RELATIONSHIP BETWEEN TRAFFIC FLOW AND SAFETY OF FREEWAYS IN CHINA: A CASE STUDY OF JINGJINTANG FREEWAY

Liande ZHONG, PhD
Associate professor, Certified Safety Engineer,
Research Institute of Highway, Ministry of Transport of P. R. China
8 Xitucheng Rd, Beijing, China, 100088, E-mail: zhongliande@126.com

Xiaoduan SUN. PhD and P.E.
Professor,
Civil Engineering Department, University of Louisiana
Lafayette, LA 70504, E-mail: xsun@louisiana.edu

Submitted to the 3rd International Conference on Road Safety and Simulation, September 14-16, 2011, Indianapolis, USA

ABSTRACT

It is well known that traffic flow characteristics greatly affect roadway safety. Operating speed under various traffic flow rates or densities plays a role in crash occurrence and type of crash. Using crash and traffic flow data from a freeway in China, this paper explores the relationship between traffic flow and safety. In addition to flow rate, the study also reveals the relationship between safety and other traffic flow characteristics such as heavy vehicles, speed gap between types of vehicles, and V/C. It is discovered that the impact of traffic flow on safety is complicated because of the complex relationship among flow parameters. This paper developed several mathematical models that describe the combined effects of traffic flow parameters on crash occurrences, which could enable engineers to identify and implement effective crash countermeasures to improve freeway safety at present.

Keywords: freeway, safety, traffic flow parameter, crash rate.

INTRODUCTION

Due to their high design standard, freeways have somewhat different safety concerns compared to other types of roadways where geometric design elements are strongly related to crash occurrence. Most of the previous studies on freeway safety have shown that the crash rate has a close correlation with traffic flow parameters, but a weak relationship with alignment variables, especially in plain terrain (Lord, 2004; Pesaud, 1991; Zhong, 2008). Crash data and microscopic traffic operation data are crucial in conducting a crash causation analysis. Lack of detailed data (crash and traffic) and data sharing (among different agencies) are mainly blamed for insufficient freeway safety studies. Particularly, the complex relationship between traffic flow characteristics and crash occurrence has not been systematically explored.

Roadway crashes occur when the four elements of a roadway system, drivers, vehicles, road alignment and environment, do not act together logically and consistently. Being the only active member of the system, drivers' safety behavior is largely affected by prevailing traffic characteristics. This paper thoroughly investigates the intrinsic relationship between traffic flow and roadway safety on freeways. The results of the study could shed light on better crash countermeasures through traffic operation and management.

DATA

The information used for this study was from one section of Jingjintang Freeway, which includes crashes, traffic characteristics and geometric design. This section of 4-lane freeway was built on a plain terrain with consistently high design standards such as large radius (the minimum is 5500 meters) and small grade (the maximum is 2.5%). Because of its function (connecting the capital and the major port), this freeway carries a heavy volume and experiences high crash frequencies. There are prominent differences in hourly flow rate and vehicular composition throughout the day. While cars dominate traffic during the day, and the night time traffic mainly consists of heavy vehicles.

The traffic flow data were obtained from the control center of Jingjintang Freeway, derived from loop detectors at 9 locations of the Beijing section. The crash data was collected from the Department of Accidents of Traffic Management Bureau of Beijing. Traffic flow data was collected at one minute intervals, as well as the flow measurements including average vehicle counts, average speed, and lane detector occupancy. In this study, vehicles were divided into two types: small vehicles and trucks. Small vehicles include passenger cars, small vans and light trucks. Trucks are designated for all medium and large vehicles as well as trailers. Table 1 lists the definitions of vehicle types.

Table 1 Vehicle Classification by "Technical Standard of Highway Engineering in China"

Types	Description
Cars	passenger cars with seats not more than 19 or trucks with loads not heavier than 2 tons
Medium vehicle	passenger vehicles with seats more than 19 or trucks with loads between 2 and 7 tons
Large vehicle	Trucks with loads between 7 and 14 tons
Trailer	Trucks with loads heavier than 14 tons

There were a total of 1985 crashes during 2002, 2003 and 2004 in the Beijing section, including 638 property damage only accidents, 1299 injury accidents and 48 fatality accidents. There were 1244 accidents on the freeway's Beijing-bound direction, 719 accidents in the opposite direction, and missing directional information on 11 crash records. To ensure the accuracy of the results, the study considered the directional analysis units (crashes and flow conditions). Furthermore, based on interchanges and loop detector locations, the Beijing section was divided into 4 segments with regards to the existence of entrance and exit ramps, which experience different volumes and vehicle composition.

With the raw traffic data in one-minute intervals, average speed, traffic volume and proportions of small vehicles and trucks within one hour were estimated and used in the analysis. The crashes were also combined into one-hour intervals accordingly. Thus, the traffic parameters are paired with crashes on an hourly basis, which enables further analysis in this study. This type of analysis is feasible only if traffic flow characteristics are similar in different cross sections. Fortunately, such similarity has been proved (Chen, 2008). The crash rate is expressed as the number of hourly crashes per million vehicle kilometers traveled per year.

ANALYSIS AND DISCUSSION

Correlation Analysis and Descriptive Statistics

Crash occurrences are the result of many factors, and the relationship among these factors are complex, thus, independent variables should be carefully chosen to ensure their close correlation with the dependent variable and no or little correlation among themselves. For that purpose, the variables descriptive statistics and correlation analysis were done. As shown in Table 2, the crash rate is the dependent variable and 8 traffic flow parameters are the independent variables. Correlation coefficients vary between -1 and +1. The greater absolute value, the stronger the correlation.

Table 2 Coefficient of correlation analysis

	Trucks speed	Small vehicles speed	Truck volume	Small vehicles volume	Total volume	Speed difference	Truck percentage	V/C	Crash rate
Trucks speed	1.000								
Small vehicles speed	0.833	1.000							
Truck volume	-0.646	-0.416	1.000						
Small vehicles volume	0.001	-0.403	-0.197	1.000					
Total volume	-0.079	-0.462	-0.077	0.993	1.000				
Speed difference	0.074	0.613	0.173	-0.728	-0.719	1.000			
Truck percentage	-0.423	-0.130	0.399	-0.769	-0.733	0.371	1.000		
V/C	-0.178	-0.525	0.076	0.962	0.988	-0.693	-0.672	1.000	
Crash rate	0.149	0.302	-0.149	-0.452	-0.478	**0.332**	**0.436**	**-0.501**	1.000

From the last line of Table 2, it is clear the biggest correlation coefficient (absolute value) is between flow saturation (V/C) and crash rate. The total flow, small vehicle volume, truck percentage as well as the average speed difference between small vehicles and trucks also have relatively close association with crash rate. Since the volume is used in computing crash rate in a linear form, only one variable related to traffic flow could be selected. Based on the correlation coefficient analysis, three independent variables are chosen for the model, i.e., speed difference, truck percentage, and V/C. A data file was created and the descriptive statistical results of independent and dependent variables are shown in Table 3.

Table 3 Result of descriptive statistics

	Speed difference	Trucks percentage	V/C	Crash rate
Mean	18.173	0.380	0.293	1.290
Std. error	0.365	0.019	0.012	0.057
Median	19.020	0.272	0.258	1.127
Std. deviation	5.061	0.263	0.164	0.792
Variance	25.618	0.069	0.027	0.627
Skewness	0.547	-1.140	0.947	5.175
Kurtosis	-0.954	0.564	1.191	1.728
Range	24.327	0.865	0.748	5.081
Minimum	2.563	0.040	0.077	0.136
Maximum	26.890	0.904	0.825	5.217

As shown in Table 3, the average speed difference is 18.2 km/h and the maximum is 26.9 km/h. It has been widely recognized that speed gap between small and large vehicles greatly affects roadway safety. It also can be seen that the 24-hour distribution of traffic composition is uneven, in the daytime, especially the peak hour, cars make the majority, and at night, especially after midnight, trucks primarily dominate the roadway.

Combined Effects of Traffic Data on Crash Rate

Numerous studies establishing statistical links between a single traffic flow parameter and crash occurrence may be found in the literature (e.g. Zhong 2006 and 2007; Golob, 2004). However, these relationships can only be established under certain conditions to some extent. Combined effects of several traffic parameters on crash rate could better explain changes of traffic flow on the safety, which could provide more valuable information in decision-makings for traffic control strategies. Based on the data, the effects of any two variables from variables of speed difference, V/C, and truck percentage were demonstrated in a 3-D setting, as shown in Figures 1, 2 and 3.

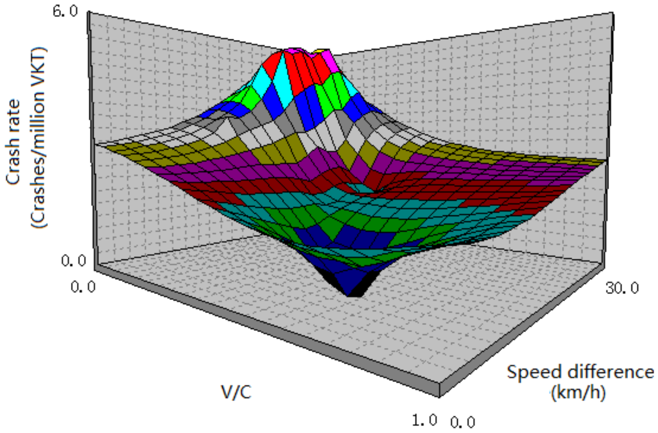

Figure 1 Crash rate vs. V/C and Speed difference

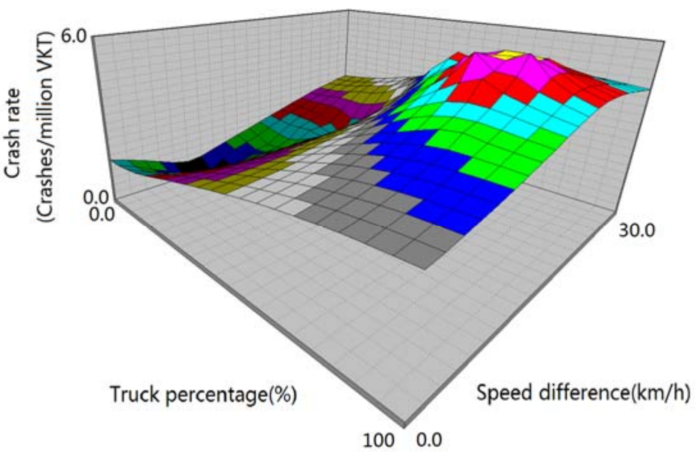

Figure 2 Crash rate vs. Truck percentage and Speed difference

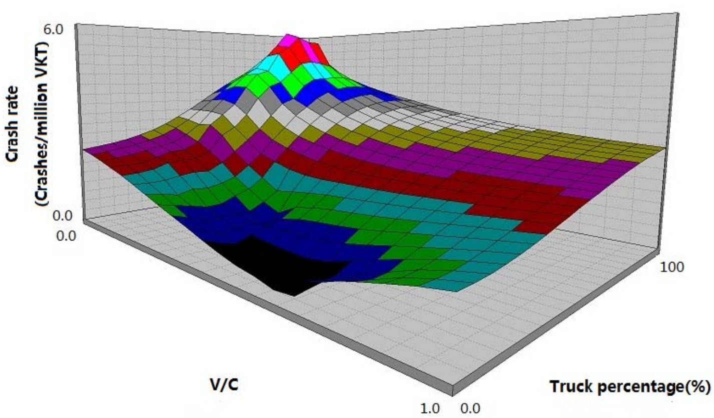

Figure 3 Crash rate vs. V/C and Truck percentage

The results from the three figures above clearly illustrate the great fluctuation of crash rate by speed difference, truck percentage and V/C. The minimum crash rate is 0.14, and the maximum is 5.22. Lower crash rates appear when V/C values are between 0.6 and 0.75, the speed differences are between 6 and 9 km/h, and the percentage of trucks is less than 15%. While higher crash rates occur when V/C values are between 0 and 0.2, the speed differences are between 20 and 27 km/h, and the percentage of trucks are between 75% and 90% of the total traffic.

The impact of a single variable varies depends on the values of other two variables. For example, the relationship between truck percentage and crash rate are different due to different levels of traffic flow. Although crash rate is positively related to percentage of trucks, under a steady flow condition (LOS B and C), crash rate increases insignificantly as trucks percentage increases. However, under free-flow condition (V/C<0.15), the crash rate increases rapidly as the percentage of trucks increases and reaches the maximum when the truck percentage reaches 90%.

Modeling Results

As presented in the above section, there is a complex relationship among crash rate and three selected independent variables. The general linear and nonlinear regression models are not suitable for modeling these complicated relationships. Thus, a multivariate ratio of polynomials method was used by a prior study to establish the mathematical function of traffic flow data and crash rate (Garber, 1991). The multivariate ratio of polynomial procedure is a heuristic process, which searches for the best model to fit the data through hundreds of potential curves. The models developed through this process offer a larger variety of surfaces than the usual polynomial models; however, care must be taken not to use the model outside of the range of the data. In addition, the model should be studied graphically to determine that the model behaves as expected between data points.

This study developed models using the Number Cruncher Statistical Systems (NCSS) software. Both the dependent and independent variables are transformed in 7 forms into the function for searching: $X,\ X^2, 1/X, 1/X^2, \sqrt{X}, 1/\sqrt{X}, \ln(X)$. Since there is no direct R^2 value for non-linear regression, a pseudo R^2 value is used by NCSS. This value is defined by the following form:

$$_pR^2 = \frac{\sum(\bar{y})^2 - \sum(\hat{y}_i)^2}{\sum(y_i)^2 - \sum(\hat{y}_i)^2} \quad (1)$$

Where:

$\hat{y}_i$: model estimates

$\bar{y}$: mean of the observations

y_i : actual observations

The $_{p}R^2$ value is used in the same manner as in a multiple regression, with values closer to 1.0 indicating more explanatory power in the model. Although the value does not have the same definition in traditional sense, it serves well for comparative purposes.

To exploit the combined effects of traffic flow variables on crash rate, speed difference, truck percentage and V/C were again used as independent variables with the crash rate as the dependent variable. Multiple functions were obtained with multivariate ratio of polynomials method considering different variable forms. The process of goodness-of-fit comparison between functions was omitted in this paper. The final the optimal function selected to study the relationship between traffic flow variables and crash rate is listed below.

SQRT(CR)=
(1.813417)-(3.890536E-04)*(Spe_differ*Spe_differ)+(8.85468E-07)*(Spe_differ*Spe_dif
fer)^2-(15.58812)*Truck_Per+(7.466595E-02)*(Spe_differ*Spe_differ)*Truck_Per-(1.45
6214E-04)*(Spe_differ*Spe_differ)^2*Truck_Per+(38.66152)*Truck_Per^2-(.1738746)*(
Spe_differ*Spe_differ)*Truck_Per^2+(2.359875E-04)*(Spe_differ*Spe_differ)^2*Truck_
Per^2+(2.346613)*(LN(VC_Ratio))-(2.113333E-04)*(Spe_differ*Spe_differ)*(LN(VC_R
atio))-(1.177809E-05)*(Spe_differ*Spe_differ)^2*(LN(VC_Ratio))-(26.17044)*Truck_Pe
r*(LN(VC_Ratio))+(9.667372E-02)*(Spe_differ*Spe_differ)*Truck_Per*(LN(VC_Ratio))
-(1.3658E-04)*(Spe_differ*Spe_differ)^2*Truck_Per*(LN(VC_Ratio))+(54.67259)*Truck
_Per^2*(LN(VC_Ratio))-(.2231054)*(Spe_differ*Spe_differ)*Truck_Per^2*(LN(VC_Rati
o))+(2.679846E-04)*(Spe_differ*Spe_differ)^2*Truck_Per^2*(LN(VC_Ratio))+(1.25242
1)*(LN(VC_Ratio))^2-(4.01137E-04)*(Spe_differ*Spe_differ)*(LN(VC_Ratio))^2-(7.6617
27E-06)*(Spe_differ*Spe_differ)^2*(LN(VC_Ratio))^2-(10.1993)*Truck_Per*(LN(VC_R
atio))^2+(3.020683E-02)*(Spe_differ*Spe_differ)*Truck_Per*(LN(VC_Ratio))^2-(2.2974
44E-05)*(Spe_differ*Spe_differ)^2*Truck_Per*(LN(VC_Ratio))^2+(18.68738)*Truck_P
er^2*(LN(VC_Ratio))^2-(6.816921E-02)*(Spe_differ*Spe_differ)*Truck_Per^2*(LN(VC_
Ratio))^2+(6.603523E-05)*(Spe_differ*Spe_differ)^2*Truck_Per^2*(LN(VC_Ratio))^2
(2)

Where:
CR: crash rate
Spe_differ: average operation speed difference between small vehicles and trucks
Truck_Per: truck percentage of traffic
VC_Ratio: traffic saturation (V/C)

It can be seen from the regression function (2), a total of 3 independent variables and 26 coefficients are included. The left side of the function is the crash rate in the square root form, which could avoid the impossible situation that the crash rate acquires a negative value. In the process of searching for the optimal function, the independent variables are transformed into different forms, such as speed difference takes the square form, V/C adopts the logarithmic form, and truck percentage remains in the original form. The $_{p}R^2$=0.511795, and by analyzing residual plots of crash rate corresponding to different variables shows the crash rate relatively scattered when the speed difference is larger and V/C is small.

According to function (2) the crash rate can be computed with the different traffic variables, and the combined effects of traffic flow on safety could be further analyzed. Figure 4, 5 and 6 are the 3-D diagrams depicting the impact of two variables on crash rate. It should be noted that these surface plots do not make use of actual data, but are obtained according to the model functional form with one variable fixed at its average value. This presentation is different from the 3-D charts in Figure 1 to Figure 3.

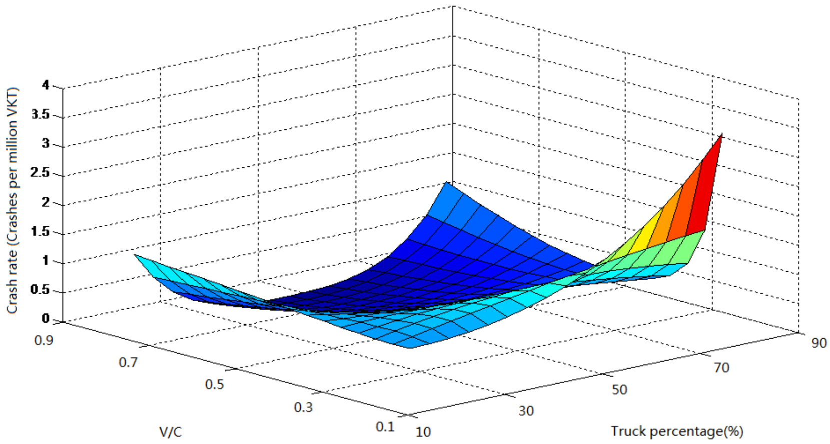

Figure 4 Crash rate vs. V/C and Truck percentage (Speed difference=18.17 km/h)

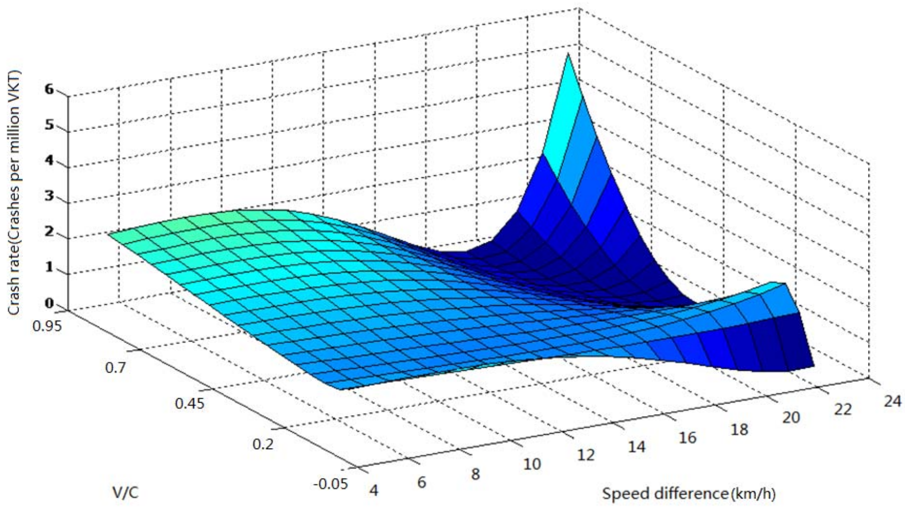

Figure 5 Crash rate vs. V/C and Speed difference (Trucks percentage=38%)

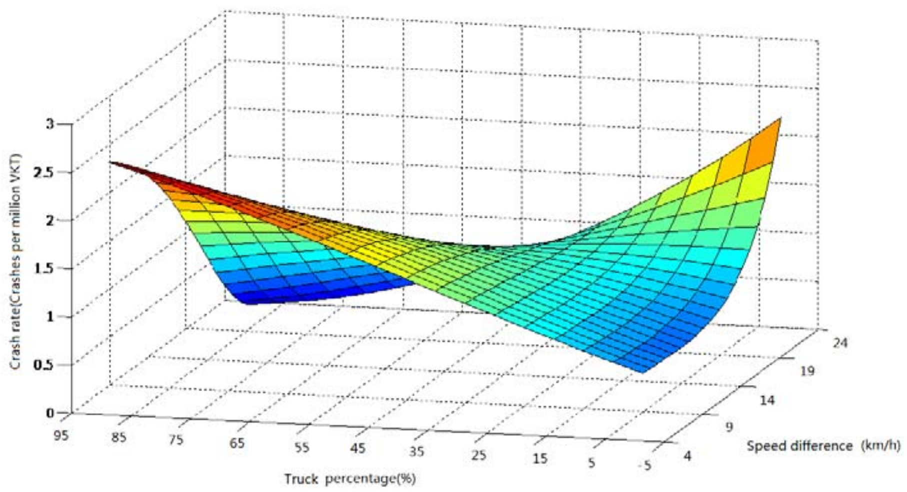

Figure 6 Crash rate vs. Truck percentage and Speed difference (V/C=0.29)

Figure 4, 5 and 6 provide the more detailed explanation. Two of the three traffic flow variables, speed difference, truck percentage and V/C, make a complicated impact on crash rate when the third variable is fixed at its average value. Just as shown in the relationship function, the influence of a single variable on the crash rate varies as one or the other two variables change.

Based on results shown in Figure 6, further analysis was done to illustrate the impact of speed difference. When V/C is at the mean (0.29) and truck percentage is 0%, the relationship between speed difference and crash rate is shown in Figure 7. It can be seen when speed difference is less than 15 km/h, the crash rate is little affected by speed difference, and it remains in a constant state, and when speed difference exceeds 15 km/h and continues to grow, then the crash rate increases by a faster rate.

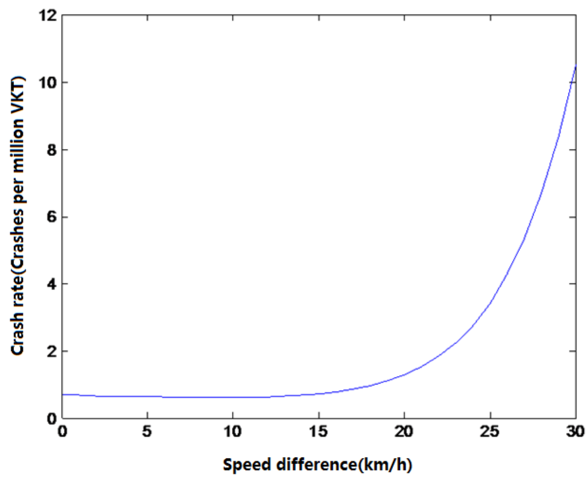

Figure 7 Crash rate vs. Speed difference (V/C=0.29, Truck percentage=0%)

Another relationship between speed difference and crash rate is shown in Figure 8, in which V/C is still 0.29 (unchanged), and trucks make up 90% of the total traffic volume. It is clear that Figure 8 is completely different from Figure 7, although both show the relationship of crash rate versus speed difference. Figure 8 reveals that the crash rate declines with increasing speed difference at first, when the speed difference reaches about 24 km/h the crash rate adopts the lowest value, then as speed difference continues to increase, the crash rate begins to soar.

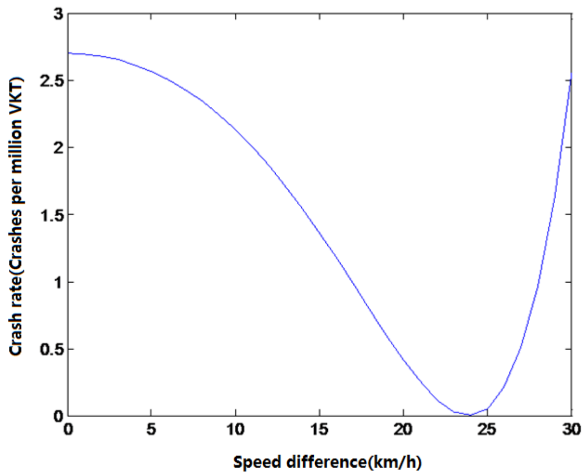

Figure 8 Crash rate vs. Speed difference (V/C=0.29, Truck percentage=90%)

FINAL WORDS

The relationship between traffic flow and safety of freeways in China was studied with the data from Jingjintang Freeway. The results prove that the speed difference, truck percentage and V/C have strong association with crash rate. This paper described the combined effects of traffic flow variables on safety from multiple perspectives. Although the model established with multivariate ratio of polynomials method is complicated and the physical effects of variables are ambiguous, several valuable conclusions are drawn. Due to the interaction among independent variables, more potential parameters contributing to highway crashes should be considered in modeling, and when a specified relationship between single parameter and crash rate is introduced, other variables' background should be clearly stated, otherwise the conclusions are meaningless.

The correctness and applicability of the findings depends on the sample size and representativeness of the data. The microscopic traffic flow data used in this study are generally difficult to obtain. This paper used the data from only one freeway as a case study to illustrate a causation analysis between traffic flow and safety, therefore, the accuracy of the function needs further validation and calibration.

ACKNOWLEDGEMENTS

This research was supported in part by grants from the National Science Foundation of China (Grant No. 50908106）and National Science and Technology Support Program (Grant No. 2009BAG13A03). The contents of this paper reflect the views of the authors who are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the sponsors.

REFERENCE

Chen Y. S., Sun X. D. and Zhong L. D. (2008). Speed Difference and Its Impact on Traffic Safety of One Freeway in China. *Transportation Research Record*, 2038:105-110.

Garber, N. J. and Ehrhart, A. A. (1991). Final Report: *The effect of speed, flow, and geometric characteristics on crash rates for different types of Virginia highways*. Virginia Transportation Research Council, Charlottesville, Virginia.

Golob, T. F. and Recker, W. W. (2004). A method for relating type of crash to traffic flow characteristics on urban freeways. *Transportation Research Part A*, 38(1): 53-80.

Lord, D., Manar, A., Vizioli, A. (2004). Modeling Crash-Flow-Density and Crash-Flow-V/C Ration Relationships for Rural and Urban Freeway Segments. *Presented at the 83rd Annual Meeting of the Transportation Research Board*, Washington, D.C.

Persaud, B N. (1991). Estimating accident potential of Ontario road sections. *Transportation Research Record*, 1327: 47-53.

Zhong L. D. *et al* (2006). Exploring the Relationships between Crash Rates and Average Speed Differentia between Cars and Large Vehicles on a Suburban Freeway. *Proceedings of IEEE ITSC 2006*, September 17-20, Toronto, Canada.

Zhong L. D. *et al* (2007). Research on the Relationship between V/C and Crash Rate on Freeway. *Journal of Beijing University of Technology*, 33(1): 37-40.

Zhong L. D. (2008). *Research on traffic accident prediction model for freeway*. Beijing: College of Architecture and Civil Engineering, Beijing University of Technology.