UCI-ITS-TS-WP-03-3

Freeway Safety as a Function of Traffic Flow

UCI-ITS-TS-WP-03-3

Thomas F. Golob [1]
Will Recker [2]
Veronica M. Alvarez [1]

[1] Institute of Transportation Studies, University of California, Irvine
tgolob@uci.edu, valvarez@uci.edu

[2] Department of Civil and Environmental Engineering and
Institute of Transportation Studies, University of California, Irvine
wwrecker@uci.edu

April 2003

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

Freeway Safety as a Function of Traffic Flow

by

Thomas F. Golob
Institute of Transportation Studies
University of California, Irvine
tgolob@uci.edu

Wilfred W. Recker
Institute of Transportation Studies and
Department of Civil and Environmental Engineering
University of California, Irvine
wwrecker@uci.edu

and

Veronica M. Alvarez
Institute of Transportation Studies
University of California, Irvine
+1.949.824.6571
valvarez@uci.edu

Abstract

In this paper we present evidence of strong relationships between traffic flow conditions and the likelihood of traffic accidents (crashes), by type of crash. Traffic flow variables are measured using standard monitoring devices such as single inductive loop detectors. The key traffic flow elements that affect safety are found to be mean volume and speed, and temporal variations in volume and speed, where variations need to be distinguished by freeway lane. We demonstrate how these relationships can form the basis for a tool that monitors the real-time safety level of traffic flow on an urban freeway. Such a safety performance monitoring tool can also be used in cost-benefit evaluations of projects aimed at mitigating congestion, by comparing the levels of safety of traffic flows patterns before and after project implementation.

Institute of Transportation Studies
University of California, Irvine

April 17, 2003

INTRODUCTION

A common aim of transportation management and control projects on urban freeways is to increase productivity by reducing congestion. Reducing congestion ostensibly leads to reductions in travel time, vehicle emissions and fuel usage, and improved travel time reliability. Tools have been recently implemented to measure the real-time performance of any instrumented section of freeway in terms of throughput: travel time per vehicle, average speed or total delay (Chen, *et al.*, 2001; Choe, Skabardonis, Varaiya, 2002; Varaiya, 2001). The inputs to these tools are typically total flows and mean speeds computed from volume and occupancy data from single inductive loop detectors, typically for intervals of 30-seconds or more. Increasingly, such single loop detectors are distributed throughout the freeway system. Data from more accurate but less ubiquitous sensors, such as double loops and video cameras, is sometimes used to adjust or calibrate single loop measurements, but the primary source of real-time surveillance data for traffic management is likely to remain the single loop detector for the foreseeable future.

Reduced congestion and smoothed traffic flow are also likely to improve safety, as well as reduce psychological stress on drivers. Concentrating on the safety issue, our objective in this paper is to demonstrate that researchers are beginning to understand the relationship between safety and improved traffic flow. Recent developments indicate that the time is right to refine and implement analytical tools that can be used in real-time monitoring of the safety level of the traffic flow on any instrumented section of freeway. As opposed to tools that measure freeway performance in terms of throughput or travel time, we found that the key elements of traffic flow affecting safety are not only mean volume and speed, but also variations in volume and speed. We further determined that it is important to capture variations in speed and flows separately across freeway lanes, and that such information is useful in differentiating types of crashes.

In addition to real-time monitoring of safety levels, a safety performance tool can be used in project evaluation and planning. The safety aspects of costs and benefits can be assessed by comparing the levels of safety estimated by the tool for traffic flows before and after implementation of a treatment, such as a component of an intelligent transportation system (ITS) or infrastructure project. Such a tool can also be used in planning by applying it in forecasting the levels of safety for simulated traffic flows. In the remainder of this paper, we present some evidence that supports relationships between traffic flow and likelihood of traffic accidents (crashes).

PREVIOUS STUDIES

A number of studies have used aggregate measures of traffic flow, such as hourly traffic counts and volume-to-capacity measures, to estimate functional relationships between crash rates per vehicle mile of travel, and traffic flow and speed, conditional upon roadway characteristics (e.g., Aljanahi, *et al.*, 1999; Cedar and Livneh, 1982; Feng, 2001; Frantzeskakis and Iordanis, 1987; Garber and Gadiraju, 1990; Gaudry and Lassarre, 2000; Gwynn, 1967; Gwynn and Baker, 1970; Hall and Pendleton, 1989; Jadaan and Nicholson, 1992; Maher and Summersgill, 1996; Sandhu and Al-Kazily, 1996; Stokes and Mutabazi, 1996; Sullivan, 1990; Sullivan and Hsu, 1988; Turner and Thomas, 1986; and Zhou and Sisiopiku, 1997). Complementary studies have dealt with quantification of the safety component of the marginal costs of roadway use, as a function of traffic speed and density or flow (Dickerson, Peirson and Vickerman, 2000; Jansson, 1994; Johansson, 1996; Jones-Lee, 1990; Newberry, 1988; O'Reilly, *et al.*, 1994; Shefer and Rietveld, 1997; Vickery, 1969; and Vitaliano and Held, 1991). Relationships between aggregate measures of flow and safety are useful mainly in an historical context, as they are not very useful for real-time monitoring of precursors of safety. In addition, Mensah and Hauer (1998) point out two fundamental problems with using such aggregate data: an "argument averaging" problem, caused by using aggregate traffic flow data, rather than data measuring traffic conditions at the time of the crash, and a "function" averaging problem, caused by using the same functional relationship for all types of crashes under all conditions.

In the last few years, studies have uncovered results that demonstrate how the likelihood of different types of crashes varies with patterns of traffic flow prevailing at the time of the crash. These studies (Golob and Recker, 2003a; 2003b; Golob, Recker and Alvarez, 2002; 2003; Lee, Saccomanno and Hellinga, 2002; Lee, Hellinga and Saccomanno, 2003; Oh, Oh and Chang, 2001; and Oh, *et al.,* 2001; Oh, Oh, Ritchie and Chang, 2001) all use archived data from traffic monitoring devices, combined with historical crash records to describing the traffic flow conditions that prevailed just prior to the time of each crash. By using lane-by-lane traffic flow data measured on short time intervals (e.g., 20 or 30 seconds), these investigators were able to relate expected numbers of crashes by type of crash to traffic flow in terms of central tendencies *and* variations in volumes, densities, and speed, potentially differentiated across freeway lanes. These studies lay the groundwork necessary for implementing a real-time safety performance evaluation tool that uses readily available traffic monitoring data.

THE FITS (Flow Impacts on Traffic Safety) PROTOTYPE

In the remainder of this paper we will describe how a real-time safety monitoring tool might emerge based on recent results generated through testing of a prototype software tool called FITS (Flow Impacts on Traffic Safety) (Golob, Recker, and Alvarez, 2002). FITS uses a data stream of 30-second observations from single inductance loop detectors to forecast the types of crashes that are most likely to occur for the flow conditions being monitored. The FITS algorithms, in their present form, are based on analyses of crash characteristics of more than 1,000 crashes on six major freeways in Orange County California in 1998 as a function of traffic flow conditions for a thirty-minute time period immediately preceding the crashes. Orange County is an urban area of about three million population located between Los Angeles and San Diego.

Data

FITS was calibrated based on crash data for 1998 drawn from the Traffic Accident Surveillance and Analysis System (TASAS) database (Caltrans, 1993), which covers all police-reported, on the California State Highway System. Crash typology is defined according to three primary crash characteristics: (1) crash type, based on the type of collision (rear end, sideswipe, or hit object), the number of vehicles involved, and the movement of these vehicles prior to the crash, (2) the crash location, based on the location of the primary collision (e.g., left lane, interior lanes, right lane, right shoulder area, off-road beyond right shoulder area), and (3) crash severity, in terms of injuries and fatalities per vehicle. These variables are described in Table 1, together with their breakdown for the data on which the current version of the tool is calibrated.

Table 1 Crash Characteristic with Breakdown of Sample of 1998 Crashes on Dry Orange County Freeways (N = 1192)

Crash Characteristic	Percent	Crash Characteristic	Percent
Crash type		Crash Location	
Single vehicle hit object or overturn	14.2	Off-road, driver's left	13.8
Multiple vehicle hit object or overturn	5.9	Left lane	25.8
Two-vehicle weaving crash [a]	19.3	Interior lane(s)	32.7
Three-or-more-vehicle weaving crash [a]	5.5	Right lane	19.3
Two-vehicle straight-on rear end	33.8	Off road, driver's right	8.3
Three-or-more-vehicle rear end	21.3		
		Severity	
		Property damage only	71.9
		Injury or fatality	28.1

[a] Sideswipe or rear end crash involving lane change or other turning maneuver

To relate these characteristics to traffic flow conditions, FITS uses raw detector data that provide information on two variables: count and occupancy for each thirty-second interval. Although these two variables can be used (under very restrictive assumptions of uniform speed and average vehicle length, and taking into account the physical installation of each loop) to infer estimates of point speeds, we avoid making any such assumptions, and use only these direct measurements in the analyses. However, in interpreting the results, where such relative terms as means and variances are employed, we routinely assume that the ratio flow/occupancy is proportional to and a surrogate for mean speed.

Four blocks of three variables (one variable for each of the three lanes: left, interior, and right) were found to be related to crash typology. The variables are listed in Table 2. The first block comprises the median of the ratio of volume to occupancy for each of the three lanes, and measures the central tendency of (density), an approximate proportional indicator of space mean speed. Median, rather than mean, is used in order to avoid the influence of outlying observations that can be due to failure of the loop detectors or unusual vehicle mixes. The second block comprises the difference of the 90^{th} percentile and 50^{th} percentile in the ratio of volume to occupancy (density) for each lane, and represents the temporal variation of this ratio. Here we use the percentile differences because we wish to minimize the influence of outlying observations.

Table 2 Traffic Flow Variables

Block 1 Central tendency of speed	Median volume/occupancy - left lane Median volume/occupancy - interior lane Median volume/occupancy - right lane
Block 2 Variation in speed	Difference between 90^{th} and 50^{th} percentiles of volume/occupancy - left lane Difference between 90^{th} and 50^{th} percentiles of volume/occupancy - interior lane Difference between 90^{th} and 50^{th} percentiles of volume/occupancy - right lane
Block 3 Central tendency of volume	Mean volume - left lane Mean volume - interior lane Mean volume - right lane
Block 4 Variation in volume	Standard deviation of volume - left lane Standard deviation of volume - interior lane Standard deviation of volume - right lane

The third block of traffic flow variables comprises the mean volumes for all three lanes taken over the entire 27.5-minute period preceding the accident. Volume alone is not as sensitive to outliers as the ratio of volume to occupancy is, so mean, rather than median, is used as a measure of central tendency. (4) Finally, the fourth block is composed of the standard deviations of the 30-second volumes for all three lanes as a measure of variation in volume over the 27.5-minute period.

In order to reduce any effects of multicollinearity among the traffic flow measures (particularly among the three variables in each of the four blocks), principal components analysis was applied to extract a sufficient number of factors to identify independent "composite" traffic flow variables while simultaneously discarding as little of the information in the original variables as possible. A reduction from twelve original variables to six factors accounted resulted in a loss of only about 13% of the variance in the original twelve variables. One variable, highly correlated with the factor, was then selected to represent each of the six factors and used as input to FITS. These six variables are listed in Table 3, together with the factor that they represent.

Table 3 Loop Detector Variables Used to as Input to the FITS Tool

Specific Traffic Flow Variable	Flow Factor Represented
Median volume/occupancy interior lane	Central tendency of speed - all lanes
90^{th}%tile - 50^{th}%tile of volume/occupancy interior lane	Variation in speed – all lanes but right
90^{th}%tile - 50^{th}%tile of volume/occupancy right lane	Variation in speed – right lane
Mean volume left lane	Central tendency of flow – all lanes
Standard deviation of volume interior lane	Variation in flow – all lanes but right
Standard deviation of volume right lane	Variation in flow – right lane

Determining Traffic Flow Regimes According to Differences in Crash Typology

Calibration of FITS using the limited 1998 data was based upon application of a series of multivariate statistical methods that determine optimal patterns between crash rates by type of crash and traffic flow characteristics (Golob and Recker, 2003b). Two of these methods used are well known: (a) principal components analysis, the most common form of factor analysis, and (b) cluster analysis. Principal components analysis was used to eliminate problems with redundancy among traffic flow variables by reducing the dataset to a smaller number of variables with minimum loss of information. Cluster analysis is a method of grouping observations based on similar data structure. In the calibration process, cluster analysis was used to find homogenous groups of traffic flow conditions, which are called "traffic flow regimes."

A perplexing problem in cluster analysis is to determine the best number of "natural" clusters. In the calibration, we used a unique method for finding the optimal number of clusters by comparing how well each clustering solution for traffic flow regimes explains crash typology. To measure the strengths of the relationships between different clustering solutions and crash characteristics, we employed a third type of multivariate analysis: nonlinear (nonparametric) canonical correlation analysis (NLCCA).

Because it is not commonly used in transportation research, NLCCA needs some explanation. Conventional linear canonical correlation analysis (CCA) can be viewed as an expansion of regression analysis to more than one dependent variable; there are two sets of variables, and the objective is to find a linear combination of the variables in each set so that the correlation between the linear combinations is as high as possible. The linear combinations are defined by optimal variable weights. Depending on the number of variables in each set and their scale types, further linear combinations (canonical variates, similar to principal components in factor analysis) can be found that have maximum correlations subject to the conditions that all canonical variates are mutually independent. Nonparametric, or nonlinear CCA is designed for problems with variable sets that contain categorical or ordinal (nonlinear, or nonparametric) variables. The linear combinations can be defined only when there is a metric to quantify the categories of each nonlinear variable. NLCCA simultaneously determines both (1) optimal re-scaling of the categories of all categorical and ordinal variables and (2) component loadings (variable weights), such that the linear combination of the weighted re-scaled variables in one set has the maximum possible correlation with the linear combination of weighted re-scaled variables in the second set. The NLCCA method we use is based on the alternating least squares (ALS) algorithm, which is described in detail in De Leeuw (1985), Gifi (1990), Michailidis and de Leeuw (1998), Van der Burg (1988), van Buren and Heiser (1989) and van der Boon, 1996). In ALS both the variable weights and optimal category scores are determined by minimizing a meet-loss function derived from lattice theory. The solution is a particular kind of singular decomposition (eigenvalue) problem (Israëls, 1987).

Based on results that demonstrated how safety patterns were affected by weather and lighting conditions (Golob and Recker, 2003a), the FITS tool has been calibrated for three different environmental segments: (1) daylight and dusk-dawn conditions on dry roads, (2) nighttime conditions on dry roads, and (3) wet roads under all lighting conditions (Golob, Recker and Alvarez, 2002). For demonstration purposes, we report only on FITS results for the largest segment: daylight and dusk-dawn conditions on dry roads.

FITS Calibration for Daylight and Dry Road Conditions

Using 1998 data, cluster analyses were performed in the space of the six principal traffic flow variables in Table 3 in order to establish relatively homogenous traffic flow regimes. The objective is to determine the best grouping of observations into a specified number of clusters, such that the pooled within groups variance is as small as possible compared to the between group variance given by the distances between the cluster centers. The criteria used to select the optimal number of clusters involved how well each of the clustering explained differences in crash typology, as determined by the performance of each clustering scheme using NLCCA. For prevailing traffic conditions for crashes on dry roads during daylight in 1998, we found that we needed eight clusters of traffic flow conditions, which we called "Regimes." The eight traffic flow regimes are defined based on the location of their cluster centers in the six-dimensional space of the traffic flow variables.

The eight Traffic Flow Regimes can be visually compared using radar diagrams that display all dimensions simultaneously. The dimensions are standardized (origin set at system mean, and scale in standard deviation units) for easy comparison among the dimensions. Figure 1 is a key to the radar diagrams. Using compass orientation, mean speed is measured on the north axis; mean flow is measured on the opposing south axis; speed variances are on the two east axes; and flow variances are on the two west axes. Variations on the right lane are measured on the opposing northwest and southeast axes; and variations on all the other lanes are measured on the opposing southeast and northeast axes.

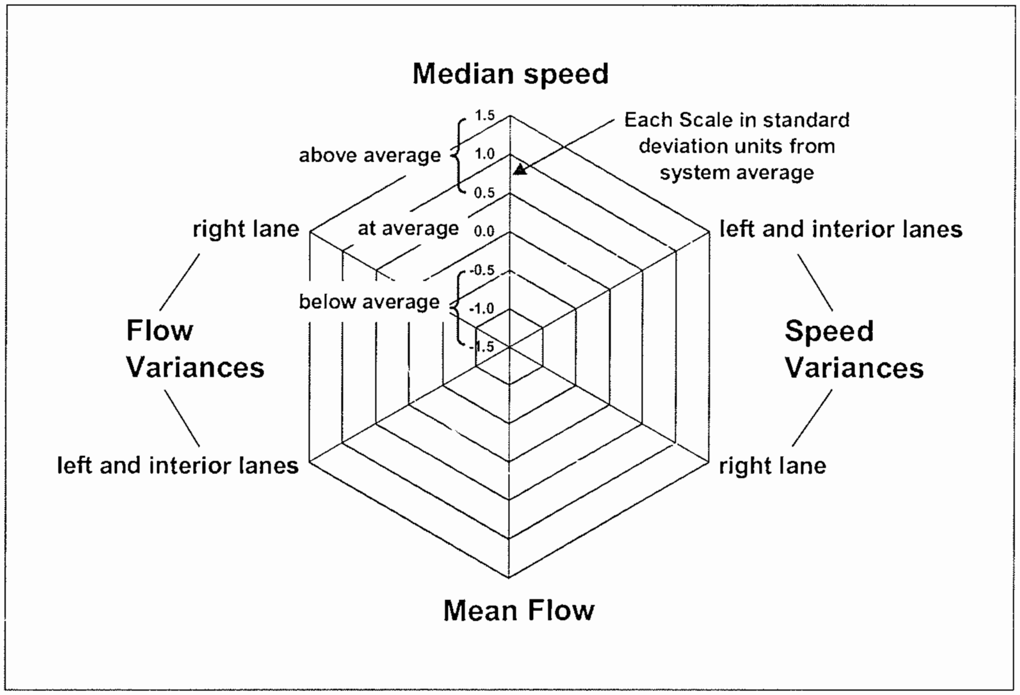

Figure 1. Key to Radar Diagram Used to Describe Traffic Flow Regimes in Terms of Centroid Locations for Six Traffic Flow Variables

The eight dry-daylight Regimes are graphed in Figures 2a and 2b. They are numbered in order of increasing demand for road space, as described in the next Section.

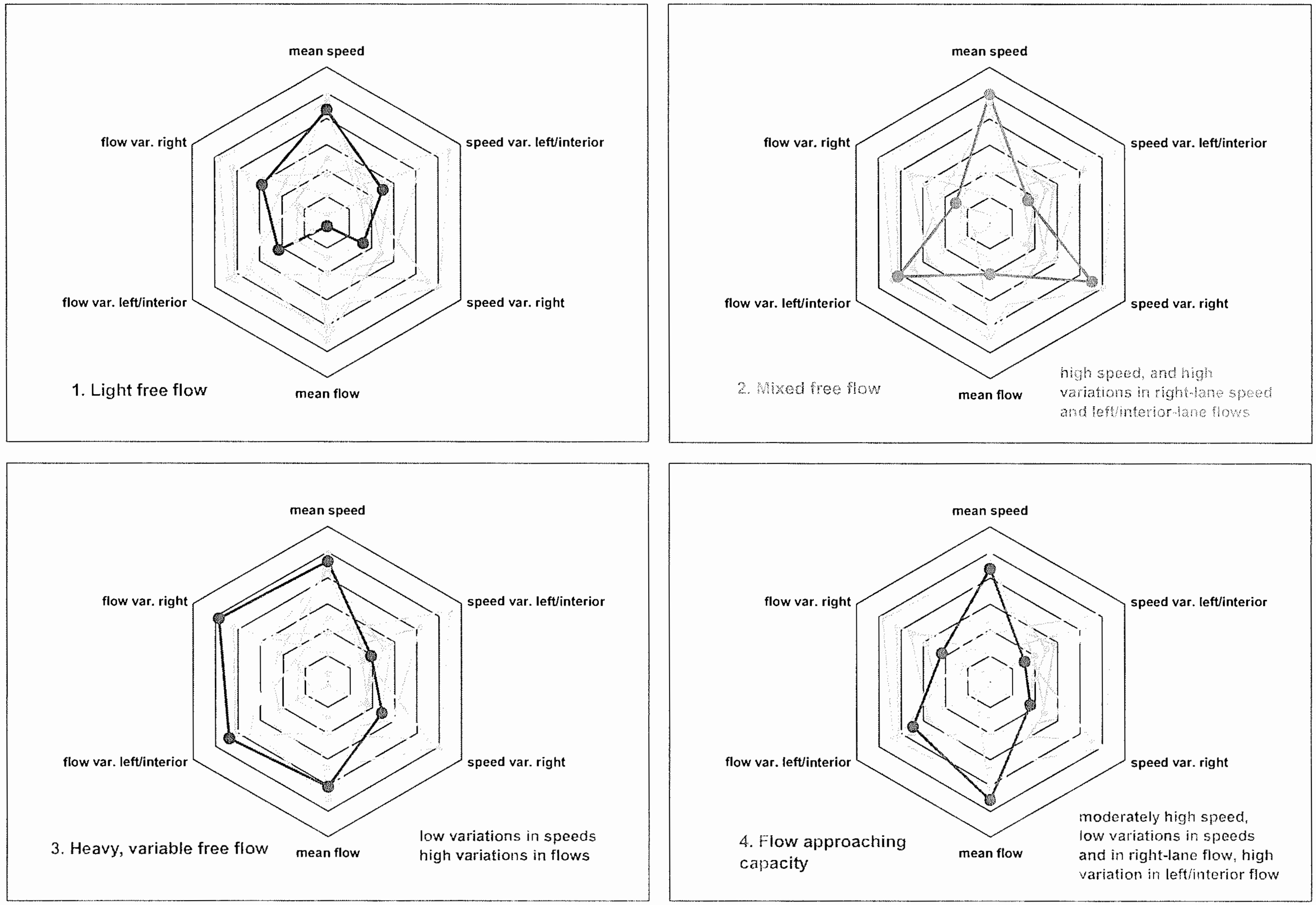

Figure 2a Radar Diagrams of Traffic Flow Regimes for Daylight and Dry Road Conditions (Regimes 1 through 4)

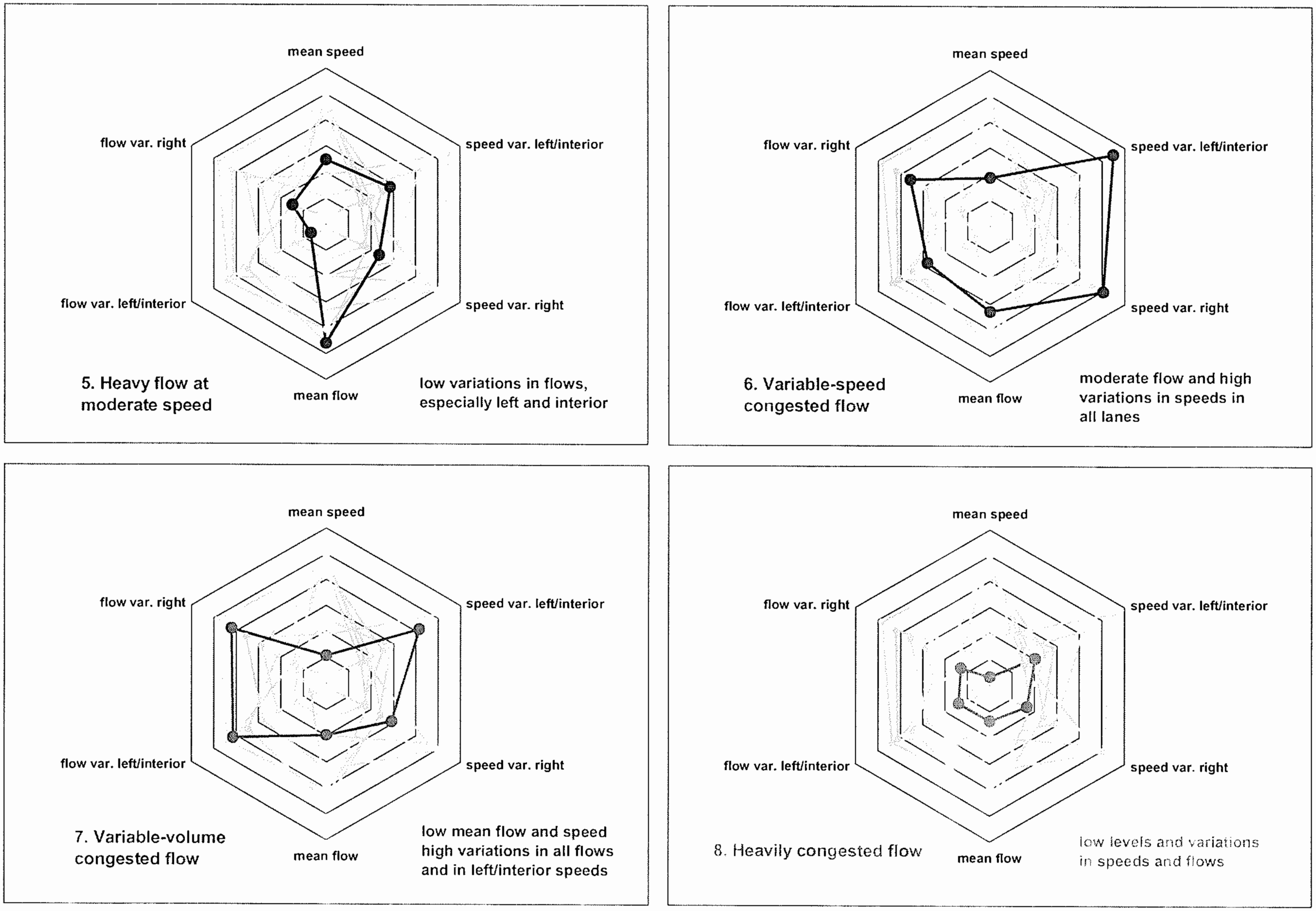

Figure 2b Radar Diagrams of Traffic Flow Regimes for Daylight and Dry Road Conditions (Regimes 5 through 8)

Relationship between Regimes and Speed – Flow Curves

It is instructive to plot the eight Regime centroids in the space of just two of the six variables: mean speed and mean flow (Figure 3). (As stated previously, for purposes of discussion, we assume that flow/occupancy is a surrogate for speed.) These centroids trace a speed-flow curve that is a familiar concept in traffic engineering (Roess, McShane and Prassas, 1998). The curve has three distinct branches: (1) a top nearly horizontal convex segment, generally known as "free flow," (2) a vertical segment near maximum observable flow, known as "queue discharge," and (3) a bottom segment known as "congested flow" or "within the queue" (Hall, Hurdle and Banks, 1992). The shape traced by our Regimes is similar to that found in many empirical studies (Pushkar, Hall and Acha-Daza, 1994; Schoen, at al., 1995). Conceptually, as demand for road space increases, you move clockwise through the curve. The Regimes are numbered in this manner.

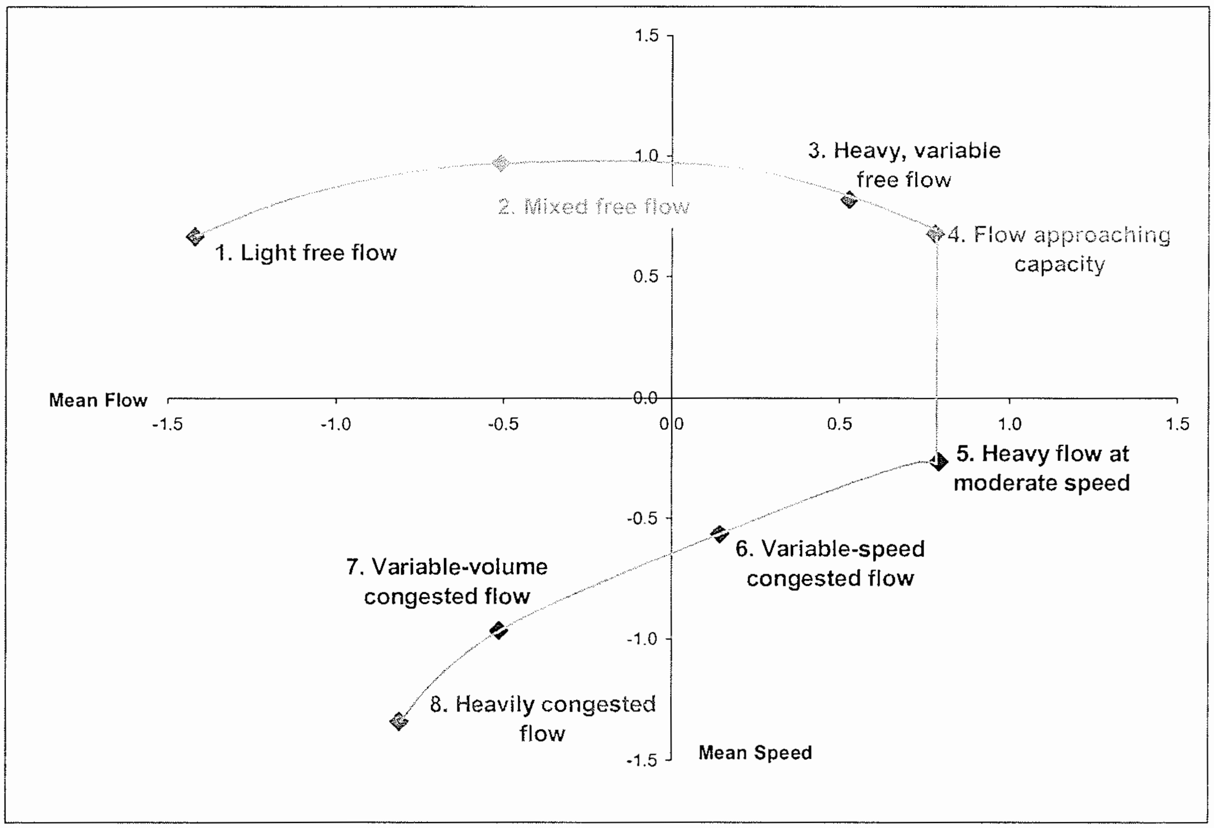

Figure 3 Speed-Flow Curve Implied by Locations of the Eight Traffic Flow Regimes in Standardized Speed-Flow Space

On the free flow segment, traced by the first four Regimes, speed at first increases with demand. One plausible (albeit unsubstantiated) explanation for this observation may be that, as individual vehicles become less exposed, drivers feel that they are less vulnerable to enforcement of speed limits, anecdotally, a common perception among southern California's commuters. Speed then decreases with demand in the free flow branch as driver behavior begins to be influenced by flow density. On the congested segment of the implied speed-flow curve traced by the remaining four Regimes, speeds decrease with decreasing flows as demand increases.

By superimposing the six-dimensional plots of the eight traffic flow regimes in the space of speed-flow, we can see that each of the regimes is quite different in terms of the four remaining dimensions (Figure 4). Beginning with the lowest level of demand, which can occur either at off-peak times or at any time downstream of a bottleneck, Regime1 is characterized by light flow, with very low to moderate variations in speeds and flow. The second Regime, "Mixed free flow," exhibits the highest mean speed and very high variations in right-lane speed and high variation in left- and interior-lane 30-second volumes. This Regime, which has been purported to capture the behavior of traffic with heterogeneous freeway trip lengths, is more often observed in the 10AM to 1PM period on weekdays and immediately before 10AM on Saturdays (Golob, Recker and Alvarez, 2002). Regime 3, "Heavy, variable free flow," is similar to Regime 4, "Flow approaching capacity" in terms of mean speed and only slightly below Regime 4 in terms of mean flow. Regimes 3 and 4 are also similar in terms of low variations in speeds, but they differ substantially in terms of variation in right-lane flow. As shown in the next Section 3.6., despite their similarities on all but a single flow dimension, this difference in flow variance leads to differences in the safety profile of these two Regimes.

As nominal capacity is exceeded, the speed-flow curve moves from maximum "stable" flow at relatively high speeds (Regime 4) to a similar level maximum "unstable" flow (Regime 5), with a substantial reduction in mean speed. These two Regimes are connected by what traffic engineers designate as the "queue discharge" (nearly vertical) section of the speed-flow diagram. In moving from Regime 4 to Regime 5, Figure 4 shows that the radar diagram becomes squashed from the top (indicating reduced speed), but the variation in flow also decreases dramatically, especially in all lanes except the right lane. Regime 5 might be considered as predominantly "synchronized flow" (Kerner and Rehborn, 1996a; 1996b).

On the congested branch of the implied speed-flow diagram, both flow and speed variances increase with decreasing mean flows and decreasing mean speeds, as one moves toward ever more congested flow (Regimes 6, 7 and 8). Regimes 6 and 7 both represent stop-and-go traffic characterized by shock wave dynamics and bunching. Regime 6 is characterized mostly by very high variances in speeds, in both lane groupings. Regime 7 is characterized more by high variances in volumes. Further study is required to determine how our results are related to theories about waves of rising and falling vehicle density, particularly to the six phases proposed by Helbing and his colleagues (Helbing, Hennecke, and Treiber, 1999; Helbing, and Huberman, 1998; and Helbing and Schreckenberg, 1999). These phases are distinguished by how often waves pass through the stream of vehicles and how much the density drops off between waves. In a phase called a "pinned localized cluster," for instance, an enduring but very localized bunching haunts the immediate vicinity of an on ramp. Daganzo, Cassidy, and Bertini (1999) provide possible explanations of Helbing's phases that could prove useful in identifying the best theoretical explanation. Aside from theoretical explanation, these differences in variances in speeds and flow, by lanes, explain differences in safety profiles for different types of congested flow that cannot be explained simply in terms of mean speeds and flows. These accident profiles are described in the next Sections.

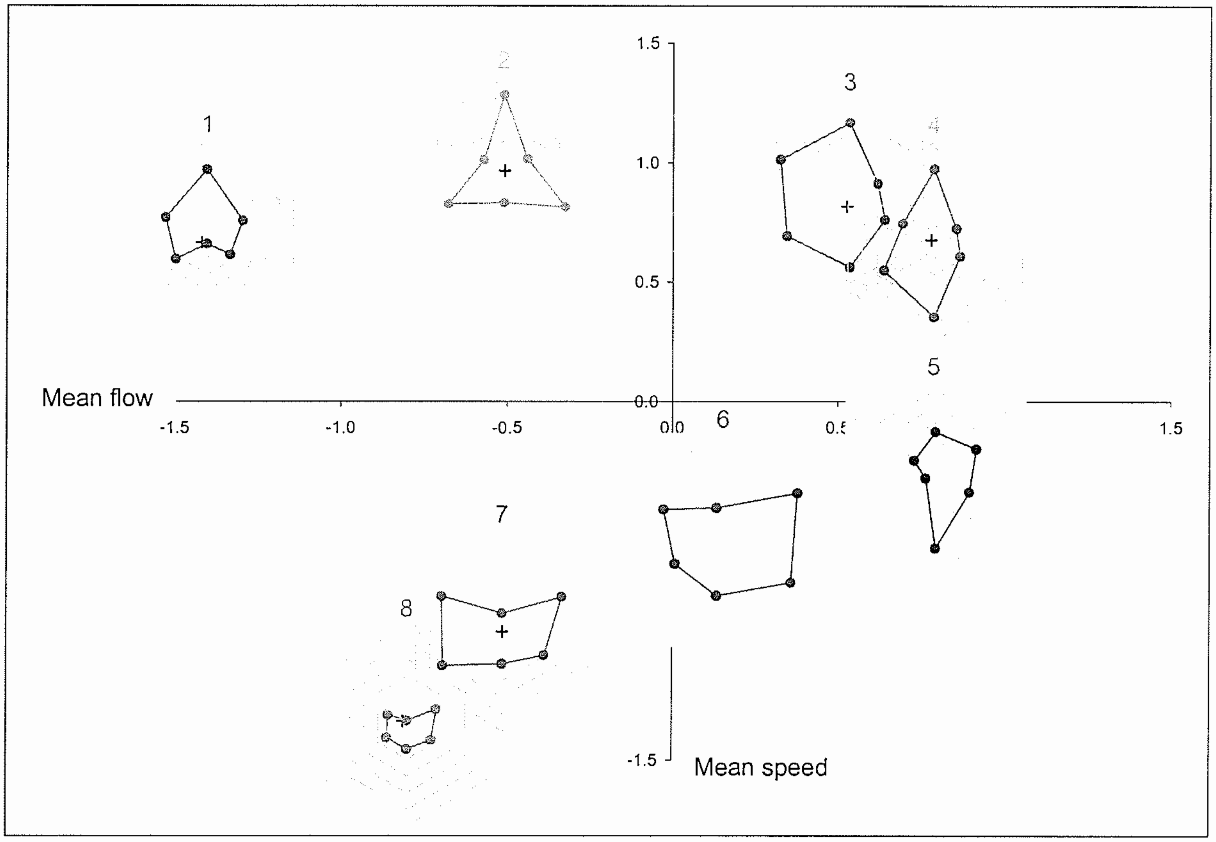

Figure 4 Six-Dimensional Radar Plots for the Eight Traffic Flow Regimes Plotted in Standardized Speed-Flow Space

FITS Demonstration on 1998 AM Peak Hour Data

For purposes of initially demonstrating FITS, we drew a random sample of traffic flow measurements to estimate vehicle exposure to each of the dry-road traffic flow Regimes for the six major Orange County freeways for the AM peak hours (6:00 AM to 9:00 AM inclusive) for all of calendar year 1998 (Golob, Recker and Alvarez, 2003). Because of systematic biases introduced by non-reporting loop stations in 1998, the following is intended for demonstration purposes only; no claim is made that the results are representative of actual conditions. However, these estimates should be in the right ballpark, and they demonstrate what might be learned from a full-scale implementation of such a safety performance analysis tool. The estimated temporal distribution of the eight Regimes during the AM peak period in 1998 is graphed in Figure 5.

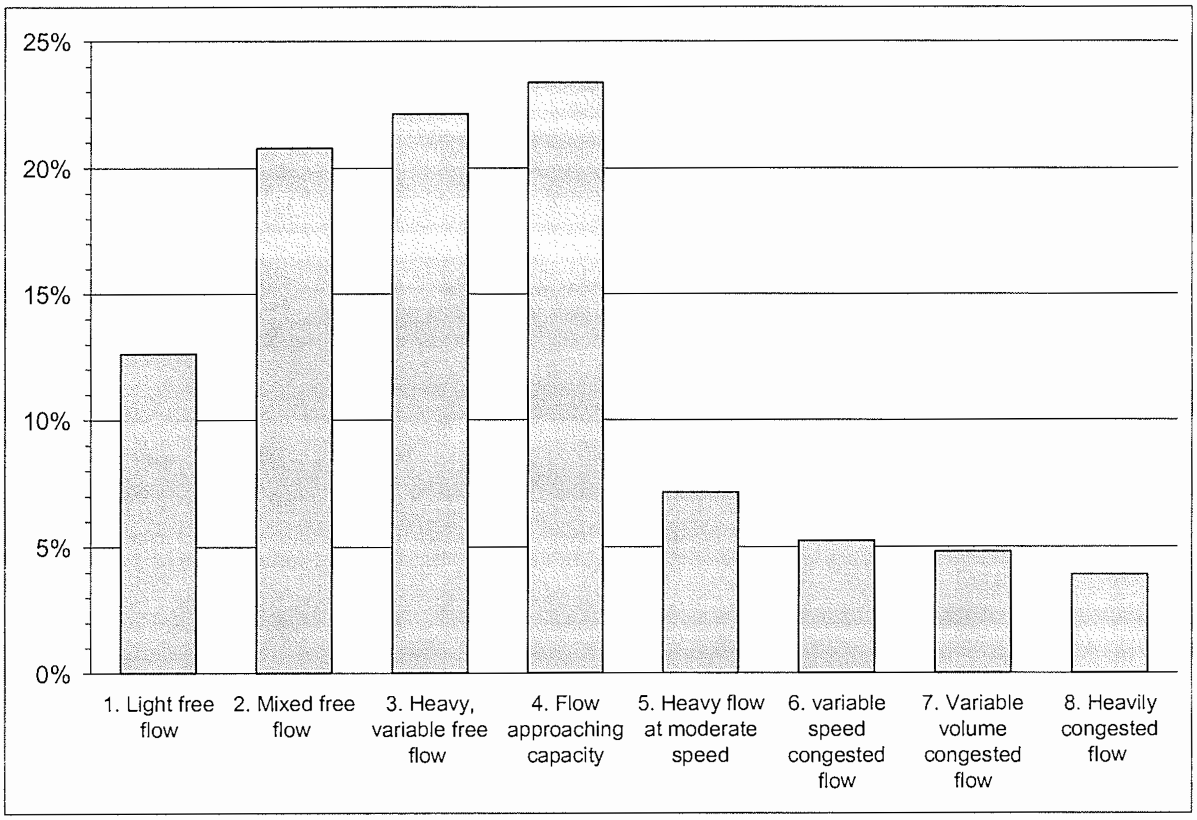

Figure 5 Temporal Distribution of Daylight, Dry Road Regimes on Six Orange County Freeways for 1998 AM Weekday Peak Hours

Regimes 1 through 4 are on the free-flow branch of the speed-flow curve (Figures 3 and 4), and approximately eighty percent of the time the Orange County freeway system operated in one of these four free-flow Regimes during AM peak hours. The remaining twenty percent of the time the system operated in one of the four Regimes on the congested-flow branch of the curve. Among the four free-flow Regimes, Regime 1 is not as likely as the others, but (for the time period under consideration) it is representative of conditions downstream of a bottleneck.

The four Regimes on the congested flow branch of the speed-flow curves, Regimes 5 through 8, together account for approximately 20% of all time periods. As in the case of the four Regimes on the free-flow branch of the speed-flow curve, the likelihood of observing any one of the congested-flow Regimes is an increasing function of mean volume.

FITS was then used to estimate the distribution of 1998 AM peak period crashes across the eight Regimes. Total volumes associated with each observed Regime occurrence were calculated from total 30-second volumes across all freeway lanes. These estimates are for demonstration purposes only; additional research is needed before we can confidently assign safety levels to different traffic flow conditions. An estimate of total crashes per million exposed vehicles per Regime is graphed in speed-flow space in Figure 6.

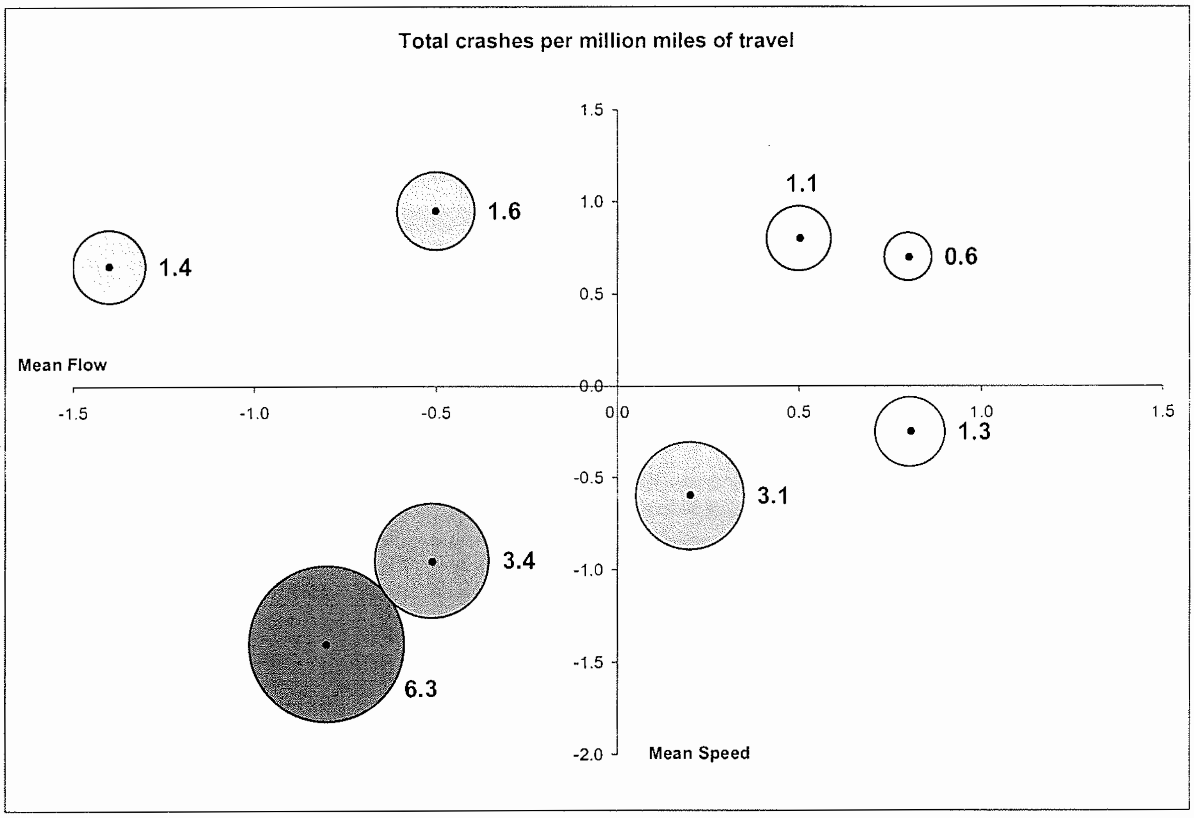

Figure 6 Estimated Total Crashes per Million Vehicle Miles of Travel for the Eight Traffic Flow Regimes During AM Peak Hours, Plotted in Standardized Speed-Flow Space

Crash rates estimated in this manner are highest along the congested-flow branch of the curve. These preliminary results demonstrate that crash rates for the same levels of flow are approximately double in congested versus free flow conditions: 0.74 crashes per million vehicles for Regime 2 “Mixed free flow” versus 1.59 for Regime 7 “Variable volume congested flow;” and 0.28 for Regime 4 “Flow approaching capacity” versus 0.62 for Regime 5 “Heavy flow at moderate speeds.” If these “demonstration” results hold up under a full-scale implementation, we will be able to directly quantify the safety benefits of improved traffic flow.

The Variation of Crash Type with Traffic Flow

Not all crashes are the same in terms of severity and effects on the system in terms of non-recurrent congestion. Crashes involving fatalities and serious injuries represent a much greater social and economic cost than do property damage only (PDO) crashes, and the costs of PDO crashes are a function of the extent of damage and the number of vehicles involved. Injury crashes also produce a greater incident effect due to needs for emergency medical attention and investigation requirements. Among PDO crashes, those involving multiple vehicles and those located in interior lanes, potentially interact with higher traffic flow volumes to cause the greatest impact on system performance. In recognition of the importance of crash typology, one of the objectives in developing the FITS tool was to analyze the relationship between type of crash and traffic flow. The test implementation of FITS for 1998 AM peak hour traffic revealed that the eight Regimes for daylight and dry road conditions were characterized by different patterns of crash types. The prevailing crash type for each Regime is arranged by mean speed and mean flow in Figure 7. By identifying the types of crashes that are most likely to occur under different traffic conditions, then identifying where and when on the freeway system these conditions occur, a safety performance tool can aid planners looking to identify and relieve dangerous conditions.

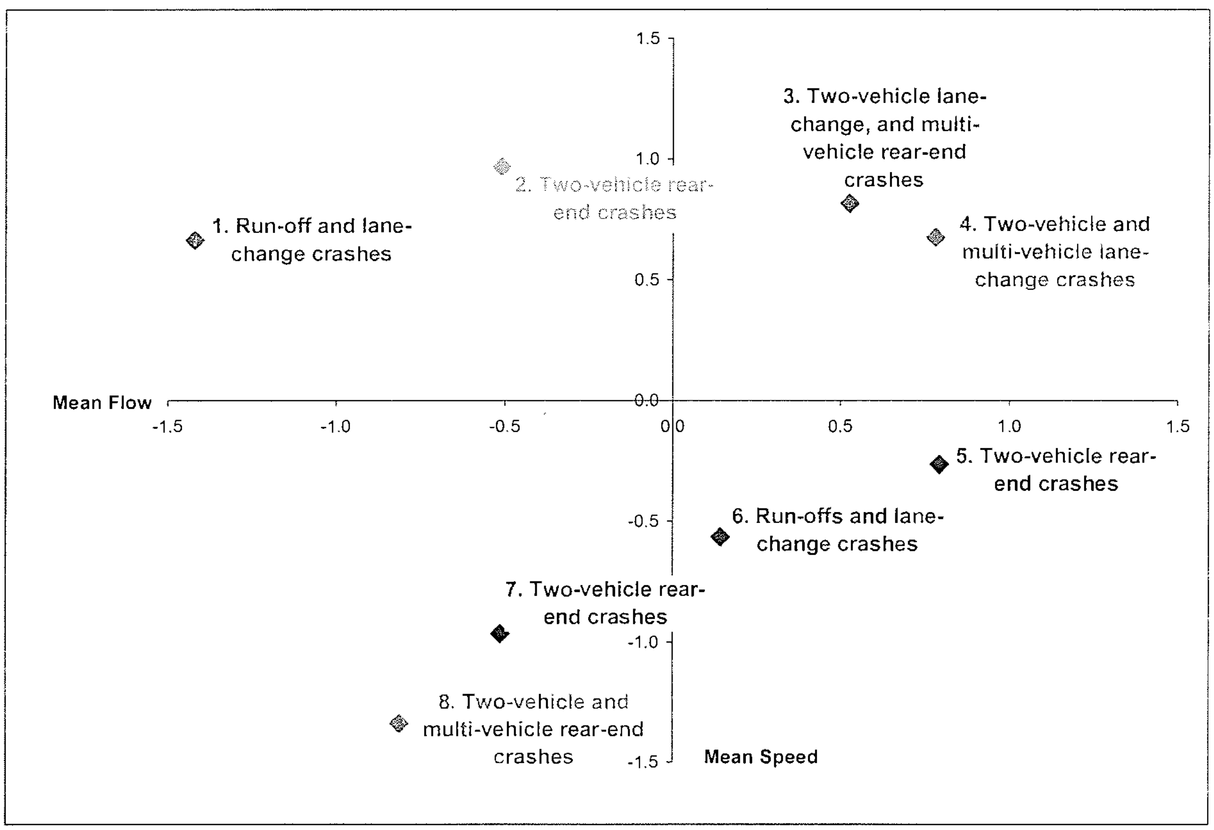

Figure 7 Prevailing Crash Types for the Eight Traffic Flow Regimes During AM Peak Hours, Plotted in Standardized Speed-Flow Space

The Effects of Flow Turbulence

The FITS tool captures flow turbulence by four variables: (1) variation in speeds in the left and interior lanes, (2) variation in speed in the right lane, (3) variation in flow in the left and interior lanes, and (4) variation in flow in the right lane. Likewise, Oh, Oh and Chang (2001) and Oh, Oh, Ritchie and Chang (2001) capture turbulence using a different measure of speed variation. Results suggest that all of these variables are effective in explaining some aspect of safety. As an illustration, the estimated number of crashes per million exposed vehicles for the 1998 AM peak period on Orange County freeways is plotted in Figure 8 as a function of variation in flow in the right lane versus variation in speeds in the left and interior Lanes. With the notable exception of Regime 8 "Heavily congested flow" which has a high crash rate and low variations, the next two highest crash rates are for the two Regimes (6 and 7) with the highest levels of turbulence. This demonstrates how reducing variations in speed and flow should lead to safer conditions.

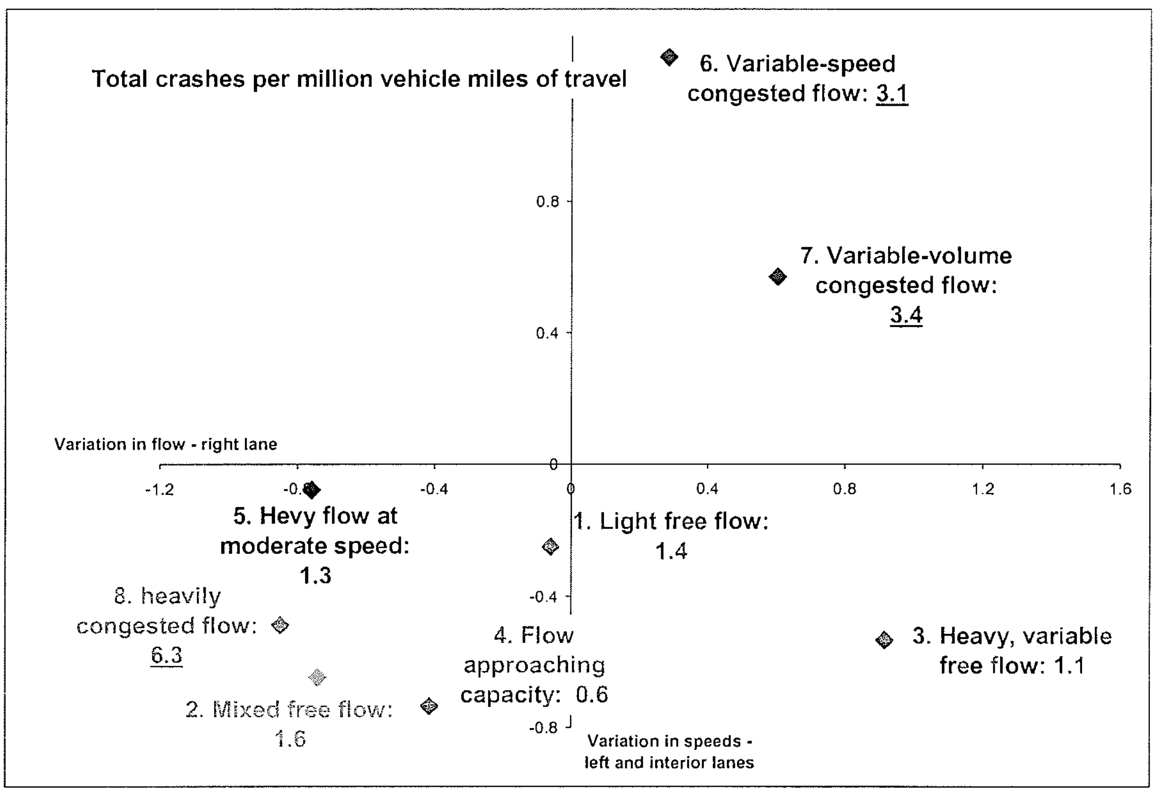

Figure 8 Estimated Total Crashes per Million Vehicle Miles of Travel by Traffic Flow Regimes Plotted in Standardized Space of (x) Variation in Flow in Right Lane *versus* (y) Variation in Speeds in Left and Interior Lanes

CONCLUSIONS

The purpose of this paper is to demonstrate that we are on the threshold of an ability to fully develop and implement tools for use in real-time assessment of the level of safety of any pattern of traffic flow on an urban freeway. The only input that such a tool requires is a stream of 30-second (or similar interval) observations from single inductive loop detectors. This stream is processed to provide a continual assessment of safety, updated every 30-seconds, based on central tendencies of flow and speed, and variations in flow and speed for different lanes of the freeway.

Traffic safety monitoring complements existing performance monitoring by adding real-time assessment of precursors of traffic safety to performance criteria that typically involve travel times, speeds and throughput. A safety performance monitoring tool can also be used as part of any evaluation that compares before and after traffic flow data. Such an evaluation might involve assessing the benefits of ATMS operations or any other ITS implementation. Another application will be to forecast the safety implications of proposed projects by evaluating the levels of safety implied by traffic simulation model outputs.

We are not purporting that the FITS methodology is the only basis upon which a real-time safety monitoring tool can be developed. Indeed, there are at least two other promising methods: the Real-time Estimation of Accident Likelihood (REAL) technique that incorporates the nonparametric density estimation technique and Bayesian probabilistic model (Oh, Oh, Ritchie and Chang, 2001), and the method of Lee, Hellinga, and Saccomanno (2003). These methods differ in terms of how estimates of accident likelihood are calculated from real-time traffic data. Future research is needed to compare the several analytical approaches that have been recently developed, with the objective of implementing a safety performance evaluation tool that incorporates the best features of each approach.

ACKNOWLEDGEMENTS

This research was funded in part by the California Partners for Advanced Transit and Highways (PATH) and the California Department of Transportation (Caltrans). 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 University of California, California PATH, or the California Department of Transportation.

REFERENCES

Aljanahi, A.A.M., Rhodes, A.H. and Metcalfe, A.V., 1999. Speed, speed limits and road traffic accidents under free flow conditions. *Accident Analysis and Prevention*, 31: 161-168.

Boon, P. van der, 1996. *A Robust Approach to Nonlinear Multivariate Analysis*. DSWO Press, Leiden.

Burg, E. van der, 1988. *Nonlinear canonical Correlation and Some Related Techniques*. DSWO Press. Leiden.

Caltrans (1993). *Manual of Traffic Accident Surveillance and Analysis System.* California Department of Transportation, Sacramento.

Cedar, A. and Livneh, L., 1982. Relationship between road accidents and hourly traffic flow. *Accident Analysis and Prevention*, 14: 19-44.

Chen, C; K. F. Petty, A. Skabardonis, P.P. Varaiya, and Jia, Z. (2001). Freeway performance measurement system: mining loop detector data. *Transportation Research Record 1748*: 96-102.

Choe, T., Skabardonis, A., and Varaiya, P.P., 2002. Freeway performance measurement system (PeMS): an operational analysis tool. Presented at Annual Meeting of Transportation Research Board, January 13-17, Washington, DC.

Coifman, B., 2001. Improved velocity estimation using single loop detectors. *Transportation Research, Part A*, 35: 863-880.

Daganzo, C.F., Cassidy, M.J. and Bertini, R.L., 1999. Possible explanations of phase transitions in highway traffic. *Transportation Research, Part A*, 33: 365-379.

De Leeuw, J., 1985. The Gifi system of nonlinear multivariate analysis. In E. Diday, et al., eds., *Data Analysis and Informatics, IV: Proceedings of the Fourth International Symposium.* North Holland, Amsterdam.

Dickerson, A., Peirson, J. and Vickerman, R., 2000. Road accidents and traffic flows: An econometric investigation. Economica, 67: 101-121.

Edwards, J.B., 1998, The relationship between road accident severity and recorded weather. *Journal of Safety Research,* 29: 249-262.

Feng, C., 2001. Synthesis of studies on speed and safety. *Transportation Research Record* 1779: 86-92.

Frantzeskakis, J.M. and Iordanis, D.I., 1987. Volume-to-capacity ratio and traffic accidents on interurban four-lane highways in Greece. *Transportation Research Record,* No. 1112: 29-38.

Fridstrøm, L., Ifver, J., Ingebrigtsen, S., Kulmala, R. and Thomsen, L.K., 1995. Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Accident Analysis and Prevention*, 27: 1-20.

Garber, N.J. and Gadiraju, R. 1990. Factors influencing speed variance and its influence on accidents. *Transportation Research Record* 1213: 64-71.

Gaudry. M. and Lassarre, S. (Eds.), 2000. *Structural Road Accident Models*. Pergamon, Oxford.

Gifi, A., 1990. *Nonlinear Multivariate Analysis*. Wiley, Chichester.

Golob, T.F. and Recker, W.W., 2003a. Relationships among urban freeway accidents, traffic flow, weather and lighting conditions. *ASCE Journal of Transportation Engineering,* in press. Available as PATH Working Paper UCB-ITS-PWP-2003-1.

Golob, T.F. and Recker, W.W., 2003b. A Method for relating type of cash to traffic flow characteristics on urban freeways. Presented at the Annual Meeting of the Transportation Research Board, January 12-16, Washington, D.C.

Golob, T.F., Recker, W.W. and Alvarez, V.M., 2002. *Freeway Safety as a Function of Traffic Flow: The FITS Tool for Evaluating ATMS Operations.* Final Report prepared for California Partners for Advanced transit and Highways (PATH). Institute of Transportation Studies, University of California, Irvine, CA.

Golob, T.F., Recker, W.W. and Alvarez, V.M., 2003. A Tool to Evaluate the Safety Effects of Changes in Freeway Traffic Flow. *ASCE Journal of Transportation Engineering,* in press.

Gwynn, D.W., 1967. Relationship of accident rates and accident involvement with hourly volumes. *Traffic Quarterly*, 21: 407-418.

Gwynn, D.W. and Baker, W.T. 1970. Relationship of accident rates with hourly volumes. *Traffic Engineering*, 40: 42-47.

Hall, F.L., Hurdle, V.F. and Banks, J.H., 1992. Synthesis of recent work on the nature of speed-flow and flow-occupancy (or density) relationships on freeways. *Transportation Research Record*, 1365: 12-18.

Hall, J.W. and Pendleton, O.J., 1989. Relationship between V/C ratios and Accident Rates. Report FHWA-HPR-NM-88-02, U.S. Department of Transportation, Washington, DC.

Helbing, D., Hennecke, A., and Treiber, M., 1999. Phase diagram of traffic states in the presence of inhomogeneities. Physical Review Letters, 82: 4360-4363.

Helbing, D. and Huberman, BA., 1998. Coherent moving states in highway traffic. *Nature*, 396: 738-740.

Helbing, D. and Schreckenberg, M., 1999. Cellular automata simulating experimental properties of traffic flow. *Physical Review E*. 59: R2505-R2508.

Hurdle, V.F., Merlo, M.I. and Robertson, D., 1997. study of speed-flow relationships on individual freeway lanes. *Transportation Research Record* 1591: 7-13.

Israëls, .Z., 1987. *Eigenvalue Techniques for Qualitative DATA.* DSWO Press, Leiden.

Jadaan, K.S. and Nicholson, A.J., 1992. Relationship between road accidents and traffic flows in an urban network. *Traffic Engineering and Control,* 33: 507-511.

Jansson, J.O., 1994. Accident externality charges. Journal of transport Economics and Policy, 28: 31-43.

Jia, Z., Chen, C., Coifman, B. and Varaiya, P., 2001. The PeMS algorithms for accurate, real-time estimates of *g*-factors and speeds from single-loop detectors. Presented at the 4th International IEE Conference on Intelligent Transportation Systems (ITSC), Aug. 25-28, Oakland, CA

Johansson, P., 1996. Speed limitation and motorway casualties: A time series count data regression approach. *Accident Analysis and Prevention*, 28: 73-87.

Jones-Lee, M.W., 1990. The value of transport safety. *Oxford Review of Economic Policy,* 6: 39-60.

Kerner, B.S. and Rehborn, H., 1996a. Experimental properties of complexity in traffic flow. *Physical Review E.* 53: R4275-R4278.

Kerner, B.S. and Rehborn, H., 1996b. Experimental features and characteristics of traffic jams. *Physical Review E.* 53: R1297-R1300.

Lee, C., Saccomanno, F., and Hellinga, B., 2002. Analysis of crash precursors on instrumented freeways. *Transportation Research Record,* 1784: 1-8.

Lee, C., Hellinga, B. and Saccomanno, F., 2003. Real-Time Crash Prediction Model for Application to Crash Prevention in Freeway Traffic. Presented at the Annual Meeting of the Transportation Research Board, January 12-16, Washington, D.C.

Maher, M.J. and Summersgill, I., 1996. A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis and Prevention*, 28: 281-296.

Mensah, A. and Hauer, E., 1998. Two problems of averaging arising from the estimation of the relationship between accidents and traffic flow. *Transportation Research Record* 1635: 37-43.

Michailidis, G. and de Leeuw, J., 1998. The GIFI system of descriptive multivariate analysis. *Statistical Science*, 13: 307-336.

Newberry, D., 1988. Road user charges in Britain. *Economic Journal*, 98: 161-176.

Oh, C., Oh, J-S. and Chang, M., 2001. An advanced freeway warning information system based on accident likelihood. Presented at the 9th World Conference on Transport Research, July 22-27, 2001, Seoul, Korea.

Oh, C., Oh, J-S., Ritchie, S.G. and Chang, M., 2001. Real-time estimation of freeway accident likelihood. Presented at the Annual Meeting of the Transportation Research Board, January 8-12, Washington D.C.

O'Reilly, D., Hopkin, J., Loomes, G., Jones-Lee, M., Philips, P., McMahon, K., Ives, D., Sobey, B., Ball, D. and Kemp, R., 1994. The value of road safety: UK Research on the valuation of preventing on-fatal injuries. *Journal of Transport Economics and Policy,* 28: 45-59.

PeMS Development Group, 2001, *PeMS Calculations with Loop Detectors*, University of California, Berkeley. http://pems.eecs.berkeley.edu/Papers/agg_calcs.pdf

Prigogine, I. And Herman, R., 1971. *Kinetic Theory of Vehicular Traffic.* American Elsevier, New York.

Pushkar, A., F.L. Hall and J.A. Acha-Daza, 1994. estimation of speeds from single-loop freeway flow and occupancy data using cusp catastrophe theory model. *Transportation Research Record* 1457: 149-157.

Roess, R.P, McShane, W.R. and Prassas, E.S., 1998. *Traffic Engineering*, 2nd Ed. Prentice Hall, Upper Saddle River, NJ.

Sandhu, B. and Al-Kazily, J., 1996. Safety impacts of freeway traffic congestion. Presented at Annual Meeting of Transportation Research Board, January 7-11, Washington, DC.

Shefer, D. and Rietveld, P., 1997. Congestion and safety on highways: Towards an analytical model. Urban Studies, 34: 679-692.

Stokes, R.W. and Mutabazi, M.I., 1996. Rate-quality control method of identifying hazardous road locations. *Transportation Research Record,* No. 1542: 44-48.

Sullivan, E.C., 1990. Estimating accident benefits of reduced freeway congestion. *Journal of Transportation Engineering*, 116: 167-180.

Sullivan, E.C. and Hsu, C-I., 1988. *Accident rates along congested Freeways.* Research Report UCB-ITS-RR-88-6, Institute of Transportation Studies, University of California, Berkeley.

Turner, D.J. and Thomas, R., 1986. Motorway accidents: an examination of accident totals, rates and severity and their relationship with traffic flow. *Traffic Engineering and Control,* 27: 377-387.

van Buren, S. and Heiser, W.J., 1989. Clustering N-objects into K-groups under optimal scaling of variables. *Psychometrika*, 54: 699-706.

Varaiya, P. P., 2001. Freeway Performance Measurement System, PeMS V3, Phase 1: Final Report. Report UCB-ITS-PWP-2001-17, California PATH Program, Institute of Transportation Studies, University of California, Berkeley, CA.

Vickery, W., 1969. Congestion theory and transport investment. *American Economic Review,* 59: 251-260.

Vitaliano, D.F. and Held, J., 1991. Road accident external effects: An empirical assessment. *Applied Economics*, 23: 373-378.

Zhou, M. and Sisiopiku, V.P., 1997. Relationship between volume-to-capacity ratios and accident rates. *Transportation Research Record,* No. 1581: 47-52.