CORRELATION OF TRUCK ACCIDENTS WITH HIGHWAY GEOMETRY

by

Yusuf M Mohamedshah

Thesis submitted to the Faculty of the

Virginia Polytechnic Institute and State University

in partial fuflfillment of the requirements for the degree of

Master of Science

in

Civil Engineering

APPROVED:

Dr. Antoine G. Hobeika, Chairman

Richard D. Walker
Antonio Trani

Dr. Richard D. Walker
Dr. Antonio A. Trani

August 5, 1991

Blacksburg, Virginia

CORRELATION OF TRUCK ACCIDENTS WITH HIGHWAY GEOMETRY

by

Yusuf M Mohamedshah

Dr. Antoine G. Hobeika, Chairman

Civil Engineering

(ABSTRACT)

Growth trends in vehicle transportation for the year 1989 showed that truck travel has increased from 400 billion vehicle miles of travel to 600 billion vehicle miles from 1980 to 1989, a staggering 50% increase. If this trend continues, then truck travel will reach 800 billion vehicle miles by the end of the year 2000. This increase in truck travel poses operational and safety problems for both passenger vehicles and trucks. To improve the existing highway facilities for trucks as well as to determine the design standards for new truck facilities, an understanding of the relationship between truck accidents and highway geometry is required. A number of models have been developed in the past but none of them consider all of the geometric features of the highway which are crucial for truck travel and the causation of truck accidents.

The objectives of this study were to identify the roadway variables that affect truck accidents and to develop mathematical models which would determine truck involvement rates, per mile, per year.

Data from the Highway Safety Information System (HSIS) was used in this analysis. The HSIS is a new data base developed by FHWA which contains accident, roadway and traffic data from five States. Models for truck accidents on Interstates, 2 lane rural roads, and for over turning accidents on Interstates were developed. The models indicate that truck accidents are primarily affected by horizontal curvature and vertical gradient albeit their values are different for Interstates and 2 lane rural roads. The number of truck accidents decreases on 2 lane rural roads as the shoulder width increases, and the model indicates that gradient has no effect on truck accidents on these roads and this, may be due to the inadequacy of the data. The Interstate model indicates that the higher the degree of curvature and the percentage of gradient, the greater the number of truck accident, as well as overturning truck accident involvement rates.

to my parents,
brother,
and wife

Acknowledgements

I would like to thank Dr. Antoine G. Hobeika, Director of the Center for Transportation Research at Virginia Tech and Chairman of my thesis committee and Mr. Jeffery F. Paniati, Highway Research Engineer and advisor at FHWA, for their constant guidance and encouragement during the progress of the project. I would like to acknowledge Dr. Antonio A. Trani and Dr. Richard D. Walker for serving on my thesis committee.

I would like to extend my special thanks to Dr. Sam Tignor, Mr. Justine True, Mr. John Viner and Mr. Harry Lum of FHWA for their constructive suggestions during the project and to Mr. Forrest Council and his staff and Mr. Charlie Zeeger at Highway Safety Research Center at North Carolina for their invaluable support in the program development for this project. A very special thanks to Carol Conley in helping me understand SAS and the basic nature of people and Reza Saremi for a constant supply of humor and popcorn. Also I would like to thank Mike Griffith for letting me stay in his royal house when I was a novice in D. C. Also I would like to thank Madeline Shaw and my colleagues at the Information and Behavioral Systems Division at FHWA and at the Center for Transportation Research at Virginia Tech.

Last but not least, many thanks to my parents for conceiving me, my wife for being with me and my brother for encouraging me to pursue graduate studies in the USA.

Table of Contents

List of Figures

List of Tables

Research is not only getting the facts, not only experimentation,
not only building experiment--although these may be
indispensable aids; research is the conception of a new
relationship between the variables entering into the problem.

Unknown

1.0 Introduction

1.1 Background of the Problem

The economy of the USA is largely based on freight transportation and 90% of this freight movement takes place through highways by means of trucks. Growth trends in vehicle transportation for the year 1989 showed that truck travel has increased from 400 billion vehicle miles of travel to 600 billion vehicle miles from 1980 to 1989, a staggering 50% increase[16]. If this trend continues then it will reach 800 billion vehicle miles by the end of this century. This increase in truck travel causes a number of operational and safety problems on the highway as well as increasing congestion and causing extensive deterioration of the roads. The increase in pressure from the trucking industry to change the regulations would allow the industry to transport more goods in larger, heavier trucks, which poses an even greater threat to this nations highway.

In order to repair of the highways and construct new facilities for trucks that would improve the safety of the existing highway, so that trucks and passenger vehicles could exist in harmony, a clear understanding of truck accidents with other vehicles and the geometry of the highway is required. To achieve this, a mathematical relationship between truck accident rates and geometric variables is necessary to quantify the roadway variables causing truck accidents.

Though a number of models have been developed in the past, they are either single variable models or the data used for their development is only of one or two years duration, or the total mileage of roadway information used is minimal. Hence they are not able to explain the truck accidents very clearly. Also, most of the research was done in mid 60s; however, the traffic patterns have drastically changed since then, and the data bases used for the model development is outdated.

1.2 Research Objectives

The objectives of this research are:

- Study and analyze the Highway Safety Information System (HSIS) data base,
- Study truck accident characteristics, and
- Develop a mathematical relationship between truck accident

involvement rates and key roadway geometric variables.

To fulfill these objectives following tasks will be performed:

Task A: Conduct Literature Review - This concerns conducting a literature review which will determine what variables are necessary to develop a meaningful, mathematical relationship between truck accidents and roadway variables, as well as determine what type of models are most suitable for the study.

Task B: Learn HSIS Operation - This concerns learning how to handle large data bases and write suitable SAS programs to access them. Also, to develop programs which would merge the accident, vehicle and roadway file so that a master file could be formed to contain all of the information of an accident along with the roadway information.

Task C: Select Variables/State - This task will be performed to determine what variables are available in each HSIS States and identify those variables which are most suitable for model development. Then, the State that has the required variables will be decided and some statistics regarding roadway variables of that State will be developed.

Task D: Develop Merged Files - After the State is selected, various files such as accident and vehicle files will be merged together and will be used to attach the accident information on each road segment along with curve and grade information.

Task E: Test and Select Models - This task concerns the development of different models and performing statistical tests on each of them in order to select the most appropriate model.

Task F: Analysis Final Results - Finally the best model will be analyzed and the usage of the model will be shown for certain sections of the roadway and the results will be documented.

1.3 Organization of Work

The next chapter presents the result of an extensive literature review conducted as per task A. It discusses the basic issues like truck accident experience, critical variables affecting truck accidents and accident models which have been developed in the past. The third chapter deals with the HSIS data base. The first part gives a brief outline of the requirement of a good accident analysis data base and how HSIS conforms to that definition. Then it goes on to describe various files in HSIS, various States present, and the amount of data points available along with the types of variables in each State. Finally a detailed description of the data from the State of Utah is introduced, along with some statistical tables.

The fourth chapter deals with the data analysis of Utah, discussing its roadway and accident characteristics and determining the variables which will be used for the model development. Also an attempt has been made to compare the accident occurrence pattern throughout three of the HSIS States namely Utah, Michigan and Minnesota.

The fifth chapter is devoted to the development and the selection of different type of models and the definition of various variables selected for the particular model.

The sixth chapter discusses the use of the selected models and presents the conclusions drawn from them along with the recommendations for future research. After that, references are followed by appendices A and B which list the SAS programs used for this study and the bar charts developed for comparing accident characteristics for three States respectively.

2.0 Literature Review

The objective of the literature review is to obtain sufficient knowledge on the past research to establish a causal relationship between various highway elements and vehicle accidents in general and truck accidents exclusively in particular. Another objective is to determine the best highway variables which affect the truck accidents. This is an important step in the development of an accident model because considerable time and effort is required to decide which variables, would best describe the occurrence of an accident and explain maximum variability in the model and how they should be incorporated in the model. To fulfill these two objectives the literature review has been concentrated in the following areas:

- Truck accident experience;
- Critical variables affecting truck accidents; and
- Accident models developed in the past.

Though the main objective of this study is to establish a relationship between geometric variables and truck accidents, the knowledge of accident occurrence, the severity of the accident, and the factors which affect the accident, are extremely important in understanding the inter-relationships developed in the model. Truck accidents are very much different from passenger car accidents because trucks differ from a passenger car in a number of aspects which are outlined below:[15]

1. Trucks are much heavier and larger in dimension compared to passenger cars;
2. Trucks have poorer acceleration capabilities compared to passenger cars and they have greater difficulty maintaining their speeds on upgrades;
3. Trucks have slower deceleration rate in response to braking than do passenger cars;
4. The average operating and maintenance costs, per mile are higher for trucks than for cars;
5. Trucks operating on the street system are restricted to those locations where geometric design elements are sufficient for its passage;
6. The property damage costs for truck-involved accidents tend to be higher than for accidents involving only passenger cars; and
7. The average truck is 2 years older than the average car, with the result that new vehicle design standards require more time for implementation in the truck population.

Due to the above differences, truck accidents are more severe and the damage caused to passenger cars is much higher as compared to *passenger* *cars only* accidents. Thus, it is important to quantify the highway variables which affect truck accidents in order to develop a much more rational design criteria for future development of highways. Also, due to special loading and vehicle characteristics of trucks, they cause greater deterioration of highways, particularly on two-lane rural roads. Upgrading of these highways also requires accurate information on the relationships between accidents and various geometric and roadside designs.

After reviewing a number of articles, papers and publications on truck accidents modelling, truck accidents experience, and general accident modelling revealed that many results are in conflict with each other for different roadway elements. To give a few examples one study by Belmont in California[2,3] and Blensly and Head in Oregon[4] indicates that accident rates increase by wider shoulder conditions while investigations by Perkins[20], Taragin and Ekhardt[27], Raff[22] and Foody and Long[12] suggest that shoulder effects on accidents are marginal and insignificant. Yet another study by Stohner[26], Jorgensen[11] and Zeeger and Mayes[28] found a reduction in accident rates due to wider shoulders. These examples show that various conclusions are reached for the effect of a single roadway characteristic on accident occurrence and some of the reasons for such conclusions and variability are that only single variables were used for the model development; accidents were not further categorized as run-off-road, rear-end, sideswipe, head-on etc and small samples were used, the data used for the analysis was very old etc. In order to explain these variabilities in the above results and to have a common basis for comparison between different studies, a special criteria is developed by Zeeger and Perkins which is listed below:[29]

1. Type of data analysis and statistical testing performed*;
2. Reliability of the accident data sample;
3. Characteristics of roadway sections used; and
4. Accident types** used in the study.

This criteria is adopted in the literature review conducted and the results are described in the following sections.

2.1 Truck Accident Experience

This literature search was conducted

- to determine the causes of truck accidents;
- to gain an insight as to how truck accidents occur; and

---

*Statistical Tests considered are χ^2 test and t tests.

**Accident Types such as single vehicle, run-off-road, head-on etc.

- to determine the special characteristics of truck accidents.

A number of studies have been done in the past which give an idea of the previous three points. In a technical report published by FHWA in 1989[17] on improving truck safety, the University of Michigan Transportation Research Institute (UMTRI) has identified six major design deficiencies in interchanges which cause truck accidents on interchanges due to rollover and/or jackknifing of the trucks. These are: 1) poor transition to superelevation, 2) abrupt changes in compound curves, 3) short deceleration lanes preceding a tight-radius exit, 4) steep downgrade at the exit ramp, 5) provision of the curb on the outside of the entrance ramp and 6) friction levels on a high speed ramp. In another report for larger trucks[19] it was found that large trucks encroach over the edgeline and into other lanes especially on curves because the width provided is not adequate, also they have difficulty in negotiating turns at intersections which cause delays in traffic flow and there is an abrupt speed drop and shift in lateral placement by oncoming vehicles.

In an important study done by FHWA[21] regarding the transportation of hazardous materials, a number of factors have been identified which directly affect truck accidents. This study outlines the state-of-the-art of transportation of hazardous materials movement on highways. It outlines the literature review of the truck accidents in general and hazardous material accidents in particular. The geometric features which directly affect the truck accidents as identified by the report are:

- Number of lanes;
- Lane width;
- Shoulder width;
- Median width;
- Horizontal and vertical alignment;
- Surface condition; and
- Pavement condition and type.

The truck accidents are affected by vertical alignment because downgrades cause an excessive increase in the truck speed which causes runaway accidents and the rear ending of slow moving vehicles. However, while on the upgrade truck moves at a slow speed causing the rear ending of trucks by fast moving vehicles posing a potential accident threat. Usually passing zones are provided for cars and they are often inadequate for trucks and trailer combination. Some of the studies that show a comparison between a single unit truck and tractor-trailer combination* indicate that doubles are consistently over-involved in accidents by a factor of 2 or 3[18] regardless of the

---

*referred as articulated in this report.

driver's age, hours of driving, cargo weight or type of fleet. The same conclusion was reached by T. Chira-chavala and D. E. Cleveland[5] when they investigated the accident involvement rates for single unit trucks and tractor-trailers by using a multivariate discreet model showed that tractor trailer were highly involved in an accident as compared to single unit trucks. This model was developed using trailer style, vehicle configuration, number of axles of power unit, trip length, road class, road surface condition, loading status, day/night condition, driver experience and driver age as input variables. Also in a study conducted by using 1977 BMCS[7] data, it was found that: wet snowy pavements raised the accident rates of all trucks on all roads, wet-snowy conditions at night were specially hazardous, urban roads have higher accident rates than rural roads and doubles usually showed higher accident rate on undivided rural roads than on divided rural roads. However, one study conducted by Jovanis et.al.[18] indicates contrary results, that doubles have better safety performance than singles except during the transition year 1984 and large trucks travelling on non-access-controlled highways have consistently higher accident rates than those on access-controlled* highways.

One significant factor determined from the literature search was the fact that truck exposure data is the weakest link in any truck accident study and

---

*Controlled by a stop sign or a traffic signal.

analysis. This is because no databases have specific exposure data for trucks in general and for truck types* in particular. Also, the determination of truck accident rates is erroneous because any accident involving at least one truck is classified as a truck accident, and consequently, dividing those accidents by truck ADT to determine the truck accident rate, which causes the truck accident rates to be of a larger magnitude than they would be if the nontruck exposure was also considered. The reason behind this, is that multivehicle accidents involving trucks and nontrucks are only counted as truck accidents. To get a true picture of truck accident occurrence, truck involvement rates should be used instead. This is calculated as the total number of trucks involved in an accident divided by truck ADT.

For example[6] Michigan recorded a total of 84,640 truck accidents during 1977; where a truck accident is defined as one that involves at least one truck. These truck accidents involved approximately 90,000 trucks and 63,000 nontrucks. The truck accident rate for this example would be higher as compared to truck involvement rates and thus, the traditional method of accident rate determination artificially causes high truck accident rates.

---

*Such as straight trucks with or without trailer, tractor/semitrailer, western double, rocky mountain double or turnpike double.

In a recently published report by TRB[9] on data requirements for monitoring truck safety, the emphasis has been placed to have a greater quality control in collecting truck travel data to get a better truck exposure. According to this report there is currently no consensus among existing programs on the truck travel estimates, which are required in computing accident rates. Also there are no truck travel data that are consistent over a period of time and provide details by type of truck, road class, and geographical area. Some of the current data bases which give information on truck travel are *Truck Inventory and Use Survey* (TIUS), *The National Truck Activity and Commodity Survey* (NTACS), *National Truck Trip Information Survey* (NTTIS), *Highway Performance Monitoring System* (HPMS) and, *Annual Truck Weight Study* (ATWS). All of these data bases do not capture current truck populations because of the time lag between the sampling and survey years.

From the above discussion it is apparent that TRB is still in the process of formulating a policy for creating an ideal data base which could be used for truck safety studies and which would give a good measure of truck exposure. Due to the above limitations of the data bases, the current study will try to make the optimum use of the HSIS data base so that the results could be viewed with a confiecnt degree of certainty. The exposure data will be used to calculate the truck involvement rate rather than the truck accident rate which is erroneous by the fact that the method of determining it causes truck accident rates to be of larger magnitude.

2.2 Critical Variables Affecting Truck Accidents

The first task in developing the model is to determine the types of accidents that are related to trucks. Although a number of studies have been done to determine the most probable variables affecting accidents, very few have been done exclusively for trucks. Several studies analyzed total accidents by accident severity, pavement condition or by time of the day. However, detailed accident types i. e. run-off-road, head-on, rear-end, right angle etc was analyzed in studies by Rinde (1977)[23], Zeeger et. al. (1979)[9], and Rogness et. al. (1982).[24] Based on the results of these three studies, there is strong evidence that the run-off-road and opposite direction accidents are primary types affected by lane and/or shoulder improvements. Another study[30] on rural-two lane roadway, using a caravan of three control vehicles (trucks) and a camera mounted on it for taking the photographs of an on coming vehicle, concluded that lane width, shoulder width and the presence of curvature affected the operation of large trucks, while degree of curvature had little effect on the lateral placement of trucks, but large degrees of curvature, greater then

7°, did cause opposing vehicles to slow down while passing large trucks.

In a 1978 study done by Roy Jorgensen Associates[25] on determining the relationship between accident frequency and shoulder width, it was found that on tangents, as the right shoulder width increases beyond the width necessary to accommodate a parked vehicle, the safety benefits become insignificant. As the right shoulder width increases on curves, the accident rate decreases, and also paved right shoulders produce fewer accidents than unpaved right shoulders. Another study by Zeeger et.al.[31] in Illinois indicates that widening of 18-ft pavement to 22-ft causes a reduction of 39% in accidents per million vehicle miles. A literature review conducted by Zeeger et. al.[32] concluded that numerous geometric, traffic and roadway variables interrelate with lane and shoulder widths and shoulder type in affecting accidents on two-lane rural roads. These include roadside characteristics, horizontal and vertical curvature, traffic volume, access points, intersections, and others. Thus these might be important independent variables affecting the accidents.

The analysis of the literature review indicates that most of the studies have been done for all accidents in which very few have attempted to categorize accidents by its types. The most important variables that emerge from the review and affect the accidents most are: shoulder width, lane width, and shoulder type, while other variables such as the degree of horizontal and vertical curves, traffic volume and roadside characteristics may act as supplementary variables which would be able to explain the remaining variability in accidents. Also, the above studies do not indicate which variables affect the truck accidents exclusively. This is very difficult to determine because an accident is a very complex phenomenon which involves vehicle, road, and driver. Yet there are certain characteristics which are different in an accident involving passenger cars and an accident involving trucks.

2.3 Accident Models Developed in the Past

Considerable effort has gone into the development of a causal model which quantifies various geometric and accident variables causing accidents, however very few studies have been conducted which tried to develop a mathematical model for truck accidents exclusively. An all accident model was developed by Zeeger et. al[31]. in which roadway, accident, vehicle and traffic data was collected for 5,000 miles of two lane roads from the following seven State highway agencies:

- Alabama
- Michigan

- Montana
- North Carolina
- Utah
- Washington
- West Virginia

The data consisted of a wide distribution of geographic characteristics, climate conditions, roadway designs, terrain conditions, traffic conditions, and other factors. The regression model developed had an R^2 of 0.456 and is given below:

$$AO/M/Y = 0.0019(ADT)^{.8824}(0.8786)^{W}(0.9192)^{PA}(0.9316)^{UP}(1.2365)^{H}$$

$$(0.8822)^{TER1}(1.3221)^{TER2} \tag{2.1}$$

where

AO/M/Y = Single vehicle plus opposite direction head-on, opposite direction sideswipe, and same direction sideswipe accidents per-mile-per-year

ADT = average daily traffic

W = lane width (feet)

PA = average paved shoulder width (feet)

UP = average gravel/stabilized/earth/grass shoulder width (feet)

H = median roadside (or hazard) rating

TER1 = terrain condition (1 if flat and 0 otherwise)

TER2 = terrain condition (1 if mountainous and 0 otherwise)

This model showed that the first foot of lane widening (i.e. two feet of pavement widening) corresponds to a 12 percent reduction in related (AO) accidents; two feet of lane widening (e.g. from 9 to 11 feet) results in a 23 percent reduction, three feet results in a 32 percent reduction, and four feet of widening would result in a 40 percent reduction, however these reductions apply only for lane widths between 8 and 12 feet. Also for shoulder widths between 0 and 12 feet, the percent reduction in related (AO) accidents due to adding paved shoulders is 16 percent for 2 feet widening, 29 percent for 4 feet widening, and 40 percent for 6 feet widening. Adding an unpaved shoulder would result in 13 percent, 25 percent, and 35 percent reductions in related accidents for 2, 4, and 6 feet of widening respectively.

In one of the models developed by Nicholas Garber[13] to investigate the effect of truck strategies such as trucks in one particular lane, impacts on traffic volumes due to differential speed limit regulations etc., which were eventually used to determine truck accident involvement rates due to the imposition of these strategies showed that the model had only one independent variable, TRVOL (truck volume), and one dependent variable, TRINV (truck involvement), which is:

$$TRINV = 8.27 + 0.00278 \times TRVOL \tag{2.2}$$

In this equation, the truck accident rate is not used, instead, the truck involvement rate is used, which means that if there were two trucks involved in an accident that accident will be counted twice. Hence, it takes care of some of the bias which occurs by considering just the truck accident rates. In another study[14] done by the same author for the Virginia Department of Transportation, different models were developed for estimating the truck accident rate and its involvement using linear and Poisson regression methods. It was found that the log linear model using Poisson regression was able to explain truck accidents much more reliably than the linear models. The Poisson model consisted of slope change rate (SCR), ADT and truck percentage (TPERCNT) as independent variables and truck accident involvements (TRINVOL) as dependant variable. Akaike's Information Criteria (AIC) was used for selecting the best model. The model developed had an AIC of 62.06 and is shown below:

$$TINVOL = .015237(SCR)^{.0577}(ADT)^{.5024}(TPERCNT)^{.5731} \quad (2.3)$$

This model does not have variables for lane width and road segment length because all the sites used for the development of this model had a segment length of about 2 miles. Also, it was found that there was some correlation between ADT and lane width; hence, the lane width was excluded from the model. In developing a relationship between rural highway geometry and accident rates in Louisiana, Dart and Mann[10] found that the percentage of trucks, traffic volume ratio, lane width, shoulder width, pavement cross slope, horizontal alignment, vertical alignment, percentage of continuous obstructions, marginal obstructions per mile, and traffic access points per mile were significant variables in the model. The study was conducted on approximately 1,000 miles of rural highway. The model used the total number of accidents, accidents on wet and dry roads, accidents during day and night time, total injuries and total fatalities as the dependant variables. The model for the total number of accidents had an R^2 of 0.46 indicating that 46% of the variation in accident rates could be explained by using the geometric variables included in the model, while 54% of the variation was due to driver or vehicle characteristics or variables not included in the model.

2.4 Conclusions from the Past Research

The literature review reveals that considerable work has been done in the past which tried to quantify various geometric and traffic variables to explain the occurrence of an accident. But very few studies have been conducted exclusively for trucks and the results from them are conflicting in nature. Very few of the above studies provide an accurate prediction of accidents and none provide the prediction of truck accidents. The present study of developing a regression model for trucks is undertaken to fill this void and hopefully come with an equation which would give concise definitions and measures to better describe traffic and geometric variables for highway safety, construction and maintenance purpose. The relevance of the above studies, however, is immense as they have tried to at least identify the variables which affect the accidents and have narrowed down to handful of them which are most important. The important variables which emerge from the above studies are shoulder width, shoulder type, median width, median type, ADT and lane width supplemented by the variables indicating the curvature and gradient of the roadway segment.

Slowly but surely humanity achieves what its wise men have dreamed.

Anatole France

3.0 HSIS: An Overview

3.1 The Need for New Data Base

A large number of data bases exist which give information about accidents, driver characteristics, roadway segments, vehicle characteristics and roadside structure inventories; however, all data bases do not have this information. Some data bases contain only State level information while some have national statistics. Most data bases could not be used to conduct a sound study to determine roadway or vehicles safety or to formulate certain safety policies based on their analysis. The most important reason for this is that the data bases do not have all the information of an accident such as where it occurred, what were the factors involved in its occurrence, what was the geometry of the road section on which it occurred etc. Some of them do have this information but the number of data points might not be adequate to perform a statistical analysis.

Another important factor in a safety study is that the database should have information on failures as well as successes of an accident, i.e., it should have all the road sections where accidents occurred as well as those sections where accidents did not occur. Without both, it is virtually impossible to determine the factors which differentiate from one another, and that is the basic nature of safety research.

The national data bases which exists are accident-based. They provide information on the specifics of vehicles and drivers, but do not provide information about the highway system and its characteristics. They cannot identify or provide data on accident free locations and as a result, they have limited applications for highway-related analyses.

Due to the above deficiencies and short-comings of the data bases in 1983 the Federal Highway Administration began a study to develop a data base which would meet the highway safety analysis needs. *Highway Safety Information System* (HSIS) is the result of this study which was conducted by FHWA in conjunction with University of North Carolina Highway Safety Research Center (HSRC).

3.2 Description of HSIS

Highway Safety Information System has data files on accidents, vehicles, drivers, roadway segments and other roadside inventories. Five States are included in this system which are Utah, Minnesota, Michigan, Illinois and Maine and the data is available for five years from 1985 to 1989. Currently, the data from Pennsylvania is being included in the data base. Each State was contacted and the police accident reports were obtained and coded in a Statistical Analysis System* (SAS) format to form easily mergeable files. The files are named as accident file, vehicle file, driver file and roadlog file accompanied by interchange, intersection and guardrail inventory files for some States. After coding the raw data into SAS format, quality control checks were performed on a majority of the variables from each State and on each type of file.

The purpose of converting the raw data** into SAS format was to make the output more readable and directly accessable by the SAS package to obtain cross tabulations, perform regression analysis and frequency counts or perform any other SAS procedures. Each of the variables have a unique name along

---

*A data base handling software package

**Police accident records

with a 16 character label description to identify the categories of the variable. Also the data files could be merged together to perform different types of analysis. The merging of the files could be done by using certain variables which are common to two or more files being merged. The rationale of keeping the information in separate files is that different kinds of analyses require data in different fashions, and instead of using a full data base every time for various studies, small files could be used for a particular study. This makes processing the data easier and faster and also it saves computer memory.

The data from the five States is not merged together to form a huge data set because there is no common system of variable definitions applied across all the States. Also, there are differences in variable names for similar variables, and large differences in the category labels for the same variable across the States. Thus, combining the data from the same variable from different States would require adopting a lowest common denominator definition in which case a large amount of data information and specificity would be lost. Table 3.1 (see page 28) gives an idea of the quantity of the data available in HSIS for the selected States and Table 3.2 (see page 28) shows types of files available from selected States.

Table 3.1 HSIS Data Quantity

States	Accidents / year	Roadway Mileage
Illinois	160,000	16,000
Maine	40,000	22,000
Michigan	145,000	10,000
Minnesota	70,000	60,000
Utah	37,000	13,000
Pennsylvania	150,000	--

Table 3.2 Files Available from HSIS States

Type of Files	Utah	Minnesota	Illinois	Maine	Michigan
Accident	X	X	X	X	X
Roadway Inventory	X	X	X	X	X
Traffic Volumes	X	X	X	X	X
Roadway Geometrics	X		X		X
Intersection Data	X	X			X
Guardrail					X

Special Guidebooks are available to give the user information about the data present in the HSIS for each State. The Guidebooks provide a listing of all the variables with a brief description of each category in the variable. Each Guidebook is divided into four sections which are briefly described below:

Section A: *Basic description:* A general description of the State data is provided with an overview of the types of data available. In addition, there is a detailed description regarding the quality control checks performed and a comment on which variables are more reliable to use as compared to others.

Section B: *SAS formats:* SAS format names and category labels for each variable in each State file are included in this section.

Section C: *Single variable tabulations:* This section gives the single variable tables described in section B. These tables give row and column percentages along with the percentages of each category in a particular variable, thus, giving an overall idea of the total data points available for the majority of the variables.

Section D: *Computer programs:* Basic programs to process and merge the variables are included in this section. In addition, this section also contain programs which combine files to calculate basic accident rates.

3.3 Analysis of HSIS Data Base

The major task in the analysis of the data base was to determine the presence of variables that affect an accident and the adequacy of the data points in performing safety research exclusively for trucks.

A preliminary study of the data base showed that it had enough variables and data points for all accidents as well as for truck accidents to do safety research. To get a better feel of the data, single variable tables present at the end of the Guidebooks were thoroughly studied for each State and it was found that the State of Utah had all the variables required for the study. Particularly the curve and grade files had useful information on the degree of curvature and the gradient of the roadway segment which is not present in other States. Tables 3.3 and 3.4 (see pages 31, 33) show the different variables present in each State's file, their availability and the name by which they could be identified. The State of Utah shows that it has a maximum availability of accident and roadway variables while other States have a large number of cells

Table 3.3 Variables classified according to States

Variable	Utah	Maine	Minnesota	Michigan	Illinois
Accident milepost	A_MILEPT*	YRCASE*	ACCNUM* REF_PNT*	-------	MILEPOST*
Accident route number	A_ROUTE*	C_LINK*	VEH_NBR* REF_PST* RTE_NBR*	CASENUM* RTE_NBR	COUNTY* RTE_NBR*
Type of accident	ACC_TYPE	ACCTYPE	ACCTYPE	ANALYS	N/A
Severity of accident	ACCSEV	N/A	SEVERITY	SEVERITY INJ_SEV	ACCSEV SEV_CDE
Accident year	ACCYR	ACCYR	ACCYR	ACC_YR	YEAR
Collision type	COLLTYPE	N/A	ACCDIGM	ACCTYPE	COL_TYPE
Light	LIGHT	LIGHT	LIGHTCON	N/A	LIGHT
Road char	RDCHAR	RDCHAR	RDCHAR	RD_ALIGN	N/A
Road condition	RDSURF RDEFECT	RDCOND	RDSURF	RDSURF RD_DEF	RD_DEFEC RD_SURF
Total injured	TOTNINJ TOTOINJ	OCC_INJ TOT_NON	INJURED	TOT_INJ	TOT_INJ
Total killed	TOTNK TOTOK	TOT_K	KILLED	TOT_KILL	TOT_KILL
Weather	WEATHER	WEATHER	WEATHER	WEATHER	WEATHER
Impact speed	IMPT_SPD	N/A	N/A	N/A	N/A
Posted speed	SPDLIMIT	SPDLMT	SPEED DESG_SPD	SPD_LIMT	SPD_LIMT
Travel speed	TRVL_SPD	N/A	N/A	N/A	N/A

* This variable is used to merge various files.

Table 3:3 (contd) Variables classified according to States

Variable	Utah	Maine	Minnesota	Michigan	Illinois
Vehicle Type	VEHTYPE	VEHTYPE	TYPE_VEH	VEH_TYP	VEHTYPE
Truck ADT	N/A	N/A	HAADT	COM_ADT	COM_VOL
End Milepost	ENDMP	END_PT	ENDMP	END_SEG ENDMP	END
Begin Milepost	BEGMP	BEG_PT	BEGMP	BEG_SEG BEGMP	BEGIN
Func Class	FUNC	FUNC	FUNC_CLS	N/A	FUNC_CLS
Roadway Type	ONEWAY	ONEWAY	DIV_CODE	RDWY_TYP	DIR_OPER
Lane Width	LANEWID	PAVEWID	SURF_WD1	LANE_WD	SURF_WD
Median Type	MEDTYPE	N/A	MED_TYPE	MED_TYPE	MED_TYPE
Median Width	MEDWID	MEDIAN	MED_WIDT	N/A	MED_WIDT
No. of Lanes	NO_LANE	LANES	NBRLN_DM NBRLN_IM	NBR_LANS	NBR_LNS
Pavement Type	PAVETYPE	SUFTYPE	SURF_TY1	N/A	SURF_TYP
Shoulder Width	SHLDWIDL SHLDWIDR	LSHLDW RSHLDW	LSHL_WD1 RSHL_WD1	PAV_WIDL PAV_WIDR	SHLD_WD
Shoulder Type	SHLDTYP	LSHLDT RSHLDT	LSHL_TY1 RSHL_TY1	CUR_TYPR	IN_SHDTY
Terrain	TERRAIN	N/A	N/A	N/A	N/A
Section Length	SECT_LNG	SEG_LNG	SEG_LNG	N/A	SEG_LNGT
Degree of Curve	DEG_CURV	N/A	N/A	CURV_DEG	DEF_ANGL
Percent Grade	PCT_GRD	N/A	N/A	N/A	N/A

* This variable is used to merge various files.

Table 3.4 Availability of Variables for All States

Description of Variables	Utah	Maine	Minnesota	Michigan	Illinois
Type of Accident	Available	Available	Available 20% coded "other"	Available as Accident analysis	N/A
Type of Collision	Available	N/A	Available	Available as Accident Type	Available
Number of Vehicles	Available	Available	Available	Moving Vehicles	Available
Road Char	Available	Available	Available	Available	N/A
Type of Vehicle	Available	Available	Available	Available	Available
AADT	Available	Available	Available	Available	Available
Commercial AADT	Peak Percentage	N/A	Available	Available for 1985	Available
Func class	Available	Available	Available	N/A	Available
Lane Width	38% uncoded	Pavement Width	N/A	Available	N/A
Median Type	Available	N/A	Available	Available 76% uncoded	Available
Median Width	Errors in coding	Available	Available	N/A	Available
Number of Lanes	Errors in coding	Use With Caution	Available	Available Data Suspect	Available
Pavement Type	Available	Surface Type	Surface Type	N/A	N/A
Shoulder Type	Available	Available	Available	N/A	Available
Shoulder Width	Available	Available	Available	Available	Available
Horizontal Curves	Available	N/A	N/A	Available	Available
Vertical Grades	Available	N/A	N/A	N/A	Available

in the table showing N/A (not available). Also for a truck accident study, the vertical grade and horizontal curve information is required and it is present in the Utah and Illinois files only, but Illinois does not have adequate curve and gradient information. Thus, the State of Utah is chosen to develop the mathematical model while the other States will be used to develop the bar charts to determine similarities and dissimilarities in accident occurrence across the States.

Another factor which weighed heavily in favor of using the Utah file was the variable VEHTYPE. This variable describes the type of vehicle involved in an accident. As the mathematical model has to be developed exclusively for trucks, the variable VEHTYPE should have the category showing trucks and non-trucks. Tables 3.5, 3.6, 3.7, 3.8 and 3.9 (see pages 35, 36, 37, 38, 39) show various categories available in VEHTYPE variable for each State. From the tables, it is seen that the States of Utah, Maine and Minnesota have a good classification of trucks, but Maine has certain categories of trucks which overlap and it will be difficult to segregate them, while Minnesota has good truck classification but it does not have curve and grade information. Thus the State of Utah is the best candidate for development of the models.

Table 3.5 Vehicle Types Broken Down by Percentage for State of Utah

Vehicle Type	Total	Percentage
Error Code	125	0.00
Passenger Cars	121,065	30.10
Car/Compact	143,593	35.70
Car and House Trailer	234	0.10
Car & Boat	143	0.00
Car & Other Trailer	149	0.00
Car/Public Own	795	0.20
Pickup/Panel	101,475	25.20
Pickup & House	450	0.10
Pickup & Boat	266	0.10
Pickup & Other	1,027	0.30
Pickup/Pub Own	823	0.20
Pickup With Camper	547	0.10
One Unit Truck	5,393	1.30
Truck & Trailer	782	0.20
Semi (Bobtail)	402	0.10
Semi & Trailer	5,566	1.40
Comm. bus	947	0.20
School Bus	526	0.10
Motorcycle	6,022	1.50
Motorcycle/Public	67	0.00
Motor Scooter	103	0.00
Ambulance/Non	25	0.00
Ambulance/Emergency	59	0.00
Ambulance/Pub	16	0.00
Farm Tractor	172	0.00
Special Equipment	465	0.10
Truck + Mobile Home	310	0.10
Other Vehicle	681	0.20
Moped	82	0.00
Truck+2 Short Trailers	37	0.00
Truck + Long Trailer	174	0.00
Semi + 2 Short Trailers	253	0.10
Semi + 2 Trailers	274	0.10
Semi + 2 Long Trailers	85	0.00
Semi + Long/Short	219	0.10
Semi + 3 Short	90	0.00
Hit & Run Vehicle	8,922	2.20
Total	402,364	100.00

Table 3.6 Vehicle Type Broken Down by Percentage for State of Maine

Vehicle Type	Total	Percentage
2 Door	87,773	32.25
4 Door	74,258	27.28
Convertible	97	0.04
Station Wagon	22,604	8.30
Van/Campers	9,951	3.66
Pickup	43,029	15.81
Truck	4,786	1.76
Truck Trailer	236	0.09
Semi Trailer	239	0.09
Semi Tank	20	0.01
Bus	296	0.11
School Bus	602	0.22
Motor Home	122	0.04
Motorcycle	3,736	1.37
Moped	128	0.05
Motor Bike	25	0.05
Bicycle	1,507	0.55
Snowmobile	94	0.03
Pedestrian	3,498	1.29
2 Axle Truck 2/Tire	3,402	1.25
2 Axle S/Axle Semi	89	0.03
2 Axle T/Axle Semi	396	0.15
All Terrain Vehicle	153	0.06
2 Axle 1S2A Trailer	3	0.00
3 Axle/One Unit	1,581	0.58
3 Axle S/Axle/Semi	42	0.02
3 Axle T/Axle Semi	2,834	1.04
3 Axle Truck/Axle Semi	474	0.17
3 Axle 1AS2A Trailer	3	0.00
3 Axle 2AS2A Trailer	2	0.00
3 Axle 2AS3A Trailer	9	0.00
3 Axle 2AS3A Trailer/R	2	0.00
3 Axle 2AS4A Trailer	8	0.00
4 Axle/One Unit	309	0.11
4 Axle Truck/T Axle	19	0.01
3 & 4 Axle N/List	117	0.04
5 Axle N/List	10	0.00
All Other Semi	19	0.01
Farm Vehicle	43	0.02
Unknown	6,303	2.32
Total	272,202	100.00

Table 3.7 Vehicle Type Broken Down by Percentage for State of Minnesota

Vehicle Type	Total	Percentage
Invalid Data	2	0.00
Automobile	289,401	71.26
Auto with Trailer	704	0.17
Truck Tractor/ Trailer	7,978	1.96
Tru/Tract W/Semi	6,969	1.72
Tru/Tract W/Twin	110	0.03
Tru/Tract W/Other	1,047	0.26
Pickup Truck	43,337	10.67
Van	11,612	2.86
Motorcycle	6,391	1.57
Moto Scooter	139	0.03
Moped	218	0.05
ATV	158	0.04
School Bus	1,339	0.33
Other Bus	1,109	0.27
Motor Home/Camp	441	0.11
Snowmobile	131	0.03
Farm Tract/Equip	521	0.13
Taxicab	546	0.13
Hit & Run Vehicle	22,324	5.50
Police Vehicle	1,077	0.27
Fire Dept Vehicle	98	0.02
Ambulance	102	0.03
Military Veh	48	0.01
Rd Main Vehicle	679	0.17
Oth Pub Own Vehicles	1,006	0.25
Oth Priv Own Vehicles	240	0.06
Bicyclist	3,872	0.95
Pedestrian	4,534	1.12
92	13	0.00
Total	406,146	100.00

Table 3.8 Vehicle Type Broken Down by Percentage for State of Michigan

Vehicle Type	Total	Percentage
Under 1500 Lb	1,642	0.21
1500 To 2499 Lb	459,924	59.75
2500 To 3500 Lb	52,761	6.85
+ 3500 Lb	47,655	6.19
Carryall	18,424	2.39
Jeep Type	7,431	0.97
Pickup	86,169	11.20
State / Dump Truck	16,436	2.14
Truck Tractor	22,058	2.87
Other or Not Known	57,190	7.43
Total	769,690	100.00

Table 3.9 Vehicle Type Broken Down by Percentage for State of Illinois

Vehicle Type	Total	Percentage
Not Stated	12,384	1.30
Passenger Car Large	460,439	48.59
Passenger Car Small	283,091	29.87
Single Unit Truck	25,362	2.68
Trailer Tractor W/Semi	35,600	3.76
Van	38,684	4.08
Pickup	77,562	8.18
Farm Equipment	433	0.05
Bus	4,181	0.44
Motocycle	8,412	0.89
Other	1,523	0.16
Total	947,635	100.00

3.4 Analysis of Utah Data Files

3.4.1 General Description of Utah Files

The Guidebook of Utah shows the following files:

- Accident Data
  - Accident Subfile
  - Vehicle Subfile
  - Occupant Subfile
- Roads File
- Horizontal Curve File
- Vertical Grade File
- RR Grade Crossing File
- Bridge File
- Materials File (Pavement)

Accident subfile can be linked to Vehicle subfile and Occupant subfile by using an accident control number (YRCASE) and an accident year (ACCYR), while Accident and other major files can be linked using route-milepoint system variables and accident year with the mile points for a given route beginning at the State Line. The Roads file has information on shoulders, medians, pavement type, width and the number of lanes for 13,000 miles of road, as well as having AADT. Horizontal Curve and Vertical Grade files can be merged with Road Files and Accidents Files by route-milepoint key.

3.4.2 Brief Description of each File

Accident Subfile: It contains approximately 37,000 accidents per year involving 665,000 vehicles and 110,000 occupants/pedestrians. The occupant subfile contains information about all occupants whether injured or not. Seventy percent are property damage accidents, 0.6% are fatal accidents, 72% are multivehicle accidents and 28% are single vehicle accidents. The accident reporting threshold is $400 total damage and/or personal injury.

Vehicle Subfile: It contains the information related to vehicle such as vehicle type, driver characteristics, year of accident, speed limit of the highway and the contributing factors for an accident. Some of the variables are not properly coded but those variables are insignificant for the present study.

Occupant Subfile: This file contains information regarding the occupant of a vehicle involved in an accident. It describes injuries suffered by the occupant, sex of the occupant, age of the occupant and the safety equipment, such as safety belt, used by the occupant at the time of the accident.

Road File: It has approximately 10,000 records each year for 13,000 miles of road comprising of 70% of primary roads. Approximately 84% of the roadway mileage on the file are two-lane sections, and approximately 10% have four or more lanes. Most of the road segments on the file are low volume with approximately 60% of the mileage having ADT's of 500 vehicles per day or less.

Horizontal Curve File: It has 13,659 records covering 5820 miles of roadway and contains data such as the degree of the curve, the direction of the curve and the type of roadway on which the tangent and curved sections fall. The file included in HSIS has data only for the year 1987.

Vertical Grade Curve: It has 5,540 records covering 5850 miles of roadway and contains data such as percent grade, direction of grade +ve or -ve and road type.

RR Grade Crossing File: This file contains information on all grade-crossings in the State, and is prepared and maintained according to FHWA requirements.

Bridge File: It contains information on bridge structures across the State. The data is quite accurate since it is based on federal bridge inventory.

Pavement File: This file contains information describing pavement characteristics such as thickness, type, design load and pavement condition such as skid number, Pavement Serviceability Index, amount of cracking, and ruts for various classes of roadways.

The last three files are not yet converted into SAS format; hence, they are only available in raw format.

'The time has come,' the Walrus said,
'To talk of many things:
Of shoes--and ships--and sealing wax--
Of caggabes--and kings--
And why the sea is boiling hot--
And whether pigs have wings.'

Through The Looking Glass
(Alice in Wonderland)

4.0 Results of Data Analysis

The analysis of the data was done to determine the basic characteristics of the data base. The analysis is divided into the following categories:

- Data Base Characteristics
- General Accident Experience
- Determination of Key Variables

4.1 Data Base Characteristics

4.1.1 Roadway Characteristics

The Utah data base has five years of roadway information from 1985 to 1989. The single variable tables at the end of the Guidebook indicated very little change in the roadway characteristics over the five year period and hence, for simplicity, the road file from 1987 is considered for further analysis.

The 1987 road file contains 11149.7 miles of rural roadways (5770 sections) and 1786.47 miles of urban roadways[*] (4056 sections), for a total of 12936.2 miles with 9826 sections. The average section length for rural roads is 1.9 miles while for urban roads it is 0.44 miles with 1.32 miles overall. Data consists of information on shoulder width ranging from 0 to 28 feet and lane width ranging from 0 to 24 feet as shown in the table 4.1 (see page 47). Table 4.2 (see page 48) shows the ADT for different lane widths. Approximately 29% of the sections have ADT of 500 or less while 5112 sections have ADT ranging from 501 to 50,000.

A summary of the number of sections and mileage (in parenthesis) of the data base by area type and speed limit is shown in Table 4.3 (see page 49). The speed limit varies from 15 to 65 mph with 47% of the sections having a speed limit of zero which indicates that those sections are not properly coded. Out of 9826 sections, 4698 sections have uncoded speed limits while from the remaining sections 1911 (approximately 20%) of them have speed limits of 55 mph. Also 1489 sections with a speed limit of 55 mph lies in rural segments while only 422 sections are in urban areas. Table 4.4 (see page 50) shows the road sections classified according to the terrain condition. This table shows that

---

[*] This is a combined category having *small urban* and *urbanized* sections.

Table 4.1 Descriptive Statistics for Utah Roadlog File (Lane Width by Shoulder Width)

Number of Sections, with total mileage in parenthesis							
Shoulder Width(ft)	Lane Width (ft)						Total
	0 - 8	9	10	11	12	13 - 24	
0 - 1	1483 (3641)	92 (215.56)	258 (739.56)	385 (499.85)	3029 (2849.8)	74 (103.13)	5321 (8048.9)
2 - 3	3 (0.35)	2 (0.7)	4 (5.28)	22 (36.62)	1304 (1806.7)	1 (0.33)	1336 (1850)
4 - 5	1 (0.25)	0 (0.00)	3 (15.49)	8 (11.16)	836 (889.26)	0 (0.21)	848 (916.16)
6 - 7	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	684 (723.5)	1 (0.53)	685 (724.03)
8 - 9	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	484 (296.69)	0 (0.00)	484 (296.69)
> = 10	0 (0.00)	0 (0.00)	0 (0.00)	3 (1.4)	1148 (1098.92)	1 (0.10)	1153 (1100.46)
Total	1487 (3641.6)	94 (216.26)	265 (760.33)	418 (549.03)	7485 (7664.87)	77 (104.09)	9826 (12936.2)

Table 4.2 Descriptive Statistics for Utah Roadlog File (Lane Width by AADT)

Number of Sections, with Total Mileage in Parenthesis							
AADT	Lane Width (ft)						Total
	0 - 8	9	10	11	12	13 - 24	
Unknown	192 (535.8)	0 (0.00)	1 (2.25)	0 (0.00)	14 (20.54)	0 (0.00)	207 (558.63)
500 or less	1057 (2913.2)	92 (214.21)	215 (699.51)	197 (387.52)	1213 (2573.7)	28 (52.69)	2802 (6840.8)
501 - 1000	87 (89)	1 (0.19)	28 (46.99)	35 (50.62)	847 (1376.3)	11 (12.96)	1009 (1578)
1001 - 2000	73 (36.78)	0 (0.00)	4 (2.21)	69 (50.44)	988 (1049.4)	10 (15.16)	1144 (1153.9)
2001 - 10000	76 (65.45)	1 (1.86)	6 (6.36)	50 (32.24)	2801 (1987.5)	25 (20.45)	2959 (2113.8)
> = 10001	2 (1.35)	0 (0.00)	11 (3.01)	67 (28.21)	1622 (657.50)	3 (0.83)	1705 (690.90)
Total	1487 (3641.6)	94 (216.26)	265 (760.33)	418 (549.03)	7485 (7664.87)	77 (104.09)	9826 (12936.2)

Table 4.3 Descriptive Statistics for Utah Roadlog File (Area Type by Speed Limit)

Number of Sections, with Total Mileage in Parenthesis			
Speed Limit (mph)	Area Type		Total
	Rural	Urban	
0	2509 (6116.8)	2189 (1155.1)	4698 (7271.9)
15 - 25	65 (104.71)	168 (45.48)	233 (150.19)
30	252 (261.8)	315 (66.9)	567 (328.7)
35	256 (256.94)	226 (71.22)	482 (328.16)
40	448 (426.09)	344 (102.81)	792 (528.9)
45	100 (139.65)	169 (53.29)	269 (192.94)
50	181 (251.91)	171 (57.75)	352 (309.66)
55	1489 (2830.8)	422 (213.04)	1911 (3043.8)
65	470 (761.02)	52 (20.91)	522 (781.93)
Total	5770 (11149.7)	4056 (1786.47)	9826 (12936.2)

Table 4.4 Descriptive Statistics for Utah Roadlog File (Terrain Condition)

Terrain	Number of Sections	Mileage
Flat	801	1157.08
Rolling	594	1043.78
Mountainous	239	466.41
Uncoded	8189	10266.48
Total	9823	12933.75

83% of the sections (14555) are uncoded while 1157.08 miles are of flat terrain (801 sections), 1043.78 miles are of rolling (594 sections) and 466.41 miles are of mountainous (239 sections) terrain.

4.1.2 Accident Characteristics

There are 185,341 total reported accidents in the data base out of which 124,161 are property damage accidents, 44,178 are injury accidents and 17,002 are fatal accidents (this category is a combination of fatal and incapacitating accidents). Other accident characteristics are shown in table 4.5 (see page 51). The variable *collision type* (last row of the table) indicates that single vehicle accidents (27.20%) are most reported followed by rearend (23.80%), angle (14.60%), and intersection related accidents (10.70%). The average accident rate as indicated by the table is 2.87 accidents per mile, per year. Another variable *accident type* (second to last row of the table) shows that 70.50% were motor vehicle accidents followed by 8.80% run-off-road accidents while accidents involving animals, fixed and other objects, pedestrians, bicycle, train and overturn were less then 15%.

In order to get the *vehicle type* into the accident file to identify the type of vehicle involved in an accident, a merge was done between the accident file

Table 4.5 Summary of Accident Statistics for Utah

Variable Name	No. of Accs	Acc/100 MVM	Acc/Mi/Yr	% of Total Accs
Total Accs.	185341.00	24.75	2.87	100.00
PDO Accs.	124161.00	16.58	1.92	66.90
Injury Accs.	44178.00	5.90	0.68	23.80
Fatal Accs.	17002.00	2.32	0.27	9.30
People Injured	60805.00	8.12	0.94	N/A
People Killed	1254.00	0.17	0.02	N/A
Daylight Accs.	125185.00	16.72	1.94	67.40
Dark no lights.	24683.00	3.30	0.38	13.30
Dark with lights.	24022.00	3.21	0.37	12.90
Dawn or Dusk Accs.	9866.00	1.32	0.15	5.30
Dry Accs.	138836.00	18.54	2.15	74.80
Wet Accs.	24709.00	3.30	0.38	13.30
Snow Accs.	11161.00	1.49	0.17	6.00
Icy Accs.	9866.00	1.32	0.15	5.30
Muddy/Oily Accs.	322.00	0.04	0.00	0.20
Motor Vehicle	130697.00	17.46	2.02	70.50
ROR	16386.00	2.19	0.25	8.80
Animals	11340.00	1.51	0.18	6.10
Fixed & Other Obj	8590.00	1.15	0.13	4.60
Ped/Bic	6094.00	0.81	0.09	3.30
ROR - Median	5745.00	0.77	0.09	3.10
Overturn	2160.00	0.29	0.03	1.20
Train	151.00	0.02	0.00	0.10
Single Vehicle	50476.00	6.74	0.78	27.20
Rearend	44186.00	5.90	0.68	23.80
Approach Angle	27153.00	3.63	0.42	14.60
Intersection	19845.00	2.65	0.31	10.70
Turning	12255.00	1.64	0.19	6.60
Parked Vehicle	7378.00	0.99	0.11	4.00
Passing	6876.00	0.92	0.11	3.70
Sideswipe-Pass	5884.00	0.79	0.09	3.20
Backing	3281.00	0.44	0.05	1.80
Rearend-Pass	2850.00	0.38	0.04	1.50
Sideswipe-Opp	2440.00	0.33	0.04	1.30
Head-On	2372.00	0.32	0.04	1.30

Table 4.6 Summary of Accident Statistics for Utah (Vehicle based)

Variable Name	Trucks	Non Trucks	% of Trucks	% of Non Trucks
Total Accs.	11060.00	327794.00	3.26	96.74
PDO Accs.	7762.00	218055.00	2.29	64.35
Injury Accs.	2074.00	81685.00	0.61	24.11
Fatal Accs.	1224.00	28054.00	0.36	8.28
People Injured	3221.00	108836.00	N/A	N/A
People Killed	170.00	1742.00	N/A	N/A
Daylight Accs.	8389.00	231998.00	2.48	68.47
Dark no Lights.	1605.00	31555.00	0.47	9.31
Dark with lights	541.00	44244.00	0.16	13.06
Dawn or Dusk Accs.	454.00	16714.00	0.13	4.93
Dry Accs.	8336.00	244009.00	2.46	72.01
Wet Accs.	1189.00	46829.00	0.35	13.82
Snow Accs.	848.00	18816.00	0.25	5.55
Icy Accs.	626.00	16372.00	0.18	4.83
Muddy/Oily Accs.	23.00	476.00	0.01	0.14
Motor Vehicle	7519.00	269760.00	2.22	79.76
ROR	942.00	16952.00	0.28	5.00
Fixed & Other Obj	732.00	9223.00	0.22	2.73
ROR-Median	421.00	5815.00	0.12	1.72
Animals	386.00	11143.00	0.11	3.29
Overturn	362.00	1954.00	0.11	0.58
Ped/Bic	89.00	6411.00	0.03	1.90
Train	16.00	199.00	0.00	0.06
Single Vehicle	3063.00	49088.00	0.90	14.49
Rearend	2248.00	96658.00	0.66	28.53
Turning	1729.00	23253.00	0.51	6.86
Approach Angle	860.00	54994.00	0.25	16.23
Sideswipe-Pass	792.00	11532.00	0.23	3.40
Parked Vehicle	560.00	15055.00	0.17	4.44
Intersection	478.00	40560.00	0.14	11.97
Backing	384.00	61494.00	0.11	1.81
Passing	282.00	13740.00	0.08	4.06
Sideswipe-Opp	245.00	4908.00	0.07	1.45
Head-On	177.00	4787.00	0.05	1.41
Rearend-Pass	155.00	5705.00	0.05	1.68

and the vehicle file by vehicle type. The SAS program for this is given in Appendix A as program 1 (see page 95). Table 4.6 (see page 53) is the result of this merge and it shows accident characteristics for trucks and non-trucks. It shows that out of 338,854 vehicles involved in an accident in Utah, 3.26% were trucks and 96.74% were non-trucks. The variable *collision type* shows that after single vehicle accidents for trucks (0.90%), rearend accidents (0.66%) and turning accidents (0.51) are the highest, while non-trucks show the lowest percentage of turning accidents. Also the variable *accident type* shows that after motor vehicle accidents (2.22%) for trucks, ran-off-road accidents (0.28%) and fixed and other objects accidents (0.22%) are significantly higher than other accident types. While in the same variable, the non-trucks also show the same characteristics. Various differences in the characteristics of truck and non-truck accidents are discussed in detail in section 4.3.

4.2 Determination of Key Variables

The development of a truck accident predictive model requires the identification of independent variables which most affect the truck accidents. To achieve this objective, extensive literature review was done, and the results were included in the previous chapter. That exercise gave an invaluable insight into the causes of vehicle accidents in general and truck accidents in particular. After that identification the important task is to locate those variables in the data base from which the model has to be developed.

The first task was to merge the accident and vehicle files and obtain the cross tabulations of relevant accident variables with each vehicle type. The vehicle file of Utah has a variable which identifies the type of vehicle involved in an accident. This variable has various categories as described earlier and those categories were grouped into three categories viz. non-trucks, straights and articulated(trucks). The straights are the trucks which have a single body without any hitch between the trailer and the front driving mechanism (bobtail) while articulated trucks are all those trucks which have a bobtail in front and is attached by single or double trailers which are also commonly called tractor-trailer combinations. The category non-truck, has all the vehicles which do not fall into the above two definitions. The merging was done by using the variable YRCASE which identifies a particular accident in the Accident file as well as the Vehicle file. The SAS program for this is listed in Appendix A (see page 94).

The main aim of this exercise was to determine the number of involvements of straight trucks, articulated trucks and non-trucks with each type of variable. The variables considered for this purpose are accident type, accident severity, collision type, light condition, road type, weather condition, object struck by the vehicle, road characteristics on which the accident occurred, design of the road section, surface of the road, and people involved, killed and injured in an accident. The cross tables are obtained by using the procedure FREQ of SAS by crossing each of the above variables with each vehicle type. These cross tables were used to develop bar charts for each variable and these bar charts are shown in Appendix B (see page 130). A similar analysis is done for the States of Minnesota and Michigan and an attempt has been made to compare the accident experience across the three States. The purpose of this comparison is that if some similarity could be found from the accident characteristics in all the three States, then the results of the regression model developed from Utah could be applied to other HSIS States, showing the same accident characteristics without developing the model for each State separately.

4.3 HSIS Accident Experience

The comparison of bar charts for variable *accident type* and *collision type* for Utah, Minnesota and Michigan shows that the maximum accident involvement occurs between motor vehicles (see pages 131, 132, 133). Other accident type categories such as accidents with animals, pedestrians and bicycles show a similar trend in all the three States with non trucks having a higher percentage than the trucks. Collisions with fixed and other objects show a higher percentage for articulated trucks followed by straight trucks and non-trucks in Utah and Minnesota while Michigan shows a higher percentage for straight trucks. One striking contrast is overturning accidents, Utah and Michigan show higher percentages for articulated trucks (4.4% and 5.90% respectively) while Minnesota has a higher percentage for straight trucks(6.7%). Run-off-road accidents also show similar trends in Utah and Minnesota with articulated trucks having a higher percentage than straights and non-trucks in that order.

The *collision type* variable (see pages 134, 137) shows some striking similarities such as passing sideswipe, opposite sideswipe and approach angle accidents which have the same trend. Articulated trucks have a higher percentage of the above three catergories followed by straights and non-trucks. However, head-on and rearend accidents do not show similarities. In Utah and Minnesota, rearend accidents have a higher percentage for non-trucks followed by straights and articulated while Michigan shows exactly opposite trend. Also one important observation is that the rearend accidents are highest in all the three States after motor vehicle accidents.

The reason for a higher involvement of articulated trucks in sideswipe accidents is due to its braking characteristics. Articulated trucks experience locking of their rear wheels while braking and also trailer swing which would contribute to sideswipe accidents.

The *roadway alignment* or *roadway characteristics* variable (see pages 138, 139, 140) indicate on which portion of the roadway the accident occurred i.e. whether the road was straight or curved or leveled or at grade or the combination of all. All three States show the same trend with straight and level roads having the highest percentage of total accidents followed by straight grade, curve grade and curve level in Utah and Minnesota. The State of Michigan does not have as fine categories as the State of Utah and Minnesota. The percentage of truck accidents is higher on a curve grade, curve level and straight grade roads indicating that trucks are much more likely to be involved in an accident on curved and graded highways. The reason for the larger involvement of articulated trucks could be attributed to their braking and accelerating capabilities. On the upgrades, trucks cannot accelerate as quickly, and as a result, a speed differential occurs between trucks and motor cars causing rearend accidents for cars and this is also evident in the graph of *collision type* by *vehicle type* (see pages 134, 137). On the downgrade trucks cannot brake as efficiently as cars and thus are involved in run-of-road accidents, which is confirmed by the previous bar charts.

Road surface and *weather condition* play an important part in the causation of accidents. The charts show (see pages 141, 142, 143, 144, 145, 146) that the most accidents occurr on dry roads, and this may be due to the fact that weather conditions are dry for about 70% of the time. But the important thing to notice is that in dry conditions, the articulated trucks have higher accident involvement than in any other category for all of the three States. Non-trucks show a higher percentage of accidents in wet conditions followed by straights and articulated trucks, and this fact is corroborated by a study done by Bowman and Hummer[6]. The study done on BMCS[7] data indicated that large trucks have lower involvement in accidents during rainy conditions and this is confirmed by HSIS data as well.

The *light condition* variable (see pages 147, 148) shows the approximate time of the day when the accident occurred. Once again, the trend for all of the categories across Utah and Minnesota is same. Daylight conditions show maximum accident occurrences with straight trucks experiencing the maximum number of accidents followed by articulated trucks and non-trucks. Non-trucks show a high involvement during darkness without street lights, which indicates that trucks travel on freeways and interstates at a higher frequency during the night, while non-trucks travel on lighted roads. Also daylight condition show lower percentage for articulated trucks indicating that they travel more at night compared to non-trucks and straights. This is true because trucks are generally used for long haul journies. The BMCS[7] data analysis also showed that 40% of trucks are involved in *darkness with no light* conditions and another report by Joseph and Hummer[6] showed that large trucks are more involved in an accident at night as compared to any other vehicles.

The *accident severity* variable (see pages 149, 150, 151) indicates the type of severity of an accident such as fatal*, injury or property damage only. From the charts, it is seen that all three States have similar severity characteristics. Articulated trucks show a higher involvement for fatal accidents which is also confirmed by a another variable which shows the number of people killed in an accident and it indicates that articulated trucks have a higher fatality as compared to non-trucks. The injury accidents show higher involvements for non-trucks which is also corroborated by another variable showing the number of people injured in an accident (see pages 152, 153, 154) which has a higher percentage of people injured for non-trucks as compared to other vehicles. The observation of articulated trucks having

---

*Fatal accidents are obtained by grouping *fatal* and *incapacitating* categories.

higher accident involvement for fatal accidents is confirmed by a study[6] of truck accident rates using Michigan accident data. It concluded that tractor-trailers (articulated trucks) had a higher rate of fatal accidents than other vehicles.

From the HSIS accident characteristics, it could be concluded that the articulated and straight trucks show very similar characteristics across the three HSIS States. Most of the truck accidents occur on dry roads and in daylight conditions. Also the horizontal curvature and vertical grade play an important part in truck accidents and these variables demand a careful investigation. The rearend accidents are the maximum after the motor vehicle accidents, and articulated trucks are involved in more accidents as compared to other vehicles. The similarity of accidents across these three States is quite striking apart from some characteristics, which are unique to each State.

There are three kinds of lies: lies, damned lies, and statistics.

Disraeli

5.0 Development of Models

5.1 Description of Variables

After investigating the Utah data files, developing a number of cross tables for different variables and based on the conclusions from the literature review, the following is a list of variables which will be considered for model development:

Truck involvement rate per mile, per year (TINVOL/M/Y)

Average Daily Traffic per lane (AADT)

Shoulder width in feet (SHLDWID)

Truck adt per lane(TRUCKADT)

Percentage of the sections having a certain horizontal curvature (HCUV)

Percentage of the sections having a certain vertical gradient (GGRD)

The variable TINVOL/M/Y is the dependant variable in the model and is calculated by dividing the total number of trucks involved in an accident by the number of years for which the data is collected and the length of the section on which the accident occurred. The variable TRUCKADT indicates the average truck traffic per lane on the section of the roadway. This variable is not present in the Utah files but it was created by multiplying the truck percentage by the average daily traffic on the section. The truck percentage variable indicates the percentage of trucks in the traffic stream. When the roadway data files were analyzed, it was found that all of the segments did not have values in the truck percentage variable; therefore, a SAS program was written to fill in the blanks based on the variables route number (RTNUM) and the values of the truck percentage known. The listing of this program can be found in Appendix A, program 8 (see page 127). This program matches the route numbers of each segment, and if the value of truck percentage is zero for that matched segment, it replaces that zero by the value of truck percentage of the previous segment. The inherent assumption in this procedure is that the volume of the truck progresses in a forward direction without any loss due to the presence of interchanges, intersections or any other route. This procedure was adopted in consultation with Mr. Forrest Council and Mr. Charlie Zeeger from the Highway Safety Research Center at the University of North Carolina, who are authorities in conducting safety research, and with Mr. Jeff Paniati at FHWA and Dr. A. G. Hobieka at Virginia Tech. According to them, the exposure data for trucks is not available directly from the State, therefore this is the best available method of determining truck volume on the road segments using HSIS data.

The variables indicating the horizontal curvature and vertical gradient are available in two separate files and they had to be attached to the roadway segments for the development of the models. The variables that are available in the files indicate the degree of curvature and percentage of gradient for a segment of the roadway, however, these segments do not match the segments available in the road file. Due to this fact, the curve and grade data cannot be attached directly to the road segments in road file.

The sections of the grade and curve file do not match with that of the roadway segments because of the way in which the data is coded. Each of these files have segments with beginning and ending mile posts which define a roadway section. And these sections are created whenever there is a change in any geometric feature such as shoulder width, lane width, a change in the number of lanes in the road file, a change in the degree of curvature in a curve file, or a change in the percentage of gradient in the grade file. Since these files are separate, the changes do not coincide with each other and it so happens that the length of each section having a particular degree of curvature may be larger than the length of the corresponding section in the road file. This may be true for the road and the grade file. In other words, a particular road section in the road file may have different values of curvature and gradient for different lengths in the same section. In order to take this fact into account separate variables for gradient and curvature are created which indicate the percentage of the road segment having a particular percentage of gradient and degree of curvature.

The variables set up for this particular study are divided into different categories for Interstate roads and 2 lane rural roads based on the design values given in the AASHTO manual[1]. Figure 5.1 (see page 67) shows the variables created for Interstates and figure 5.2 (see page 67) shows variables created for 2 lane rural roads. For Interstates the curvature is divided into HCUV1, HCUV2 and HCUV3 indicating the percentage of each section having degree of curvatures between 1° and 2.5°, 2.5° and 4° and greater than or equal to 4° respectively and the vertical gradient is divided into GGRD1, GGRD2 and GGRD3 indicating the percentage of section as having percentage of gradients between 1% and 3%, 3% and 5% and greater than or equal to 5% respectively.

Consider a one mile section of the Interstate roadway as shown in figure 5.1. Assuming that a one mile section had four 0.25 miles subsection each with a degree of curvature for the first 0.25 mile as 3°, for next 0.25 mile as 2° while rest of the section having 0° of curvature. Also the percentage of grade as 2.5% for first 0.25 miles, 0% for next 0.25 miles followed by 6% for another 0.25 miles and 0% for the remaining 0.25 miles, then the values of the new curvature and gradient variables will be as follows:***

HCUV1 = 25%, HCUV2 = 25%, HCUV3 = 0%, GGRD1 = 25%, GGRD2 = 0% and GGRD3 = 25%. Similarly the variables for 2 lane rural roads are as explained in the figure 5.2. The distribution of these variables by roadway mileage is shown in tables 5.1 and 5.2 (see page 67).

To calculate the values of the curve and grade variables, a series of programs had to be developed and the listing of these programs is given in appendix A, programs 2 to 6 (see page 97, 98, 99, 106, 112). These programs are developed in consultation with SAS programmers from the University of North Carolina Highway Safety Research Center.

Before running the series of programs which calculated the curvature and gradient variables, it was decided that the sections which were less than 1 mile should be deleted. The reason for this being that the State of Utah has a number of small sections varying from 0.01 mile up to 1 mile and the truck accident rate, if calculated for these small sections, would be very high if a single accident occurred, say on 0.01 mile section. Also it would not be appropriate to include this value in regression models as it would represent a false picture of accident occurrence. Also, after running the program which

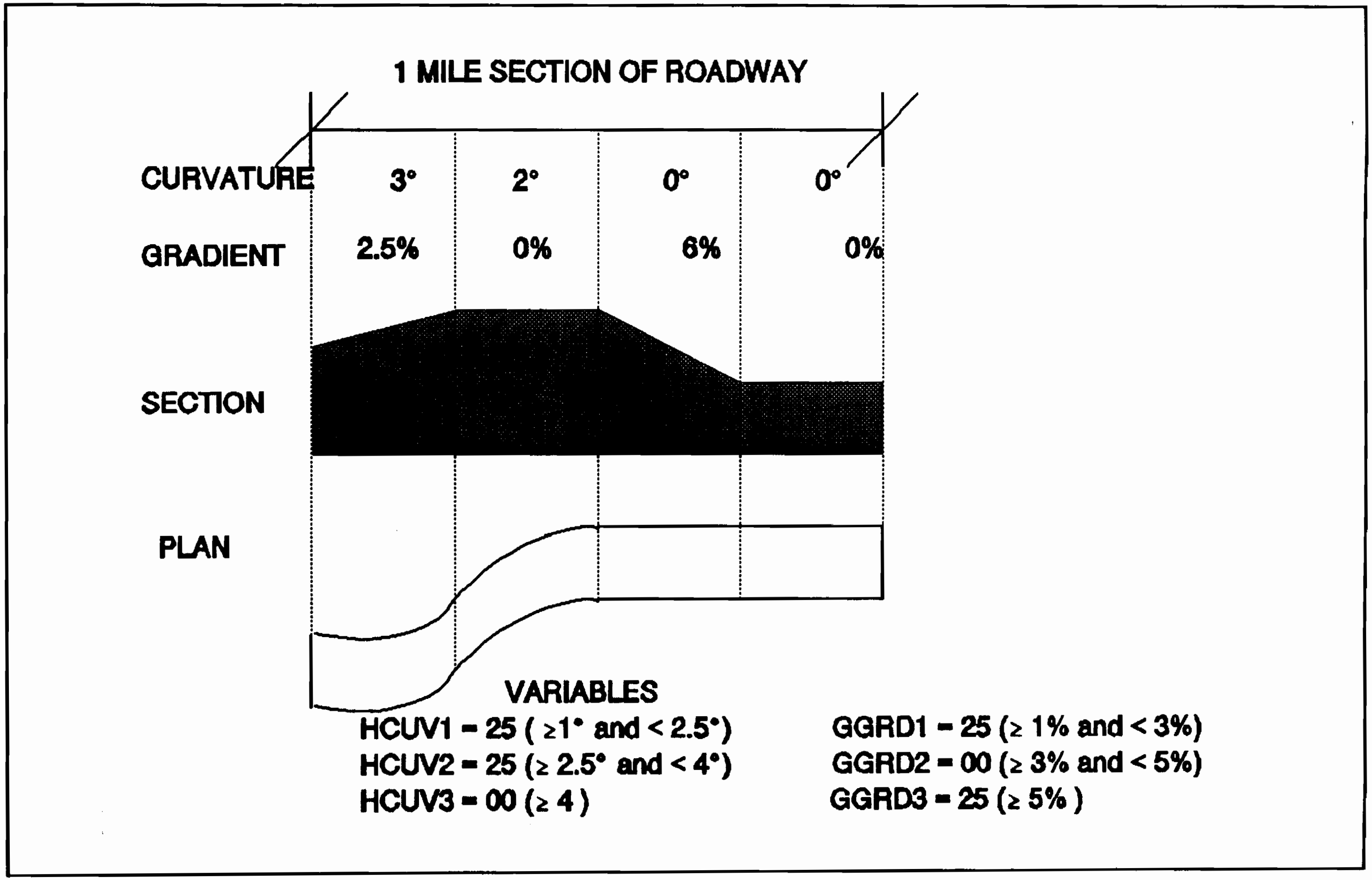
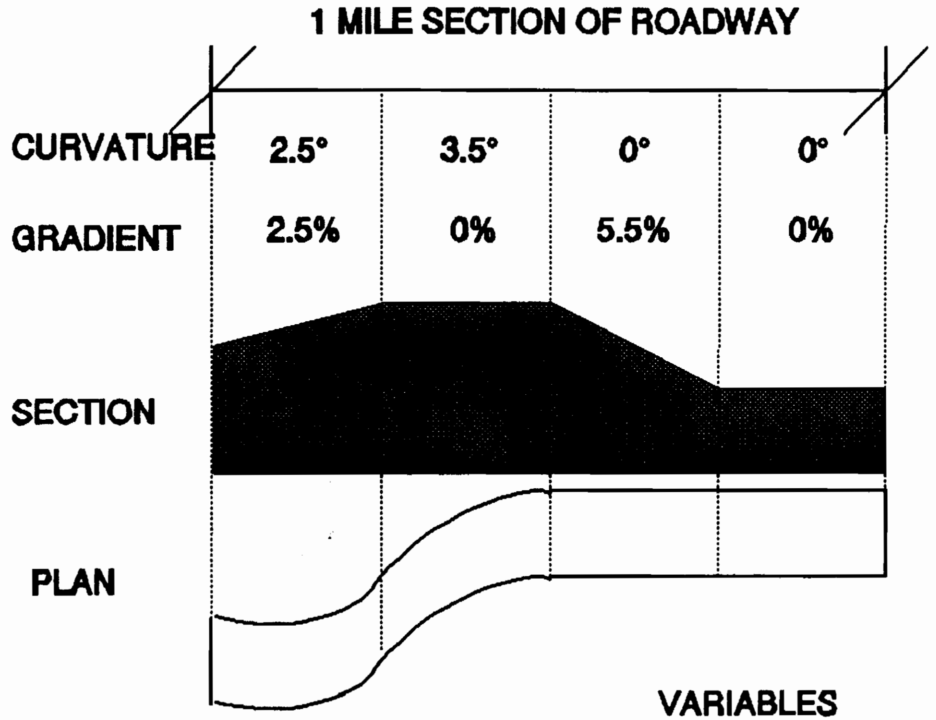


Figure 5.1 Curve and Grade variables used for Interstates



VARIABLES

HCUV1 = 25% (≥1° and < 3°)
HCUV2 = 25% (≥ 3° and < 6°)
HCUV3 = 00% (≥ 6° and < 10°)
HCUV4 = 00% (≥ 10°)

GGRD1 = 25% (≥ 1% and < 3%)
GGRD2 = 00% (≥ 3% and < 5%)
GGRD3 = 25% (≥ 5% and < 7%)
GGRD4 = 00% (≥ 7%)

Figure 5.2 Curve and Grade variables for 2 Lane Rural Roads

Table 5.1 Distribution of curvature and gradient by mileage for Interstates

Degree of curvature	Roadway Mileage
> = 1° and < 2.5°	228
> = 2.5° and < 4.0°	67
> = 4.0°	125
Percentage of Gradient	**Roadway Mileage**
> = 1.0% and < 3.0%	617
> = 3.0% and < 5.0%	311
> = 5.0%	111

Table 5.2 Distribution of curvature and gradient by mileage for 2 lane rural road

Degree of Curvature	Roadway Mileage
> = 1.0° and < 3.0°	154
> = 3.0° and < 6.0°	2167
> = 6.0° and < 10.0°	1414
> = 10.0°	1011
Percentage of Gradient	**Roadway Mileage**
> = 1.0% and < 3.0%	3254
> = 3.0% and < 5.0%	2268
> = 5.0% and < 7.0%	890
> = 7.0%	653

determines the missing values of the truck percentage on a particular, route it was found that some routes did not have any truck percentages indicating that there is no truck travel on those routes. Some of the sections did not have any curve and grade variables and those sections were eliminated to obtain the final file. These facts caused the shrinking of the data base and thus the final data file consists of 2073 sections covering 7290 miles of roadway. Some of the cross tables developed from this file are as shown in tables 5.3, 5.4, and 5.5 (see pages 72, 73, 74). Table 5.3 shows that most of the data lies in the Rural area and very little roadway mileage classified as Local roads is available. Also Urban Freeways show very little mileage (41 miles out of 7291 miles) because the majority of the roads in Utah are classified as either Interstates, or arterial collectors or local roads. Hence, these categories were not considered for model development. Table 5.4 shows that most (99%) of the approximately 6400 miles of roadway classified as primary arterials and arterial collectors are two lane roads. While table 5.5 shows that all the Interstates have 10 feet shoulders while the primary arterials and collectors have a fairly even distribution of shoulder width varying from 0 to 11 feet. Based on these observations, two models were selected for development: Interstates without shoulder width and lane width variables and 2 lane rural arterials and collectors. The data used for the truck Interstate accident model includes 264 road sections ((718.71 miles) and 1787 total trucks involved in accidents. The 2

Table 5.3 Table of Functional Class by Rural Urban Designation for Final Utah file

Number of Sections, with mileage in parenthesis			
Functional Class	**Rural Urban Designation**		**Total**
	Rural	**Urban**	
Interstate	224 (655.47)	40 (63.24)	264 (718.71)
Primary Arterial	197 (501.53)	29 (39.36)	226 (540.89)
Arterial Collector	1489 (5882.1)	73 (108.01)	1562 (5990.1)
Local	8 (22.18)	11 (15.41)	19 (37.59)
Urban Freeways	0 (0.00)	2 (3.42)	2 (3.42)
Total	1918 (7061.28)	155 (229.44)	(2073 7290.72)

Table 5.4 Table of Functional Class by Number of Lanes for Final Utah File

Number of Sections, with mileage in parenthesis					
Functional Class	Number of lanes				Total
	2	3	4	6	
Interstate	12 (29.01)	6 (9.77)	210 (621.4)	36 (58.53)	264 (718.71)
Primary Arterial	166 (442.1)	34 (59.2)	25 (38.53)	1 (1.06)	226 (540.89)
Arterial Collector	1519 (5924.5)	17 (29.74)	25 (34.74)	1 (1.1)	1562 (5990.1)
Local	19 (37.59)	0 (0.00)	0 (0.00)	0 (0.00)	19 (37.59)
Urban Freeways	0 (0.00)	0 (0.00)	2 (3.42)	0 (0.00)	2 (3.42)
Total	1716 (6433.23)	57 (98.71)	262 (698.01)	38 (60.69)	2073 (7290.72)

Table 5.5 Table of Functional Class by Shoulder width for Final Utah file

Number of Sections, with mileage in parenthesis								
Functional Class	Shoulder width in feet							Total
	0-1	2-3	4-5	6-7	8-9	10	> = 11	
Interstate	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	264 (718.71)	0 (0.00)	264 (718.71)
Primary Arterials	27 (50.28)	24 (46.12)	66 (173.88)	56 (142.32)	32 (66.66)	14 (51.49)	7 (10.14)	226 (540.89)
Arterial Collectors	898 (3553.8)	348 (1389.8)	143 (492.34)	122 (425.85)	40 (102.47)	5 (14.69)	6 (11.15)	1562 (5990.1)
Local	13 (20.54)	3 (8.47)	2 (7.21)	0 (0.00)	1 (1.37)	0 (0.00)	0 (0.00)	19 (37.59)
Urban Freeways	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	0 (0.00)	1 (1.9)	1 (1.52)	2 (3.42)
Total	938 (3624.63)	375 (1444.39)	211 (673.43)	178 (568.17)	73 (170.5)	284 (786.79)	14 (22.81)	2073 (7290.7)

lane rural road model includes 1614 sections (6259.17 miles) with 1313 total trucks involved in accidents.

5.2 Selection of Models

The occurrence of highway accidents is best described by the Poisson Process and hence the Poisson probability equation is used for truck accident models. The general form of Poissons model is as follows:

$$P(k) = \frac{e^{-y}\, y^{k}}{k!} \tag{5.1}$$

where

p(k) is probability of k accidents occurring per year, and the value of y could be predicted using various models which are as below:

$$y = \beta_0 \ (A_1)^{\beta_1}(\beta_2)^{A_2}(\beta_3)^{A_3}(\beta_4)^{A_4}\ldots\epsilon \tag{5.2}$$

$$y = \beta_0 \ (\beta_1)^{A_1}(\beta_2)^{A_2}(\beta_3)^{A_3}(\beta_4)^{A_4}\ldots\epsilon \tag{5.3}$$

$$y = \beta_0 + A_1\beta_1 + A_2\beta_2 + A_3\beta_3 + A_4\beta_4\ldots + \epsilon \tag{5.4}$$

Where

β_0 is the intercept, β_1, β_2, β_3, β_4 are regression coefficients and A_1, A_2, A_3, A_4 are geometric variables and y is accidents per mile per year.

A two step process is used to determine the model which would best predict the value of y in the above equations. In the first step, *stepwise* procedure was employed to determine which variables are significant at $\alpha = 0.05$. In the second step, the values of beta obtained from the *stepwise* procedure were used as initial values for procedure NLIN, which is a non linear equation fitting procedure. This procedure produces the least-square estimates of the parameters through the Marquardt iterative method, where the residuals are regressed onto the partial derivatives of the model with respect to the parameters, until the iteration converges. Using this procedure, it was found that there was no significant improvement in the parameter estimates and the linear models were much better than non linear models.

A number of models were tried for Interstates and 2 lane rural roads for all truck accidents, single vehicle truck accidents and over turning truck accidents. From this combination of models, only three models were found to be significant based on their R^2 values and the F test performed on them. Those models are:

All Truck Accident Model for Interstates

$$TINVOL/M/Y = -0.1777 + 0.0002AADT + 0.0006TRUCKADT + 0.0053HCUV2 + 0.0098HCUV3 + 0.0022GGRD2 + 0.0048GGRD3 \tag{5.5}$$

with $R^2 = 0.713$.

All Truck Accident Model for 2 Lane Rural Roads

$$TINVOL/M/Y = 0.0027 + 0.00009AADT + 0.0004TRUCKADT - 0.0025SHLDWID + 0.0011HCUV3 + 0.0007HCUV4 \tag{5.6}$$

with $R^2 = 0.415$.

Overturning Truck Accident Model for Interstates

$$TINVOL/M/Y = -0.0131 + 0.00008TRUCKADT + 0.0102HCUV4 + 0.0038HCUV3 + 0.0031GGRD4 \tag{5.7}$$

with $R^2 = 0.317$.

All of the variables in the models are significant at $\alpha = 0.05$ and the variables AADT, TRUCKADT, CURVATURE and GRADIENT show positive sign indicating that as the values of these variables increases the truck involvement rate would increase. Also, in equation 5.5 the values of the coefficients for HCUV3 and GGRD3 are larger than HCUV2 and GGRD2 respectively showing that a road section with a degree of curvature greater than or equal to 4 will cause more accidents compared to the degree of curvature between 2.5 and 4. A section of roadway with a percentage of gradient greater than or equal to 5 will have higher truck accidents as compared to the percentage of gradient between 3 and 5. However, this is not true for 2 lane rural model as shown in equation 5.6 and also it does not have any variables for gradient which may be because of the inadequacy of the data. The shoulder width in this model shows a negative sign indicating that as the width of the shoulder is increased, the truck involvement rate would decrease.

The model in equation 5.7 is for over turning truck accidents on Interstates. The data for 2 lane rural roads was inadequate for overturning trucks and hence the model developed for that category had a very low R^2 and is not included in the final models. The variables for curvature and gradient used for over turning trucks are different from the other two models. Curvatures and gradients are divided into four categories each from HCUV1 to HCUV4 and GGRD1 to GGRD4 and each category is defined as a percentage of the sections having a degree of curve between 1° and 2.5°, 2.5° and 5°, 5° and 7°, greater than or equal to 7°; and percentage of gradient between 1% and 2.5%, 2.5% and 5%, 5% and 7%, and greater than or equal to 7%. This was done because the models developed with previous categories for Interstates did not show any significant variables in the model for overturning truck accidents.

5.3 Model Validation

To validate the Interstate models, the data was divided in half with even and odd numbered observations. Even numbered observations were used to develop the model for all truck accident involvement rates for Interstates, and the following model resulted:

$$\begin{aligned} TINVOL/M/Y &= -0.3014 + 0.0002AADT + 0.0009TRUCKADT \\ &\quad + 0.0067HCUV2 + 0.0139HCUV3 + 0.0042GGRD2 \\ &\quad + 0.004GGRD3 \end{aligned} \tag{5.8}$$

with $R^2 = 0.756$. The coefficients in this equation are close to the coefficients obtained in equation 5.5. This model was then used to determine the predicted value of the truck involvement rate from the other half of the data. The predicted values and observed values matched very closely with the correlation coefficient of 0.844 and the result of t test on the values of predicted and observed values of truck involvement rate at $\alpha = 0.05$ also indicated that the validation was successful. The same exercise was carried out for over turning truck accident model for Interstates, and it was found that the developed model was valid.

Also the same procedure was employed to validate the model for 2 lane rural roads. The resulting model from even numbered data is as follows:

$$TINVOL/M/Y = 0.0025 + 0.00009AADT + 0.0004TRUCKADT - 0.0023SHLDWID + 0.0008HCUV3 + 0.001HCUV4 \tag{5.9}$$

with $R^2 = 0.431$. In this case also the coefficients are very close to that obtained in equation 5.6. The coefficient of correlation for predicted and observed the truck involvement rate was 0.631 and in this case, t test at $\alpha = 0.05$ was successful in validating the model.

The figure 2.2 children per adult female was felt to be in some respects absurd, and a Royal Commission suggested that the middle classes be paid money to increase the average to a rounder and more convenient number.

Punch

6.0 Application of Models

The models developed for truck accident involvement rates could be used to compare different sections of the existing roadway and the type of improvements required to reduce the truck accidents could be determined. Also these models will be very useful in determining the impact of changing a particular geometric variable on the truck accident occurrence. Another use of these models will be in designing the roadway sections and justifying the design based on the values obtained from these models. The use of these models is demonstrated in the following sections by assuming a one mile section of the roadway as having certain geometric variables and a certain truck accident involvement rate is then calculated for each section by varying the values of the key variables.

6.1 All Truck Accident Interstate Model

The models developed for truck accident involvement could be used for comparing various sections of the roadway and the expected percentage decrease in accidents could be determined by using these models. The following example illustrates the use of the model.

Consider a one mile section of roadway having a total average daily traffic of 4000 vehicles, 5% trucks, 4 lanes, 10 feet shoulders and 12 feet wide lanes. Let 40% of the section have 3° curvature and 60% have 6° curvature and let 50% have a gradient of 4% and another 50% have a gradient of 6%. Using equation 5.6 the truck involvement rate can be calculated for this section. Table 6.1 (see page 84) shows the truck involvement rates for the above base section and for the section obtained after certain modifications.

From the table in example 1 it is seen that if the length of 6° section is increased from 60% to 80% the truck involvement rate increases, also if no curvature is present on the road then the accident rate is less as can be seen from example 3. Also, if only the curvature is present without any grade, then the accident rate is higher (example 4) as compared to only the presence of the gradient (example 3) indicating that truck accidents are affected more by the presence of curvature than gradient. In example 2, if the length of 6% gradient section is increased from 50% to 70% the truck involvement rate increases. Example 5 demonstrates the potential reduction in truck involvements if all curves and grades were eliminated. Thus, in general, it can be concluded that the combination of curvature and gradient causes more truck accidents compared to flat roadway sections (example 5) and the presence of curvature has a greater affect on the truck accidents as compared to the presence of gradient.

Table 6.1 Predicted Number of All Truck Involvement Rates Using Interstate Model

Examples	Non-Truck AADT/Lane	TRUCK AADT/Lane	HCUV2	HCUV3	GGRD2	GGRD3	TINVOL/M/Y
Base section	950	50	40	60	50	50	1.192
1	950	50	20	80	50	50	1.282
2	950	50	0	0	50	50	0.392
3	950	50	40	60	30	70	1.244
4	950	50	40	60	0	0	0.842
5	950	50	0	0	0	0	0.042

6.2 All truck accident model for 2 lane rural roads

The model for 2 lane rural roads is weaker than the Interstate model by the fact that it does not have any gradient variable in it. The main reason for this could be the inadequacy of the data. Because of these limitations, the model can only be used for determining the reduction in truck accidents by increasing the shoulder width. Consider a one mile long section of a rural 2 lane roadway with 50% of the section having 6° curve and another 50% having 8° curve and shoulder widths being 0 feet, 3 feet and 6 feet. Using equation 5.7 and substituting the values of HCUV3 = 50 and HCUV4 = 50 the truck involvement rates decreases as shoulder width increases. The truck involvement rate increases as AADT increases which can be seen in table 6.2 (see page 87).

6.3 Overturning truck accident model for Interstates

The R^2 for this model is very low but it shows an interesting relationship between variables which causes truck overturning accidents. The model does not have AADT, instead it has only truck AADT which is true because trucks are involved in overturning accidents most of the time. The reason for such a low R^2 is that there are only 95 overturning accidents in the final file. This model could be used in the similar manner as all truck interstate model. Consider a section one mile long, having a total average daily traffic of 4000 vehicles 5% percent trucks, 4 lanes, 10 feet shoulders and 12 feet wide lanes. Let 60% of the section have 6° curve and 40% of section have 4° curve and let the entire section have a gradient of 6%. Table 6.3 (see page 88) shows the overturning truck involvement rates for various modified sections calculated by using equation 5.7. The conclusions from the table are similar to that for all truck accidents on Interstates. The highest involvement rate is for example 1 which has a combination of horizontal curve and vertical gradient while a flat section (example 5) has the lowest rate. Also, according to this model the truck overturning is more affected by sections having a horizontal curvature greater than or equal to 5° and gradients with percentage greater than or equal to 7%.

Table 6.2 Predicted Number of All Truck Accidents Using 2 Lane Rural Model

TOTAL AADT		2000			3000			4000		
NON TRUCK AADT/LANE		970	950	900	1455	1425	1350	1940	1900	1800
TRUCK AADT/LANE		30	50	100	45	75	150	60	100	200
Shoulder Width	0	0.192	0.196	0.211	0.239	0.248	0.272	0.289	0.301	0.332
	3	0.185	0.188	0.204	0.231	0.241	0.264	0.281	0.294	0.325
	6	0.177	0.181	0.196	0.224	0.226	0.257	0.274	0.286	0.317

Table 6.3 Predicted Number of Overturning Truck Accidents Using Overturning Truck Interstate Model

EXAMPLES	TOTAL AADT	NON TRUCK AADT/LANE	HCUV4	HCUV3	GGRD4	TINVOL/M/L
Base section	4000	38	60	40	100	1.505
1	4000	38	80	20	100	1.633
2	4000	38	60	0	100	1.353
3	4000	38	0	0	100	0.741
4	4000	38	60	40	0	1.195
5	4000	38	0	0	0	0.431

7.0 CONCLUSIONS

The study of the accident characteristics showed that trucks, both straight and articulated, are more involved in accidents as compared to non trucks. Also, the truck accident characteristics are similar across the three HSIS States. The severity of truck accident is higher than the non truck accidents. The development of models for truck accidents showed that horizontal curvature and vertical gradients are the main variables affecting the truck accidents. However, there is a difference in the degree of curvature and percentage of gradients which affect these accidents and their values depend upon the roadway type on which the truck travels. For all truck accidents on Interstates, the significant degree of curve is ≥ 2.5°, on 2 lane rural roads it is ≥ 6° and for overturning accidents on Interstates it is ≥ 7°. In case of vertical gradients, for all truck accidents on Interstates, the significant percentage is ≥ 3% and for overturning accidents it is ≥ 7%. The 2 lane rural model shows that truck accidents are not affected by the presence of gradient because the gradient variable does not appear in the model. But the appearance of grade variables in the other two model strongly suggests that if the data was adequate for 2 lane rural roads, the grade variable would have appeared in the model. The shoulder width variable shows a negative coefficient for 2 lane rural roads which indicates that as the shoulder width is increased there will be savings in the truck accident occurrence. This fact could be used to develop the accident reduction factors for different sections of the roadway which would be an invaluable tool for comparing various roadway sections by varying the width of the shoulder.

7.1 Recommendations for Future Research

It is recommended that data collection methods should be improved and an emphasis should be given to the collection of truck data by truck type. Further, models should be developed by truck type to get a better comparison between single unit truck accident occurrence and multi unit truck accident occurrence. Also, the models developed in this study should be used to do economic analysis' of the road sections and cost/benefit analysis' should be done to determine the feasibility of constructing an exclusive truck facility.

Bibliography

1. **A Policy on Geometric Design of Highway and Streets,** American Association of State Highway and Transportation Officials, 1990.

2. Belmont, D. M., **Effects Of Shoulder Width On Accidents On Two Lane Tangents,** HRB, Bull 91, 1954.

3. Belmont, D. M., **Accidents Versus Width Of Paved Shoulders On California Two Lane Tangents: 1951 and 1952,** HRB, Bull 117, 1956, pp 1-16.

4. Blensly, R. C. and Head, J. A., **Statistical Determination Of Effects Of Paved Shoulder Width,** HRB, Bull 151, 1956.

5. Bowman B. L., and Hummer J. E., **Examination of Truck Accidents on Urban Freeways.,** FHWA, Washington D. C., 1989, FHWA-RD-89-201.

6. Bowman, B. L. and Hummer, J. E., **Examination of Truck Accidents on Urban Freeway.** FHWA-RD-89-201, Federal Highway Administration, Mclean, VA.

7. Bureau of Motor Carrier Safety., **Accidents of Motor Carriers of Property** 1980-1981, Federal Highway Administration, Washington D. C. August 27, 1982.

8. Chira-Chavala T., and Cleveland D. E., **Causal Analysis of Accident Involvements for the Nation's Large Trucks and Combination Vehicles.,** Transportation Research Record No. 1047, Washington D. C.

9. Committee for the Truck Safety Data Needs Study, **Data Requirements for Monitoring Truck Safety,** Special Report 228, Transportation Research Board, 1990, Washington D. C.

10. Dart O. K., and Mann L. Jr., **Relationship of Rural Highway Geometry to Accident Rates in Louisiana.,** Transportation Research Record No. 312, TRB, pp. 1-16, 1970.

11. **Evaluation of Criteria For Safety Improvements On The Highway,** Roy Jorgensen and Associates, Inc., Gaithersburg, MD, 1966.

12. Foody, T. J. and Long, M. D., **The Identification Of Relationships Between Safety And Roadside** Obstructions, Ohio Department of Transportation, Columbus, 1974.

13. Garber N. J., and Gadiraju R., **Effects of Truck Strategies on Traffic Flow and Safety on Multilane Highways.**, Trucking Issues, TRR No. 1256, Washington D. C., 1990.

14. Garber N. J., and Joshua S. C., **Estimating Truck Accident Rate and Involvements using Linear and Poisson Regression Model.**, University of Virginia, Published in UK, 1989.

15. Hall, J. W. and Dickinson, L. V., **Truck Speeds And Accidents On Interstate Highways,** Transportation Research Record No. 486, Washington D. C. 1974.

16. **Highway Fact Book,** Highway Users Federation and Automotive Safety Federation. Washington, D. C., 1989

17. **Improving Truck Safety at Interchanges.,** Technical Report, FHWA, Sept 1989, FHWA-IP-89-024.

18. Jovanis P. P., Hsin-Li Chang, and Zabaneh I., **Comparison of Accident Rates for Two Truck Configurations.,** Transportation Research Record No. 1249, Washington D. C., 1989.

19. **Operating Large Trucks on Roads with Restrictive Geometry.,** Summary Report, FHWA, Sept 1989, FHWA-IP-89-025.

20. Perkins, E. T., **Relationship Of Accident Rate To Highway Shoulder Width,** HRB, Bull 151, 1956.

21. **Present Practices of Highway Transportation of Hazardous Materials.,** Technical Report, May 1990, FHWA-RD-89-013.

22. Raff, M. S., **Interstate Highway Accident Study,** HRB, Bull 74, 1953.

23. Rinde, E. A., **Accident Rates vs. Shoulder Width,** Report No. CA-DOT-TR-3147-1-77-01, California DOT, FHWA, September 1977.

24. Rogness, R. O., Fambro, D. B., and Turner D. S., **Before-After Accident**

Analysis for Two Shoulder Upgrading Alternatives, Transportation Research Record No. 855, TRB, 1982.

25. Roy Jorgensen Associates Inc., **Cost and Safety Effectiveness of Highway Design Elements.,** NCHRP, Rept 197, 1978.

26. Stohner, W. R., **Relation Of Highway Accidents To Shoulder Width On Two-Lane Rural Highways In New York State,** Proc., HRB, Vol 35, 1965, pp 500-504.

27. Taragin, A. and Ekhardt, H. G., **Role Of Highway Shoulders In Traffic Operations,** HRB, Bull 151, 1957.

28. Zeeger, C. V. and Mayes, J. G., **Cost Effectiveness Of Lane And Shoulder Widening On Rural, Two-Lane Road In Kentucky.** Kentucky department of Transportation, Frankfort, Res. Rept. 999, 1979.

29. Zeeger, C. V. and Perkins, D. D., **Effect of Shoulder Width and Condition on Safety: A Critique of Current State of the Art,** TRR No. 757, Transportation Research Board, Washington D. C., 1980.

30. Zeeger, C. V., Hummer, J. and Hanscom, F., **Operational Effects of Larger Trucks on Rural Roadways,** Offered for publication to the Transportation Research Board, Washington D. C., 1990.

31. Zeeger C. V., Dean R. C., and Mayes J. G., **Effect of Lane and Shoulder Widths on Accident Reduction on Rural, Two-lane Roads.,** Transportation Research Record 806, Washington D. C., 1981.

32. Zeeger C. V., Hummer J., Reinfurt D., Herf, L., and Hunter, W., **Safety Effects of Cross Section Design For Two-Lane Roads,** FHWA, Washington D. C., 1987.

Appendix A

SAS Programs

Program 1

```
/*This program merges Accident and Vehicle Files. YRCASE is the common
variable used for merging. */

LIBNAME UTAH '.';
DATA ACCIDENT;
SET UTAH.ACCCONCA;/*Reading Utah accident file for 1985-89*/

     /*Following statements deletes bad data from the file */

     TEST1 = SUBSTR(A_ROUTE, 1, 1);
     TEST2 = SUBSTR(A_ROUTE, 3, 1);
     IF TEST1 NE 'A' AND TEST3 GE '5' THEN DELETE;

PROC SORT;
     BY YRCASE;

DATA TRUCK;
SET UTAH.TRUCKS;/*Reading Utah vehicle file for 1985-89*/

     /*This file has only truck information*/

DATA TRUCKS;
SET UTAH.TRUCKS(KEEP = YRCASE);
     BY YRCASE;
     IF FIRST.YRCASE;

DATA TRKACCD;
MERGE ACCIDENT (IN = A)
TRUCKS (IN = T);
     BY YRCASE;
     IF A AND T;

DATA UTAH.TRUCKACC;/*Creating a permanent merged file*/
MERGE TRKACCD (IN = A)/*Merging accident and vehicle file*/
TRUCK;
     BY YRCASE;
     IF A;
PROC FREQ;
```

```
TABLES VEHTYP;
PROC PRINT DATA=UTAH.TRUCKACC(OBS=50);
RUN;
```

Program 2

```
/*This program is used to create a grade file 'GRADMOD' which is to be used
in subsequent program */

LIBNAME DD1 '.';
OPTIONS REPLACE PS=75 LS=75;
DATA DD1.GRADMOD(DROP=SECTLNG YR DIR_GRAD CAPACITY TEST1)
TEST1 TEST3);/*Creating a permanant gradient file 'GRADMOD' */
SET  DD1.UTGRAD89;/*Reading 1989 Utah gradient file */
IF PCT_GRAD  < 10 THEN DELETE;

     /*Following statements deletes bad data from the file */

     TEST1=SUBSTR(RTNUM,1,1);
     IF TEST1 NE 'A' THEN DELETE;

RENAME    SECT_LNG = GRD_SECT/*Renaming the variables as same */
          RTNUM = GRD_RTE/*variables exist in the Roadlog */
          BEGMP = G_BEGMP/*File */
          ENDMP = G_ENDMP;

PROC SORT DATA=DD1.GRADMOD;
     BY GRD_RTE G_BEGMP;
RUN;
```

Program 3

```
/*This program is used to create a curve file 'CURVMOD' which is to be used
in subsequent program */

LIBNAME DD1 '.';
OPTIONS REPLACE PS=80 LS=80;
DATA   DD1.CURVMOD(DROP=SECTLNG   TEST1   TEST3);/*Creating   a
permanant gradient file 'GRADMOD' */
SET  DD1.UTCURV87;/*Reading 1987 Utah Curve File*/
IF DEG_CURV  < 10 THEN DELETE;

     /*Following statements deletes bad data from the file */

     TEST1=SUBSTR(RTNUM,1,1);
     IF TEST1 NE 'A' THEN DELETE;

RENAME    SECT_LNG = CUV_SECT/*Renaming the variables as same */
          RTNUM = CUV_RTE/*variables exist in the Roadlog */
          BEGMP = C_BEGMP/*File*/
          ENDMP = C_ENDMP;
PROC SORT DATA=DD1.CURVMOD;
     BY CUV_RTE C_BEGMP;
RUN;
```

Program 4

```
/*This program is used to create a permanent roadlog file 'UTRDXT87'. This
program deletes the road sections which are less than 1 mile and creates
homogeneous sections of length greater than 1 mile. This file is than used for
subsequent program to attach the gradient and curvature data to these road
segments.*/

      /*STEP 1*/

LIBNAME DD1 'E:\UTAH';
DATA UTRD87EX/*Creating temperory road file*/
(KEEP= TFED_AID TFUNC_CL TRUR_URB TSHLDTYP TPAVE_TY TRTNUM
TBEGMP TENDMP SECT_LNT TAADT TLANEWID TMEDTYP TNO_LANE
TOFFPEAK T_ONEWAY TPEAKTRK TROADYR TSHLDWID TTERRAIN
TMEDWID);
SET DD1.UTROAD87;/*Reading Utah 1987 Roadlog File*/
IF SECT_LNG < 1.00 THEN DELETE;

      /*Following statements deletes bad data from the file */

      TEST1 = SUBSTR(RTNUM,1,1);
      TEST3 = SUBSTR(RTNUM,3,1);
      IF TEST1 NE 'A' AND TEST3 GE '5' THEN DELETE;

      /*Following statements groups FEDAID variable into three catergories*/

IF FEDAID = '1' THEN FED_AID = 1;
      ELSE IF FEDAID = '2' OR FEDAID='3' OR FEDAID = '4' THEN
            FED_AID = 2;
      ELSE IF FEDAID = '8' THEN FED_AID = 3;

      /*Following statements groups FUNC variable into five catergories*/

IF FUNC = '01' OR FUNC = '11' THEN FUNC_CL = 01;
      ELSE IF FUNC = '02' OR FUNC = '14' THEN FUNC_CL = 02;
      ELSE IF (FUNC GE '06' AND FUNC LE '08') OR (FUNC GE '16' AND
            FUNC LE '17') THEN FUNC_CL = 03;
      ELSE IF FUNC = '09' OR FUNC = '19' THEN FUNC_CL = 04;
      ELSE IF FUNC = '12' THEN FUNC_CL = 05;
```

```
        /*Following statements groups RURURB variable into two catergories*/

IF RURURB = '1'  THEN RUR_URB = 1;
        ELSE IF (RURURB  GE '2' AND RURURB LE '3') THEN RUR_URB = 2;

RENAME
        FED_AID= TFED_AID
        FUNC_CL= TFUNC_CL
        RUR_URB= TRUR_URB
        SHLDTYP= TSHLDTYP
        PAVETYPE= TPAVE_TY
        RTNUM   = TRTNUM
        BEGMP   = TBEGMP
        ENDMP   = TENDMP
        SECT_LNG  = SECT_LNT
        AADT    = TAADT
        LANEWID= TLANEWID
        MEDTYP  = TMEDTYP
        NO_LANE= TNO_LANE
        OFFPEAK= TOFFPEAK
        ONEWAY= T_ONEWAY
        PEAKTRK= TPEAKTRK
        ROADYR  = TROADYR
        SHLDWIDR= TSHLDWID
        TERRAIN  =TTERRAIN
        MEDWID  = TMEDWID;

PROC SORT;
        BY TRTNUM TBEGMP;

        /*STEP 2*/

DATA  DD1.UTR1XT87(KEEP=FED_AID  FUNC_CL  RUR_URB  SHLDTYP
PAVETYPE RTNUM BEGMP ENDMP  AADT LANEWID MEDTYP NO_LANE
OFFPEAK ONEWAY PEAKTRK ROADYR SHLDWIDR TERRAIN CHNG_FLG SECT
SECT_CNT MEDWID);

RETAIN FED_AID FUNC_CL RUR_URB SHLDTYP PAVETYPE RTNUM BEGMP
ENDMP AADT LANEWID MEDTYP MEDWID NO_LANE OFFPEAK ONEWAY
PEAKTRK ROADYR SHLDWIDR TERRAIN CHNG_FLG SECT SECT_CNT;
```

```
SET UTRD87EX END=EOF;

        /*Checking segments of same route*/

IF TRTNUM  = RTNUM THEN
        DO;
                IF TBEGMP  =ENDMP THEN/*Segments adjacent*/
                        DO;/*Check for changes*/
                                LINK CHECKREC;
                                LINK WRITREC;
                        END;
                ELSE IF ENDMP GT TBEGMP THEN/*Overlaps*/
                        DO;/*SET FLAG*/
                                CHNG_FLG   = CHNG_FLG + 1;  /*Output rec*/
                                SECT_CNT  = SECT_CNT + 1;
                                LINK WRITREC;
                        END;
                ELSE  /*Gaps-sections not touching*/
                        DO;
                                CHNG_FLG = CHNG_FLG  + 1;  /*Output rec*/
                                SECT_CNT  = SECT_CNT + 1;
                                LINK WRITREC;
                        END;
        END;
ELSE IF TRTNUM NE RTNUM THEN /* Routes have changed*/
        DO;
                IF _N_ NE 1 THEN
                        DO;
                                CHNG_FLG = CHNG_FLG  + 1;  /*Output rec*/
                                SECT_CNT  = SECT_CNT + 1;
                                LINK WRITREC;
                        END;
                ELSE DO;
                        LINK WRITREC;
                END;
        END;
RETURN;

CHECKREC:    /*Check for changes*/
        IF TFED_AID NE FED_AID THEN CHNG_FLG  = CHNG_FLG + 1;
        IF TFUNC_CL NE FUNC_CL THEN CHNG_FLG  = CHNG_FLG + 1;
```

```
        IF TRUR_URB NE RUR_URB THEN CHNG_FLG = CHNG_FLG + 1;
        IF TPAVE_TY NE PAVETYPE THEN CHNG_FLG = CHNG_FLG + 1;
        IF TSHLDTYP NE SHLDTYP THEN CHNG_FLG = CHNG_FLG + 1;
        IF TAADT NE AADT THEN CHNG_FLG = CHNG_FLG + 1;
        IF TLANEWID NE LANEWID THEN CHNG_FLG = CHNG_FLG + 1;
        IF TMEDTYP  NE MEDTYP  THEN CHNG_FLG = CHNG_FLG + 1;
        IF TNO_LANE NE NO_LANE THEN CHNG_FLG = CHNG_FLG + 1;
        IF TOFFPEAK NE OFFPEAK THEN CHNG_FLG = CHNG_FLG + 1;
        IF TPEAKTRK NE PEAKTRK THEN CHNG_FLG = CHNG_FLG + 1;
        IF T_ONEWAY NE ONEWAY THEN CHNG_FLG = CHNG_FLG + 1;
        IF TROADYR NE ROADYR  THEN CHNG_FLG = CHNG_FLG + 1;
        IF TSHLDWID NE SHLDWIDR THEN CHNG_FLG = CHNG_FLG+1;
        IF TTERRAIN NE TERRAIN THEN CHNG_FLG = CHNG_FLG + 1;
        IF TMEDWID NE MEDWID  THEN CHNG_FLG = CHNG_FLG + 1;
        IF CHNG_FLG  = 0  THEN
            DO;           /*If no changes then add to the*/
                ENDMP  = TENDMP;/*Length of the roadway segment*/
                SECT= SECT + SECT_LNT;  /*Do not write rec*/
                SECT_CNT  = SECT_CNT  + 1; /*Update saved rec*/
            END;/*Endmp & sect_lngt*/
        ELSE DO;
            SECT_CNT  = SECT_CNT  + 1;
        END;
    RETURN;

    WRITREC:
        IF _N_ NE 1 THEN
            DO;
                IF EOF THEN   /*Last record processing*/
                    DO;
                        OUTPUT DD1.UTR1XT87;
                        LINK SAVREC;
                        LINK LABELS;
                        OUTPUT DD1.UTR1XT87;
                    END;
                ELSE DO;
                    IF CHNG_FLG  GT  0   THEN   /*If changes output
                                      record*/
                        DO;
                            OUTPUT DD1.UTR1XT87;
                            LINK SAVREC;
```

```
                              LINK LABELS;
                              SECT_CNT=0;
                              CHNG_FLG = 0;
                         END;
               END;
          END;
     ELSE DO;            /*First rec processing*/
          SECT_CNT =0;
          SECT = SECT + SECT_LNG; /*Add first section length*/
          LINK SAVREC;
          LINK LABELS;
          CHNG_FLG = 0;
     END;
RETURN;

SAVREC:
     FED_AID = TFED_AID;
     FUNC_CL = TFUNC_CL;
     RUR_URB =  TRUR_URB;
     SHLDTYP = TSHLDTYP;
     PAVETYPE= TPAVE_TY;
     RTNUM    = TRTNUM;
     BEGMP    = TBEGMP;
     ENDMP    = TENDMP;
     SECT     = SECT_LNT;
     SECT_LNG=SECT_LNT;
     AADT     = TAADT;
     LANEWID = TLANEWID;
     MEDTYP  = TMEDTYP;
     NO_LANE = TNO_LANE;
     OFFPEAK = TOFFPEAK;
     ONEWAY  = T_ONEWAY;
     PEAKTRK = TPEAKTRK;
     ROADYR  = TROADYR;
     SHLDWIDR= TSHLDWID;
     TERRAIN =TTERRAIN;
     MEDWID  = TMEDWID;
RETURN;

LABELS:
     LABELFED_AID ='MODIFIED FEDERAL AID SYSTEM'
```

```
        FUNC_CL= 'MODIFIED FUNCTIONAL CLASS'
        RUR_URB= 'MODIFIED RURAL URBAN DESIGNATION'
        SHLDTYP= 'MODIFIED SHOULDER TYPE'
        PAVETYPE= 'MODIFIED PAVEMENT TYPE'
        RTNUM   = 'ROUTE NUMBER'
        BEGMP   = 'BEGINNING MILEPOST'
        ENDMP   = 'ENDING MILEPOST'
        SECT    = 'SECTION LENGTH EXTENDED'
        SECT_LNG= 'SECTION LENGTH ORIGINAL'
        SECT_CNT= 'NBR OF SECTIONS EXTENDED'
        CHNG_FLG= 'HOMEGEN SECTION BREAK INDICATOR'
        AADT    = 'AADT'
        LANEWID= 'LANE WIDTH'
        MEDTYP  = 'MEDIAN TYPE'
        NO_LANE= 'NUMBER OF LANES'
        OFFPEAK= '% TRUCKS IN OFF PEAK PER'
        ONEWAY  = 'ONE-WAY OR TWO-WAY FAC.'
        PEAKTRK= '%COMMERC VEH IN PEAK PRD'
        ROADYR  = 'YEAR'
        SHLDWIDR= 'RIGHT SHOULDER WIDTH'
        TERRAIN = 'PERDOMINANT TERRIAN TYPE'
        MEDWID  = 'MEDIAN WIDTH';

RETURN;

PROC SORT DATA=DD1.UTR1XT87;
        BY RTNUM BEGMP;

PROC MEANS DATA=DD1.UTR1XT87;
        VAR SECT ;
TITLE 'MEANS OF MODIFIED OUTPUT WITH ALL SECTIONS';

PROC MEANS DATA=UTRD87EX;
        VAR SECT_LNT;
TITLE 'MEANS OF ORIGINAL FILE WITH ALL SECTIONS';

PROC PRINT DATA=UTRD87EX(OBS=50);
VAR     TFED_AID TFUNC_CL TRUR_URB TSHLDTYP TPAVE_TY TRTNUM
        TBEGMP TENDMP TAADT TLANEWID TMEDTYP TNO_LANE TOFFPEAK
        T_ONEWAY TPEAKTRK TROADYR TSHLDWID TTERRAIN TMEDWID;
TITLE 'UTAH 1987 ORIGINAL DATA WITH ALL SECTIONS';
```

```
PROC PRINT DATA=DD1.UTR1XT87(OBS=50);
VAR  FED_AID FUNC_CL RUR_URB SHLDTYP PAVETYPE RTNUM AADT
     LANEWID MEDTYP NO_LANE OFFPEAK ONEWAY PEAKTRK ROADYR
     SHLDWIDR TERRAIN MEDWID BEGMP ENDMP SECT SECT_CNT ;
TITLE1 'UTAH 1987 ROADWAY WITH EXTENDED HOMOGENOUS SECTIONS';

PROC CONTENTS DATA=DD1.UTR1XT87;
RUN;
```

Program 5

```
/*This program creates five curvature variables and adds them to the
corresponding road sections. It uses the roadlog file 'UTRIXT87' created in
program 4 and curvature file 'CURVMOD' created in program 3. The permanent
file 'UTCUV1XT' created is used in subsequent program.*/

LIBNAME DD1 'E:\';

DATA DD1.UTCUV1XT(DROP=CUV_RTE C_BEGMP C_ENDMP DEG_CURV
TMP_LGN LOG_CNT CUV_CNT DIR_CURV YR FLAG TYP1_SEC TYP2_SEC
TYP3_SEC CUV_SECT TMP1 TMP2 TMP3);

/*Initializing the counters zero */

TYP1_CNT=0; TYP2_CNT=0; TYP3_CNT=0; CUV_CNT=0; LOG_CNT=0;
TYP1_SEC=0; TYP2_SEC=0; TYP3_SEC=0; H_CUV1=0; H_CUV2=0;
H_CUV3=0; H_CUV4=0; H_CUV5=0; TMP_LGN=0; FLAG=0; TMP1=0;
TMP2=0; TMP3=0;

LINK READLOG;
LINK READCUV;

LOOP: IF RTNUM = 'AAAAAA' THEN STOP;

    /* Checking if routes are same*/

    IF RTNUM > CUV_RTE THEN  /*Road ahead of curve*/
        DO;
            LINK READCUV;
            GO TO LOOP;
        END;

    IF RTNUM < CUV_RTE THEN /*Curve ahead of road*/
        DO;
            LINK TOTAL;
            LINK WRITREC;
            LINK READLOG;
            GO TO LOOP;
        END;
```

```
/*Routes are same now check milepost of each file*/

IF ENDMP < C_BEGMP THEN
       DO;  /*No curve rec for this road section*/
              LINK TOTAL;
              LINK WRITREC;
              LINK READLOG;
              GO TO LOOP;
       END;

/*Check beginning milepost*/

IF BEGMP > C_ENDMP THEN  /*Road ahead of curve*/
       DO;
              LINK READCUV;
              GO TO LOOP;
       END;

IF C_BEGMP > = BEGMP AND C_ENDMP < = ENDMP THEN
       DO;
              LINK CHECKREC; /*Ok-curve on road section*/
              LINK READCUV;
              GO TO LOOP;
       END;

/*Curve bigger than road section*/

IF C_BEGMP < BEGMP AND C_ENDMP > ENDMP THEN
       DO;
              LINK CHECKRE2;
              LINK TOTAL;
              LINK WRITREC;
              LINK READLOG;
              GO TO LOOP;
       END;

/*Curve overlaps beginning of of road section*/

IF C_BEGMP < BEGMP THEN
       DO;
              TMP_LGN = (C_ENDMP - BEGMP)/ 100;
```

```
                IF TMP_LGN > 0 THEN
                     DO;
                          FLAG = FLAG + 1;
                          LINK CHECKREC;
                          FLAG = 0;
                          LINK READCUV;
                          GO TO LOOP;
                     END;
                ELSE DO;
                     LINK READCUV;
                     GO TO LOOP;
                END;
           END;

      /*Curve overlaps ending of of road section*/

      IF C_BEGMP < ENDMP THEN
           DO;
                TMP_LGN = (ENDMP - C_BEGMP) / 100;
                IF TMP_LGN > 0 THEN
                     DO;
                          FLAG = FLAG + 1;
                          LINK CHECKREC;
                          FLAG=0;
                          LINK TOTAL;
                          LINK WRITREC;
                          LINK READLOG;
                          GO TO LOOP;
                     END;
                ELSE DO;
                     LINK TOTAL;
                     LINK WRITREC;
                     LINK READLOG;
                     GO TO LOOP;
                END;
           END;

      /*Curve begins and the ending milepost of the road*/
      /*Do not include get another road section*/

      IF C_BEGMP = ENDMP THEN
```

```
            DO;
                  LINK TOTAL;
                  LINK WRITREC;
                  LINK READLOG;
                  GO TO LOOP;
            END;
      PUT // 'NO CONDITION MET' _ALL_ ;
RETURN;


CHECKREC:
      IF FLAG = 0 THEN
            DO;
                  IF (DEG_CURV >= 10 AND DEG_CURV < 25) THEN
                        DO;
                              TYP1_SEC = TYP1_SEC + CUV_SECT;
                              TYP1_CNT = TYP1_CNT + 1;
                        END;
                  ELSE IF (DEG_CURV >= 25 AND DEG_CURV < 40)
                  THEN
                        DO;
                              TYP2_SEC = TYP2_SEC + CUV_SECT;
                              TYP2_CNT = TYP2_CNT + 1;
                        END;
                  ELSE IF DEG_CURV >= 50 THEN
                        DO;
                              TYP3_SEC = TYP3_SEC + CUV_SECT;
                              TYP3_CNT = TYP3_CNT + 1;
                        END;
            END;
      ELSE DO;
            IF (DEG_CURV >= 10 AND DEG_CURV < 25) THEN
                  DO;
                        TYP1_SEC = TYP1_SEC + TMP_LGN;
                        TYP1_CNT = TYP1_CNT + 1;
                  END;
            ELSE IF (DEG_CURV >= 25 AND DEG_CURV < 40) THEN
                  DO;
                        TYP2_SEC = TYP2_SEC + TMP_LGN;
                        TYP2_CNT = TYP2_CNT + 1;
                  END;
```

```
                ELSE IF DEG_CURV > = 50 THEN
                        DO;
                                TYP3_SEC = TYP3_SEC + TMP_LGN;
                                TYP3_CNT = TYP3_CNT + 1;
                        END;
        END;
RETURN;

CHECKRE2:
        IF (DEG_CURV > = 10 AND DEG_CURV < 25) THEN
                DO;
                        TYP1_SEC = TYP1_SEC + SECT;
                        TYP1_CNT = TYP1_CNT + 1;
                END;
        ELSE IF (DEG_CURV > = 25 AND DEG_CURV < = 40) THEN
                DO;
                        TYP2_SEC = TYP2_SEC + SECT;
                        TYP2_CNT = TYP2_CNT + 1;
                END;
        ELSE IF DEG_CURV > = 50 THEN
                DO;
                        TYP3_SEC = TYP3_SEC + SECT;
                        TYP3_CNT = TYP3_CNT + 1;
                END;
RETURN;

TOTAL:
        TMP1   = 100 * (TYP1_SEC / SECT);
        H_CUV1 = ROUND(TMP1,.01);
        IF H_CUV1 > 100 THEN PUT 'H_CUV1 > 100';
        IF H_CUV1 > 100 THEN PUT _ALL_;
        TMP2   = 100 * (TYP2_SEC / SECT);
        H_CUV2 = ROUND(TMP2,.01);
        IF H_CUV2 > 100 THEN PUT 'H_CUV2 > 100';
        IF H_CUV2 > 100 THEN PUT _ALL_;
        TMP3   = 100 * (TYP3_SEC / SECT);
        H_CUV3 = ROUND(TMP3,.01);
        IF H_CUV3 > 100 THEN PUT 'H_CUV3 > 100';
        IF H_CUV3 > 100 THEN PUT _ALL_;
        LABEL     H_CUV1 = '% OF CURVES > = 1.0° AND < 2.5°'
                  H_CUV2 = '% OF CURVES > = 2.5° AND < 4.0°'
```

```
                    H_CUV3 = '% OF CURVES >= 5.0°'
                    TYP1_CNT = 'NBR OF CURVES >= 1.0° AND < 2.5°'
                    TYP2_CNT = 'NBR OF CURVES >= 2.5° AND < 4.0°'
                    TYP3_CNT = 'NBR OF CURVES >= 5.0°';
RETURN;

WRITREC:
        OUTPUT DD1.UTCUV1XT;
        TYP1_CNT=0;   TYP2_CNT=0;   TYP3_CNT=0;   TMP_LGN=0;
        TYP1_SEC=0; TYP2_SEC=0; TYP3_SEC=0; H_CUV1=0; H_CUV2=0;
        H_CUV3=0; TMP_LGN=0; FLAG=0; TMP1=0; TMP2=0; TMP3=0;
RETURN;

READLOG:
        IF EOFLOG THEN DO;
                RTNUM = 'AAAAAA';
                PUT LOG_CNT = 'TOTAL ROADLOG RECS READ';
        END;
        ELSE DO;
                LOG_CNT = LOG_CNT + 1;
                SET DD1.UTR1XT87 END=EOFLOG;
        END;
RETURN;

READCUV:
        IF EOFCUV THEN DO;
                CUV_RTE = 'AAAAAA';
                PUT CUV_CNT = 'TOTAL CURVE RECS READ';
        END;
        ELSE DO;
                CUV_CNT = CUV_CNT + 1;
                SET DD1.CURVMOD END=EOFCUV;
        END;
RETURN;

PROC SORT;
        BY RTNUM BEGMP;

PROC PRINT DATA=DD1.UTCUV1XT(OBS=50);
TITLE 'UTAH ROAD SECTIONS MERGED WITH UTAH CURVES';
RUN;
```

Program 6

```
/*This program creates five gradient variables and adds them to the
corresponding road sections. It uses the roadlog file 'UTCUVIXT' created in
program 5 and gradient file 'GRADMOD' created in program 2. The permanent
file 'UTGRADXT' created is permanent roadlog file having gradient and
curvature variables attached to the road sections.*/

LIBNAME DD1 'E:\';
OPTIONS PS=60;
DATA DD1.UTGRADXT(DROP=GRD_RTE G_BEGMP G_ENDMP PCT_GRAD
TMP_LGN LOG_CNT GRD_CNT FLAG TMP6 TMP7 TMP8 TYP6_SEC TYP7_SEC
TYP8_SEC GRD_SECT NEW_SECT);

        /*Initializing the counters zero */

TYP6_CNT=0; TYP7_CNT=0; TYP8_CNT=0; GRD_CNT=0; LOG_CNT=0;
TYP6_SEC=0; TYP7_SEC=0; TYP8_SEC=0; G_GRD1=0; G_GRD2=0;
G_GRD3=0; TMP_LGN=0; FLAG=0; TMP6=0; TMP7=0; TMP8=0;
NEW_SECT=0;

LINK READLOG;
LINK READGRD;

LOOP: IF RTNUM = 'AAAAAA' THEN STOP;

        /* Section to check routes are same*/

        IF RTNUM > GRD_RTE THEN   /*Road ahead of grade*/
                DO;
                        LINK READGRD;
                        GO TO LOOP;
                END;
        IF RTNUM < GRD_RTE THEN /*Grade ahead of road*/
                DO;
                        LINK TOTAL;
                        LINK WRITREC;
                        LINK READLOG;
                        GO TO LOOP;
                END;
```

```
/*Routes are same now check milepost of each file*/

IF ENDMP < G_BEGMP THEN
        DO;/*No grade rec for this road section*/
                LINK TOTAL;
                LINK WRITREC;
                LINK READLOG;
                GO TO LOOP;
        END;

/*Check beginning milepost*/

IF BEGMP > G_ENDMP THEN/*Road ahead of grade*/
        DO;
                LINK READGRD;
                GO TO LOOP;
        END;
IF G_BEGMP > = BEGMP AND G_ENDMP < = ENDMP THEN
        DO;
                LINK CHECKREC; /*Ok-grade on road section*/
                LINK READGRD;
                GO TO LOOP;
        END;

/*Grade bigger than road section*/

IF G_BEGMP < BEGMP AND G_ENDMP > ENDMP THEN
        DO;
                LINK CHECKRE2;
                LINK TOTAL;
                LINK WRITREC;
                LINK READLOG;
                GO TO LOOP;
        END;

/*Grade overlaps beginning of of road section*/

IF G_BEGMP < BEGMP THEN
        DO;
                TMP_LGN = (G_ENDMP - BEGMP)/ 100;
                IF TMP_LGN > 0 THEN
```

```
                        DO;
                                FLAG = FLAG + 1;
                                LINK CHECKREC;
                                FLAG = 0;
                                LINK READGRD;
                                GO TO LOOP;
                        END;
                ELSE DO;
                        LINK READGRD;
                        GO TO LOOP;
                END;
        END;

/*Grade overlaps ending of of road section*/

IF G_BEGMP < ENDMP THEN
        DO;
                TMP_LGN = (ENDMP - G_BEGMP) / 100;
                IF TMP_LGN > 0 THEN
                        DO;
                                FLAG = FLAG + 1;
                                LINK CHECKREC;
                                FLAG=0;
                                LINK TOTAL;
                                LINK WRITREC;
                                LINK READLOG;
                                GO TO LOOP;
                        END;
                ELSE DO;
                        LINK TOTAL;
                        LINK WRITREC;
                        LINK READLOG;
                        GO TO LOOP;
                END;
        END;

/*Grade begins and the ending milepost of the road*/
/*Do not include get another road section*/

IF G_BEGMP = ENDMP THEN
        DO;
```

```
                  LINK TOTAL;
                  LINK WRITREC;
                  LINK READLOG;
                  GO TO LOOP;
              END;
       PUT // 'NO CONDITION MET' _ALL_ ;
RETURN;

CHECKREC:
       IF FLAG = 0 THEN
              DO;
                  IF (PCT_GRAD > = 10 AND PCT_GRAD < 30) THEN
                         DO;
                                TYP6_SEC = TYP6_SEC + GRD_SECT;
                                TYP6_CNT = TYP6_CNT + 1;
                         END;
                  ELSE IF (PCT_GRAD > = 30 AND PCT_GRAD < 50)
                  THEN
                         DO;
                                TYP7_SEC = TYP7_SEC + GRD_SECT;
                                TYP7_CNT = TYP7_CNT + 1;
                         END;
                  ELSE IF PCT_GRAD > = 50 THEN
                         DO;
                                TYP8_SEC = TYP8_SEC + GRD_SECT;
                                TYP8_CNT = TYP8_CNT + 1;
                         END;
              END;
       ELSE DO;
              IF (PCT_GRAD > = 10 AND PCT_GRAD < 30) THEN
                  DO;
                         TYP6_SEC = TYP6_SEC + TMP_LGN;
                         TYP6_CNT = TYP6_CNT + 1;
                  END;
              ELSE IF (PCT_GRAD > = 25 AND PCT_GRAD < 50) THEN
                  DO;
                         TYP7_SEC = TYP7_SEC + TMP_LGN;
                         TYP7_CNT = TYP7_CNT + 1;
                  END;
              ELSE IF PCT_GRAD > = 50 THEN
                  DO;
```

```
                        TYP8_SEC = TYP8_SEC + TMP_LGN;
                        TYP8_CNT = TYP8_CNT + 1;
                   END;
          END;
RETURN;

CHECKRE2:
          IF (PCT_GRAD > = 10 AND PCT_GRAD < 30) THEN
               DO;
                    TYP6_SEC = TYP6_SEC + SECT;
                    TYP6_CNT = TYP6_CNT + 1;
               END;
          ELSE IF (PCT_GRAD > = 25 AND PCT_GRAD < 50) THEN
               DO;
                    TYP7_SEC = TYP7_SEC + SECT;
                    TYP7_CNT = TYP7_CNT + 1;
               END;
          ELSE IF PCT_GRAD > = 50 THEN
               DO;
                    TYP8_SEC = TYP8_SEC + SECT;
                    TYP8_CNT = TYP8_CNT + 1;
               END;
RETURN;

TOTAL:
          TMP6 = 100 * (TYP6_SEC / SECT);
          G_GRD1 = ROUND(TMP6,.01);
          IF G_GRD1 > 100.01  THEN
               DO;
                    NEW_SECT = (ENDMP - BEGMP) /100;
                    TMP6 = 100 * (TYP6_SEC / NEW_SECT);
                    G_GRD1 = ROUND(TMP6,.01);
               END;
          IF G_GRD1 > 100.01  THEN PUT _ALL_;
          TMP7 = 100 * (TYP7_SEC / SECT);
          G_GRD2 = ROUND(TMP7,.01);
          IF G_GRD2 > 100.01  THEN
               DO;
                    NEW_SECT = (ENDMP - BEGMP) /100;
                    TMP7 = 100 * (TYP7_SEC / NEW_SECT);
                    G_GRD2 = ROUND(TMP7,.01);
```

```
          END;
     IF G_GRD2 > 100.01 THEN PUT _ALL_;
     TMP8 = 100 * (TYP8_SEC / SECT);
     G_GRD3 = ROUND(TMP8,.01);
     IF G_GRD3 > 100.01  THEN
          DO;
               NEW_SECT = (ENDMP - BEGMP) /100;
               TMP8 = 100 * (TYP8_SEC / NEW_SECT);
               G_GRD3 = ROUND(TMP8,.01);
          END;
     IF G_GRD3 > 100.01 THEN PUT _ALL_;
     LABEL     G_GRD1 = '% GRADES > = 1.0% AND < 3.0%'
               G_GRD2 = '% GRADES > = 3.0% AND < 5.0%'
               G_GRD3 = '% GRADES > = 5.0%'
               TYP6_CNT = 'NBR GRADES > = 1.0% AND < 3.0%'
               TYP7_CNT = 'NBR GRADES > = 3.0% AND < 5.0%'
               TYP8_CNT = 'NBR GRADES > = 5.0%';
RETURN;

WRITREC:
     OUTPUT DD1.UTGRADXT;
     TYP6_CNT=0;  TYP7_CNT=0;  TYP8_CNT=0;  TMP_LGN=0;
     TYP6_SEC=0;TYP7_SEC=0;TYP8_SEC=0;G_GRD1=0;G_GRD2=0;
     G_GRD3=0; FLAG=0; TMP6=0; TMP7=0; TMP8=0; NEW_SECT=0;

RETURN;

READLOG:
     IF EOFLOG THEN DO;
          RTNUM = 'AAAAAA';
          PUT LOG_CNT = 'TOTAL ROADLOG RECS READ';
     END;
     ELSE DO;
          LOG_CNT = LOG_CNT + 1;
          SET DD1.UTCURVXT END=EOFLOG;
     END;
RETURN;

READGRD:
     IF EOFGRD THEN DO;
          GRD_RTE = 'AAAAAA';
```

```
          PUT GRD_CNT = 'TOTAL GRADE RECS READ';
     END;
     ELSE DO;
          GRD_CNT = GRD_CNT + 1;
          SET DD1.GRADMOD END=EOFGRD;
     END;
RETURN;

PROC SORT;
     BY RTNUM BEGMP;

PROC FORMAT;
     VALUE GRADF  .01 - .99 = '.01 - .99'
              1.00 - 10.99 = '1.00 - 10.99'
             11.00 - 20.99 = '11.00 - 20.99'
             21.00 - 30.99 = '21.00 - 30.99'
             31.00 - 40.99 = '31.00 - 40.99'
             41.00 - 50.99 = '41.00 - 50.99'
             51.00 - 60.99 = '51.00 - 60.99'
             61.00 - 70.99 = '61.00 - 70.99'
             71.00 - 80.99 = '71.00 - 80.99'
             81.00 - 90.99 = '81.00 - 90.99'
             91.00 - 100.00 = '91.00 - 100.00';

PROC FREQ DATA=DD1.UTGRADXT;
TABLES G_GRD1;
FORMAT G_GRD1 GRADF.;

PROC PRINT DATA=DD1.UTGRADXT(OBS=40);
TITLE '1987 UTAH EXTENDED ROADWAY SECTIONS MERGED';
TITLE2 'WITH UTAH CURVES AND GRADES';

RUN;

PROC SUMMARY DATA=DD1.UTGRADXT;
VAR  H_CUV1  H_CUV2  H_CUV3  G_GRD1  G_GRD2  G_GRD3  TPYP1_CNT
     TYP2_CNT TYP3_CNT TYP7_CNT TYP8_CNT TYP6_CNT;
OUTPUT OUT=CNTS SUM=;
PROC PRINT DATA=CNTS;

RUN;
```

Program 7

```
/*This program first matches the accidents from the merged accident-vehicle
file with the road segments from road log file created by program 6 and than
counts each type of accident and attaches a counter to the corresponding road
segments. The outcome of this program is a permanent file 'UTCOUNTS' which
has road segments having accident counters and curve and grade information
attached to them.*/

LIBNAME UTAH '.';

DATA UTAH.UTCOUNTS/*Creating a permanent file*/
(KEEP = AADT BEGMP ENDMP FED_AID FUNC_CL LANEWID MEDTYP
MEDWID NO_LANE OFFPEAK ONEWAY PAVETYPE PEAKTRK ROADYR
RTNUM RUR_URB SECT SHLDTYP SHLDWIDR TERRAIN A_ROUTE ACCD_CNT
MV_AW_AD BICPEDTR MV_MV ROR OTURN OTHER COL_CNT HEADON INT
RR_PASS REAREND OPP_SS PASS_SS TURNING APP_ANG PASSING
BACKING PK_VEH SIG_VEH OTH ROD_CNT S_LEVEL S_GRADE S_CREST
S_DIP C_LEVEL C_GRADE C_CREST C_DIP CLIM_CNT RAINING SNOWING
CLEAR FOG VEH_CNT STRAIGHT ARTICUL G_GRD1 G_GRD2 G_GRD3
H_CUV1 H_CUV2 H_CUV3 TYP7_CNT TYP1_CNT TYP2_CNT TYP3_CNT
TYP6_CNT TYP8_CNT TYP7_CNT );

        /* This step matches accident data with roadlog segments.*/


LENGTH RTNUM A_ROUTE $6;

        /*Initializing Counters*/

ACCD_CNT = 0; MV_AW_AD = 0; BICPEDTR = 0; MV_MV = 0; ROR = 0;
OTURN = 0; OTHER = 0; COL_CNT = 0; HEADON = 0; INT = 0; RR_PASS
= 0; REAREND = 0; OPP_SS = 0; PASS_SS = 0; TURNING = 0; APP_ANG
= 0; PASSING = 0; BACKING = 0; PK_VEH = 0; SIG_VEH = 0; OTH = 0;
ROD_CNT = 0; S_LEVEL = 0; S_GRADE = 0; S_CREST = 0; S_DIP = 0;
C_LEVEL = 0; C_GRADE = 0; C_CREST = 0; C_DIP = 0; CLIM_CNT = 0;
RAINING = 0; SNOWING = 0; CLEAR = 0; FOG = 0; VEH_CNT = 0;
STRAIGHT = 0; ARTICUL = 0; RTMATCH = 0; DROPAC = 0; RTLOW = 0;

LABEL     ACCD_CNT = 'TOTAL ACCIDENT COUNTER'
          MV_AW_AD = 'ANIMALS ACCIDENT COUNTER'
```

```
          MV_MV    = 'MOTOR VEHICLE ACCIDENT COUNTER'
          BICPEDTR  =  'BICYCLE  PEDESTRAIN  TRAIN  ACCIDENT
                    COUNTER'
          ROR      = 'RUN OFF THE ROAD ACCIDENT COUNTER'
          OTURN    = 'OVER TURNING ACCIDENT COUNTER'
          OTHER    = 'OTHER ACCIDENT TYPE ACCIDENT COUNTER'
          COL_CNT  = 'TOTAL COLLISION ACCIDENT COUNTER'
          HEADON   = 'HEAD ON ACCIDENT COUNTER'
          INT      = 'INTERSECTION ACCIDENT COUNTER'
          RR_PASS  = 'REAR END PASSING ACCIDENT COUNTER'
          REAREND  = 'REAREND ACCIDENT COUNTER'
          OPP_SS   = 'OPPOSITE SIDESWIPE ACCIDENT COUNTER'
          PASS_SS  = 'PASSING SIDESWIPE ACCIDENT COUNTER'
          TURNING  = 'TURNING ACCIDENT COUNTER'
          APP_ANG  = 'APPROACH ANGLE ACCIDENT COUNTER'
          PASSING  = 'PASSING ACCIDENT COUNTER'
          BACKING  = 'BACKING ACCIDENT COUNTER'
          PK_VEH   = 'PARKED VEHICLE ACCIDENT COUNTER'
          SIG_VEH  = 'SINGLE VEHICLE ACCIDENT COUNTER'
          OTH      = 'OTHER COLLISION ACCIDENT COUNTER'
          ROD_CNT  = 'TOTAL ROAD CHARACTERISTICS COUNTER'
          S_LEVEL  = 'STRAIGHT LEVEL ROAD COUNTER'
          S_GRADE  = 'STRAIGHT GRADE ROAD COUNTER'
          S_CREST  = 'STRAIGHT CREST ROAD COUNTER'
          S_DIP    = 'STRAIGHT DIP ROAD COUNTER'
          C_LEVEL  = 'CURVE LEVEL ROAD COUNTER'
          C_GRADE  = 'CURVE GRADE ROAD COUNTER'
          C_CREST  = 'CURVE CREST ROAD COUNTER'
          C_DIP    = 'CURVE DIP ROAD COUNTER'
          CLIM_CNT = 'TOTAL WEATHER CONDITION COUNTER'
          RAINING  = 'RAINING CONDITION COUNTER'
          SNOWING  = 'SNOWING CONDITION COUNTER'
          CLEAR    = 'CLEAR CONDITION COUNTER'
          FOG      = 'FOGY CONDITION COUNTER'
          VEH_CNT  = 'TOTAL TRUCKS COUNTER'
          STRAIGHT = 'STRAIGHT TRUCKS COUNTER'
          ARTICUL  = 'ARTICULATED TRUCKS COUNTER';

LINK READLOG;
LINK READACC;
```

```
LOOP:
     IF RTNUM = 'AAAAAA' THEN STOP;

     /* This step varifies state route numbers are the same.*/

     IF RTNUM > A_ROUTE THEN
          DO;
               RTLOW + 1;
               LINK READACC;
               GO TO LOOP;
          END;

     IF RTNUM < A_ROUTE THEN
          DO;
               LINK TOTAL;
               LINK READLOG;
               GO TO LOOP;
          END;

     /*At this point there is a match on route. The next step is to compare*/
     /*a_milept on accident file with begmp and endmp on roadlog file.*/

     IF A_MILEPT > = BEGMP AND A_MILEPT < ENDMP THEN
          DO;
               LINK COUNTS;
               LINK READACC;
               GO TO LOOP;
          END;
     ELSE IF A_MILEPT < BEGMP THEN
          DO;
               DROPAC + 1;
               LINK READACC;
               GO TO LOOP;
          END;
     ELSE IF A_MILEPT > = ENDMP THEN
          DO;
               LINK TOTAL;
               LINK READLOG;
               GO TO LOOP;
          END;
RETURN;
```

```
READLOG:
       IF EOFLOG THEN DO;
              RTNUM = 'AAAAAA';
              PUT DROPAC = 'MISSING SECTIONS';
              PUT RTLOW = 'ACCS ON MISSING ROUTES';
              PUT RTMATCH = 'ACCIDENTS COUNTED';
       END;
       ELSE
              SET UTAH.NEWGRD1 END=EOFLOG;
RETURN;

READACC:
       IF EOFACC THEN DO;
              A_ROUTE = 'AAAAAA';
       END;
       ELSE
              SET UTAH.UTGROUP END=EOFACC;
RETURN;

COUNTS:
       RTMATCH + 1;

       /*Developing various accident type variable counters.*/

       IF ACCTYPE = 1 THEN
              DO;
                     MV_AW_AD = MV_AW_AD + 1;
              END;
       ELSE IF ACCTYPE = 2 THEN
              DO;
                     BICPEDTR = BICPEDTR + 1;
              END;
       ELSE IF ACCTYPE = 3 THEN
              DO;
                     FDOOB = FDOOB + 1;
              END;
       ELSE IF ACCTYPE = 4 THEN
              DO;
                     MV_MV = MV_MV + 1;
              END;
       ELSE IF ACCTYPE = 5 THEN
```

```
          DO;
               ROR = ROR + 1;
          END;
     ELSE IF ACCTYPE = 6 THEN
          DO;
               OTURN = OTURN + 1;
          END;
     ELSE IF ACCTYPE = 7 THEN
          DO;
               OTHER = OTHER + 1;
          END;

     /*Developing various collision type variable counters.*/

     IF COL_TYPE = 1 THEN
          DO;
               HEAD_ON = HEAD_ON + 1;
          END;
     ELSE IF COL_TYPE = 2 THEN
          DO;
               INT = INT + 1;
          END;
     ELSE IF COL_TYPE = 3 THEN
          DO;
               RR_PASS = RR_PASS + 1;
          END;
     ELSE IF COL_TYPE = 4 THEN
          DO;
               REAREND = REAREND + 1;
          END;
     ELSE IF COL_TYPE = 5 THEN
          DO;
               OPP_SS = OPP_SS + 1;
          END;
     ELSE IF COL_TYPE = 6 THEN
          DO;
               PASS_SS = PASS_SS + 1;
          END;
     ELSE IF COL_TYPE = 7 THEN
          DO;
               TURNING = TURNING + 1;
```

```
        END;
ELSE IF COL_TYPE = 8 THEN
        DO;
                APP_ANG = APP_ANG + 1;
        END;
ELSE IF COL_TYPE = 9 THEN
        DO;
                PASSING = PASSING + 1;
        END;
ELSE IF COL_TYPE = 10 THEN
        DO;
                BACKING = BACKING + 1;
        END;
ELSE IF COL_TYPE = 11 THEN
        DO;
                PK_VEH = PK_VEH + 1;
        END;
ELSE IF COL_TYPE = 12 THEN
        DO;
                SIG_VEH = SIG_VEH + 1;
        END;
ELSE IF COL_TYPE = 13 THEN
        DO;
                OTH = OTH + 1;
        END;

/*Developing road characteristics variable counters.*/

IF RD_CHAR = 1 THEN
        DO;
                S_LEVEL = S_LEVEL + 1;
        END;
ELSE IF RD_CHAR = 2 THEN
        DO;
                S_GRADE = S_GRADE + 1;
        END;
ELSE IF RD_CHAR = 3 THEN
        DO;
                S_CREST = S_CREST + 1;
        END;
ELSE IF RD_CHAR = 4 THEN
```

```
          DO;
                    C_LEVEL = C_LEVEL + 1;
          END;
ELSE IF RD_CHAR = 5 THEN
          DO;
                    C_GRADE = C_GRADE + 1;
          END;
ELSE IF RD_CHAR = 6 THEN
          DO;
                    C_CREST = C_CREST + 1;
          END;
ELSE IF RD_CHAR = 7 THEN
          DO;
                    S_DIP = S_DIP + 1;
          END;
ELSE IF RD_CHAR = 8 THEN
          DO;
                    C_DIP = C_DIP + 1;
          END;

/*Developing climate condition variable counters.*/

IF CLIMATE = 99 THEN
          DO;
                    CLEAR = CLEAR + 1;
          END;
ELSE IF CLIMATE = 1 THEN
          DO;
                    RAINING = RAINING + 1;
          END;
ELSE IF CLIMATE = 2 THEN
          DO;
                    SNOWING = SNOWING + 1;
          END;
ELSE IF CLIMATE = 3 THEN
          DO;
                    FOG = FOG + 1;
          END;

/*Developing vehicle type variable counters.*/
```

```
    IF VEH_TYPE = 1 THEN
        DO;
            STRAIGHT = STRAIGHT + 1;
        END;
    ELSE IF VEH_TYPE = 2 THEN
        DO;
            ARTICUL = ARTICUL + 1;
        END;

RETURN;

TOTAL:
    ACCD_CNT = MV_AW_AD + BICPEDTR + MV_MV + ROR + OTURN
                + OTHER;
    COL_CNT = HEADON + INT + RR_PASS + REAREND + OPP_SS +
               PASS_SS + TURNING + APP_ANG + PASSING +
               BACKING + PK_VEH + SIG_VEH + OTH;
    ROD_CNT = S_LEVEL + S_GRADE + S_CREST + S_DIP + C_LEVEL
               C_GRADE + C_CREST + C_DIP;
    CLIM_CNT = CLEAR + RAINING + SNOWING + FOG;
    VEH_CNT = STRAIGHT + ARTICUL;
    OUTPUT UTAH.UTCOUNTS;
    ACCD_CNT = 0; MV_AW_AD = 0; BICPEDTR = 0; MV_MV = 0; ROR
    = 0; OTURN = 0; OTHER = 0; COL_CNT = 0; HEADON = 0; INT =
    0; RR_PASS = 0; REAREND = 0; OPP_SS = 0; PASS_SS = 0;
    TURNING = 0; APP_ANG = 0; PASSING = 0; BACKING = 0; PK_VEH
    = 0; SIG_VEH = 0; OTH = 0; ROD_CNT = 0; S_LEVEL = 0; S_GRADE
    = 0; S_CREST = 0; S_DIP = 0; C_LEVEL = 0; C_GRADE = 0;
    C_CREST = 0; C_DIP = 0; CLIM_CNT = 0; RAINING = 0; CLEAR =
    0; SNOWING = 0; FOG = 0; VEH_CNT = 0; STRAIGHT = 0; ARTICUL
    = 0;
RETURN;

PROC SUMMARY;
VAR  ACCD_CNT CLIM_CNT VEH_CNT COL_CNT ROD_CNT;
OUTPUT OUT=COUNTS SUM=;

PROC PRINT DATA=COUNTS;
PROC CONTENTS DATA=UTAH.UTCOUNTS;
PROC PRINT DATA=UTAH.UTCOUNTS(OBS=50);
RUN;
```

Program 8

```
/*This program removes all the zero values of variable 'PEAKTRK' and replaces
it with the values of preceeding or succeeding segments as the case maybe and
remove all the segments which do not have any peaktrk percentage. The output
of this program is the final master file with truck exposure variable attached to
each roadway segment and this file is to be used for development of models.*/

LIBNAME UTAH '.';

        /*Following statements renames the two variables in the roadlog file.*/

DATA EXPRENAM;
SET UTAH.TRUCKCNT; /* This file is output of program 7.*/
RENAME      RTNUM = SRTNUM
            PEAKTRK = SPEAKTRK;
RUN;

        /* This program creates permanent file "exposure" which has the
        computed peaktrk precentage as "speaktrk" and another permanent file
        "noexpose" which has the segments on which there is no peaktrk
        percentage for any segment on that entire route. */

DATA   UTAH.EXPOSURE(DROP=RTNUM   PEAKTRK)   UTAH.NOEXPOSE
(DROP=RTNUM PEAKTRK);

LENGTH RTNUM SRTNUM $6.;

        /*Initializing counters*/

Z_CNT = 0; MATCH = 0; NOMATCH = 0; UNMATCH = 0; ORIGINAL = 0;

LINK READLOG1;
LINK READLOG2;

LOOP:
        IF SRTNUM = 'AAAAAA' THEN STOP;
        IF SRTNUM > RTNUM THEN
                DO;
                        LINK READLOG2;
                        GO TO LOOP;
```

```
          END;

      IF SRTNUM  < RTNUM THEN
          DO;
               UNMATCH + 1;
               OUTPUT UTAH.NOEXPOSE;
               LINK READLOG1;
               GO TO LOOP;
          END;

      IF SRTNUM  = RTNUM THEN
          DO;
               IF SPEAKTRK  = 0 AND PEAKTRK > 0 THEN
                    DO;
                         SPEAKTRK  = PEAKTRK;
                         Z_CNT + 1;
                         MATCH + 1;
                         OUTPUT UTAH.EXPOSURE;
                         LINK READLOG1;
                         GO TO LOOP;
                    END;
               ELSE IF SPEAKTRK  = 0 AND PEAKTRK  = 0 THEN
                    DO;
                         NOMATCH + 1;
                         LINK READLOG2;
                         GO TO LOOP;
                    END;
          END;

      IF SPEAKTRK > 0 THEN
          DO;
               ORIGINAL + 1;
               OUTPUT UTAH.EXPOSURE;
               LINK READLOG1;
               GO TO LOOP;
          END;
RETURN;

READLOG1:
      IF EOFLOG1 THEN
          DO;
```

```
                SRTNUM = 'AAAAAA';
                PUT   UNMATCH   =   'SEGMENTS   DOES   NOT   HAVE
                          PEAKTRK';
                PUT Z_CNT = 'ZERO PEAKTRK ENCOUNTERED';
                PUT MATCH = 'PEAKTRK MATCHED';
                PUT NOMATCH = 'PEAKTRK DID NOT MATCHED';
            END;
        ELSE
            SET EXPRENAM END= EOFLOG1;
RETURN;

READLOG2:
        IF EOFLOG2 THEN
            DO;
                RTNUM = 'AAAAAA';
            END;
        ELSE
            SET   UTAH.TRUCKCNT(KEEP=RTNUM   PEAKTRK)   END   =
            EOFLOG2;
RETURN;

RUN;
```

Appendix B

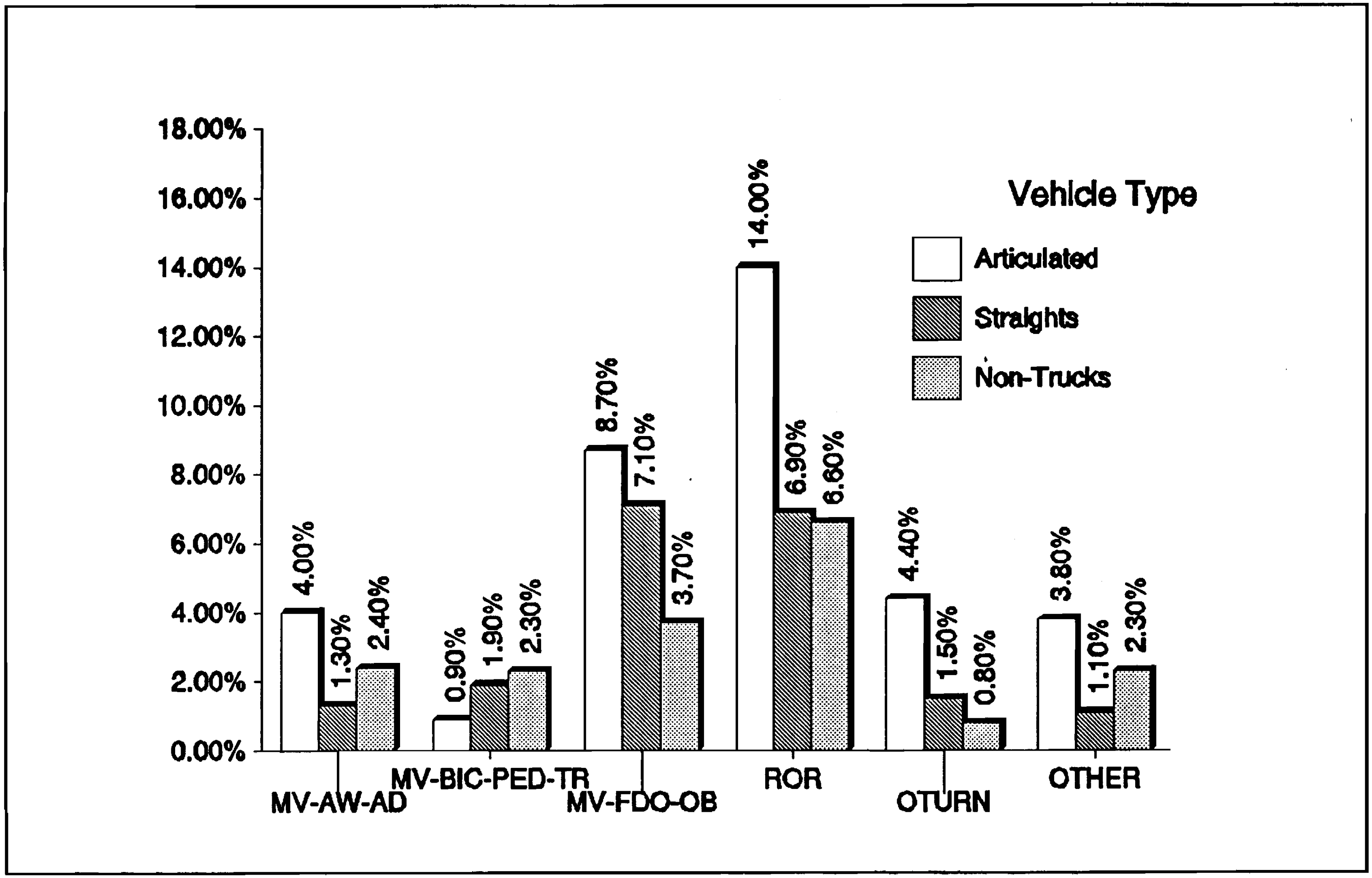


Figure B-1 Accident Type by Vehicle type (Utah)

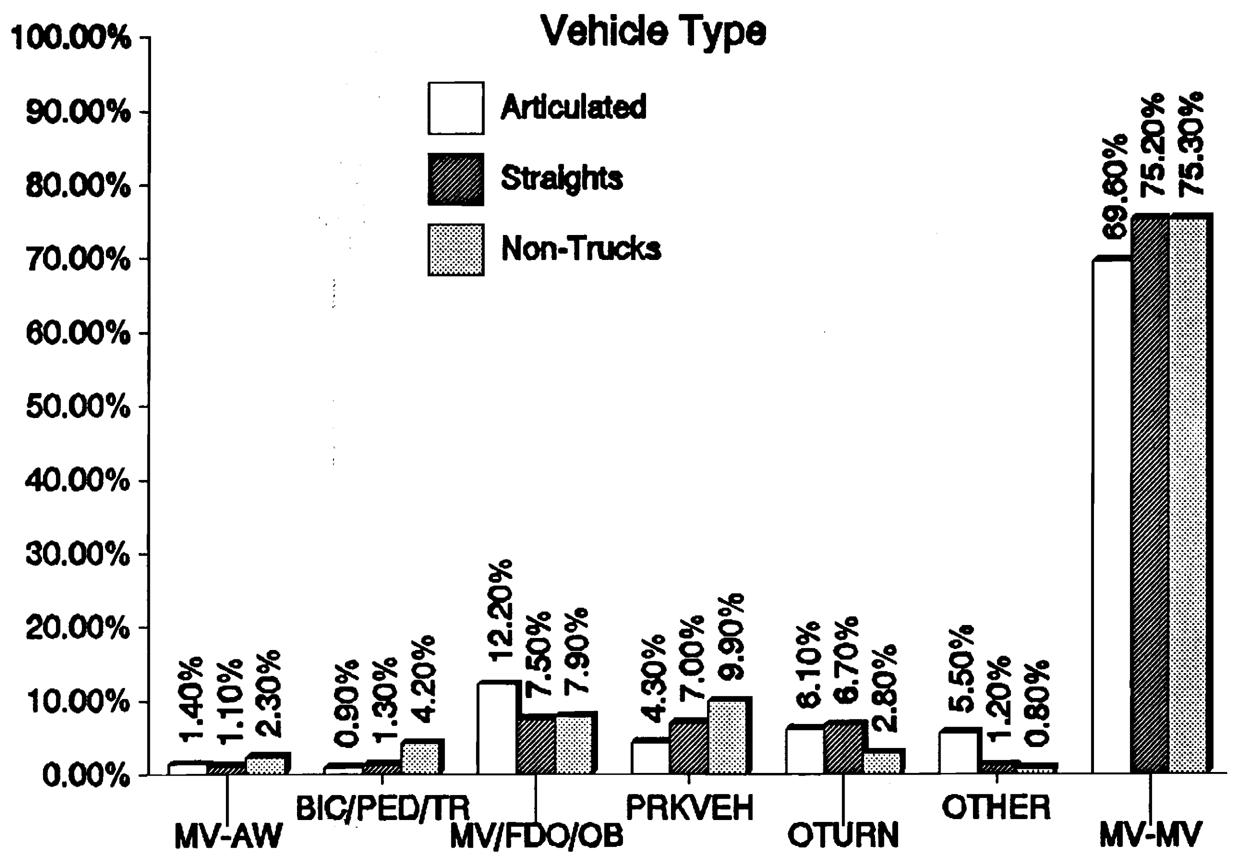


Figure B-2 Accident Type by Vehicle Type (Minnesota)

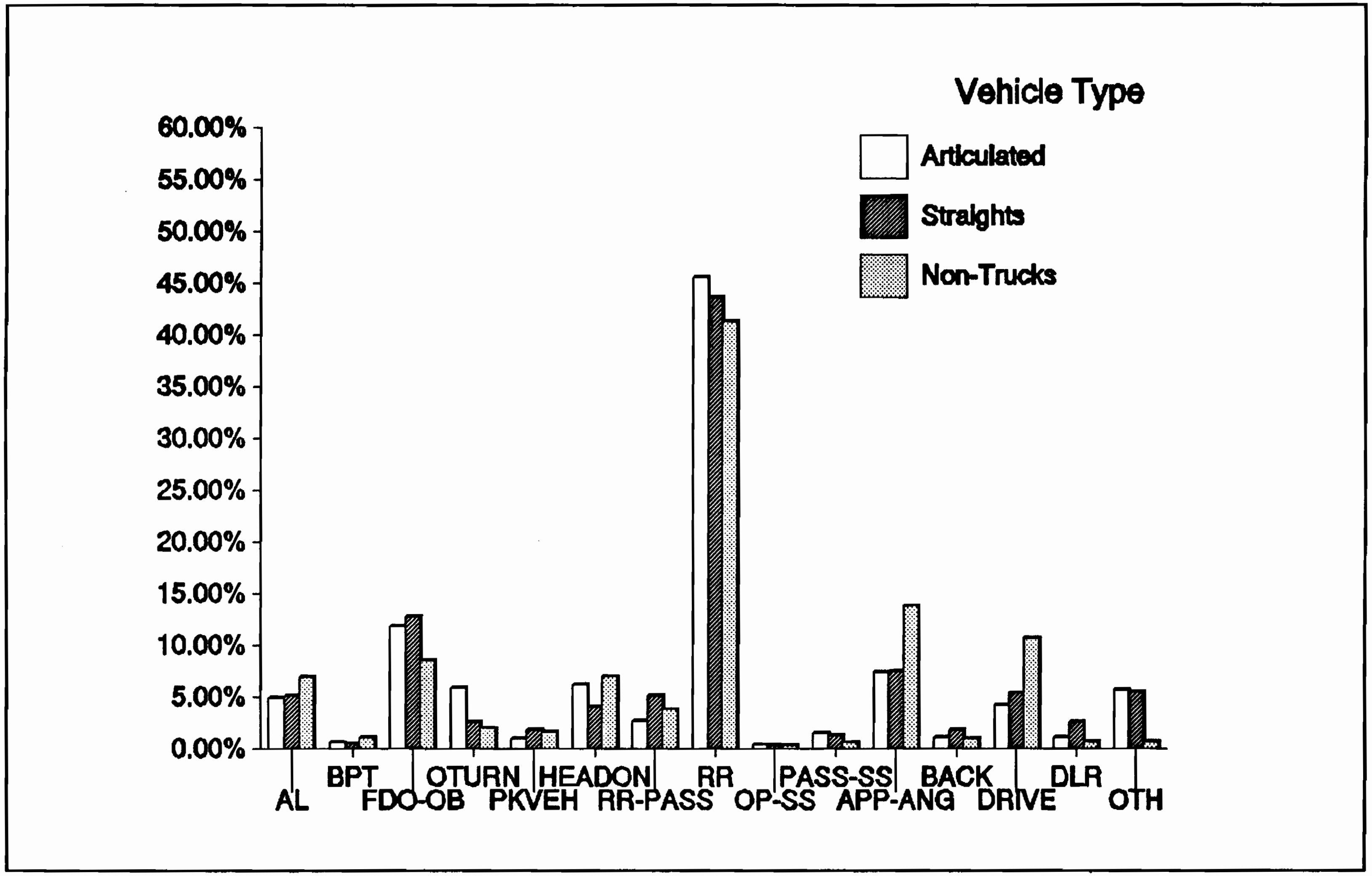


Figure B-3 Accident Type by Vehicle Type (Michigan)

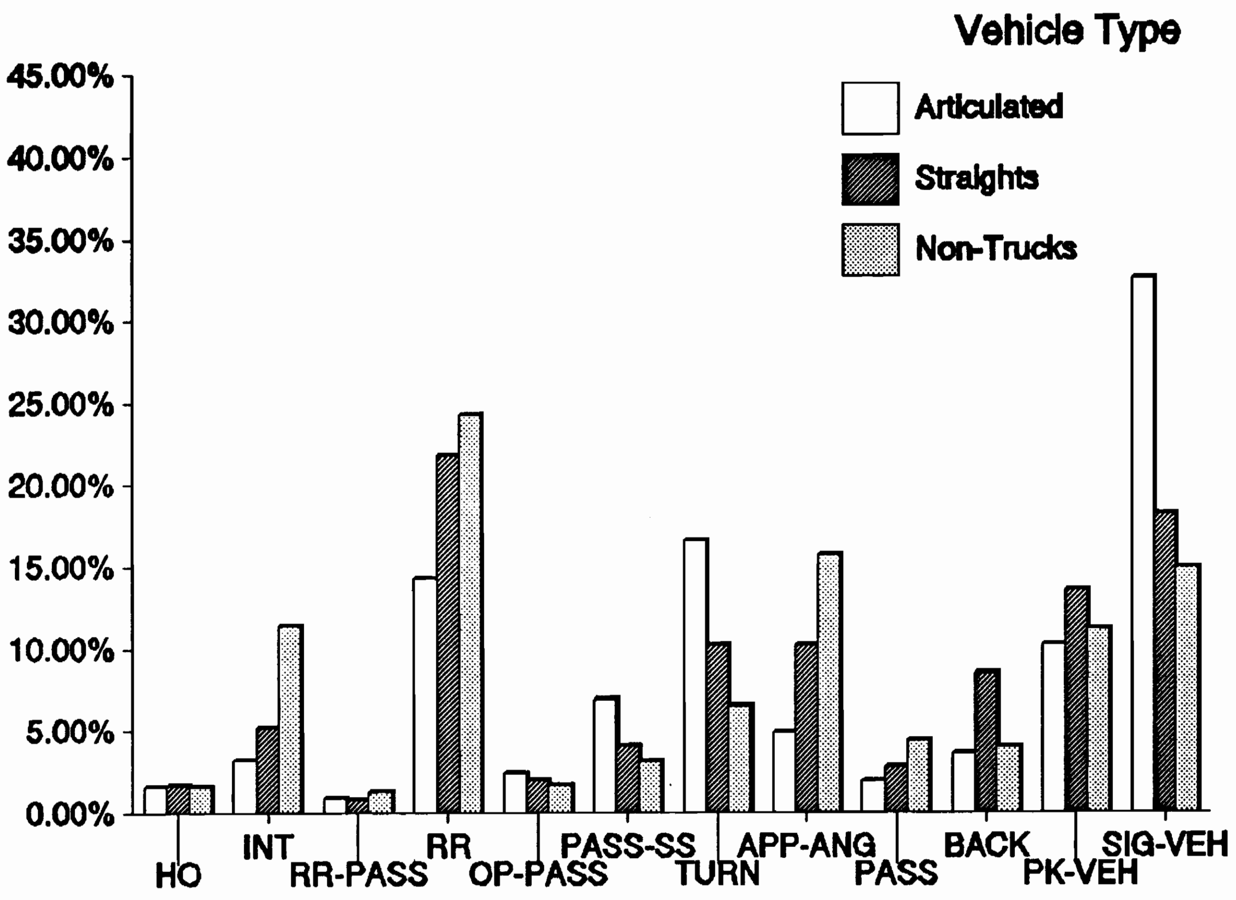


Figure B-4 Collision Type by Vehicle Type (Utah)

Table B-1 Accident Type by Vehicle Type (Michigan)

Accdent Type	Articulated	Straights	Non-Trucks
AL	4.90	5.10	6.90
BPT	0.60	0.50	1.10
FOOB	11.80	12.80	8.50
OT	5.90	2.60	2.00
PKVH	1.00	1.80	1.60
HO	6.20	4.10	7.00
RR-PASS	2.70	5.10	3.80
RR	45.60	43.70	41.30
OPSS	0.40	0.30	0.30
PASSS	1.50	1.30	0.60
APAG	7.40	7.50	13.80
BACK	1.10	1.80	1.00
DRIVE	4.20	5.40	10.70
DLR	1.10	2.60	0.70
OTH	5.70	5.50	0.70
TOTAL	100.00	100.00	100.00

Table B-2 Collision Type by Vehicle Type (Utah)

Accident Type	Articulated	Straight	Non Trucks
HO	1.6	1.7	1.6
INT	3.2	5.2	11.4
RR_PASS	0.9	0.8	1.3
RR	14.3	21.8	24.3
OP_PASS	2.4	2.00	1.7
PASS_SS	6.9	4.1	3.1
TURN	16.6	10.2	6.5
APP_ANG	4.9	10.2	15.7
PASS	1.9	2.8	4.4
BACK	3.6	8.5	4
PK_VEH	10.2	13.6	11.2
SIG_VEH	32.6	18.2	14.9
Total	100.00	100.00	100.00

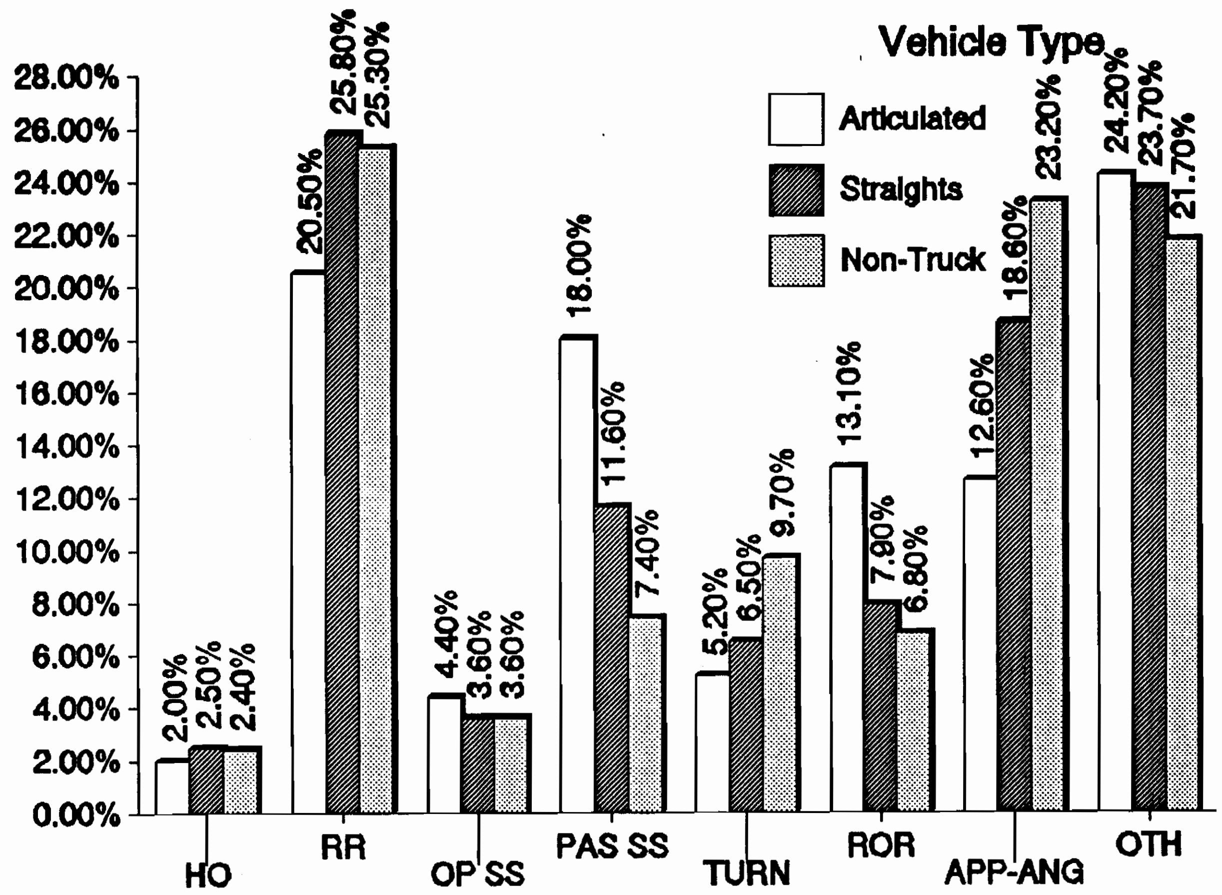


Figure B-5 Collision Type by Vehicle Type (Minnesota)

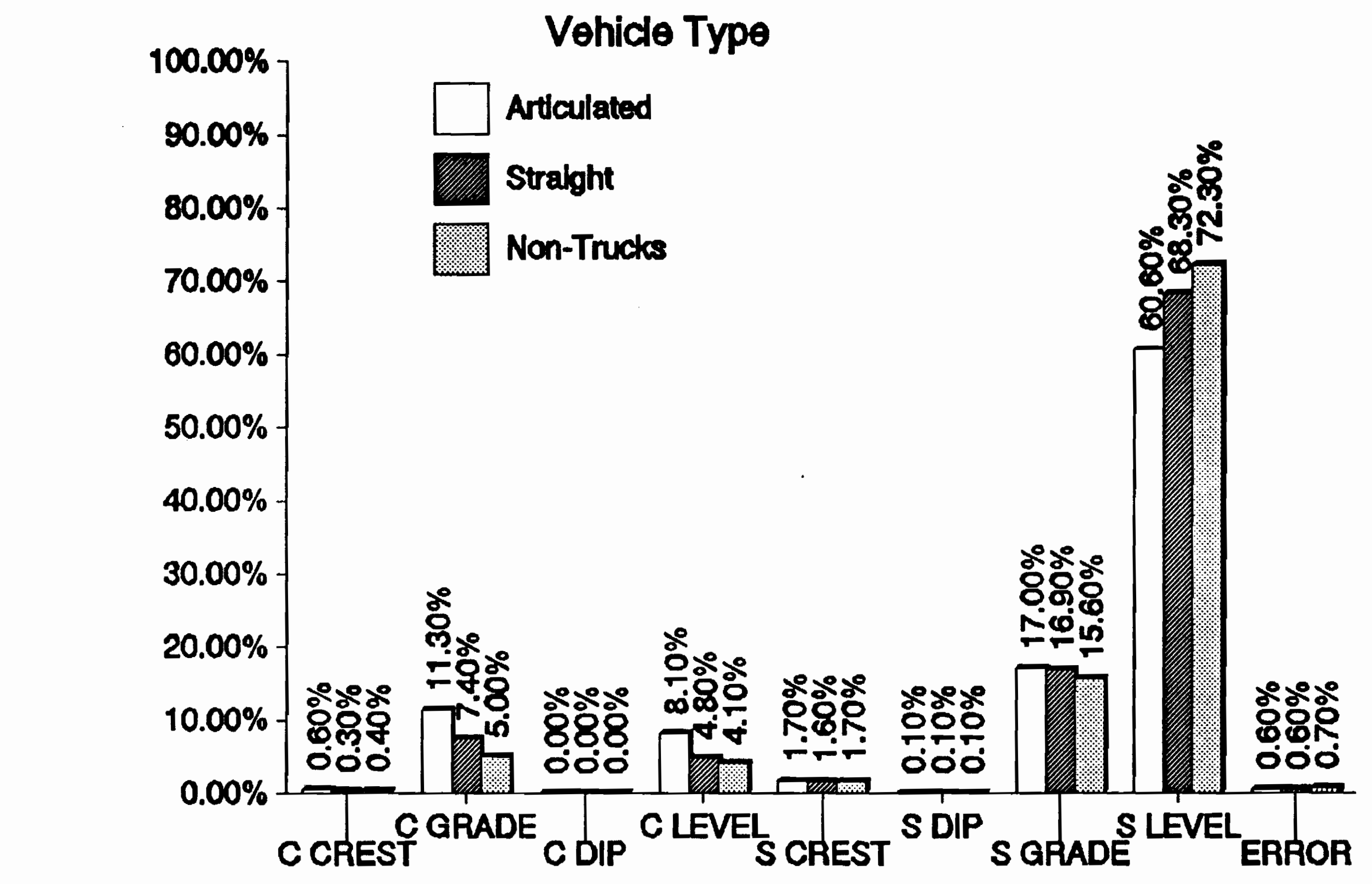


Figure B-6 Road Characteristics by Vehicle Type (Utah)

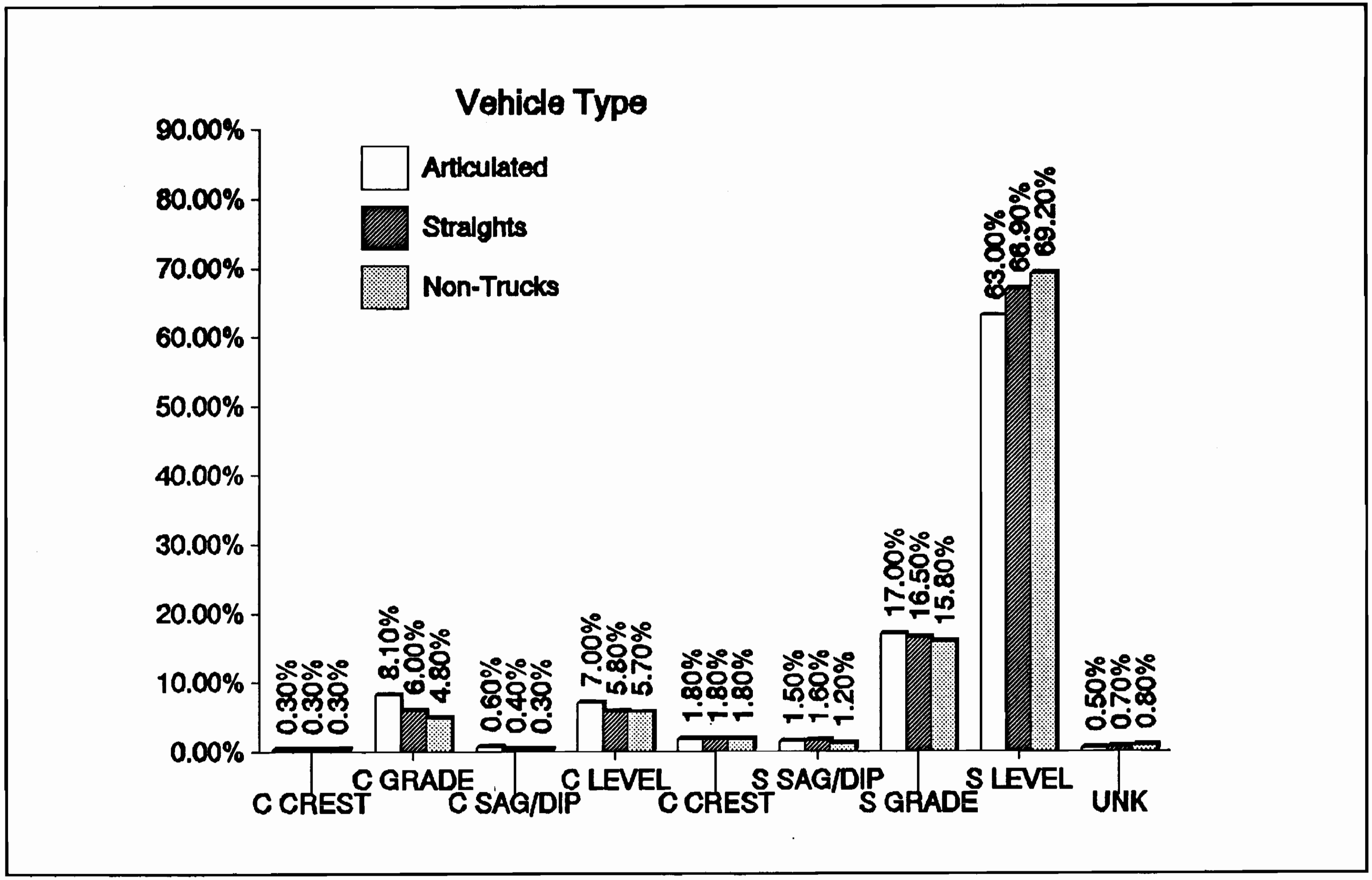


Figure B-7 Road Characteristics by Vehicle Type (Minnesota)

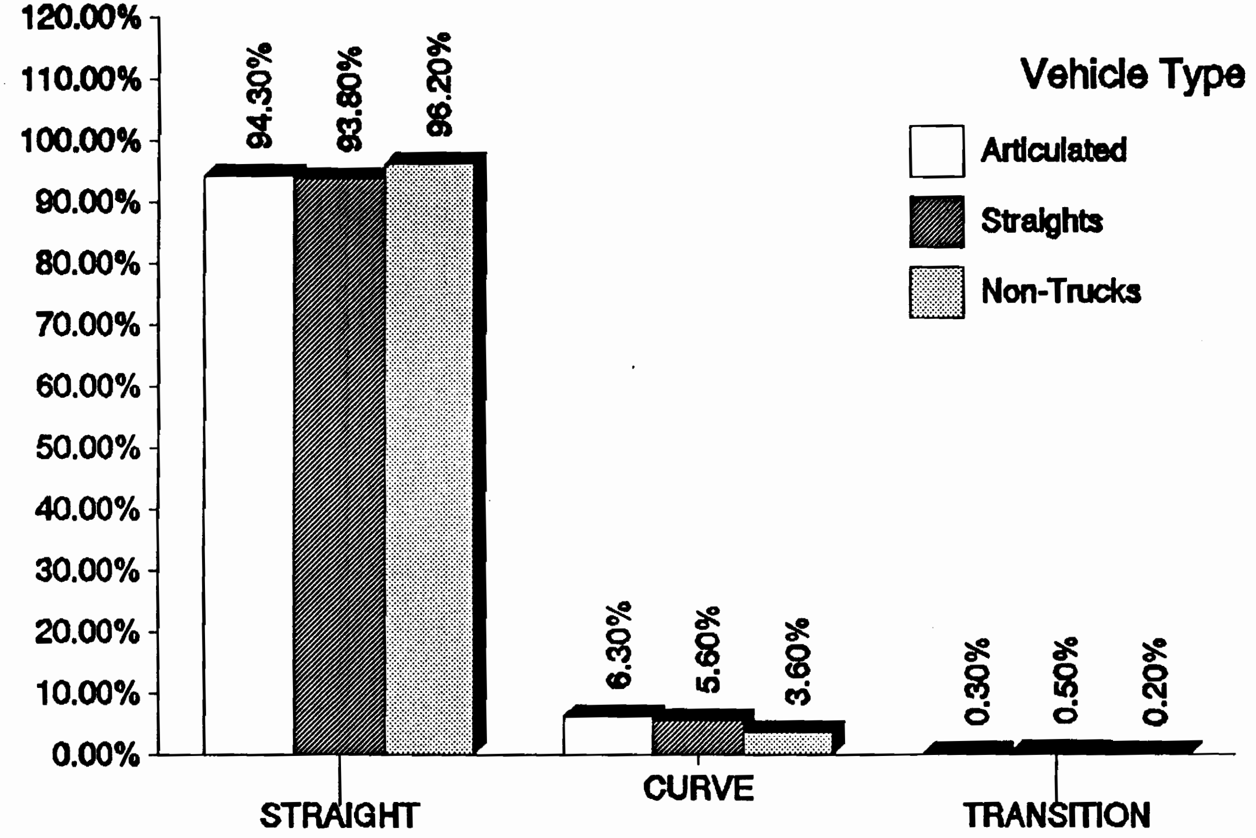


Figure B-8 Road Characteristics by Vehicle Type (Michigan)

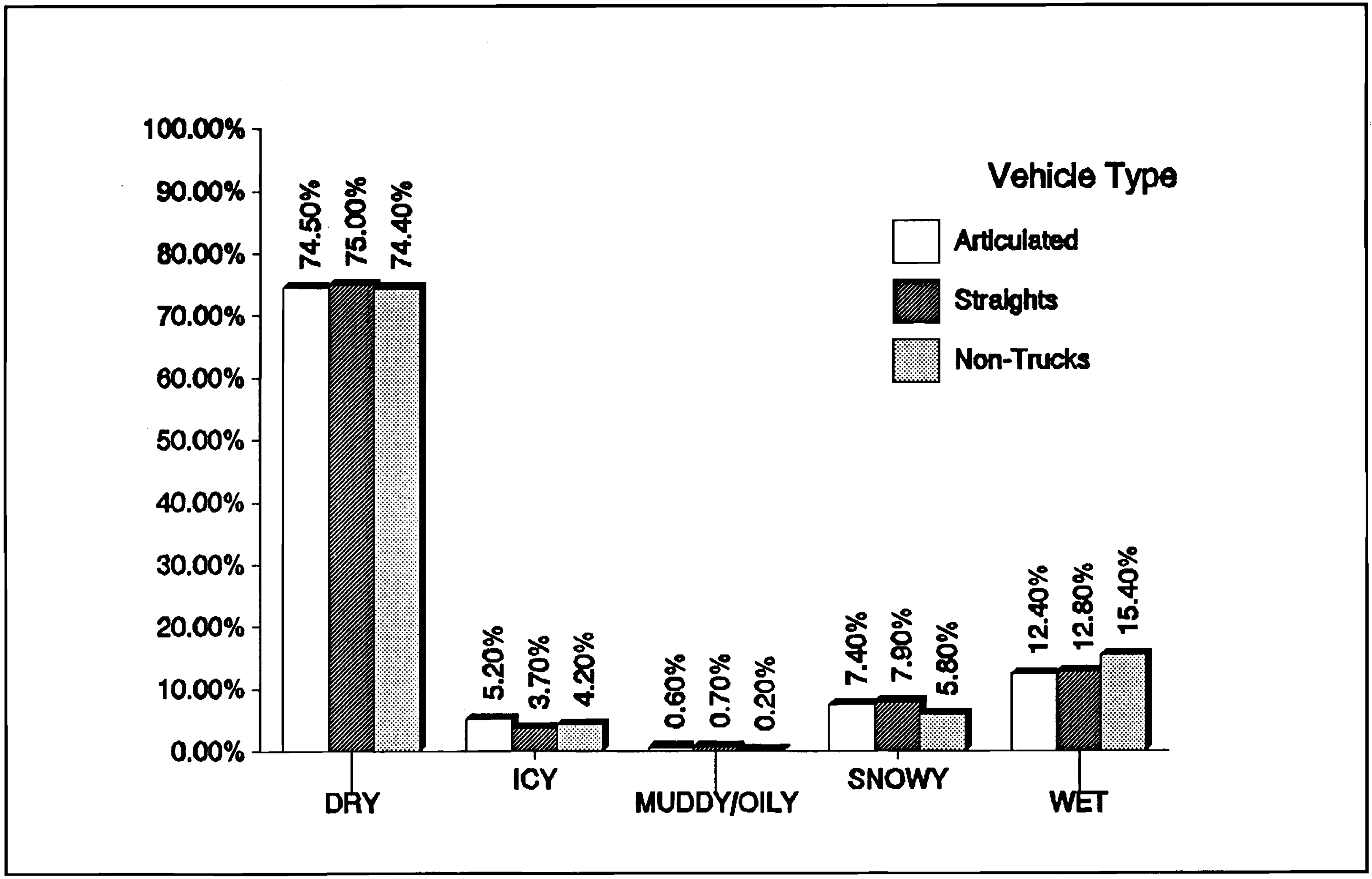


Figure B-9 Road Surface by Vehicle Type (Utah)

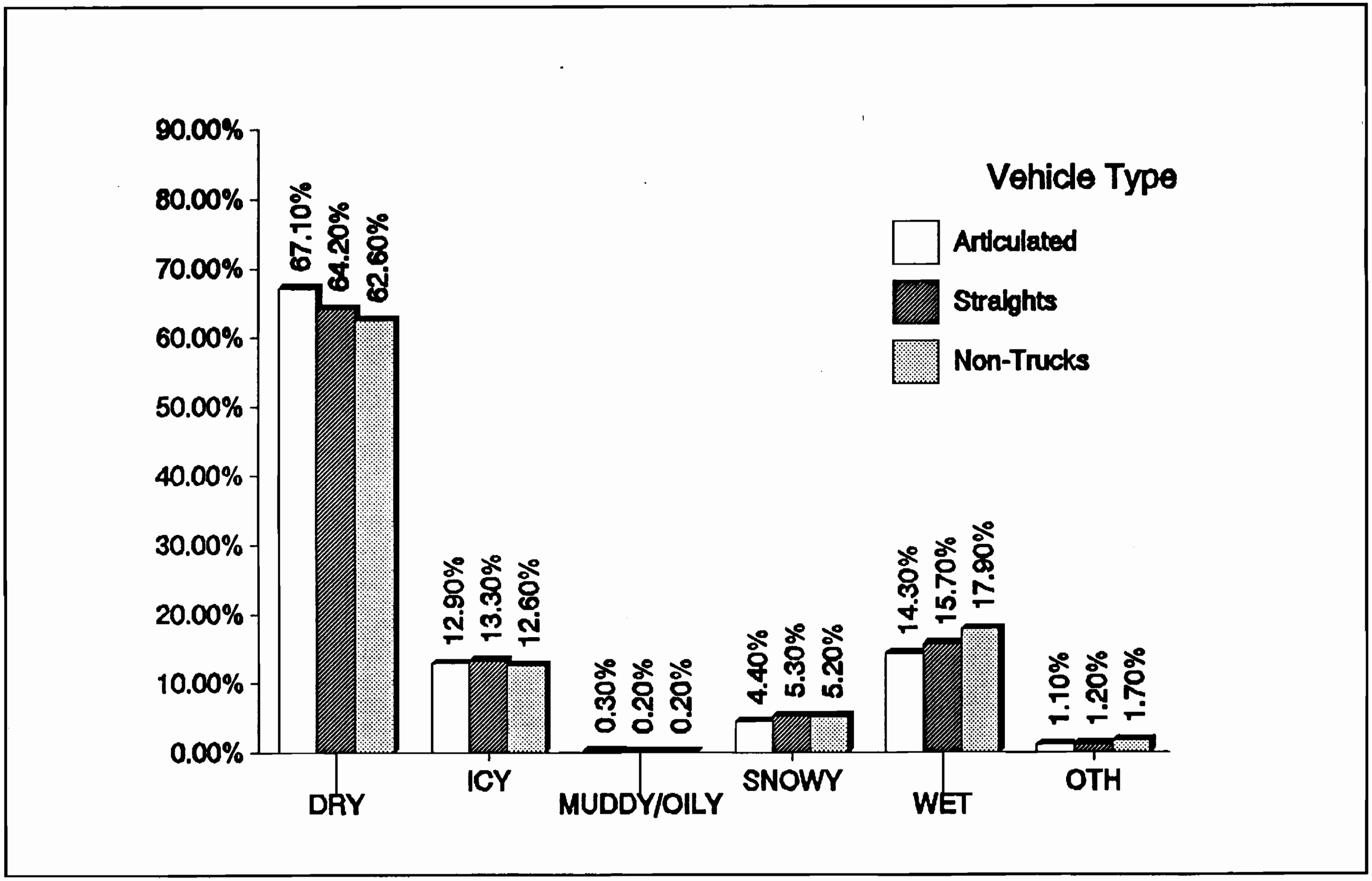


Figure B-10 Road Surface by Vehicle Type (Minnesota)

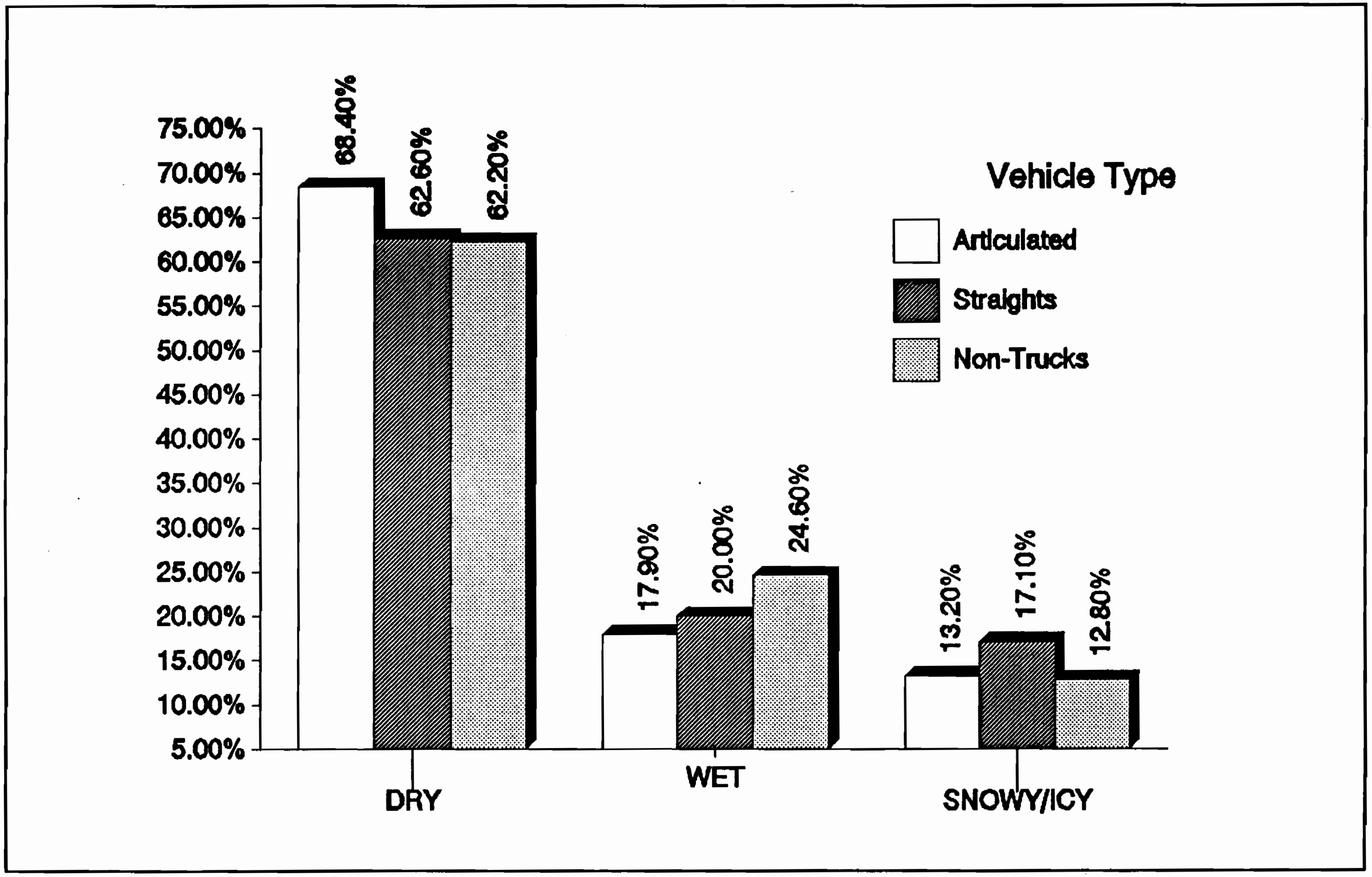


Figure B-11 Road Surface by Vehicle Type (Michigan)

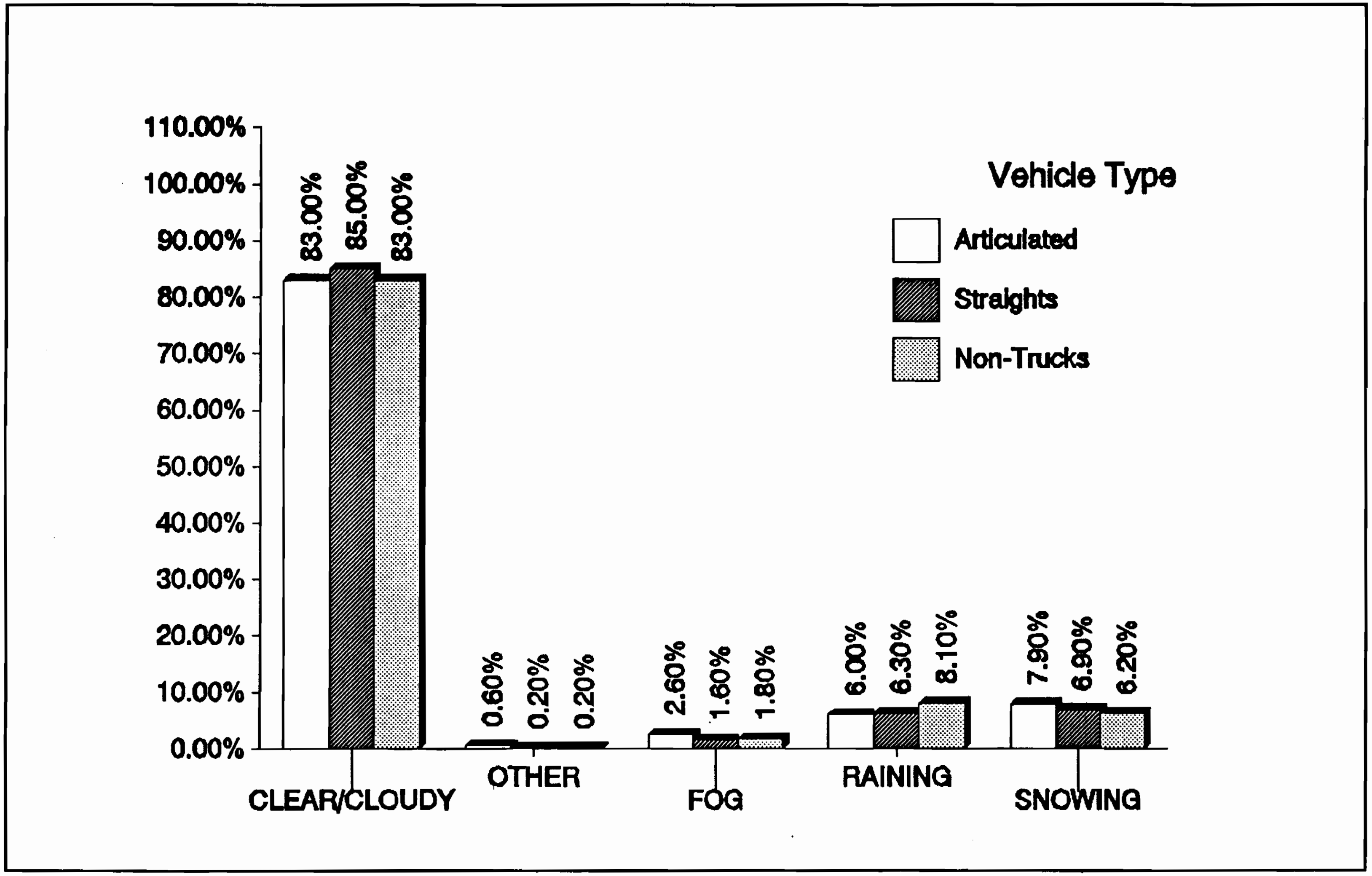


Figure B-12 Weather Condition by Vehicle Type (Utah)

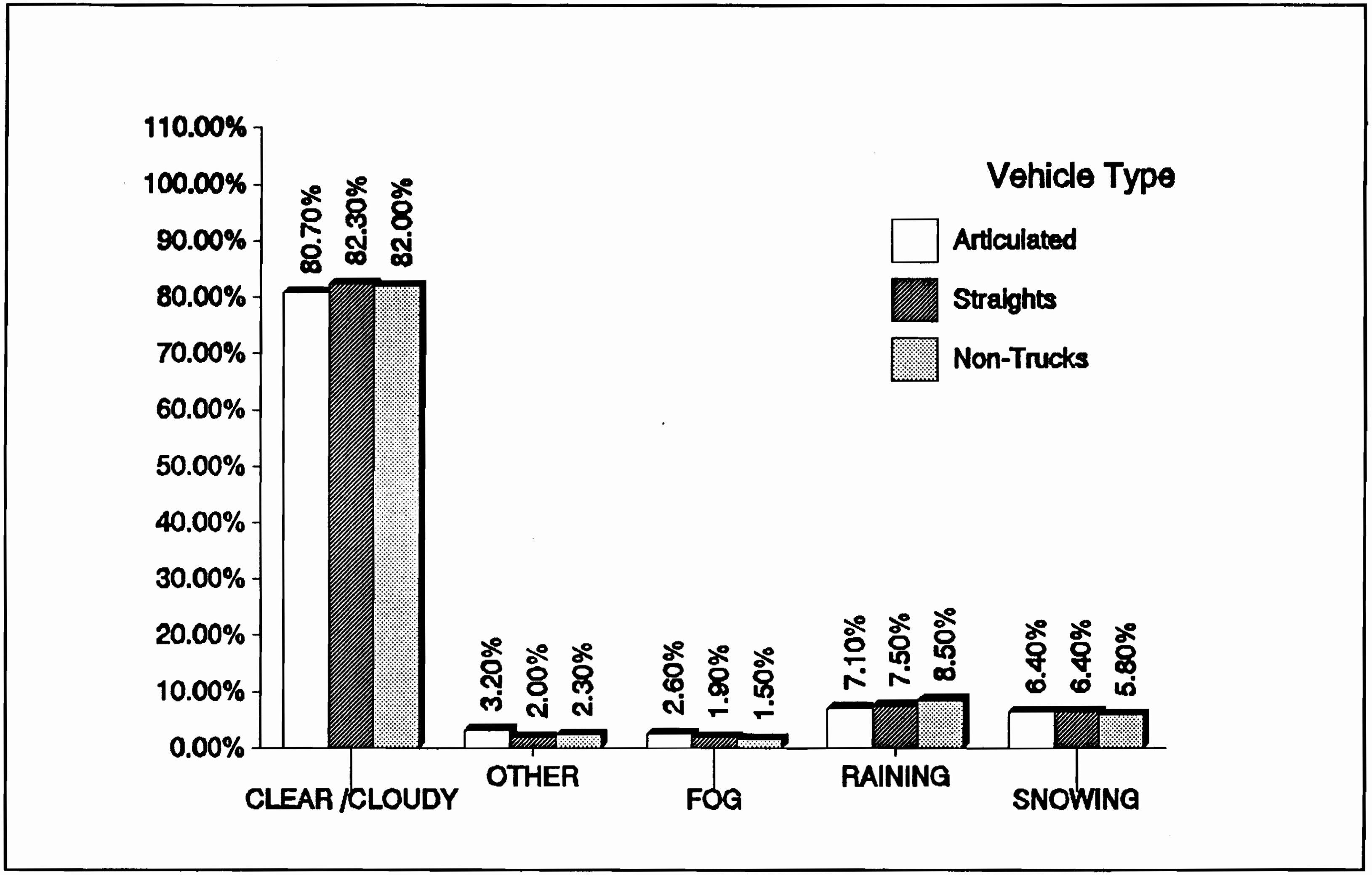


Figure B-13 Weather Condition by Vehicle Type (Minnesota)

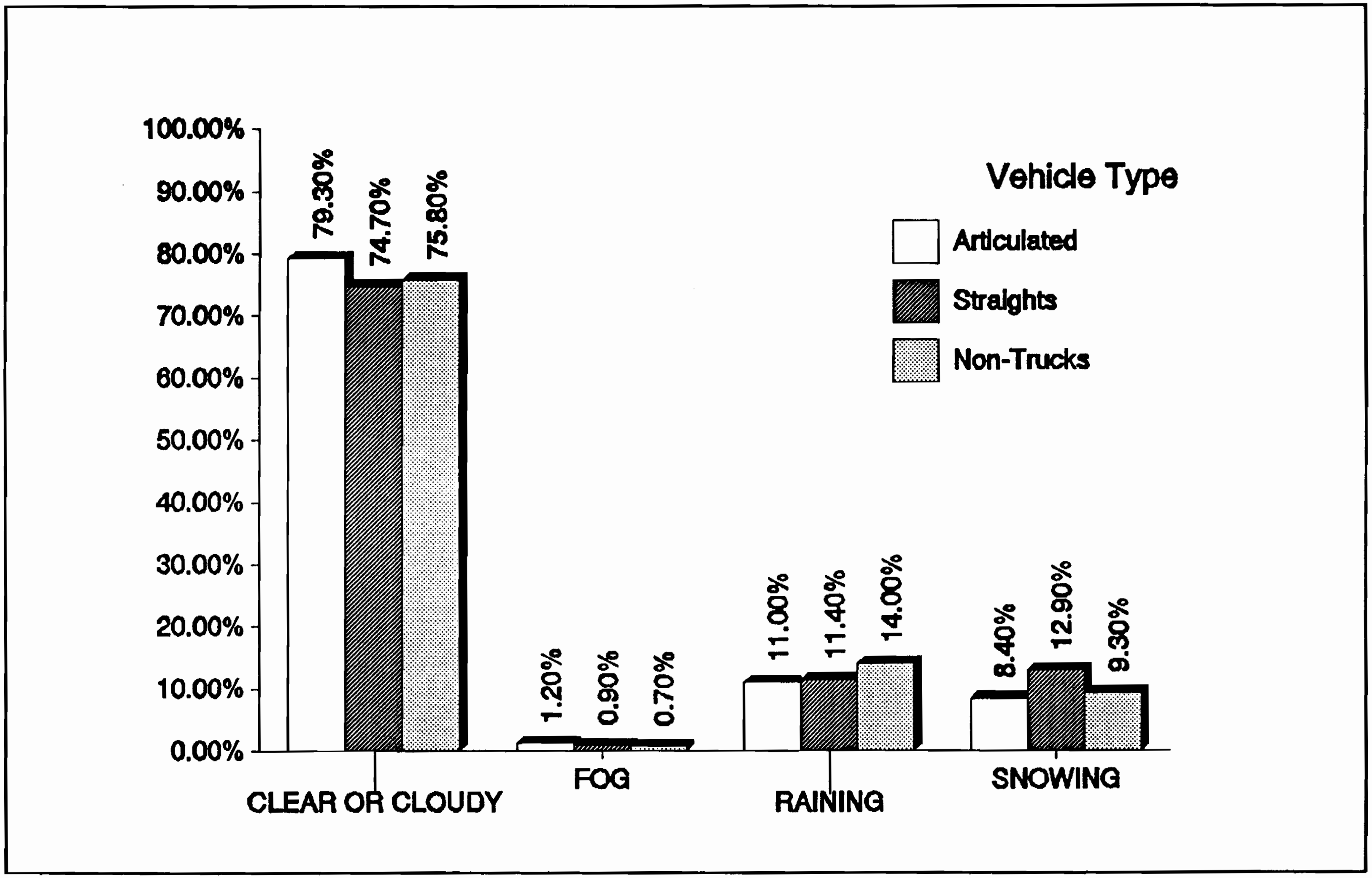


Figure B-14 Weather Condition by Vehicle Type (Michigan)

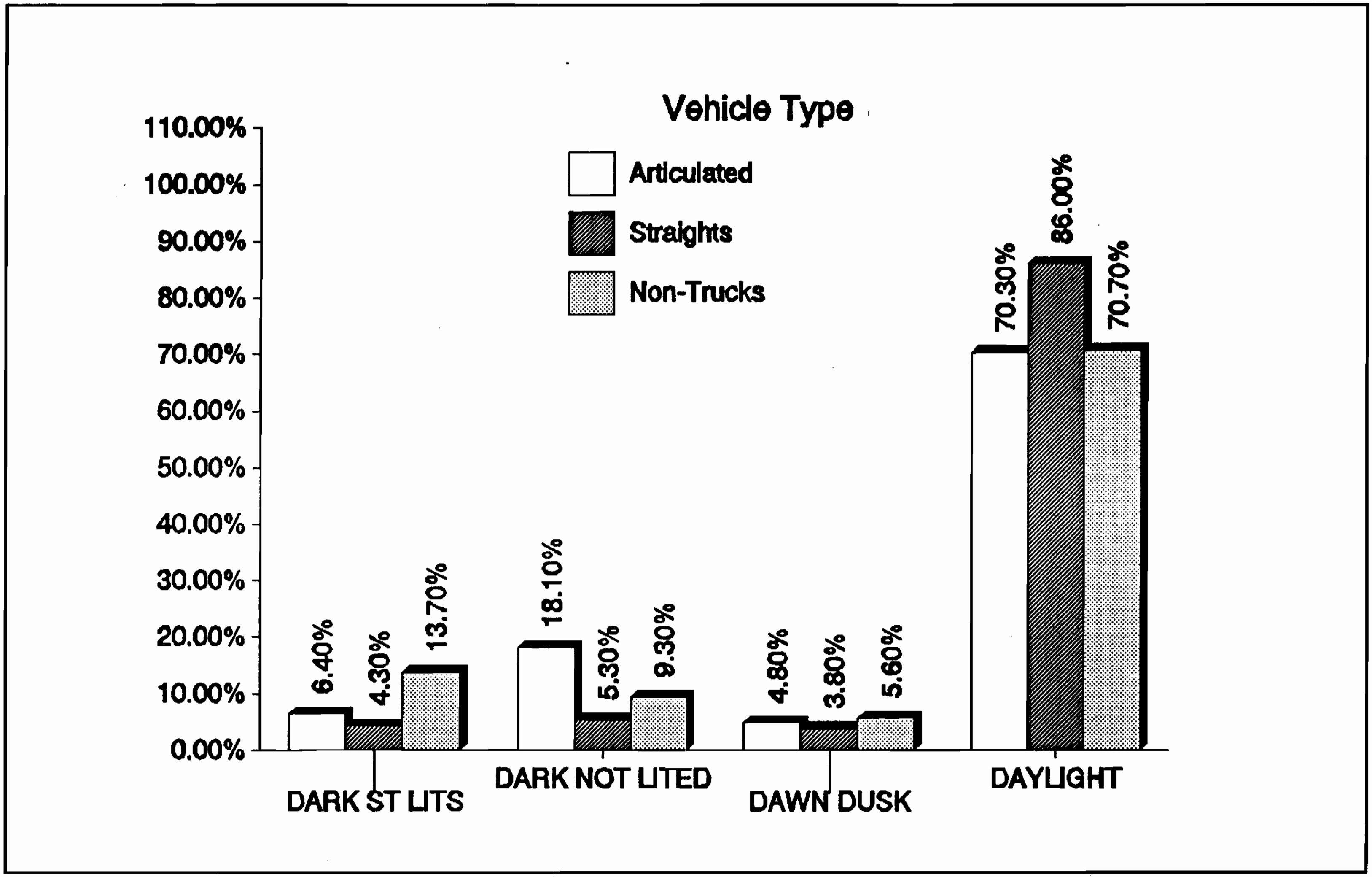


Figure B-15 Light Condition by Vehicle Type (Utah)

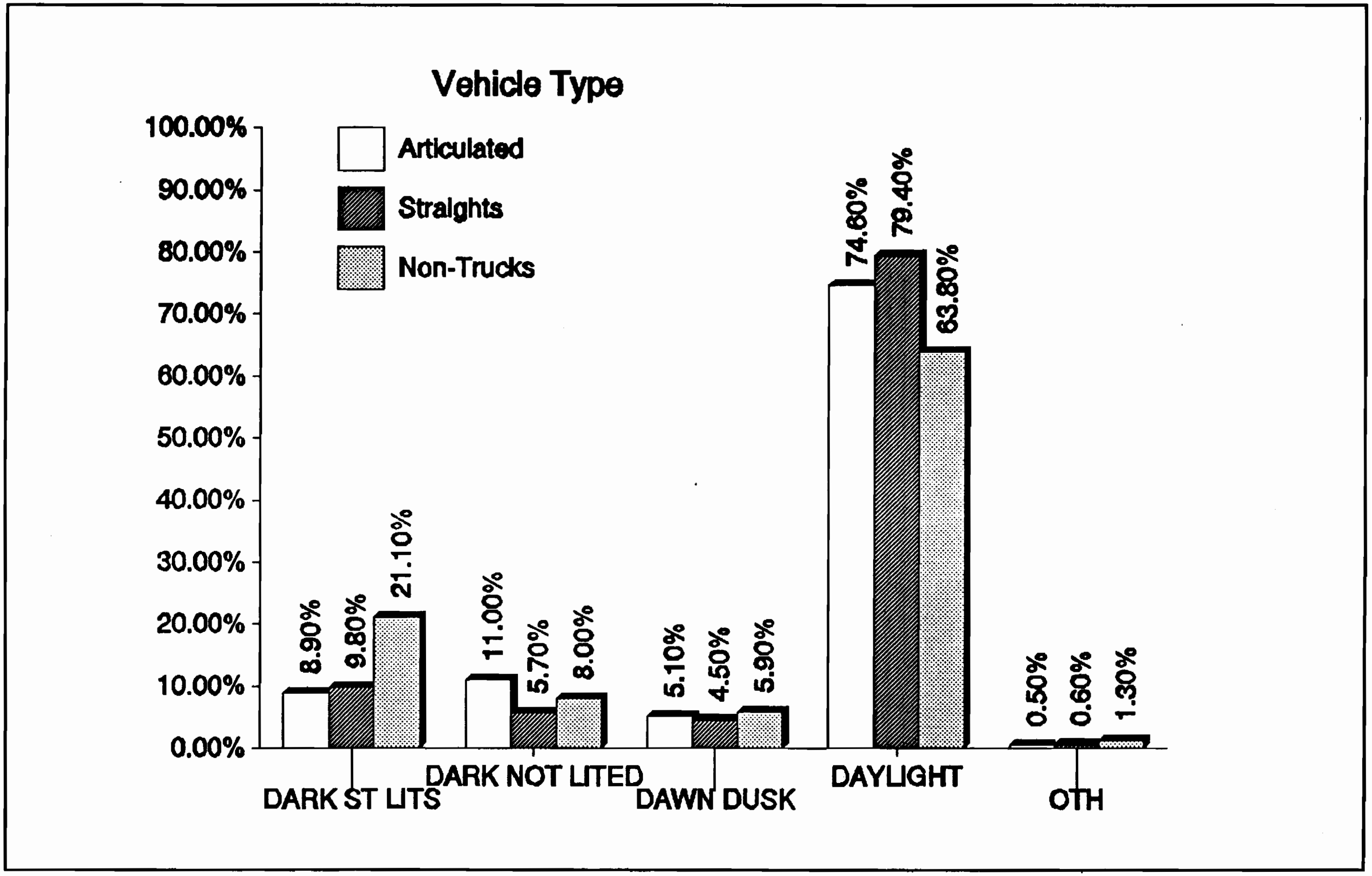


Figure B-16 Light Condition by Vehicle Type (Minnesota)

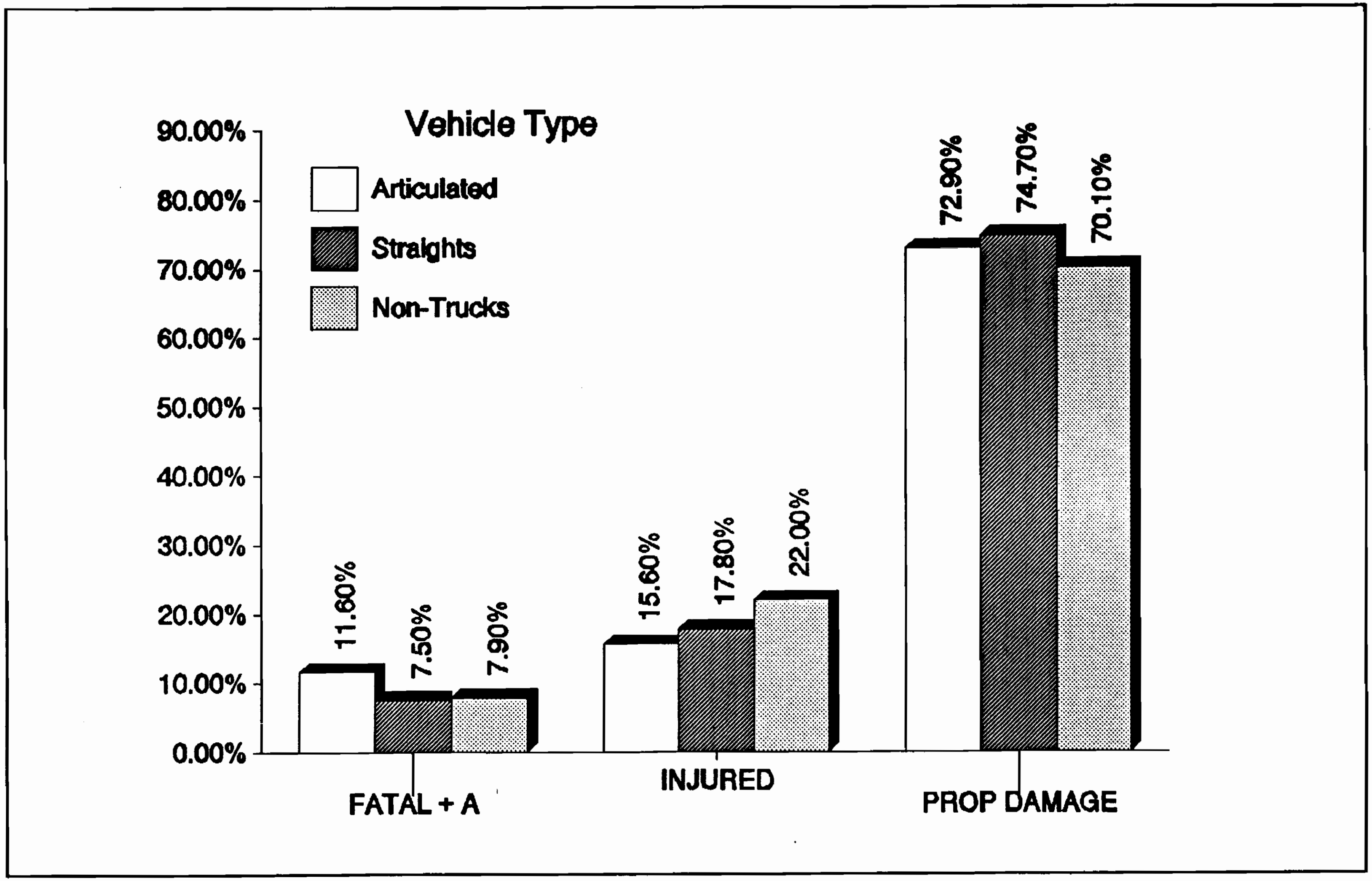


Figure B-17 Accident Severity by Vehicle Type (Utah)

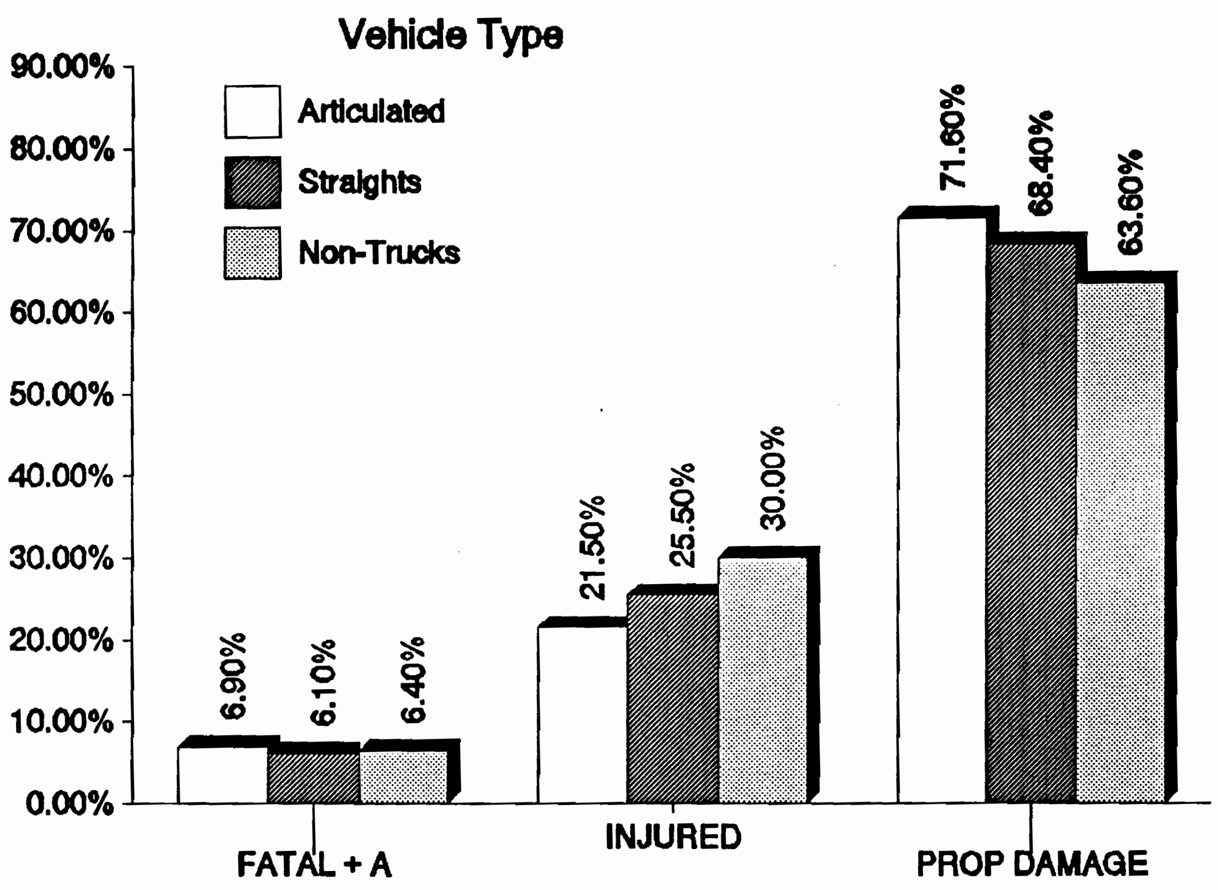


Figure B-18 Accident Severity by vehicle Type (Minnesota)

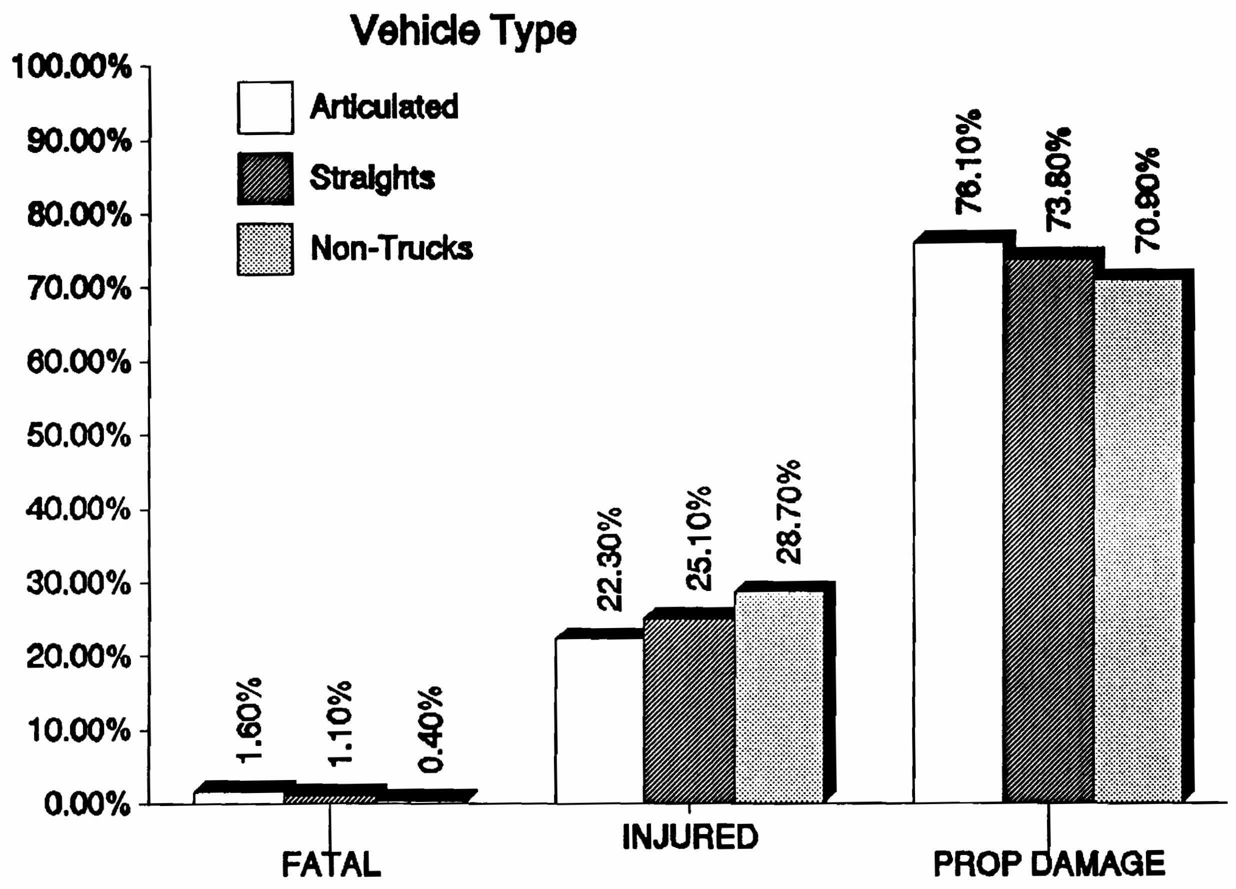


Figure B-19 Accident Severity by Vehicle Type (Michigan)

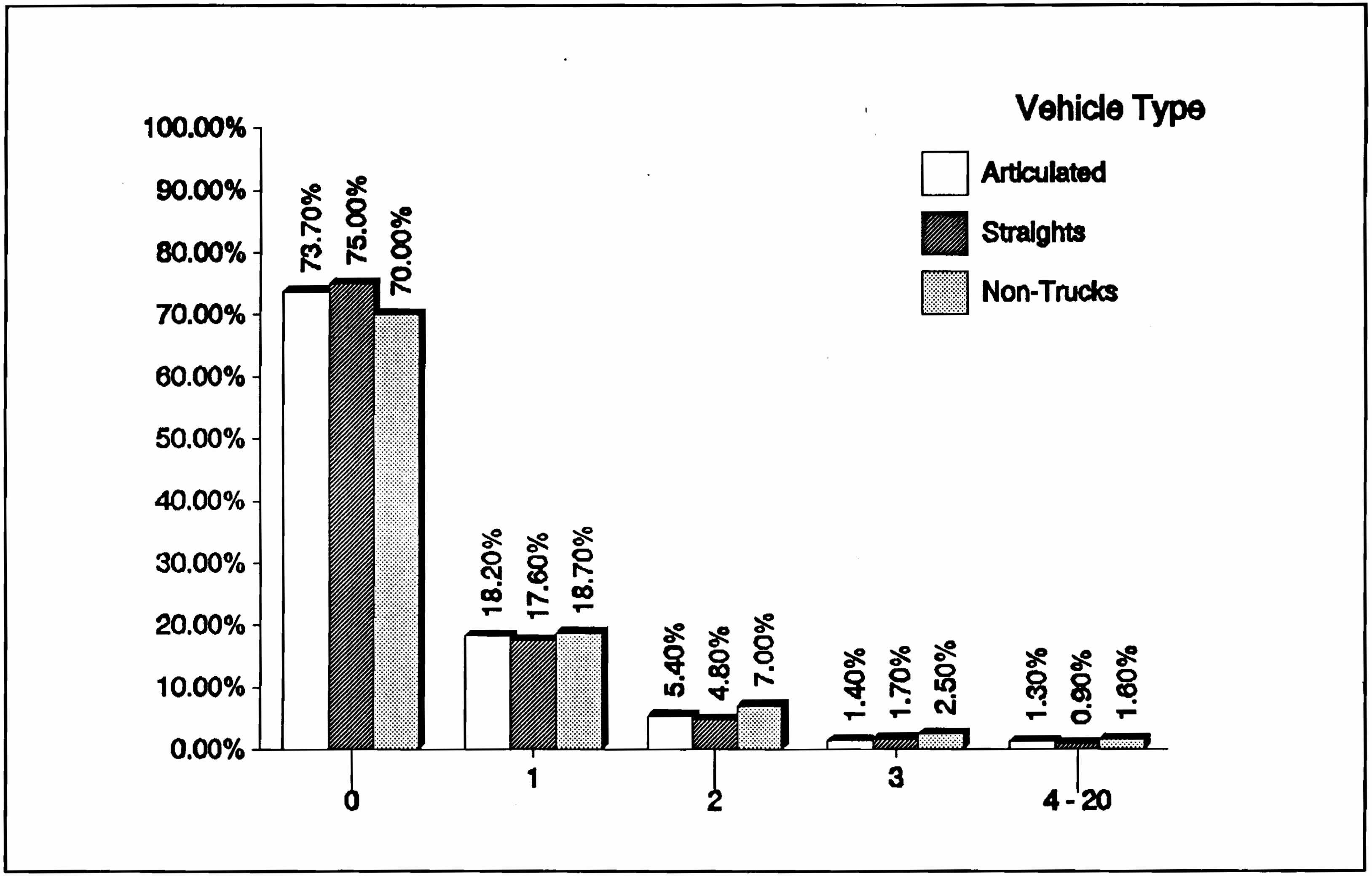


Figure B-20 People Injured by Vehicle Type (Utah)

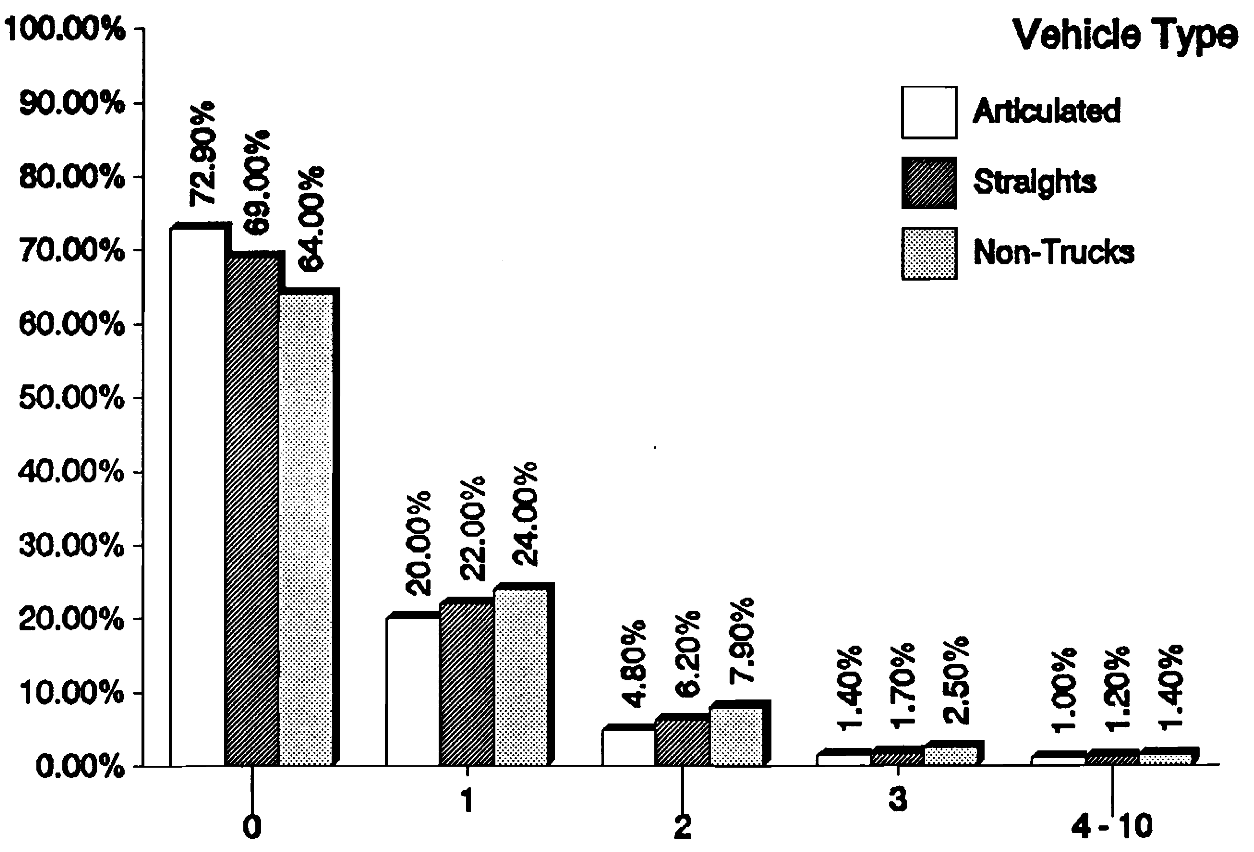


Figure B-21 People Injured by Vehicle Type (Minnesota)

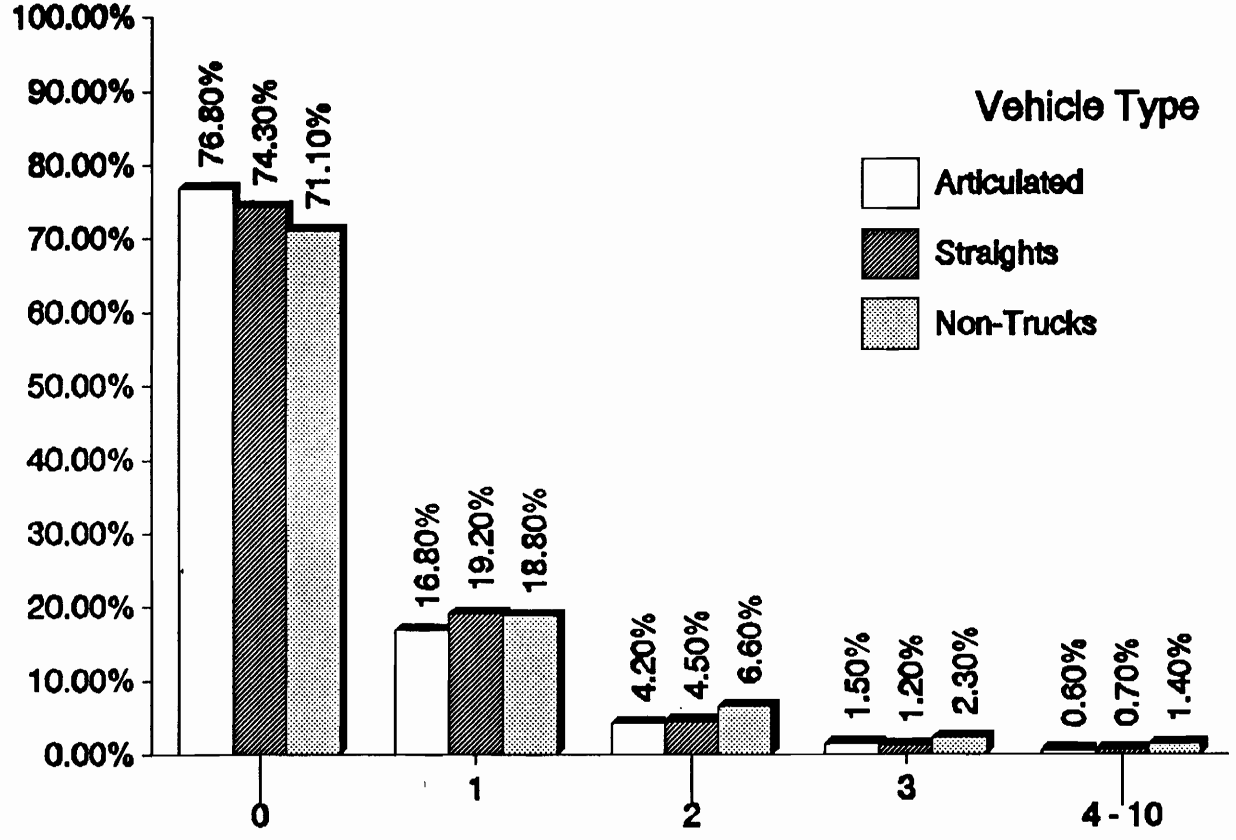


Figure B-22 People Injured by Vehicle Type (Michigan)

Vita

Mr. Yusuf Mohamedshah was born on March 25, 1965 in Bombay, India. He grew up in the neighborhoods of Bombay, the industrial capital of India. He completed his schooling in various parts of the city and finally joined University of Bombay to pursue a career in Engineering. He obtained his Bachelor's degree in Civil Engineering in July 1988. After graduating, he started working as a site engineer with a private contractor in Bombay. His life was going on peacefully, but one fine morning he decided to become a transportation engineer and came to United States and joined the graduate program at Virginia Polytechnic and State University in Fall 1989. Upon his graduation he will work as transportation engineer/planner/researcher in a reputed firm and once again will live in peace.