ws.Cells(CrashNum + 1, 12) = Varia
        ws.Cells(CrashNum + 1, 13) = Vol
        NewDate = DateAdd("d", -7, sNewDate) 'Calculating the data one week before
        If Year(NewDate) < 1998 Then
            ws.Cells(CrashNum + 1, 9) = " -->"
            ws.Cells(CrashNum + 1, 19) = "No 1997 Data Available"
            GoTo quit3
        End If
        WeekBefore = 111
        GoTo quit4
    End If
End If
End If

quit3: Loop Until IsEmpty(ws.Cells(CrashNum + 2, 3))
wb.Save
wb.Close
myExl.Quit
Set ws = Nothing
Set wb = Nothing
Set myExl = Nothing

End Sub
Private Sub Command3_Click()
    Unload Form1
End Sub

Private Sub ATRProcess(ErrMsg As String, MeanSpeed As Single, StdDev As Single, _
            Variance As Single, TotalSC As Single, LaneNum As Integer)

    Dim NN, N, sATR As Integer
    Dim ExcelOpen As Integer
    Dim StringRow, tempFileName, sFileName, sFolder As String
    Dim p As Long
    Dim i, j, CRow As Integer
    Dim HourFound, SpeedCheck, WrongSpeed, CurrentRow, CurrentColumn As Integer
    Dim SpeedCount(20), MidBin(20), CountMultiplyBin(20), TotalCMB As Single
    Dim Diff(20), DifSqCount(20), SumDSC As Single

    Dim FSObj As Object
    Set FSObj = CreateObject("Scripting.FileSystemObject")

   ErrMsg = ""
    On Error GoTo OpenError
    ExcelOpen = 0

    NN = 0
    Open qFileName For Input As #1
    On Error Resume Next 'Without this, the code will crash when encountering abnormal 'characters in the index file
    Do Until EOF(1)
        Line Input #1, StringRow
        sATR = Mid(StringRow, 23, 4)
        If qATR = sATR Then
            If NN = 0 Then
                NN = 999
            Else
                NN = 99
                tempFileName = Mid(StringRow, 2, 12)
            Exit Do
            End If
        End If
    Loop
    Close #1

    On Error GoTo 0
    If NN = 0 Then
ErrMsg = "Invalid ATR Number"
Exit Sub
End If
If NN = 999 Then
ErrMsg = "No Speed File Available"
Exit Sub
End If

Dim exWBook As Object
Dim exSheet As Object
Set ex = CreateObject("Excel.Application")
Set exWBook = ex.Workbooks().Add
Set exSheet = exWBook.Worksheets("sheet1")
exWBook.Visible = True
N = 0
On Error GoTo OpenError
sFileName = "c:\ATR\" + qYear + ";" + qYear + qMonth + "ATR\" + tempFileName
ExcelOpen = 9
Open sFileName For Input As #2
p = 35
For i = 1 To 10
  p = p + 5
  ex.Cells(5, i + 4) = p
  MidBin(i + 4) = p - 2.5
Next
ex.Cells(5, 15) = 147
MidBin(15) = 87.5
On Error Resume Next
Do Until EOF(2)
  Line Input #2, StringRow
  N = N + 1
  If N <= 6 Then GoTo quit1
  ex.Cells(N, 1) = CInt(Mid(StringRow, 1, 2))
  ex.Cells(N, 2) = CInt(Mid(StringRow, 4, 2))
  ex.Cells(N, 3) = CInt(Mid(StringRow, 7, 1))
  ex.Cells(N, 4) = CInt(Mid(StringRow, 9, 4))
  For i = 1 To 11
    p = 9 + i * 5
    ex.Cells(N, i + 4) = CInt(Mid(StringRow, p, 4))
  Next
quit1: Loop
Close #2 'close that *.PRN file
On Error GoTo 0
ex.Cells(N + 2, 1) = "Speed Count(an hour before the crash)"
ex.Cells(N + 3, 3) = "Mid Pt"
ex.Cells(N + 4, 3) = "No. * Bin Speed"
ex.Cells(N + 5, 3) = "Difference"
ex.Cells(N + 6, 3) = "DifSq * Count"
ex.Cells(N + 8, 5) = "Mean Spd"
ex.Cells(N + 8, 6) = "Std Dev"
ex.Cells(N + 8, 7) = "Variance"
ex.Cells(N + 8, 8) = "Volume"

CRow = 7
HourFound = 0
LaneNum = 0
WrongSpeed = 0
For j = 1 To 20
  SpeedCount(j) = 0
Next
Do
  If ex.Cells(CRow, 4) = IntCrashTime Then
    LaneNum = LaneNum + 1
    ' Check the correctness of the original speed data, if SUM(each column)<>the value of first cell?
    SpeedCheck = 0
    CurrentColumn = 5
    Do
      SpeedCheck = SpeedCheck + ex.Cells(CRow, CurrentColumn)
      CurrentColumn = CurrentColumn + 1
    Loop Until IsEmpty(ex.Cells(CRow, CurrentColumn))
    If ((ex.Cells(CRow, 5) = SpeedCheck) And (SpeedCheck <> 0)) Then
      WrongSpeed = WrongSpeed + 1
    End If
  HourFound = 99
  ' CRow = CRow + Slot - 1 'the 15-minute slot includes the crash time
  For j = 5 To 15
    ' SpeedCount(j) = ex.cells(CRow, j)
    ' For i = CRow - 1 To CRow - 3 Step -1
    ' For i = CRow To CRow + 3
      SpeedCount(j) = SpeedCount(j) + ex.Cells(CRow, j)
    Next
  Next
  Exit Do
End If
CRow = CRow + 1
Loop Until IsEmpty(ex.Cells(CRow, 4))

If HourFound = 0 Then
    ErrMsg = "No Speed Data Found in that Hour"
    ex.Cells(N + 9, 5) = "No Speed Data Found in that Hour"
    'Have to save the original speed data like this, otherwise two Excel objects will conflict.
    GoTo quit5
End If

' If original data doesn't have correct data format ...
If (CurrentColumn < 16) Or (WrongSpeed >= 2) Then
    ErrMsg = "Wrong Speed Data"
    ex.Cells(N + 9, 5) = "Wrong Speed Data"
    GoTo quit5
End If

'Calculate the speed distribution
TotalSC = 0
TotalCMB = 0
For j = 5 To 15
    ex.Cells(N + 2, j) = SpeedCount(j)
    TotalSC = TotalSC + SpeedCount(j)
    ex.Cells(N + 3, j) = MidBin(j)
    CountMultiplyBin(j) = SpeedCount(j) * MidBin(j)
    ex.Cells(N + 4, j) = CountMultiplyBin(j)
    TotalCMB = TotalCMB + CountMultiplyBin(j)
Next
ex.Cells(N + 2, 16) = TotalSC

'Check the original speed distribution, in case divided by zero.
If TotalSC = 0 Then
    ErrMsg = "Total Speed Count = 0"
    ex.Cells(N + 9, 5) = "Total Speed Count = 0"
    GoTo quit5
End If

'Continue calculating the speed distribution
ex.Cells(N + 4, 16) = TotalCMB
MeanSpeed = Round(TotalCMB / TotalSC, 2)
ex.Cells(N + 9, 5) = MeanSpeed

SumDSC = 0
For j = 5 To 15
    Diff(j) = MidBin(j) - MeanSpeed
    ex.Cells(N + 5, j) = Diff(j)
DifSqCount(j) = (Dif(j))^2 * SpeedCount(j)
SumDSC = SumDSC + DifSqCount(j)
ex.Cells(N + 6, j) = DifSqCount(j)
Next
ex.Cells(N + 6, 16) = SumDSC
StdDev = Round((SumDSC / TotalSC)^0.5, 2)
Variance = Round(SumDSC / TotalSC, 2)
ex.Cells(N + 9, 6) = StdDev
ex.Cells(N + 9, 7) = Variance
ex.Cells(N + 9, 8) = TotalSC
GoTo quit5

OpenError:
' MsgBox "Error " & Format$(Err.Number) & " opening file." & vbCrLf & _
Err.Description
If ExcelOpen = 9 Then
   ErrMsg = "Error " & Format$(Err.Number) & _
     ".There should be Speed Data file. But, " & Err.Description
   GoTo quit6
Else
   ErrMsg = "Error " & Format$(Err.Number) & _
     ".opening ATR Index file." & Err.Description
End If
Set FSObj = Nothing
Exit Sub 'do not allow an error handler to continue to the routine's End statement'
'Use an Exit statement to leave the routine.

'quit5: sFileName = "c:\ATRProcessed\" + CStr(qATR) + "+qMonth + qDay + 
   CStr(IntCrashTime) + "+" + qYear + ".xls"
quit5: sFolder = "c:\ATRProcessed\" + CStr(qATR) + \" + qYear + \\n'Only one subfolder can be created at one time If FSObj.FolderExists(sFolder) = False Then
   FSObj.CreateFolder sFolder
End If
' MkDir (sFolder) ' Another way to create a new Excel folder 
sFileName = sFolder + CStr(qATR) + "+qMonth + qDay + "+CStr(IntCrashTime)
+ "+qYear + ".xls"
exWBook.SaveAs sFileName
quit6: ex.Quit 'quit Excel
Set exWBook = Nothing
Set exSheet = Nothing
Set ex = Nothing

Set FSObj = Nothing
End Sub
ATR by 24 Hours

Private Declare Function APIBeep Lib "kernel32" Alias "Beep" (ByVal dwFreq As Long, ByVal dwDuration As Long) As Long
Private qYear, qMonth, qDay, qFileName As String
Private IntCrashTime As Integer
Private tempATR, qATR As String
Private SpeedCount(5, 20) As Long
Private RawSpeed(5, 20) As Integer

Private Sub Command1_Click()
    Dim TotalCount(S) As Long
    Dim TotalATR, ATRNum As Integer
    Dim StringRow, tempFileName, sFileName, stFileName, sFolder As String
    Dim MidBin(20), CountMultiplyBin(20) As Single
    Dim TotalCMB(5) As Single
    Dim MeanSpeed(S), Var(5) As Single
    Dim Pre85(5), EightyFifth(5) As Single
    Dim AccumCount(5) As Long
    Dim byHour As Integer
    Dim LaneNum As Integer
    Dim temp As Integer
    ' Dim Sheet_count As Variant
    ' Dim MyPos
    Dim LaneN, NN, N As Integer
    Dim MSpeed As Single, SDev As Single, Varia As Single, Vol As Single
    Dim SqDiff(5, 20), SumDSC(5) As Single
    Dim ATRDay As Integer
    Dim HourCount(S) As Integer
    Dim ErrType As Integer
    Dim qDate, NewDate, sNewDate, ExcelFileName As String
'Cannot write it like this:
' Dim MSpeed, SDev, Varia, Vol As Single
' Otherwise type missing
    Dim StartPos, tempMonth, tempDay As Integer
    Dim WeekBefore As Integer
    Dim CrashNum, k, qtime As Integer
    Dim CrashYear, CrashMonth, CrashDay, CrashTime As Integer
    Dim ErrorMsg As String
    Dim ii As Integer
    Dim FTestResult As String
' select an Excel file to process ATR one by one
    CommonDialog1.ShowOpen
    ExcelFileName = CommonDialog1.FileName

    Dim myExl As Excel.Application
    Dim wb As Workbook
    Dim ws As Worksheet
    " Dim var As Variant
    Set myExl = New Excel.Application
    Set wb = myExl.Workbooks.Open(ExcelFileName)
    Set ws = wb.Worksheets("Base")
Set ws = wb.Worksheets("sheet1")
" var = ws.Range("A1").Value
" or
" var = ws.Cells(1, 1).Value
" ws.Cells(1, 1).Font.Bold = True
myExl.Visible = True
TotalATR = 0
Do
    TotalATR = TotalATR + 1
    Loop Until IsEmpty(ws.Cells(TotalATR + 1, 1))
    TotalATR = TotalATR + 3
ATRNum = 1
Do '—>Loop Until IsEmpty(ws.Cells(ATRNum, 1))
    qATR = ws.Cells(ATRNum, 1)
    ws.Cells(TotalATR, 1) = "ATR_" + qATR
    ws.Cells(TotalATR, 3) = "by 24 Hours"
    ws.Cells(TotalATR + 2, 1) = "Time of Day"
    ws.Cells(TotalATR + 1, 3) = "Mean Spd"
    ws.Cells(TotalATR + 1, 4) = "Mean Spd"
    ws.Cells(TotalATR + 1, 5) = "Mean Spd"
    ws.Cells(TotalATR + 1, 6) = "Mean Spd"
    ws.Cells(TotalATR + 2, 3) = "Lane 1"
    ws.Cells(TotalATR + 2, 4) = "Lane 2"
    ws.Cells(TotalATR + 2, 5) = "Lane 3"
    ws.Cells(TotalATR + 2, 6) = "Lane 4"
    ws.Cells(TotalATR + 1, 8) = "85% Spd"
    ws.Cells(TotalATR + 1, 9) = "85% Spd"
    ws.Cells(TotalATR + 1, 10) = "85% Spd"
    ws.Cells(TotalATR + 1, 11) = "85% Spd"
    ws.Cells(TotalATR + 2, 8) = "Lane 1"
    ws.Cells(TotalATR + 2, 9) = "Lane 2"
    ws.Cells(TotalATR + 2, 10) = "Lane 3"
    ws.Cells(TotalATR + 2, 11) = "Lane 4"
    ws.Cells(TotalATR + 1, 13) = "Var"
    ws.Cells(TotalATR + 1, 14) = "Var"
    ws.Cells(TotalATR + 1, 15) = "Var"
    ws.Cells(TotalATR + 1, 16) = "Var"
    ws.Cells(TotalATR + 2, 13) = "Lane 1"
    ws.Cells(TotalATR + 2, 14) = "Lane 2"
    ws.Cells(TotalATR + 2, 15) = "Lane 3"
    ws.Cells(TotalATR + 2, 16) = "Lane 4"
    ws.Cells(TotalATR + 1, 18) = "Valid Hrs"
    ws.Cells(TotalATR + 1, 19) = "Valid Hrs"
    ws.Cells(TotalATR + 1, 20) = "Valid Hrs"
    ws.Cells(TotalATR + 1, 21) = "Valid Hrs"
    ws.Cells(TotalATR + 2, 18) = "Lane 1"
    ws.Cells(TotalATR + 2, 19) = "Lane 2"
    ws.Cells(TotalATR + 2, 20) = "Lane 3"
    ws.Cells(TotalATR + 2, 21) = "Lane 4"
For byHour = 100 To 2400 Step 100 'Create a Excel file for each hour
" Do '—>Loop Until IsEmpty(ws.Cells(ATRNum, 1))
ws.Cells(TotalATR + 2 + byHour / 100, 1) = byHour

' Create a new Excel file for this ATR —>
Dim FSObj As Object
Set FSObj = CreateObject("Scripting.FileSystemObject")

sFolder = "c:\ATRProcessed" + qATR + "_N\"
If FSObj.FolderExists(sFolder) = False Then
    FSObj.CreateFolder sFolder
End If
sFolder = sFolder + qATR + "_Hourly\"
If FSObj.FolderExists(sFolder) = False Then
    FSObj.CreateFolder sFolder
End If

sFileName = sFolder + qATR + CStr(byHour) + ".xls"
If FSObj.fileExists(sFileName) = False Then
    Dim exWBook As Object
    Dim exSheet As Object
    Set ex = CreateObject("Excel.Application")
    Set exWBook = ex.Workbooks().Add
    Set exSheet = exWBook.Worksheets("sheet1")
    ex.Visible = True
    ' ex.ActiveSheet.Name = "POLK 135"
    ' ex.Sheets.Add
    ' ex.ActiveSheet.Name = "WRIGHT"
    ' Sheet_count = ex.ActiveWorkbook.Sheets.Count
    ' MyPos = ActiveSheet.Index
    ' Sheets(MyPos + 2).Activate

    ex.Cells(2, 1) = "Date"
    ex.Cells(2, 2) = "Error Types"
    For i = 5 To 15
        ex.Cells(1, i) = "Lane1"
        Next i
    For i = 17 To 27
        ex.Cells(1, i) = "Lane2"
        Next i
    For i = 29 To 39
        ex.Cells(1, i) = "Lane3"
        Next i
    For i = 41 To 51
        ex.Cells(1, i) = "Lane4"
        Next i

    p = 40
    For i = 5 To 14
        ex.Cells(2, i) = p
        p = p + 5
    Next
    ex.Cells(2, 15) = 147
    p = 40
    For i = 17 To 26
        ex.Cells(2, i) = p
        p = p + 5
    Next
    ex.Cells(2, 27) = 147
    p = 40
    For i = 29 To 38
ex.Cells(2, i) = p
p = p + 5
Next
ex.Cells(2, 39) = 147
p = 40
For i = 41 To 50
ex.Cells(2, i) = p
p = p + 5
Next
ex.Cells(2, 51) = 147

If Len(qATR) = 1 Then tempATR = " " + qATR
If Len(qATR) = 2 Then tempATR = " " + qATR
If Len(qATR) = 3 Then tempATR = " " + qATR
If Len(qATR) >= 4 Then tempATR = Right(qATR, 4)

ATRDay = 0  'initialization -->
For i = 1 To 4
HourCount(i) = 0
For j = 1 To 15
SpeedCount(i, j) = 0
Next j
Next i

For CrashYear = 1998 To 2003 'begin searching in the original files -->
For CrashMonth = 1 To 12
For CrashDay = 1 To 31
qYear = CStr(Right(CrashYear, 2))
If CrashMonth < 10 Then
qMonth = "0" + CStr(CrashMonth)
Else
qMonth = CStr(CrashMonth)
End If
If CrashDay < 10 Then
qDay = "0" + CStr(CrashDay)
Else
qDay = CStr(CrashDay)
End If

qFileName = "c:\ATR\" + qYear + "\" + qYear + qMonth + "ATR\"_" + "D" + qMonth + qDay + qYear + ".IND"  'ATR index file
If FSObj.fileExists(qFileName) = False Then GoTo quit1  '--> Check if the file exist
Open qFileName For Input As #1
NN = 0
Do Until EOF(1)
Line Input #1, StringRow
sATR = Mid(StringRow, 23, 4)
If tempATR = sATR Then
If NN = 0 Then
NN = 999
Else
To find the second row which has the same name of speed data file

' the ATR speed data file name was found -->
NN = 0
tempFileName = Mid(StringRow, 2, 12)
Close #1

stFileName = "c:\ATR" + qYear + "]" + qYear + qMonth + "ATR\" + tempFileName 'the *.pm files
If FSObj.fileExists(stFileName) = False Then GoTo quit1 1 -- Check if the file exist
' ExcelOpen = 9

ATRDay = ATRDay + 1
ex.Cells(ATRDay + 3, 1) = qMonth + "]" + qDay + "]" + qYear

For i = 1 To 4
For j = 1 To 16
RawSpeed(i,j) = 0
Next j
Next i

LaneNum = 0
Open stFileName For Input As #2
N = 0
Do Until EOF(2)
  Line Input #2, StringRow
  N = N + 1
  If N = 4 Then 'verify the availability of original speed-bin data 40 45 50 55 60 65 this is AN IMPORTANT STEP.
    If (CInt(Mid(StringRow, 14, 4)) = 40) And (CInt(Mid(StringRow, 19, 4)) = 45) Then GoTo quit1
Close #2
  ex.Cells(ATRDay + 3, 2) = "Wrong speed-bin data file"
  GoTo quit1
  End If
  If N <= 6 Then GoTo quit1
  temp = Clnt(Mid(StringRow, 9, 4))
  If (LaneNum > 0) And (temp <> byHour) Then GoTo quit1
If temp = byHour Then
  LaneNum = LaneNum + 1
  If CInt(Mid(StringRow, 4, 2)) <> LaneNum Then GoTo quit1 'deal with a dozen 24Hr
End If

If (temp <= 100) And (temp <> 2400) Then GoTo quit15
'In ATR8040 3/29/2000 .pm file, recorded hours are wrong. There are over a dozen 2400Hr.
'Also, it has 5500Hr, 0002Hr and 1626Hr.
'Don't have to use Fix(byHour/100) to see if it has a decimal part, because we only search for 100, 200, 300, ... til 2400.
  If temp = byHour Then
    LaneNum = LaneNum + 1
    If CInt(Mid(StringRow, 4, 2)) <> LaneNum Then GoTo quit15 'deal with a dozen 24Hr
  End If
  RawSpeed(LaneNum, 2) = Clnt(Mid(StringRow, 4, 2))
  For i = 1 To 11
    p = 9 + i * 5
    RawSpeed(LaneNum, i + 4) = Clnt(Mid(StringRow, p, 4))
    On Error Resume Next 'ATR2030 03/13/1998 .prn lost a speed bin
  Next
End If
quit1: Loop
quit15: Close #2 'close that *.PRN file

For i = 1 To LaneNum
  SpeedCheck = 0
  For j = 5 To 15
    SpeedCheck = SpeedCheck + RawSpeed(i, j)
  Next
  If SpeedCheck > 2100 Then
    ex.Cells(ATRDay + 3, 3) = "Too High Counts"
    GoTo quit10
  End If
Next

For j = 5 To 15
  If ((RawSpeed(i, j) > SpeedCheck * 0.95) And (SpeedCheck > 30)) Then
    ex.Cells(ATRDay + 3, 2) = "Possible Spd Err/Bad Weather"
```vbnet
GoTo quitl0
End If
Next

For j = 5 To 15
    ex.Cells(ATRDay + 3, j + 12 * (i - 1)) = RawSpeed(i, j)
    SpeedCount(i, j) = SpeedCount(i, j) + RawSpeed(i, j)
    If j = 15 Then HourCount(i) = HourCount(i) + 1
Next 'For j = 5 To 15
quitl0: Next ' For i = 1 To LaneNum

GoTo quitl1
End If
Else ' If tempATR = sATR Then
    If NN = 999 Then NN = 0 To prevent the third ATR file is wrong,
    ' For example, ATR=1100, Day=990326

End If
Loop ' Do Until E0F(1)
Close #1
quitl1: Next CrashDay ' For CrashDay = 1 To 31
    Next CrashMonth ' For CrashMonth = 1 To 12
    Next CrashYear ' For CrashYear = 1998 To 2003

For j = 5 To 15
    MidBin(j) = 37.5 + 5 * (j - 5)
Next
For i = 1 To 4
    MeanSpeed(i) = 0
    SumDSC(i) = 0
    Var(i) = 0
    EightyFifth(i) = 0
    AccumCount(i) = 0
    For j = 5 To 15
        SqDiff(i, j) = 0
    Next
Next

ex.Cells(ATRDay + 5, 3) = "Total Hours"
For i = 1 To 4
    ex.Cells(ATRDay + 5, 5 + 12 * (i - 1)) = HourCount(i)
Next

ex.Cells(ATRDay + 6, 3) = "Speed Bin Count"
ex.Cells(ATRDay + 7, 3) = "Total Counts"
ex.Cells(ATRDay + 8, 3) = "Mid Pt of Each Bin"
ex.Cells(ATRDay + 9, 3) = "Counts * Mid Pt"
ex.Cells(ATRDay + 10, 3) = "Total(Mid Pt*Count)"
ex.Cells(ATRDay + 11, 3) = "Mean Speed"
ex.Cells(ATRDay + 12, 3) = "Sq Diff"
ex.Cells(ATRDay + 13, 3) = "SqDiff * Count"
ex.Cells(ATRDay + 14, 3) = "Total(SqDiff*Count)"
ex.Cells(ATRDay + 15, 3) = "Variance"
ex.Cells(ATRDay + 16, 3) = "Pre 85 (TotalCount*0.85)"
ex.Cells(ATRDay + 17, 3) = "Accumulated Count"
ex.Cells(ATRDay + 18, 3) = "85%th Speed"

For i = 1 To 4
    TotalCount(i) = 0
    For j = 5 To 15
        TotalCount(i) = TotalCount(i) + SpeedCount(i, j)
    Next
Next
```
For i = 1 To 4 'main calculation
If TotalCount(i) = 0 Then GoTo quit9

TotalCMB(i) = 0
For j = 5 To 15
    CountMultiplyBin(j) = SpeedCount(i, j) * MidBin(j)
    TotalCMB(i) = TotalCMB(i) + CountMultiplyBin(j)
Next
ex.Cells(ATRDay + 9, j + 12 * (i - 1)) = CountMultiplyBin(j)
ex.Cells(ATRDay + 10, 5 + 12 * (i - 1)) = TotalCMB(i)
MeanSpeed(i) = Round(TotalCMB(i) / TotalCount(i), 2)
ex.Cells(ATRDay + 11, 5 + 12 * (i - 1)) = MeanSpeed(i)
For j = 5 To 15
    SqDiff(i, j) = (MidBin(j) - MeanSpeed(i)) ^ 2
    ex.Cells(ATRDay + 9, j + 12 * (i - 1)) = SqDiff(i, j)
ex.Cells(ATRDay + 12, j + 12 * (i - 1)) = SqDiff(i, j) * SpeedCount(i, j)
    SumDSC(i) = SumDSC(i) + SqDiff(i, j) * SpeedCount(i, j)
Next
ex.Cells(ATRDay + 14, 5 + 12 * (i - 1)) = SumDSC(i)
Var(i) = Round(SumDSC(i) / TotalCount(i), 2)
ex.Cells(ATRDay + 15, 5 + 12 * (i - 1)) = Var(i)
Pre85(i) = Round(TotalCount(i) * 0.85, 2)
ex.Cells(ATRDay + 16, 5 + 12 * (i - 1)) = Pre85(i)
For j = 5 To 15
    AccumCount(i) = AccumCount(i) + SpeedCount(i, j)
ex.Cells(ATRDay + 17, j + 12 * (i - 1)) = AccumCount(i)
If Pre85(i) <= AccumCount(i) Then
    EightyFifth(i) = MidBin(j) - ((AccumCount(i) - Pre85(i)) * (MidBin(j) - MidBin(j - 1)) / SpeedCount(i, j))
ex.Cells(ATRDay + 18, 5 + 12 * (i - 1)) = Round(EightyFifth(i), 2)
    GoTo quit9
End If
Next
quit9: Next i 'For i = 1 To 4 main calculation

exWBook.SaveAs sFileName
ex.Quit 'quit Excel
Set exWBook = Nothing
Set exSheet = Nothing
Set ex = Nothing

For j = 3 To 6
    ws.Cells(TotalATR + 2 + byHour / 100, j) = MeanSpeed(j - 2)
Next
For j = 8 To 11
    ws.Cells(TotalATR + 2 + byHour / 100, j) = Round(EightyFifth(j - 7), 2)
Next
For j = 13 To 16
    ws.Cells(TotalATR + 2 + byHour / 100, j) = Var(j - 12)
Next
For j = 18 To 21
    ws.Cells(TotalATR + 2 + byHour / 100, j) = HourCount(j - 17)
Else If FSObj.fileExists(sFileName) = False Then
    ErrMsg = "The file has existed in the folder. No more process is needed."
End If
Set FSObj = Nothing

Next ' For byHour = 100 To 2400 Step 100

TotalATR = TotalATR + 29
ATRNum = ATRNum + 1
Loop Until IsEmpty ws.Cells(ATRNum, 1)

wb.Save
wb.Close
myExl.Quit
Set ws = Nothing
Set wb = Nothing
Set myExl = Nothing

APIBeep 1000, 300
Call Delay
APIBeep 1250, 300
Call Delay
APIBeep 1500, 300
Call Delay
APIBeep 1750, 300
Call Delay
APIBeep 2000, 300

End Sub

Private Sub Command3_Click()
    Unload Form1
End Sub

Private Sub DelayQ
    For tTimer = 1 To 1500 'Delay loop between beeps
        Next tTimer
    End Sub

ATR by 365Day

Private Declare Function APIBeep Lib "kernel32" Alias "Beep" (ByVal dwFreq As Long, ByVal dwDuration As Long) As Long
Private qYear, qMonth, qDay, qFileName As String

Private IntCrashTime As Integer
Private tempATR, qATR As String
Private SpeedCount(5, 20) As Long
Private DaySpdCount(5, 20) As Long
Private RawSpeed(5, 20) As Long

Private Sub Command1_Click()
    Dim TotalCount(5) As Long
    Dim TotalATR, ATRNum As Integer
Dim StringRow, tempFileName, sFileName, stFileName, sFolder As String
Dim MidBin(20), CountMultiplyBin(20) As Single
Dim TotalCMB(5) As Single
Dim MeanSpeed(S), Var(5) As Single
Dim Pre85(5), EightyFifth(5) As Single
Dim AccumCount(S) As Long
Dim byHour As Integer
Dim LaneNum As Integer
Dim dDate As Integer
Dim DaySum As Long
Dim YearCount As Integer
' Dim Sheet_count As Variant
' Dim MyPos
Dim LaneN, NN, N As Integer
Dim MSpeed As Single, SDev As Single, Varia As Single, Vol As Single
Dim SqDiff(5, 20), SumDSC(5) As Single
Dim ExRow As Integer
Dim HourCount(S) As Integer
Dim ErrType As Integer
Dim qDate, NewDate, sNewDate, ExcelFileName As String
'Cannot write it like this:
' Dim MSpeed, SDev, Varia, Vol As Single
' Otherwise type missing
Dim StartPos, tempMonth, tempDay As Integer
Dim WeekBefore As Integer
Dim CrashNum, k, qtime As Integer
Dim Crash Year, CrashMonth, CrashDay, CrashTime As Integer
Dim ErrorMsg As String
Dim ii As Integer
Dim FTestResult As String
'select an Excel file to process ATR one by one
CommonDialog1.ShowOpen
ExcelFileName = CommonDialog1.FileName

Dim myExl As Excel.Application
Dim wb As Workbook
Dim ws As Worksheet
" Dim var As Variant
Set myExl = New Excel.Application
Set wb = myExl.Workbooks.Open(ExcelFileName)

" Set ws = wb.Worksheets("Base")
Set ws = wb.Worksheets("sheet1")
" var = ws.Range("A1").Value
" or
" var = ws.Cells(1, 1).Value
" ws.Cells(1, 1).Font.Bold = True
myExl.Visible = True
TotalATR = 0
Do
    TotalATR = TotalATR + 1
Loop Until IsEmpty(ws.Cells(TotalATR + 1, 1))
TotalATR = TotalATR + 2
ATRNum = 1
Do '—>Loop Until IsEmpty(ws.Cells(ATRNum, 1))
qATR = ws.Cells(ATRNum, 1)

ws.Cells(TotalATR, 1) = "ATR_" + qATR
ws.Cells(TotalATR, 3) = "by 365 Days"

ws.Cells(TotalATR + 2, 1) = "Month_Day"
ws.Cells(TotalATR + 1, 3) = "Mean Spd"
ws.Cells(TotalATR + 1, 4) = "Mean Spd"
ws.Cells(TotalATR + 1, 5) = "Mean Spd"
ws.Cells(TotalATR + 1, 6) = "Mean Spd"
ws.Cells(TotalATR + 2, 3) = "Lane 1"
ws.Cells(TotalATR + 2, 4) = "Lane 2"
ws.Cells(TotalATR + 2, 5) = "Lane 3"
ws.Cells(TotalATR + 2, 6) = "Lane 4"

ws.Cells(TotalATR + 1, 8) = "85% Spd"
ws.Cells(TotalATR + 1, 9) = "85% Spd"
ws.Cells(TotalATR + 1, 10) = "85% Spd"
ws.Cells(TotalATR + 1, 11) = "85% Spd"
ws.Cells(TotalATR + 2, 8) = "Lane 1"
ws.Cells(TotalATR + 2, 9) = "Lane 2"
ws.Cells(TotalATR + 2, 10) = "Lane 3"
ws.Cells(TotalATR + 2, 11) = "Lane 4"

ws.Cells(TotalATR + 1, 13) = "Var"
ws.Cells(TotalATR + 1, 14) = "Var"
ws.Cells(TotalATR + 1, 15) = "Var"
ws.Cells(TotalATR + 1, 16) = "Var"
ws.Cells(TotalATR + 2, 13) = "Lane 1"
ws.Cells(TotalATR + 2, 14) = "Lane 2"
ws.Cells(TotalATR + 2, 15) = "Lane 3"
ws.Cells(TotalATR + 2, 16) = "Lane 4"

ws.Cells(TotalATR + 1, 18) = "Avg Vol"
ws.Cells(TotalATR + 1, 19) = "Avg Vol"
ws.Cells(TotalATR + 1, 20) = "Avg Vol"
ws.Cells(TotalATR + 1, 21) = "Avg Vol"
ws.Cells(TotalATR + 2, 18) = "Lane 1"
ws.Cells(TotalATR + 2, 19) = "Lane 2"
ws.Cells(TotalATR + 2, 20) = "Lane 3"
ws.Cells(TotalATR + 2, 21) = "Lane 4"

TotalATR = TotalATR + 3
If Len(qATR) = 1 Then tempATR = " " + qATR
If Len(qATR) = 2 Then tempATR = " " + qATR
If Len(qATR) = 3 Then tempATR = " " + qATR
If Len(qATR) >= 4 Then tempATR = Right(qATR, 4)

For CrashMonth = 1 To 12 'begin searching in the original files --->
If CrashMonth < 10 Then
    qMonth = "0" + CStr(CrashMonth)
Else
    qMonth = CStr(CrashMonth)
End If
' Create a new Excel file for this ATR —>
Dim FSObj As Object
Set FSObj = CreateObject("Scripting.FileSystemObject")

sFolder = "c:\ATRProcessed\" + qATR + ":\_N"
If FSObj.FolderExists(sFolder) = False Then
   FSObj.CreateFolder sFolder
End If
sFolder = sFolder + qATR + ":\_Daily"
If FSObj.FolderExists(sFolder) = False Then
   FSObj.CreateFolder sFolder
End If

sFileName = sFolder + qATR + ":\_M" + qMonth + ":\_Daily.xls"
If FSObj.fileExists(sFileName) = False Then
   Dim exWBook As Object
   Dim exSheet As Object
   Set ex = CreateObject("Excel.Application")
   Set exWBook = ex.Workbooks.Add
   Set exSheet = exWBook.Worksheets("sheet1")
   ex.Visible = True
   ' ex.ActiveSheet.Name = "POLK 135"
   ' ex.Sheets.Add
   ' ex.ActiveSheet.Name = "WRIGHT"
   'Sheets_count = ex.ActiveWorkbook.Sheets.Count
   'MyPos = ActiveSheet.Index
   'Sheets(MyPos + 2).Activate
   ex.Cells(1, 1) = "ATR " + qATR
   ex.Cells(2, 1) = "MonthDay"
   ex.Cells(2, 2) = "Err Type"
   For i = 5 To 16
      ex.Cells(1, i) = "Lane1"
   Next i
   For i = 18 To 29
      ex.Cells(1, i) = "Lane2"
   Next i
   For i = 31 To 42
      ex.Cells(1, i) = "Lane3"
   Next i
   For i = 44 To 55
      ex.Cells(1, i) = "Lane4"
   Next i
   p = 40
   For i = 5 To 14
      ex.Cells(2, i) = p
      p = p + 5
   Next i
   ex.Cells(2, 15) = 147
   ex.Cells(2, 16) = "Valid Hrs"
   p = 40
   For i = 18 To 27
      ex.Cells(2, i) = p
      p = p + 5
   Next i
   ex.Cells(2, 28) = 147
   ex.Cells(2, 29) = "Valid Hrs"
   p = 40
   For i = 31 To 40
```vbnet
ex.Cells(2, i) = p
p = p + 5
Next
ex.Cells(2, 41) = 147
ex.Cells(2, 42) = "Valid Hrs"
p = 40
For i = 44 To 53
ex.Cells(2, i) = p
p = p + 5
Next
ex.Cells(2, 54) = 147
ex.Cells(2, 55) = "Valid Hrs"

ExRow = 2 'initialization -->

For CrashDay = 1 To 31
qDate = 99
YearCount = 0
For i = 1 To 4
For j = 1 To 15
SpeedCount(i, j) = 0
Next j
Next i

For CrashYear = 1998 To 2003
qYear = CStr(Right(CrashYear, 2))
If CrashDay < 10 Then
    qDay = "0" + CStr(CrashDay)
Else
    qDay = CStr(CrashDay)
End If

qFileName = "c:\ATR\" + qYear + "\" + qYear + qMonth + "ATR\" + "+" + qMonth + qDay + qYear + ".IND" 'ATR index file
If FSObj.fileExists(qFileName) = False Then GoTo quit11 '—> Check if the file exist
Open qFileName For Input As #1
NN = 0
Do Until EOF(1)
    Line Input #1, StringRow
    sATR = Mid(StringRow, 23, 4)
    If tempATR = sATR Then
        If NN = 0 Then
            NN = 999
            tempFileName = Mid(StringRow, 2, 12)
            Close #1
            stFileName = "c:\ATR\" + qYear + "\" + qYear + qMonth + "ATR\" + tempFileName 'the *.prn files
            If FSObj.fileExists(stFileName) = False Then GoTo quit11 '—> Check if the file exist
            ExcelOpen = 9

            ExRow = ExRow + 1
            If qDate = 99 Then
```
ex.Cells(ExRow, 1) = qMonth + "_" + qDay
qDate = 0
ExRow = ExRow + 1
End If
For i = 1 To 4
For j = 1 To 16
RawSpeed(i, j) = 0
Next j
Next i
For i = 1 To 4
HourCount(i) = 0
For j = 1 To 15
DaySpdCount(i, j) = 0
Next j
Next i
ex.Cells(ExRow, 1) = CStr(CrashYear)

LaneNum = 0 'read data from original speed file
Open stFileName For Input As #2
N = 0
Do Until EOF(2)
    Line Input #2, StringRow
    N = N + 1
    If N = 4 Then 'verify the availability of original speed-bin data, this is AN IMPORTANT STEP.
        If (CInt(Mid(StringRow, 14, 4)) = 40) And (CInt(Mid(StringRow, 19, 4)) = 45) Then GoTo quit1
        Close #2
        ex.Cells(ExRow, 2) = "Wrong speed-bin data file"
        GoTo quit1
    End If
    If N <= 6 Then GoTo quit1
    LaneNum = CInt(Mid(StringRow, 4, 2))
    For i = 1 To 11
        p = 9 + i * 5
        RawSpeed(LaneNum, i + 4) = CInt(Mid(StringRow, p, 4))
        Next
    SpeedCheck = 0
    For j = 5 To 15
        SpeedCheck = SpeedCheck + RawSpeed(LaneNum, j)
        Next
    If SpeedCheck > 2100 Then
        ex.Cells(ExRow, 2) = "Too High Counts"
        GoTo quit1
    End If
    For j = 5 To 15
        If ((RawSpeed(LaneNum, j) > SpeedCheck * 0.95) And (SpeedCheck > 30)) Then
            ex.Cells(ExRow, 2) = "Possible Spd Err/Bad Weather"
            GoTo quit1
        End If
        Next
    DaySpdCount(LaneNum, j) = DaySpdCount(LaneNum, j) + RawSpeed(LaneNum, j)
    If j = 15 Then HourCount(LaneNum) = HourCount(LaneNum) + 1
Next j
quit1: Loop 'Do Until EOF(2)
quit3: Close #2 'close that *.PRN file
For i = 1 To 4
   If (HourCount(i) <> 0) And (HourCount(i) <> 24) Then
      ex.Cells(ExRow, 2) = "Less than 24Hr Data"
      GoTo quit1
   End If
Next
YearCount = YearCount + 1
For i = 1 To 4
   DaySum = 0
   For j = 5 To 15
      ex.Cells(ExRow, j + 13 * (i - 1)) = DaySpdCount(i, j)
      DaySum = DaySum + DaySpdCount(i, j)
      SpeedCount(i, j) = SpeedCount(i, j) + DaySpdCount(i, j)
   Next
   ex.Cells(ExRow, 4 + 13 * (i - 1)) = DaySum
   ex.Cells(ExRow, 16 + 13 * (i - 1)) = HourCount(i) ' Not always the 24hr
Next
GoTo quit1
End If 'If NN = 0 Then
Else 'If tempATR = sATR Then
   If NN = 999 Then NN = 0 'To prevent the third ATR file is wrong,
      'for example, ATR=1100, Day=990326
   End If
Loop 'Do Until EOF(l)
Close #1
quit1: Next Crash Year

For j = 5 To 15
   MidBin(j) = 37.5 + 5 * (j - 5)
Next
For i = 1 To 4
   MeanSpeed(i) = 0
   SumDSO(i) = 0
   Var(i) = 0
   EightyFifth(i) = 0
   AccumCount(i) = 0
   For j = 5 To 15
      SqDiff(i, j) = 0
   Next
   Next
For i = 1 To 4
   TotalCount(i) = 0
   For j = 5 To 15
      TotalCount(i) = TotalCount(i) + SpeedCount(i, j)
   Next j
Next i
If (TotalCount(1) = 0 And TotalCount(2) = 0 And TotalCount(3) = 0 And TotalCount(4) = 0) Then GoTo quit12

For i = 1 To 4
   For j = 5 To 15
      ex.Cells(ExRow + 1, j + 13 * (i - 1)) = SpeedCount(i, j)
      ex.Cells(ExRow + 3, j + 13 * (i - 1)) = MidBin(j)
   Next j
   ex.Cells(ExRow + 2, 5 + 13 * (i - 1)) = TotalCount(i)
Next i
' ex.Cells(ExRow + 5, 3) = "Total Hours"
' For i = 1 To 4
'   ex.Cells(ExRow + 5, 5 + 12 * (i - 1)) = HourCount(i)
' Next
ex.Cells(ExRow + 1, 3) = "Speed Bin Count"
ex.Cells(ExRow + 2, 3) = "Total Counts"
ex.Cells(ExRow + 3, 3) = "Mid Pt of Each Bin"
ex.Cells(ExRow + 4, 3) = "Counts * Mid Pt"
ex.Cells(ExRow + 5, 3) = "Total(Mid Pt*Count)"
ex.Cells(ExRow + 6, 3) = "Mean Speed"
ex.Cells(ExRow + 7, 3) = "Sq Diff"
ex.Cells(ExRow + 8, 3) = "SqDiff * Count"
ex.Cells(ExRow + 9, 3) = "Total(SqDiff*Count)"
ex.Cells(ExRow + 10, 3) = "Variance"
ex.Cells(ExRow + 11, 3) = "Pre 85 (TotalCount*0.85)"
ex.Cells(ExRow + 12, 3) = "Accumulated Count"
ex.Cells(ExRow + 13, 3) = "85%th Speed"

For i = 1 To 4 'main calculation
   If TotalCount(i) = 0 Then GoTo quit9
   TotalCMB(i) = 0
   For j = 5 To 15
      CountMultiplyBin(j) = SpeedCount(i, j) * MidBin(j)
      TotalCMB(i) = TotalCMB(i) + CountMultiplyBin(j)
   Next
ex.Cells(ExRow + 4, 5 + 13 * (i - 1)) = TotalCMB(i)
MeanSpeed(i) = Round(TotalCMB(i) / TotalCount(i), 2)
ex.Cells(ExRow + 6, 5 + 13 * (i - 1)) = MeanSpeed(i)
   For j = 5 To 15
      SqDiff(i, j) = (MidBin(j) - MeanSpeed(i))^2
      SumDSC(i) = SumDSC(i) + SqDiff(i, j) * SpeedCount(i, j)
   Next
ex.Cells(ExRow + 9, 5 + 13 * (i - 1)) = SumDSC(i)
Var(i) = Round(SumDSC(i) / TotalCount(i), 2)
ex.Cells(ExRow + 10, 5 + 13 * (i - 1)) = Var(i)
Pre85(i) = Round(TotalCount(i) * 0.85, 2)
ex.Cells(ExRow + 11, 5 + 13 * (i - 1)) = Pre85(i)
For j = 5 To 15
   AccumCount(i) = AccumCount(i) + SpeedCount(i, j)
   if Pre85(i) <= AccumCount(i) Then
      EightyFifth(i) = MidBin(j) - ((AccumCount(i) - Pre85(i)) * (MidBin(j) - MidBin(j - 1)) / SpeedCount(i, j))
      ex.Cells(ExRow + 13, 5 + 13 * (i - 1)) = Round(EightyFifth(i), 2)
   GoTo quit9
End If
Next
quit9: Next i 'For i = 1 To 4 main calculation
ExRow = ExRow + 15

ws.Cells(TotalATR, 1) = qMonth + "." + qDay

For j = 3 To 6
   ws.Cells(TotalATR, j) = MeanSpeed(j - 2)
Next
For j = 8 To 11
    ws.Cells(TotalATR, j) = Round(EightyFifth(j - 7), 2)
Next
For j = 13 To 16
    ws.Cells(TotalATR, j) = Var(j - 12)
Next
For j = 18 To 21
    If YearCount <> 0 Then
        ws.Cells(TotalATR, j) = Round(TotalCount(j - 17) / YearCount, 0)
    End If
Next
TotalATR = TotalATR + 1
quit2: Next CrashDay
    exWBook.SaveAs sFileName
    exWBook.Close
    ex.Quit 'quit Excel
    Set exWBook = Nothing
    Set exSheet = Nothing
    Set ex = Nothing
Else 'If FSObj.fileExists(sFileName) = False Then
    ErrMsg = "The file has existed in the folder. No more process is needed."
End If
    Set FSObj = Nothing
Next CrashMonth
TotalATR = TotalATR + 2
ATRNum = ATRNum + 1
Loop Until IsEmpty(ws.Cells(ATRNum, 1))
    wb.Save
    wb.Close
    myExl.Quit
    Set wb = Nothing
    Set myExl = Nothing
    APIBeep 1000, 300
    Call Delay
    APIBeep 1250, 300
    Call Delay
    APIBeep 1500, 300
    Call Delay
    APIBeep 1750, 300
    Call Delay
    APIBeep 2000, 300
End Sub

Private Sub Command3_Click()
    Unload Form1
End Sub

Private Sub DelayQ()
    For tTimer = 1 To 1500 'Delay loop between beeps
        Next tTimer
    End Sub
APPENDIX B – SUMMARY TABLES OF SPEED METRICS BY TYPE OF FACILITY

PLOTS OF PARAMETER COMPARISONS
### INTERSTATE HIGHWAY

<table>
<thead>
<tr>
<th>ATR NO.</th>
<th>CASE</th>
<th>CONT</th>
<th>CASE</th>
<th>CONT</th>
<th>CASE</th>
<th>85TH %ILE SPEED</th>
<th>CRASH</th>
<th>DEPARTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>69.47</td>
<td>69.60</td>
<td>72.21</td>
<td>73.38</td>
<td>0.21</td>
<td>1.63</td>
<td>65</td>
<td>846</td>
</tr>
<tr>
<td>102</td>
<td>69.13</td>
<td>70.52</td>
<td>73.48</td>
<td>75.01</td>
<td>0.12</td>
<td>0.52</td>
<td>65</td>
<td>428</td>
</tr>
<tr>
<td>104</td>
<td>70.23</td>
<td>70.82</td>
<td>73.20</td>
<td>73.89</td>
<td>0.12</td>
<td>0.84</td>
<td>65</td>
<td>840</td>
</tr>
<tr>
<td>105</td>
<td>68.82</td>
<td>69.50</td>
<td>71.16</td>
<td>72.81</td>
<td>0.24</td>
<td>0.98</td>
<td>65</td>
<td>512</td>
</tr>
<tr>
<td>106</td>
<td>69.76</td>
<td>70.53</td>
<td>72.91</td>
<td>74.07</td>
<td>0.11</td>
<td>0.51</td>
<td>65</td>
<td>662</td>
</tr>
<tr>
<td>109</td>
<td>69.38</td>
<td>70.13</td>
<td>72.12</td>
<td>73.33</td>
<td>0.22</td>
<td>1.31</td>
<td>65</td>
<td>623</td>
</tr>
<tr>
<td>110</td>
<td>68.20</td>
<td>69.86</td>
<td>72.13</td>
<td>73.41</td>
<td>0.17</td>
<td>1.40</td>
<td>65</td>
<td>801</td>
</tr>
<tr>
<td>111</td>
<td>69.24</td>
<td>70.71</td>
<td>72.20</td>
<td>73.49</td>
<td>0.12</td>
<td>1.41</td>
<td>65</td>
<td>1733</td>
</tr>
<tr>
<td>113</td>
<td>69.18</td>
<td>70.26</td>
<td>71.82</td>
<td>73.02</td>
<td>0.22</td>
<td>1.05</td>
<td>65</td>
<td>595</td>
</tr>
<tr>
<td>115</td>
<td>69.75</td>
<td>71.68</td>
<td>74.39</td>
<td>75.33</td>
<td>0.29</td>
<td>2.02</td>
<td>65</td>
<td>864</td>
</tr>
<tr>
<td>116</td>
<td>67.35</td>
<td>68.23</td>
<td>71.03</td>
<td>72.04</td>
<td>0.20</td>
<td>2.46</td>
<td>65</td>
<td>1578</td>
</tr>
<tr>
<td>117</td>
<td>67.56</td>
<td>68.05</td>
<td>72.38</td>
<td>72.05</td>
<td>0.30</td>
<td>7.24</td>
<td>65</td>
<td>1585</td>
</tr>
<tr>
<td>POLK 1-35</td>
<td>70.55</td>
<td>69.60</td>
<td>73.02</td>
<td>73.13</td>
<td>0.10</td>
<td>1.19</td>
<td>65</td>
<td>1593</td>
</tr>
<tr>
<td>WRIGHT</td>
<td>69.98</td>
<td>70.55</td>
<td>72.17</td>
<td>73.64</td>
<td>0.17</td>
<td>0.95</td>
<td>65</td>
<td>915</td>
</tr>
<tr>
<td>AVG.</td>
<td>69.19</td>
<td>70.00</td>
<td>72.44</td>
<td>73.47</td>
<td>0.19</td>
<td>1.68</td>
<td>970</td>
<td>0.04</td>
</tr>
</tbody>
</table>

**INTERSTATE CRASH RATE VS. CASE HOUR MEAN SPEED**

[Graph showing the relationship between case hour mean speed and crash rate]
INTERSTATE CRASH RATE VS. MEAN SPEED DIFFERENCE

INTERSTATE CRASH RATE VS. 85TH %ILE SPEED DIFFERENCE

INTERSTATE HIGHWAY CRASH RATE VS. SPEED DEPARTURE
INTERSTATE
CRASH DENSITY VS. CASE HOUR 85TH P%ILE SPEED

CASE HOUR 85TH P%ILE SPEED

INTERSTATE HIGHWAYS
CRASH DENSITY VS. CASE HOUR SPEED DEPARTURE

CASE HOUR DEPARTURE

INTERSTATE
CRASH DENSITY VS. CONTROL HOUR SPEED DEPARTURE

CONTROL HOUR DEPARTURE

(85TH P%ILE SPEED MINUS SPEED LIMIT)
### Freeways (W/O Interstates)

<table>
<thead>
<tr>
<th>ATR</th>
<th>Case</th>
<th>Cont</th>
<th>Case</th>
<th>Cont</th>
<th>Case</th>
<th>Cont</th>
<th>Case</th>
<th>Cont</th>
<th>Rate</th>
<th>Dens</th>
<th>Limit</th>
<th>Vol</th>
<th>ADT</th>
<th>Vol/ADT</th>
<th>Case</th>
<th>Cont</th>
<th>Disp</th>
</tr>
</thead>
<tbody>
<tr>
<td>206</td>
<td>66.11</td>
<td>67.40</td>
<td>69.42</td>
<td>70.60</td>
<td>0.43</td>
<td>1.79</td>
<td>65</td>
<td>535</td>
<td>11300</td>
<td>0.047</td>
<td>4.42</td>
<td>5.60</td>
<td>3.31</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>242</td>
<td>68.21</td>
<td>68.80</td>
<td>71.13</td>
<td>72.40</td>
<td>0.53</td>
<td>3.07</td>
<td>65</td>
<td>675</td>
<td>15900</td>
<td>0.042</td>
<td>6.13</td>
<td>7.40</td>
<td>2.92</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>245</td>
<td>68.33</td>
<td>68.70</td>
<td>71.85</td>
<td>72.20</td>
<td>0.49</td>
<td>1.70</td>
<td>65</td>
<td>470</td>
<td>9500</td>
<td>0.049</td>
<td>6.85</td>
<td>7.20</td>
<td>3.52</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>247</td>
<td>68.56</td>
<td>68.80</td>
<td>71.39</td>
<td>72.20</td>
<td>0.74</td>
<td>1.85</td>
<td>65</td>
<td>286</td>
<td>6800</td>
<td>0.042</td>
<td>6.39</td>
<td>7.20</td>
<td>2.83</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>249</td>
<td>64.41</td>
<td>69.30</td>
<td>66.70</td>
<td>72.20</td>
<td>0.32</td>
<td>1.90</td>
<td>65</td>
<td>732</td>
<td>16100</td>
<td>0.045</td>
<td>1.70</td>
<td>7.20</td>
<td>2.29</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>68.31</td>
<td>68.40</td>
<td>70.89</td>
<td>71.30</td>
<td>0.17</td>
<td>1.29</td>
<td>65</td>
<td>858</td>
<td>20600</td>
<td>0.042</td>
<td>5.89</td>
<td>6.30</td>
<td>2.58</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>701</td>
<td>62.21</td>
<td>62.65</td>
<td>65.90</td>
<td>65.90</td>
<td>0.20</td>
<td>4.56</td>
<td>60</td>
<td>1391</td>
<td>63400</td>
<td>0.022</td>
<td>5.90</td>
<td>5.90</td>
<td>3.69</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>704</td>
<td>60.02</td>
<td>60.80</td>
<td>66.11</td>
<td>67.00</td>
<td>0.41</td>
<td>11.30</td>
<td>55</td>
<td>3852</td>
<td>76000</td>
<td>0.051</td>
<td>11.11</td>
<td>12.00</td>
<td>6.09</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>705</td>
<td>58.26</td>
<td>61.80</td>
<td>62.80</td>
<td>65.10</td>
<td>0.38</td>
<td>6.78</td>
<td>55</td>
<td>2278</td>
<td>48400</td>
<td>0.047</td>
<td>7.80</td>
<td>10.10</td>
<td>4.54</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Avg.</td>
<td>64.94</td>
<td>66.29</td>
<td>68.47</td>
<td>69.88</td>
<td>0.41</td>
<td>3.80</td>
<td>1231</td>
<td>0.043</td>
<td>6.24</td>
<td>7.66</td>
<td>3.53</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Crash Rate vs. Case Hour Mean Speed

![Graph showing crash rate vs. case hour mean speed](image)
FREEWAY
CRASH DENSITY VS. CASE HOUR MEAN SPEED

FREEWAY
CRASH DENSITY VS. CASE HOUR 85TH %ILE SPEED

FREEWAY
CRASH DENSITY VS. CASE HOUR SPEED DEPARTURE (85TH %ILE SPEED MINUS SPEED LIMIT)
FREeways

CRASH DENSITY VS. CONTROL HOUR SPEED DEPARTURE

CRASH DENSITY, PER MILE PER YEAR

CONTROL HOUR DEPARTURE
(85TH %ILE SPEED MINUS SPEED LIMIT)

FREeways

CRASH DENSITY VS. 85TH %ILE SPEED DIFFERENCE

CRASH DENSITY, PER MILE PER YEAR

85TH %ILE SPEED DIFFERENCE - CASE MINUS CONTROL HOUR
EXPRESSWAYS

<table>
<thead>
<tr>
<th>ATR</th>
<th>CASE</th>
<th>CONT</th>
<th>CASE</th>
<th>CONT</th>
<th>RATE</th>
<th>DENS</th>
<th>LIMIT</th>
<th>VOL</th>
<th>ADT</th>
<th>VOL/ADT</th>
<th>CASE</th>
<th>CONT</th>
<th>DISP</th>
</tr>
</thead>
<tbody>
<tr>
<td>204</td>
<td>59.01</td>
<td>63.03</td>
<td>64.09</td>
<td>66.84</td>
<td>0.41</td>
<td>1.55</td>
<td>65</td>
<td>439</td>
<td>10400</td>
<td>0.042</td>
<td>-0.91</td>
<td>1.84</td>
<td>5.08</td>
</tr>
<tr>
<td>205</td>
<td>68.07</td>
<td>68.05</td>
<td>70.74</td>
<td>71.05</td>
<td>0.54</td>
<td>1.53</td>
<td>65</td>
<td>523</td>
<td>7800</td>
<td>0.067</td>
<td>5.74</td>
<td>6.05</td>
<td>2.67</td>
</tr>
<tr>
<td>217</td>
<td>53.30</td>
<td>55.33</td>
<td>60.56</td>
<td>60.41</td>
<td>0.66</td>
<td>3.48</td>
<td>55</td>
<td>776</td>
<td>14500</td>
<td>0.054</td>
<td>5.56</td>
<td>5.41</td>
<td>7.26</td>
</tr>
<tr>
<td>243</td>
<td>66.34</td>
<td>66.45</td>
<td>70.25</td>
<td>70.58</td>
<td>0.43</td>
<td>1.44</td>
<td>65</td>
<td>506</td>
<td>9100</td>
<td>0.056</td>
<td>5.25</td>
<td>5.58</td>
<td>3.91</td>
</tr>
<tr>
<td>246</td>
<td>66.59</td>
<td>64.54</td>
<td>70.71</td>
<td>68.41</td>
<td>0.62</td>
<td>2.13</td>
<td>55</td>
<td>432</td>
<td>9800</td>
<td>0.044</td>
<td>15.71</td>
<td>13.41</td>
<td>4.13</td>
</tr>
<tr>
<td>AVG.</td>
<td>62.66</td>
<td>63.48</td>
<td>67.27</td>
<td>67.46</td>
<td>0.53</td>
<td>2.03</td>
<td>535</td>
<td>0.052</td>
<td>6.27</td>
<td>6.46</td>
<td>4.61</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

EXPRESSWAY
CRASH RATE VS. CASE HOUR MEAN SPEED

EXPRESSWAY
CRASH RATE VS. CASE HOUR 85TH %ILE SPEED
EXPRESSWAY
CRASH DENSITY VS. 85TH %ILE SPEED DIFFERENCE

CRASH DENSITY, PER MILE PER YEAR

85TH %ILE SPEED DIFFERENCE - CASE MINUS CONTROL HOUR
### TWO-LANE HIGHWAYS

<table>
<thead>
<tr>
<th>ATR</th>
<th>AVG SPEED</th>
<th>85TH %ILE SPEED</th>
<th>CRASH</th>
<th>DEPARTURE</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>RATE</td>
<td>DENS</td>
<td>LIMIT</td>
<td>VOL/ADT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VOL</td>
<td>ADT</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VOL/ADT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VOL/ADT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VOL/ADT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VOL/ADT</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>VOL/ADT</td>
<td></td>
</tr>
</tbody>
</table>

**AVG:**
- 58.69
- 58.79
- 62.19
- 62.66
- 0.96
- 0.79
- 117
- 0.052
- 7.759
- 8.31
- 3.78
TWO-LANE HIGHWAYS
CRASH RATE VS. CASE HOUR MEAN SPEED

TWO-LANE HIGHWAYS
CRASH RATE VS. CASE HOUR 85TH %ILE SPEED

TWO-LANE HIGHWAYS
CRASH RATE VS. CONTROL HOUR MEAN SPEED
PROC IMPORT OUT= WORK.fifth
   DATAFILE= "C:\ATR\SAS_Freeways2.xls"
   DBMS=EXCEL REPLACE;
   SHEET="Input data$";
   GETNAMES=YES;
   MIXED=NO;
   SCANTEXT=YES;
   USEDATE=YES;
   SCANTIME=YES;
RUN;
data sixth; set fifth;
  if C = 0 then D = 1;
  if C = 1 then D = 0;
run;
proc logistic data=sixth ;
   class T W;
   model D = MeanSpeed Variance Volume T1 T2 T3 Dispersion Departure VolADT W / stepwise;
   output out=set1 L=lower95 P=phat U=upper95 / alpha=0.05;
   title 'Logistic analysis';
run;
Freeways – Using all input data – Output File

Logistic analysis 13:40 Friday, September 9, 2005

The LOGISTIC Procedure

Model Information

Data Set WORK.SIXTH
Response Variable D
Number of Response Levels 2
Model binary logit
Optimization Technique Fisher's scoring

Number of Observations Read 2275
Number of Observations Used 2275

Response Profile

Ordered Total
Value D Frequency
1 0 1156
2 1 1119

Probability modeled is D=0.

Stepwise Selection Procedure

Class Level Information

Class Value Design Variables
T1 0 1
1 -1
T2 0 1
1 -1
T3 0 1
1 -1

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 3153.218
Logistic analysis 13:40 Friday, September 9, 2005

The LOGISTIC Procedure

Residual Chi-Square Test

Chi-Square  DF  Pr > ChiSq
131.9254  9  <.0001

Step 1. Effect MeanSpeed entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>3155.218</td>
<td>3045.066</td>
</tr>
<tr>
<td>SC</td>
<td>3160.948</td>
<td>3056.526</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>3153.218</td>
<td>3041.066</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>112.1518</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>105.1453</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>94.5824</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

Chi-Square  DF  Pr > ChiSq
34.6054  8  <.0001

NOTE: No effects for the model in Step 1 are removed.

Step 2. Effect Departure entered:
The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>3155.218</td>
<td>3029.134</td>
</tr>
<tr>
<td>SC</td>
<td>3160.948</td>
<td>3046.323</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>3153.218</td>
<td>3023.134</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>130.0838</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>116.5137</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>98.1386</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>16.4880</td>
<td>7</td>
<td>0.0210</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 2 are removed.

Step 3. Effect VolADT entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.
Logistic analysis 13:40 Friday, September 9, 2005

The LOGISTIC Procedure

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>3155.218</td>
<td>3023.443</td>
</tr>
<tr>
<td>SC</td>
<td>3160.948</td>
<td>3046.362</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>3153.218</td>
<td>3015.443</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>137.7745</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>122.1878</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>101.6564</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.8662</td>
<td>6</td>
<td>0.1812</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 3 are removed.

Step 4. Effect Dispersion entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>3155.218</td>
<td>3017.569</td>
</tr>
<tr>
<td>SC</td>
<td>3160.948</td>
<td>3046.217</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>3153.218</td>
<td>3007.569</td>
</tr>
</tbody>
</table>
### Logistic analysis

Logistic analysis 13:40 Friday, September 9, 2005

The LOGISTIC Procedure

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>145.6491</td>
<td>4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>130.4524</td>
<td>4</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>108.1139</td>
<td>4</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1633</td>
<td>5</td>
<td>0.9483</td>
</tr>
</tbody>
</table>

Step 5. Effect MeanSpeed is removed:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>3155.218</td>
<td>3016.584</td>
</tr>
<tr>
<td>SC</td>
<td>3160.948</td>
<td>3039.503</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>3153.218</td>
<td>3008.584</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>144.6340</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>129.9010</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>108.1305</td>
<td>3</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1805</td>
<td>6</td>
<td>0.9024</td>
</tr>
</tbody>
</table>
The LOGISTIC Procedure

NOTE: No effects for the model in Step 5 are removed.

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection

<table>
<thead>
<tr>
<th>Effect</th>
<th>Number</th>
<th>Score</th>
<th>Wald</th>
<th>Pr &gt; ChiSq</th>
<th>Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>MeanSpeed</td>
<td>1</td>
<td>105.1453</td>
<td></td>
<td>&lt;.0001</td>
<td>MeanSpeed</td>
</tr>
<tr>
<td>Departure</td>
<td>2</td>
<td>17.9497</td>
<td></td>
<td>&lt;.0001</td>
<td>Departure</td>
</tr>
<tr>
<td>VolADT</td>
<td>3</td>
<td>7.6828</td>
<td></td>
<td>0.0056</td>
<td>VolADT</td>
</tr>
<tr>
<td>Dispersion</td>
<td>4</td>
<td>7.7381</td>
<td></td>
<td>0.0054</td>
<td>Dispersion</td>
</tr>
<tr>
<td>MeanSpeed</td>
<td>5</td>
<td>1.0129</td>
<td></td>
<td>0.3142</td>
<td>MeanSpeed</td>
</tr>
</tbody>
</table>

Type 3 Analysis of Effects

Wald

<table>
<thead>
<tr>
<th>Effect</th>
<th>DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>24.6293</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Departure</td>
<td>1</td>
<td>78.6437</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>VolADT</td>
<td>1</td>
<td>8.0841</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Error</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>-0.0431</td>
<td>0.1646</td>
<td>0.0687</td>
<td>0.7933</td>
</tr>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>0.1523</td>
<td>0.0307</td>
<td>24.6293</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Departure</td>
<td>1</td>
<td>-0.1141</td>
<td>0.0129</td>
<td>78.6437</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>VolADT</td>
<td>1</td>
<td>5.2652</td>
<td>1.8518</td>
<td>8.0841</td>
<td>0.0045</td>
</tr>
</tbody>
</table>

Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point Estimate</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>1.165</td>
<td>1.097 - 1.237</td>
</tr>
<tr>
<td>Departure</td>
<td>0.892</td>
<td>0.870 - 0.915</td>
</tr>
<tr>
<td>VolADT</td>
<td>193.483</td>
<td>5.133 - &gt;999.999</td>
</tr>
</tbody>
</table>
Logistic analysis 13:40 Friday, September 9, 2005 7

The LOGISTIC Procedure

Association of Predicted Probabilities and Observed Responses

<p>| | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>percent Concordant</td>
<td>60.8</td>
<td>Somers' D</td>
<td>0.224</td>
</tr>
<tr>
<td>percent Discordant</td>
<td>38.3</td>
<td>Gamma</td>
<td>0.226</td>
</tr>
<tr>
<td>percent Tied</td>
<td>0.9</td>
<td>Tau-a</td>
<td>0.112</td>
</tr>
<tr>
<td>Pairs</td>
<td>1293564</td>
<td>c</td>
<td>0.612</td>
</tr>
</tbody>
</table>
Freeways – non-weather crashes input file – output

18:31 Tuesday, October 4, 2005
The LOGISTIC Procedure

Model Information
Data Set WORK.SIXTH
Response Variable D
Number of Response Levels 2
Model binary logit
Optimization Technique Fisher's scoring

Number of Observations Read 1517
Number of Observations Used 1517

Response Profile
Ordered Value Total
   D Frequency
1   0    762
2   1    755

Probability modeled is D=0.

Stepwise Selection Procedure

Class Level Information
Class Value Design Variables
W 0   1
   1   -1
T 0   1   0   0   0
   1   0   1   0
   2   0   0   1
   3   -1  -1  -1

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 2102.976
Freeways Logistic Analysis No Weather

The LOGISTIC Procedure

Residual Chi-Square Test
Chi-Square DF Pr > ChiSq
237.5541 9 <.0001

Step 1. Effect W entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>2104.976</td>
<td>1826.775</td>
</tr>
<tr>
<td>SC</td>
<td>2110.301</td>
<td>1837.424</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>2102.976</td>
<td>1822.775</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>280.2015</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>228.7184</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>93.3863</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>10.5621</td>
<td>8</td>
<td>0.2278</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 1 are removed.

Step 2. Effect Variance entered:
The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>2104.976</td>
<td>1823.500</td>
</tr>
<tr>
<td>SC</td>
<td>2110.301</td>
<td>1839.473</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>2102.976</td>
<td>1817.500</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>285.4767</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>233.2443</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>98.2334</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2889</td>
<td>7</td>
<td>0.6248</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 2 are removed.

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection

<table>
<thead>
<tr>
<th>Effect Variable</th>
<th>Number</th>
<th>Score In Chi-Square</th>
<th>Wald Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step Entered</td>
<td>Removed</td>
<td>DF</td>
<td></td>
</tr>
<tr>
<td>Label</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 W</td>
<td>1</td>
<td>1</td>
<td>228.7184</td>
</tr>
<tr>
<td>2 Variance</td>
<td>1</td>
<td>2</td>
<td>5.2939</td>
</tr>
</tbody>
</table>
The LOGISTIC Procedure

Type 3 Analysis of Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>Wald DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>1</td>
<td>5.1524</td>
<td>0.0232</td>
</tr>
<tr>
<td>W</td>
<td>1</td>
<td>94.1028</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Error</th>
<th>Standard Chi-Square</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>1.3272</td>
<td>0.2187</td>
<td>36.8175</td>
<td>&lt;.0001</td>
<td></td>
</tr>
<tr>
<td>Variance</td>
<td>1</td>
<td>0.00598</td>
<td>0.00264</td>
<td>5.1524</td>
<td>0.0232</td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>0</td>
<td>-1.8834</td>
<td>0.1941</td>
<td>94.1028</td>
<td>&lt;.0001</td>
<td></td>
</tr>
</tbody>
</table>

Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point</th>
<th>95% Wald</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variance</td>
<td>1.006</td>
<td>1.001</td>
</tr>
<tr>
<td>W</td>
<td>0 vs 1</td>
<td>0.023</td>
</tr>
</tbody>
</table>

Association of Predicted Probabilities and Observed Responses

| percent Concordant | 64.4 | Somers' D | 0.317 |
| percent Discordant | 32.7 | Gamma     | 0.326 |
| percent Tied       | 2.8  | Tau-a     | 0.159 |
| Pairs              | 575310 | c        | 0.659 |
PROC IMPORT OUT= WORK.fifth
   DATAFILE= "C:\ATR\SAS_expressways2.xls"
   DBMS=EXCEL REPLACE;
   SHEET="Input data$";
   GETNAMES=YES;
   MIXED=NO;
   SCANTEXT=YES;
   USEDATE=YES;
   SCANTIME=YES;
RUN;

data sixth; set fifth;
   if C = 0 then D = 1;
   if C = 1 then D = 0;
run;
proc logistic data=sixth ;
   class T W;
   model D = MeanSpeed Variance Volume T Dispersion Departure VolADT W
      / stepwise;
   output out=set1 L=lower95 P=phat U=upper95 / alpha=0.05;
   title 'Logistic analysis';
run;
Logistic analysis 08:24 Friday, September 9, 2005

The LOGISTIC Procedure

Model Information

Data Set WORK.SIXTH
Response Variable D
Number of Response Levels 2
Model binary logit
Optimization Technique Fisher's scoring

Number of Observations Read 1633
Number of Observations Used 334

Response Profile

<table>
<thead>
<tr>
<th>Ordered Value</th>
<th>D Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>168</td>
</tr>
<tr>
<td>2</td>
<td>166</td>
</tr>
</tbody>
</table>

Probability modeled is D=0.

NOTE: 1299 observations were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Value</th>
<th>Design Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>T2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>T3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Step 0. Intercept entered:
The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 463.010

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.9865</td>
<td>9</td>
<td>0.9917</td>
</tr>
</tbody>
</table>

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>0.0120</td>
<td>0.1094</td>
<td>0.0120</td>
<td>0.9129</td>
</tr>
</tbody>
</table>
The LOGISTIC Procedure

Model Information

Data Set WORK.SIXTH
Response Variable D
Number of Response Levels 2
Model binary logit
Optimization Technique Fisher's scoring

Number of Observations Read 259
Number of Observations Used 259

Response Profile

Ordered Value D Total Frequency
1 0 130
2 1 129

Probability modeled is D=0.

Stepwise Selection Procedure

Class Level Information

Class Value Design Variables
W 0 1
1 -1

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 359.046
The LOGISTIC Procedure

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>46.5298</td>
<td>7</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

Step 1. Effect W entered:

Model Convergence Status

Quasi-complete separation of data points detected.

WARNING: The maximum likelihood estimate may not exist.
WARNING: The LOGISTIC procedure continues in spite of the above warning. Results shown are based on the last maximum likelihood iteration. Validity of the model fit is questionable.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>361.046</td>
<td>304.147</td>
</tr>
<tr>
<td>SC</td>
<td>364.603</td>
<td>311.261</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>359.046</td>
<td>300.147</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>58.8991</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>44.1914</td>
<td>1</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Wald</td>
<td>0.0023</td>
<td>1</td>
<td>0.9618</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.7613</td>
<td>6</td>
<td>0.8382</td>
</tr>
</tbody>
</table>

Step 2. Effect W is removed:
Expressways Logistic Analysis No Weather

The LOGISTIC Procedure

WARNING: The validity of the model fit is questionable.

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 359.046

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>46.5298</td>
<td>7</td>
<td>&lt;.0001</td>
</tr>
</tbody>
</table>

NOTE: Model building terminates because the last effect entered is removed by the Wald statistic criterion.

Summary of Stepwise Selection

Effect Step Entered Removed DF Score Wald Pr > ChiSq Variable
Intercept 1 W 1 1 44.1914 <.0001 W

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Error</th>
<th>Standard Chi-Square</th>
<th>Wald Pr &gt; ChiSq</th>
</tr>
</thead>
</table>
Two-Lane Highways – input files to SAS

PROC IMPORT OUT= WORK.fifth
   DATAFILE= "C:\ATR\SAS_2lanes.xls"
   DBMS=EXCEL REPLACE;
   SHEET="'Input data$'"
   GETNAMES=YES;
   MIXED=NO;
   SCANTEXT=YES;
   USEDATETIME=YES;
   SCANTIME=YES;
RUN;
data sixth; set fifth;
   if C = 0 then D = 1;
   if C = 1 then D = 0;
run;
proc logistic data=sixth;
   class T W;
   model D = MeanSpeed Variance Volume T Dispersion Departure VolADT W
   / stepwise;
   output out=set1 L=lower95 P=phat U=upper95 / alpha=0.05;
   title 'Logistic analysis';
run;
The LOGISTIC Procedure

Model Information

Data Set WORK.SIXTH
Response Variable D
Number of Response Levels 2
Model binary logit
Optimization Technique Fisher's scoring

Number of Observations Read 685
Number of Observations Used 683

Response Profile

Ordered Value D Total Frequency

<table>
<thead>
<tr>
<th>Value</th>
<th>D</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0</td>
<td>345</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>338</td>
</tr>
</tbody>
</table>

Probability modeled is D=0.

NOTE: 2 observations were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Value</th>
<th>Design Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>T1</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>T2</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
<tr>
<td>T3</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Step 0. Intercept entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.
The LOGISTIC Procedure

-2 Log L = 946.767

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>18.6282</td>
<td>9</td>
<td>0.0285</td>
</tr>
</tbody>
</table>

Step 1. Effect Departure entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>948.767</td>
<td>941.565</td>
</tr>
<tr>
<td>SC</td>
<td>953.294</td>
<td>950.618</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>946.767</td>
<td>937.565</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>9.2026</td>
<td>1</td>
<td>0.0024</td>
</tr>
<tr>
<td>Score</td>
<td>9.0801</td>
<td>1</td>
<td>0.0026</td>
</tr>
<tr>
<td>Wald</td>
<td>8.8649</td>
<td>1</td>
<td>0.0029</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.6342</td>
<td>8</td>
<td>0.2916</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 1 are removed.

Step 2. Effect Dispersion entered:
The LOGISTIC Procedure

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>948.767</td>
<td>933.174</td>
</tr>
<tr>
<td>SC</td>
<td>953.294</td>
<td>946.754</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>946.767</td>
<td>927.174</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>19.5930</td>
<td>2</td>
<td>&lt;.0001</td>
</tr>
<tr>
<td>Score</td>
<td>17.5624</td>
<td>2</td>
<td>0.0002</td>
</tr>
<tr>
<td>Wald</td>
<td>13.2945</td>
<td>2</td>
<td>0.0013</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.2400</td>
<td>7</td>
<td>0.9900</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 2 are removed.

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.

Summary of Stepwise Selection

<table>
<thead>
<tr>
<th>Effect</th>
<th>Number</th>
<th>Score Chi-Square</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
<th>Variable Label</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step Entered Removed DF In</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Departure</td>
<td>1</td>
<td>1</td>
<td>9.0801</td>
<td>0.0026</td>
</tr>
<tr>
<td>2</td>
<td>Dispersion</td>
<td>1</td>
<td>2</td>
<td>8.5203</td>
<td>0.0035</td>
</tr>
</tbody>
</table>
The LOGISTIC Procedure

Type 3 Analysis of Effects

Wald

<table>
<thead>
<tr>
<th>Effect</th>
<th>DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>5.1408</td>
<td>0.0234</td>
</tr>
<tr>
<td>Departure</td>
<td>1</td>
<td>9.0422</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Standard Error</th>
<th>Wald Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>0.3930</td>
<td>0.1274</td>
<td>9.5139</td>
<td>0.0020</td>
</tr>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>-0.0457</td>
<td>0.0202</td>
<td>5.1408</td>
<td>0.0234</td>
</tr>
<tr>
<td>Departure</td>
<td>1</td>
<td>-0.0573</td>
<td>0.0190</td>
<td>9.0422</td>
<td>0.0026</td>
</tr>
</tbody>
</table>

Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point Estimate</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>0.955</td>
<td>0.918 - 0.994</td>
</tr>
<tr>
<td>Departure</td>
<td>0.944</td>
<td>0.910 - 0.980</td>
</tr>
</tbody>
</table>

Association of Predicted Probabilities and Observed Responses

<table>
<thead>
<tr>
<th></th>
<th>percent Concordant</th>
<th>Somers' D</th>
<th>0.096</th>
</tr>
</thead>
<tbody>
<tr>
<td>percent Discordant</td>
<td>44.6</td>
<td>Gamma</td>
<td>0.097</td>
</tr>
<tr>
<td>percent Tied</td>
<td>1.3</td>
<td>Tau-a</td>
<td>0.048</td>
</tr>
<tr>
<td>Pairs</td>
<td>116610</td>
<td>c</td>
<td>0.548</td>
</tr>
</tbody>
</table>
The LOGISTIC Procedure

Model Information

Data Set WORK.SIXTH
Response Variable D
Number of Response Levels 2
Model binary logit
Optimization Technique Fisher's scoring

Number of Observations Read 575
Number of Observations Used 573

Response Profile

<table>
<thead>
<tr>
<th>Ordered Value</th>
<th>Total Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>290</td>
</tr>
<tr>
<td>1</td>
<td>283</td>
</tr>
</tbody>
</table>

Probability modeled is D=0.

NOTE: 2 observations were deleted due to missing values for the response or explanatory variables.

Stepwise Selection Procedure

Class Level Information

<table>
<thead>
<tr>
<th>Class</th>
<th>Value</th>
<th>Design Variables</th>
</tr>
</thead>
<tbody>
<tr>
<td>W</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>

Step 0. Intercept entered:

Model Convergence Status

Convergence criterion (GCONV=1E-8) satisfied.

-2 Log L = 794.261
Step 1. Effect Dispersion entered:

Model Convergence Status
Convergence criterion (GCONV=1E-8) satisfied.

Model Fit Statistics

<table>
<thead>
<tr>
<th>Criterion</th>
<th>Intercept Only</th>
<th>Intercept and Covariates</th>
</tr>
</thead>
<tbody>
<tr>
<td>AIC</td>
<td>796.261</td>
<td>788.640</td>
</tr>
<tr>
<td>SC</td>
<td>800.612</td>
<td>797.342</td>
</tr>
<tr>
<td>-2 Log L</td>
<td>794.261</td>
<td>784.640</td>
</tr>
</tbody>
</table>

Testing Global Null Hypothesis: BETA=0

<table>
<thead>
<tr>
<th>Test</th>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Likelihood Ratio</td>
<td>9.6210</td>
<td>1</td>
<td>0.0019</td>
</tr>
<tr>
<td>Score</td>
<td>8.0060</td>
<td>1</td>
<td>0.0047</td>
</tr>
<tr>
<td>Wald</td>
<td>4.5156</td>
<td>1</td>
<td>0.0336</td>
</tr>
</tbody>
</table>

Residual Chi-Square Test

<table>
<thead>
<tr>
<th>Chi-Square</th>
<th>DF</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.8132</td>
<td>6</td>
<td>0.9917</td>
</tr>
</tbody>
</table>

NOTE: No effects for the model in Step 1 are removed.

NOTE: No (additional) effects met the 0.05 significance level for entry into the model.
Two-Lane Highways Logistic Analysis No Weather
08:26 Tuesday, October 4, 2005

The LOGISTIC Procedure

Summary of Stepwise Selection

<table>
<thead>
<tr>
<th>Effect</th>
<th>Number</th>
<th>Score</th>
<th>Wald</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Removed</td>
<td>DF</td>
<td>In Chi-Square</td>
<td>Chi-Square</td>
<td>Pr &gt; ChiSq</td>
</tr>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>8.0060</td>
<td>0.0047</td>
<td></td>
</tr>
</tbody>
</table>

Type 3 Analysis of Effects

<table>
<thead>
<tr>
<th>Effect</th>
<th>DF</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>4.5156</td>
<td>0.0336</td>
</tr>
</tbody>
</table>

Analysis of Maximum Likelihood Estimates

<table>
<thead>
<tr>
<th>Parameter</th>
<th>DF</th>
<th>Estimate</th>
<th>Error</th>
<th>Chi-Square</th>
<th>Pr &gt; ChiSq</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>1</td>
<td>0.2262</td>
<td>0.1202</td>
<td>3.5411</td>
<td>0.0599</td>
</tr>
<tr>
<td>Dispersion</td>
<td>1</td>
<td>-0.0483</td>
<td>0.0227</td>
<td>4.5156</td>
<td>0.0336</td>
</tr>
</tbody>
</table>

Odds Ratio Estimates

<table>
<thead>
<tr>
<th>Effect</th>
<th>Point Estimate</th>
<th>95% Wald Confidence Limits</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dispersion</td>
<td>0.953</td>
<td>0.911</td>
</tr>
</tbody>
</table>

Association of Predicted Probabilities and Observed Responses

percent Concordant | 47.9 | Somers' D | -.004 |
percent Discordant | 48.4 | Gamma | -.004 |
percent Tied | 3.7 | Tau-a | -.002 |
Pairs | 82070 | c | 0.498 |
### APPENDIX D – ATR SITE DESCRIPTIONS

<table>
<thead>
<tr>
<th>ATR</th>
<th>HWY NO.</th>
<th>LENGTH</th>
<th>COUNTY</th>
<th>CRASH COUNT</th>
<th>WEATHER</th>
<th>ALL</th>
<th>WEATHER %</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>NON-WEATHER</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>100</td>
<td>I-29</td>
<td>3.98</td>
<td>Pottawattamie</td>
<td>39</td>
<td>11</td>
<td>50</td>
<td>22</td>
</tr>
<tr>
<td>102</td>
<td>I-29</td>
<td>7.32</td>
<td>Mills</td>
<td>23</td>
<td>15</td>
<td>38</td>
<td>39</td>
</tr>
<tr>
<td>104</td>
<td>I-35</td>
<td>5.57</td>
<td>Hamilton</td>
<td>28</td>
<td>48</td>
<td>76</td>
<td>63</td>
</tr>
<tr>
<td>105</td>
<td>I-29</td>
<td>6.44</td>
<td>Monona</td>
<td>38</td>
<td>13</td>
<td>51</td>
<td>25</td>
</tr>
<tr>
<td>106</td>
<td>I-35</td>
<td>4.53</td>
<td>Decatur</td>
<td>14</td>
<td>5</td>
<td>19</td>
<td>26</td>
</tr>
<tr>
<td>109</td>
<td>I-35</td>
<td>5.34</td>
<td>Cerro Gordo</td>
<td>42</td>
<td>30</td>
<td>72</td>
<td>42</td>
</tr>
<tr>
<td>110</td>
<td>I-80</td>
<td>4.39</td>
<td>Pottawattamie</td>
<td>37</td>
<td>40</td>
<td>77</td>
<td>52</td>
</tr>
<tr>
<td>111</td>
<td>I-80</td>
<td>4.38</td>
<td>Iowa</td>
<td>37</td>
<td>22</td>
<td>59</td>
<td>37</td>
</tr>
<tr>
<td>113</td>
<td>I-380</td>
<td>5.86</td>
<td>Benton</td>
<td>37</td>
<td>16</td>
<td>53</td>
<td>30</td>
</tr>
<tr>
<td>115</td>
<td>I-80</td>
<td>2.47</td>
<td>Cass</td>
<td>30</td>
<td>15</td>
<td>45</td>
<td>33</td>
</tr>
<tr>
<td>116</td>
<td>I-80</td>
<td>3.73</td>
<td>Dallas</td>
<td>55</td>
<td>35</td>
<td>90</td>
<td>39</td>
</tr>
<tr>
<td>117</td>
<td>I-80</td>
<td>1.96</td>
<td>Polk</td>
<td>85</td>
<td>74</td>
<td>159</td>
<td>47</td>
</tr>
<tr>
<td>201</td>
<td>US 65</td>
<td>1.93</td>
<td>Franklin</td>
<td>18</td>
<td>3</td>
<td>21</td>
<td>14</td>
</tr>
<tr>
<td>202</td>
<td>US 71</td>
<td>3.14</td>
<td>Audubon</td>
<td>11</td>
<td>4</td>
<td>15</td>
<td>27</td>
</tr>
<tr>
<td>203</td>
<td>US 63</td>
<td>3.03</td>
<td>Bremer</td>
<td>16</td>
<td>5</td>
<td>21</td>
<td>24</td>
</tr>
<tr>
<td>204</td>
<td>US 75</td>
<td>5.61</td>
<td>Plymouth</td>
<td>52</td>
<td>13</td>
<td>65</td>
<td>20</td>
</tr>
<tr>
<td>205</td>
<td>US 218</td>
<td>3.05</td>
<td>Henry</td>
<td>28</td>
<td>8</td>
<td>36</td>
<td>22</td>
</tr>
<tr>
<td>206</td>
<td>US 30</td>
<td>1.58</td>
<td>Story</td>
<td>17</td>
<td>4</td>
<td>21</td>
<td>19</td>
</tr>
<tr>
<td>207</td>
<td>US 18</td>
<td>2.73</td>
<td>Kossuth</td>
<td>10</td>
<td>2</td>
<td>12</td>
<td>17</td>
</tr>
<tr>
<td>208</td>
<td>US 6</td>
<td>2.57</td>
<td>Iowa</td>
<td>51</td>
<td>5</td>
<td>56</td>
<td>9</td>
</tr>
<tr>
<td>209</td>
<td>US 67</td>
<td>0.83</td>
<td>Scott</td>
<td>11</td>
<td>5</td>
<td>16</td>
<td>31</td>
</tr>
<tr>
<td>210</td>
<td>US 18</td>
<td>1.96</td>
<td>Osceola</td>
<td>7</td>
<td>4</td>
<td>11</td>
<td>36</td>
</tr>
<tr>
<td>211</td>
<td>US 20</td>
<td>2.69</td>
<td>Calhoun</td>
<td>10</td>
<td>5</td>
<td>15</td>
<td>33</td>
</tr>
<tr>
<td>217</td>
<td>Iowa 122</td>
<td>3.02</td>
<td>Cerro Gordo</td>
<td>63</td>
<td>15</td>
<td>78</td>
<td>19</td>
</tr>
<tr>
<td>219</td>
<td>US 34</td>
<td>1.14</td>
<td>Clarke</td>
<td>8</td>
<td>2</td>
<td>10</td>
<td>20</td>
</tr>
<tr>
<td>220</td>
<td>US 52</td>
<td>3.77</td>
<td>Allamakee</td>
<td>14</td>
<td>8</td>
<td>22</td>
<td>36</td>
</tr>
<tr>
<td>221</td>
<td>Iowa 9</td>
<td>3.92</td>
<td>Emmet</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>224</td>
<td>US 59</td>
<td>4.05</td>
<td>Shelby</td>
<td>8</td>
<td>1</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>226</td>
<td>US 18</td>
<td>3.47</td>
<td>Fayette</td>
<td>11</td>
<td>2</td>
<td>13</td>
<td>15</td>
</tr>
<tr>
<td>228</td>
<td>Iowa 117</td>
<td>3.01</td>
<td>Jasper</td>
<td>7</td>
<td>1</td>
<td>8</td>
<td>13</td>
</tr>
<tr>
<td>229</td>
<td>Iowa 78</td>
<td>2.50</td>
<td>Henry</td>
<td>7</td>
<td>3</td>
<td>10</td>
<td>30</td>
</tr>
<tr>
<td>ATR</td>
<td>HWY NO.</td>
<td>LENGTH</td>
<td>COUNTY</td>
<td>NON-WEATHER</td>
<td>WEATHER</td>
<td>ALL</td>
<td>WEATHER %</td>
</tr>
<tr>
<td>-----</td>
<td>---------</td>
<td>--------</td>
<td>--------------</td>
<td>-------------</td>
<td>---------</td>
<td>-----</td>
<td>-----------</td>
</tr>
<tr>
<td>230</td>
<td>US 69</td>
<td>2.61</td>
<td>Madison</td>
<td>12</td>
<td>0</td>
<td>12</td>
<td>0</td>
</tr>
<tr>
<td>231</td>
<td>Iowa 148</td>
<td>2.98</td>
<td>Cass</td>
<td>9</td>
<td>0</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>233</td>
<td>Iowa 110</td>
<td>3.78</td>
<td>Buena Vista</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>234</td>
<td>Iowa 12</td>
<td>4.80</td>
<td>Plymouth</td>
<td>30</td>
<td>6</td>
<td>36</td>
<td>17</td>
</tr>
<tr>
<td>235</td>
<td>Iowa 143</td>
<td>4.94</td>
<td>O'Brien</td>
<td>4</td>
<td>3</td>
<td>7</td>
<td>43</td>
</tr>
<tr>
<td>236</td>
<td>Iowa 196</td>
<td>2.10</td>
<td>Sac</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>238</td>
<td>Iowa 149</td>
<td>3.30</td>
<td>Keokuk</td>
<td>2</td>
<td>1</td>
<td>3</td>
<td>33</td>
</tr>
<tr>
<td>240</td>
<td>Iowa 2</td>
<td>3.72</td>
<td>Fremont</td>
<td>14</td>
<td>1</td>
<td>15</td>
<td>7</td>
</tr>
<tr>
<td>242</td>
<td>US 61</td>
<td>4.02</td>
<td>Scott</td>
<td>74</td>
<td>0</td>
<td>74</td>
<td>0</td>
</tr>
<tr>
<td>243</td>
<td>US 30</td>
<td>3.12</td>
<td>Clinton</td>
<td>27</td>
<td>9</td>
<td>36</td>
<td>25</td>
</tr>
<tr>
<td>244</td>
<td>Iowa 1</td>
<td>2.00</td>
<td>Johnson</td>
<td>28</td>
<td>6</td>
<td>34</td>
<td>18</td>
</tr>
<tr>
<td>245</td>
<td>US 20</td>
<td>2.15</td>
<td>Blackhawk</td>
<td>22</td>
<td>8</td>
<td>30</td>
<td>27</td>
</tr>
<tr>
<td>246</td>
<td>Iowa 141</td>
<td>1.02</td>
<td>Dallas</td>
<td>13</td>
<td>1</td>
<td>14</td>
<td>7</td>
</tr>
<tr>
<td>247</td>
<td>US 20</td>
<td>4.33</td>
<td>Hamilton</td>
<td>48</td>
<td>16</td>
<td>64</td>
<td>25</td>
</tr>
<tr>
<td>248</td>
<td>Iowa 21</td>
<td>2.55</td>
<td>Iowa</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>249</td>
<td>US 218</td>
<td>3.52</td>
<td>Johnson</td>
<td>40</td>
<td>13</td>
<td>53</td>
<td>25</td>
</tr>
<tr>
<td>250</td>
<td>US 65</td>
<td>2.72</td>
<td>Polk</td>
<td>21</td>
<td>18</td>
<td>39</td>
<td>46</td>
</tr>
<tr>
<td>300</td>
<td>Co Rd P14</td>
<td>3.10</td>
<td>Greene</td>
<td>5</td>
<td>2</td>
<td>7</td>
<td>29</td>
</tr>
<tr>
<td>301</td>
<td>Co Rd S14</td>
<td>2.50</td>
<td>Story</td>
<td>5</td>
<td>3</td>
<td>8</td>
<td>38</td>
</tr>
<tr>
<td>307</td>
<td>Co Rd S38</td>
<td>1.00</td>
<td>Wright</td>
<td>2</td>
<td>2</td>
<td>4</td>
<td>50</td>
</tr>
<tr>
<td>309</td>
<td>Co Rd K 52</td>
<td>1.98</td>
<td>Sioux</td>
<td>9</td>
<td>2</td>
<td>11</td>
<td>18</td>
</tr>
<tr>
<td>310</td>
<td>Co Rd Y 52</td>
<td>1.49</td>
<td>Clinton</td>
<td>3</td>
<td>2</td>
<td>5</td>
<td>40</td>
</tr>
<tr>
<td>311</td>
<td>Co Rd W 12</td>
<td>4.00</td>
<td>Benton</td>
<td>6</td>
<td>1</td>
<td>7</td>
<td>14</td>
</tr>
<tr>
<td>312</td>
<td>Co Rd E 26</td>
<td>5.00</td>
<td>Boone</td>
<td>8</td>
<td>5</td>
<td>13</td>
<td>38</td>
</tr>
<tr>
<td>314</td>
<td>CO 105</td>
<td>2.05</td>
<td>Worth</td>
<td>3</td>
<td>1</td>
<td>4</td>
<td>25</td>
</tr>
<tr>
<td>501</td>
<td>Iowa 5</td>
<td>4.67</td>
<td>Monroe</td>
<td>18</td>
<td>6</td>
<td>24</td>
<td>25</td>
</tr>
<tr>
<td>701</td>
<td>I-380</td>
<td>0.33</td>
<td>Linn</td>
<td>9</td>
<td>4</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>704</td>
<td>I-29/I-80</td>
<td>0.87</td>
<td>Pottawattamie</td>
<td>59</td>
<td>24</td>
<td>83</td>
<td>29</td>
</tr>
<tr>
<td>705</td>
<td>I-74</td>
<td>0.54</td>
<td>Scott</td>
<td>22</td>
<td>3</td>
<td>25</td>
<td>12</td>
</tr>
<tr>
<td>804</td>
<td>Iowa 4</td>
<td>0.31</td>
<td>Pocahontas</td>
<td>7</td>
<td>0</td>
<td>7</td>
<td>0</td>
</tr>
<tr>
<td>Polk</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>135</td>
<td>I-35</td>
<td>5.32</td>
<td>Polk</td>
<td>38</td>
<td>0</td>
<td>38</td>
<td>0</td>
</tr>
<tr>
<td>Wright</td>
<td></td>
<td>2.45</td>
<td>Wright</td>
<td>14</td>
<td>0</td>
<td>14</td>
<td>0</td>
</tr>
</tbody>
</table>

|       |       |        |              | 1381        | 556      | 1937 | 29        |