Accident Analysis and Prevention 36 (2004) 933–946

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

Freeway safety as a function of traffic flow

Thomas F. Golob [a,*], Wilfred W. Recker [a,b], Veronica M. Alvarez [a,1]

[a] Institute of Transportation Studies, University of California, 552 Social Science Tower, Irvine, CA 92697, USA

[b] Department of Civil and Environmental Engineering, University of California, Irvine, CA 92697, USA

Received 21 April 2003; received in revised form 8 September 2003; accepted 22 September 2003

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 median 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.

Keywords: Traffic safety; Accident rates; Traffic flow; Loop detectors; Speed; Traffic density; Congested flow; Fundamental diagram

1. 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 segment of freeway in terms of throughput: travel time per vehicle, average speed or total delay (Chen et al., 2001; Choe et al., 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 s 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 segment 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).

* Corresponding author. Tel.: +1-949-824-6287; fax: +1-949-824-8385.
E-mail addresses: tgolob@uci.edu (T.F. Golob), wwrecker@uci.edu (W.W. Recker), valvarez@uci.edu (V.M. Alvarez).
[1] Tel.: +1-949-824-6571.

0001-4575/$ – see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.aap.2003.09.006

2. Previous studies

Studies of relationships between crashes and traffic flow can be divided into two types: (a) aggregate studies, in which units of analysis represent counts of crashes or crash rates for specific time periods (typically months or years) and for specific spaces (specific roads or networks), and traffic flow is represented by parameters of the statistical distributions of traffic flow for similar time and space, and (b) disaggregate analysis, in which the units of analysis are the crashes themselves and traffic flow is represented by parameters of the traffic flow at the time and place of each crash. Disaggregate studies are relatively new, and are made possible by the proliferation of data being collected in support of intelligent transportation systems developments. Transportation management centers routinely archive traffic flow data from sensor devices such as inductive loop detectors and these data can, in principle, be matched to the times and places of crashes, as described in Section 4 of this paper.

Analyses based on aggregations of crashes are prone to statistical problems (Mensah and Hauer, 1998; Davis, 2002). Ecological fallacy arises whenever an observed statistical relationship between aggregated variables is falsely attributed to the units over which they were aggregated (Robinson, 1950). Methods for identifying and correcting for biases due to ecological fallacy are well developed in geography and regional science (e.g., Holt et al., 1996), but such methods are rarely applied in traffic safety research. Disaggregate analyses in principle avoid problems of ecological fallacy. Nevertheless, aggregate studies first provided compelling evidence that crashes are not simply a linear function of traffic volume, and several aggregate studies underpin the research reported here by having identified relationships between different types of crash rates and the three main traffic flow parameters: traffic volume, speed, and density.

Aggregate studies can be subdivided into macroscopic and microscopic studies (Persaud and Dzbik, 1992). Macroscopic studies typically use crash data in terms of vehicle miles of travel (VMT) accumulated over long time periods, such as a year. Microscopic aggregate studies, which have also been referred to as disaggregate studies (Sullivan, 1990), typically use data based on average hourly observations of crash rates and traffic flow, allowing comparisons to be made, for example, between congested and free-flow traffic conditions. Jovanis and Chang (1986) reflect on the scales of the aggregations over time and space in comparing the results of some of the early aggregate studies.

Ceder and Livneh (1982) and Martin (2002) observed that U-shaped curves depicting crash rates as a function of hourly traffic flow for free-flow conditions can result from the combination of two functional forms: (a) single-vehicle crashes decreasing at a decreasing rate as a function of flow, and (b) multiple-vehicle crashes increasing with flow, usually at an increasing rate. These curves were also observed to vary for day and night and for weekday and weekend. Garber and Subramanyan (2001) observed that peak accident rates do not occur at peak flow, but rather, crashes tend to increase with increasing density, reaching a maximum before the optimal density at which flow is at capacity. Garber and Ehrhart (2000) observed that crash rates involve an interaction of variation in speed and flow, with crash rates being an increasing function of the standard deviation of speed for all levels of flow. Aljanahi et al. (1999) and Garber and Gadiraju (1990) also found that crash rates were positively related to both mean speed and variation in speed. Baruya and Finch (1994) found that crash levels were an increasing function of the coefficient of variation of speed. Finally, Ceder (1982), Sullivan (1990), and Persaud and Dzbik (1992) each investigated how accident rates are different for free-flow versus congested conditions. Each study concluded that crash rates are higher under congested conditions.

Disaggregate analyses have been reported by Oh et al. (2001, 2001), and Lee et al. (2002, 2003), in addition to the research presented here and in Golob and Recker (2003, 2004). The objective in each of these studies was to identify freeway traffic flow conditions that are precursors of certain types of crashes. Using data for a freeway segment in the San Francisco Bay Area of California, Oh et al. (2001) developed probability density functions for crash potential based on the standard deviation of speed. Lee et al. (2002, 2003) focused on coefficients of variation of speeds and traffic densities compared across different freeway lanes, using data for an urban freeway in Toronto. In contrast, our approach is to develop a classification scheme by which traffic flow conditions on an urban freeway can be classified into mutually exclusive clusters that differ as much as possible in terms of likelihood of crash by type of crash. By using lane-by-lane traffic flow data measured on short time intervals (e.g., 20 or 30 s), disaggregate analysis holds the potential of being 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.

3. Overview of 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 et al., in press). FITS uses a data stream of 30 s 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 1000 crashes on six major freeways in Orange County, California in 1998 as a function of pre-crash traffic flow conditions. Orange County is an urban area of about three million population located between Los Angeles and San Diego. Pre-crash conditions were computed using

27.5 min of data from the closest loop station. Through sensitivity analysis, it was determined that approximately 30 min of traffic flow data prior to the reported time of the crash were needed to establish patterns in terms of central tendencies and variations of the variables described in the next section. The 2.5 min period immediately preceding the reported time was later discarded, because accident times were found to be typically rounded off to the nearest 5 min, resulting in a 27.5 min period for establishing pre-crash traffic conditions. The median distance from the crash site to the closest loop station was 0.12 miles.

FITS is based on analyses that capture statistical relationships between traffic flow, as measured by one of the simplest and most ubiquitous monitoring instruments, a string of single inductive loops buried in all lanes at a single point on a freeway, and crash characteristics, in terms readily available measures of overall severity and the type and location of the collisions. No reported crashes are removed from the samples. If crashes due to equipment failures or some other factors thought to be peripheral to driver behavior are independent of traffic conditions, these should not show up in the results. However, it might be possible that many factors not usually attributed to the contribution to crash likelihood vary by traffic conditions, due, for example, to differences in reaction times, wear and tear on vehicles, and stability of vehicle loads. Also, certain environmental factors (e.g., sight distances and locations of exits and entrances), are expected to be reflected in the detailed patterns of traffic flow. The exception involves daylight versus nighttime and dry versus wet conditions, which are handled through separate analyses. Investigations of whether or not relations between crash typology and traffic flow can be improved by adding detailed data on environmental conditions are in the realm of future research.

4. 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 cases 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 1192 Crashes in 1998 on Orange County freeways

Crash characteristic	Percent
Crash type	
Single-vehicle hit object or overturn	14.2
Multiple-vehicle hit object or overturn	5.9
Two-vehicle weaving crash[a]	19.3
Three-or-more-vehicle weaving crash[a]	5.5
Two-vehicle straight-on rear end	33.8
Three-or-more-vehicle rear end	21.3
Crash location	
Off-road, driver's left	13.8
Left lane	25.8
Interior lane(s)	32.7
Right lane	19.3
Off-road, driver's right	8.3
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 (flow) and occupancy for each 30 s 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 and their ratios 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 occupancy (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 90th percentile and 50th 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.

The third block of traffic flow variables comprises the mean volumes for all three lanes taken over the entire 27.5 min 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 s volumes for all three lanes as a measure of variation in volume over the 27.5 min period.

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 90th and 50th percentiles of volume/occupancy—left lane Difference between 90th and 50th percentiles of volume/occupancy—interior lane Difference between 90th and 50th 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

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
90th% tile–50th% tile of volume/occupancy interior lane	Variation in speed—all lanes but right
90th% tile–50th% 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 left lane	Variation in flow—all lanes but right
Standard deviation of volume right lane	Variation in flow—right lane

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 12 original variables to 6 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. The average correlation between each factor and its representative variable is 0.899.

5. Determining traffic flow regimes according to differences in crash typology

5.1. Methodology

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, 2004). 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 recurring objective 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).

In our applications of NLCCA, there are two sets of variables: a single categorical variable representing the results of each clustering on one side, and the three categorical crash characteristics of Table 1 on the other side. Separate analyses were conducted for each number of clusters, ranging from four to eighteen clusters. Based on results that demonstrated how safety patterns were affected by weather and lighting conditions (Golob and Recker, 2003), these analyses were repeated 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. Results for all three environmental segments are documented in Golob et al. (2002). Due to space limitations, we report only on results for the segment accounting for the most traffic: daylight and dusk–dawn conditions on dry roads. A detailed description of the NLCCA solution for this environmental segment are given in Golob and Recker (2004). In the remainder of this paper we focus on interpreting the relations between traffic flow and crash variables implied by the clustering results.

5.2. Results 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 can be 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. Fig. 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.

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

5.3. Regimes described in terms of a speed–flow curve

It is instructive to plot the eight regime centroids in the space of just two of the six variables: mean speed and mean

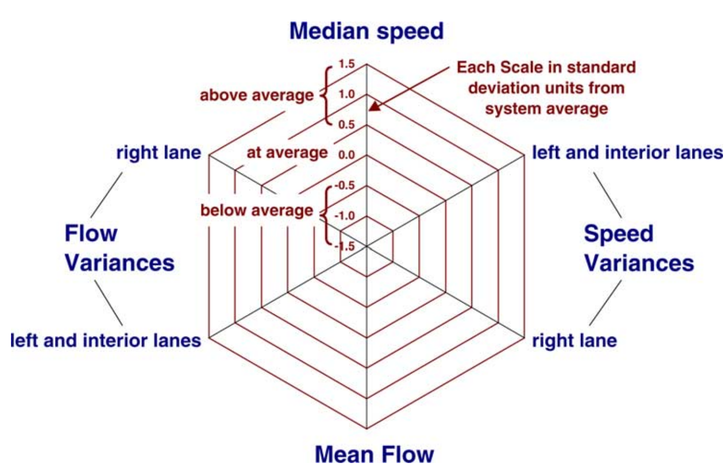

Fig. 1. Key to radar diagram used to describe traffic flow regimes in terms of centroid locations for six traffic flow variables.

flow (Fig. 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 et al., 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 et al., 1992). The shape traced by our regimes is similar to that found in many empirical studies (Pushkar et al., 1994; Schoen et al., 1995). Conceptually, as demand for road space increases, you move clockwise through the curve. The regimes are numbered in this manner.

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 (Fig. 4). Beginning with the lowest level of demand, which can occur either at off-peak times or at any time downstream of a bottleneck, Regime 1 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 s volumes. This regime, which has been purported to capture the behavior of traffic with heterogeneous freeway trip lengths, is more often observed in the 10:00 a.m. to 1:00 p.m. period on weekdays and immediately before 10:00 a.m. on Saturdays (Golob et al., 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, despite their similarities on all

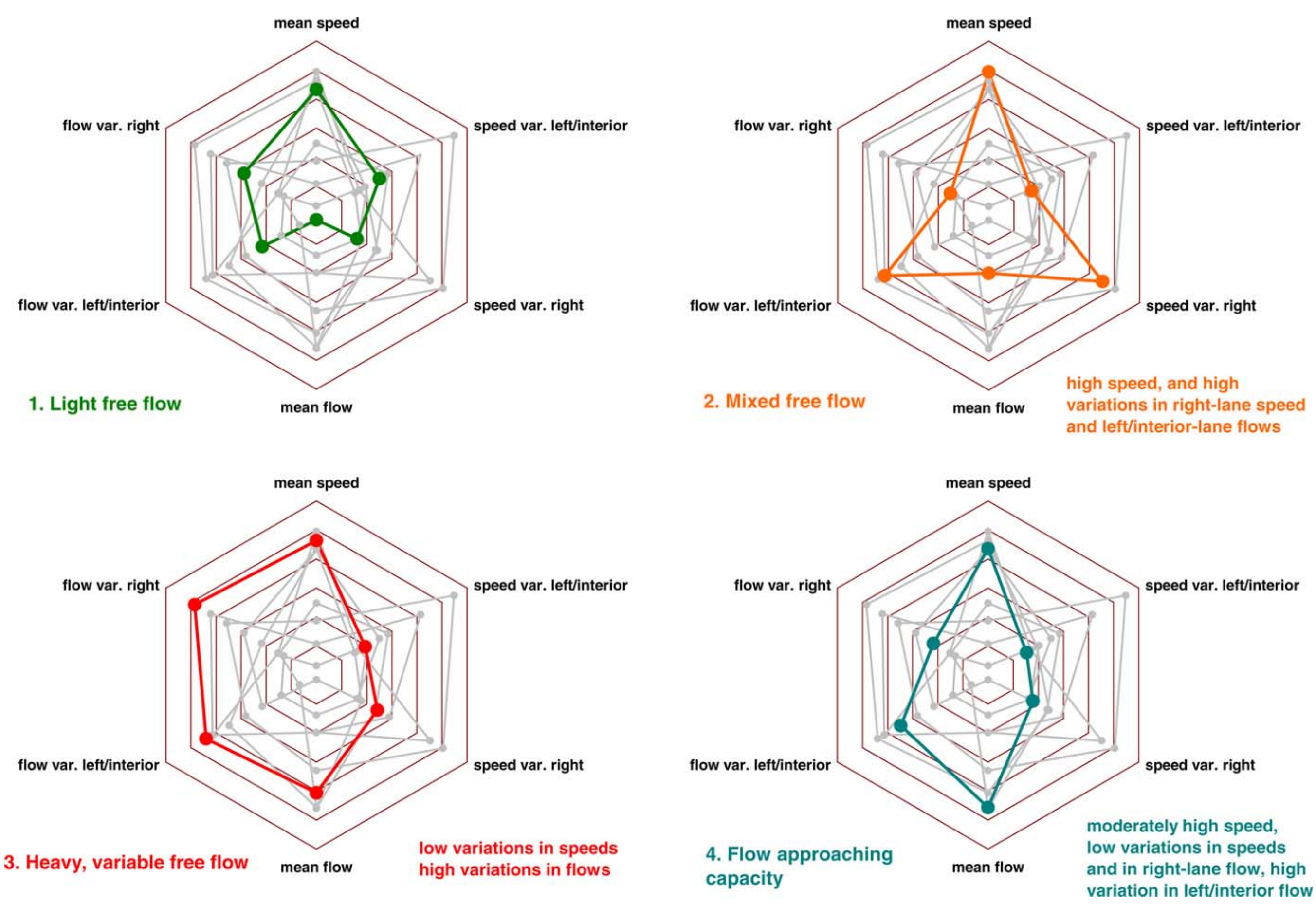

Fig. 2. (a) Radar diagrams of traffic flow regimes for daylight and dry road conditions (Regimes 1–4). (b) Radar diagrams of traffic flow regimes for daylight and dry road conditions (Regimes 5–8).

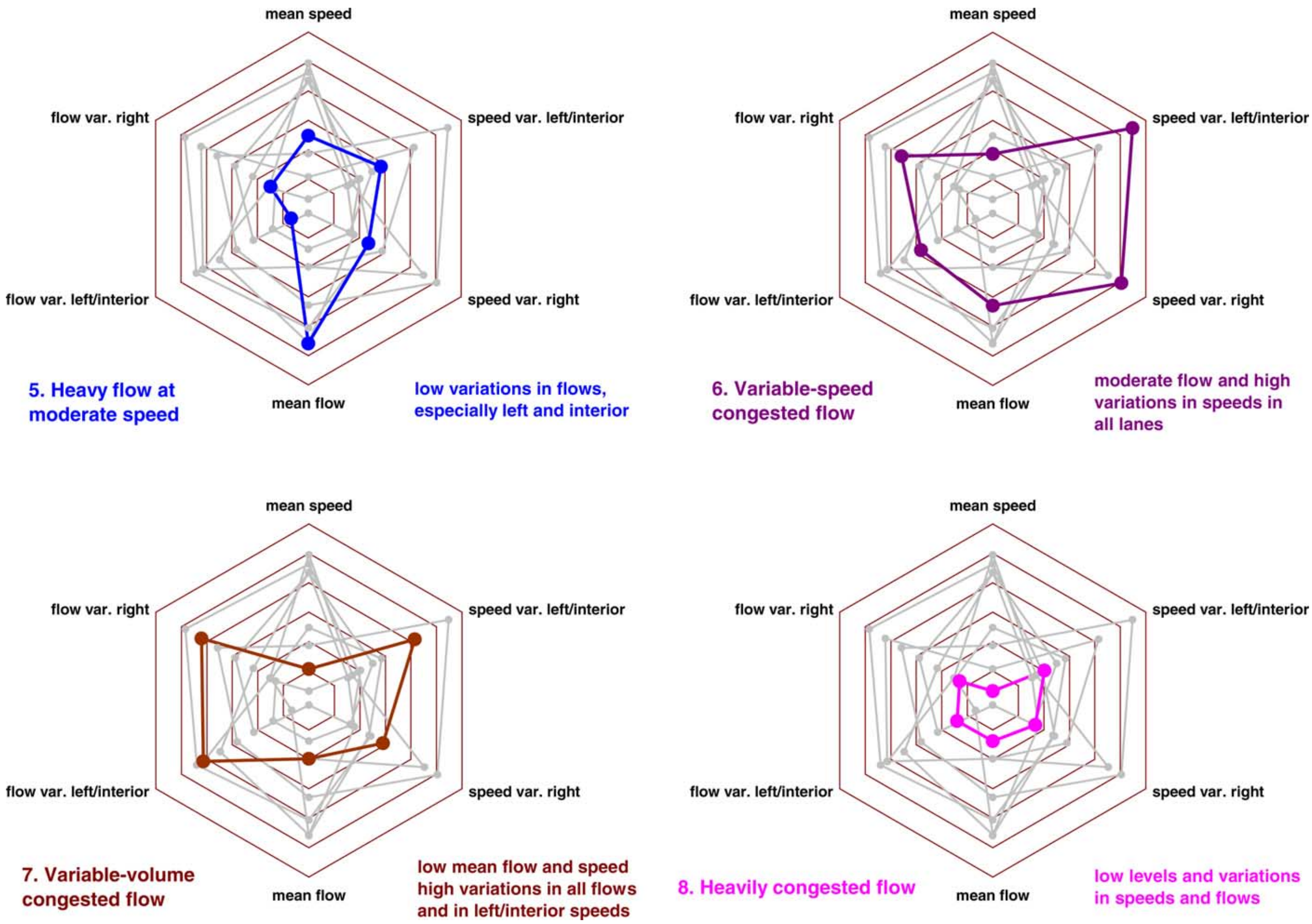

Fig. 2. (*Continued*).

but a single flow dimension, this difference in flow variance leads to differences in the safety profile of these two regimes. Regimes 2–4 exhibit similar high levels of variations in flow in all lanes except the right lane. This common trait might indicate a relatively high degree of lane-changing behavior in moderately heavy to heavy free-flow traffic.

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) portion of the speed–flow diagram. In moving from Regime 4 to Regime 5, Fig. 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,b).

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 6–8). 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 coworkers (Helbing et al., 1999; Helbing and Huberman, 1998; 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 et al. (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.

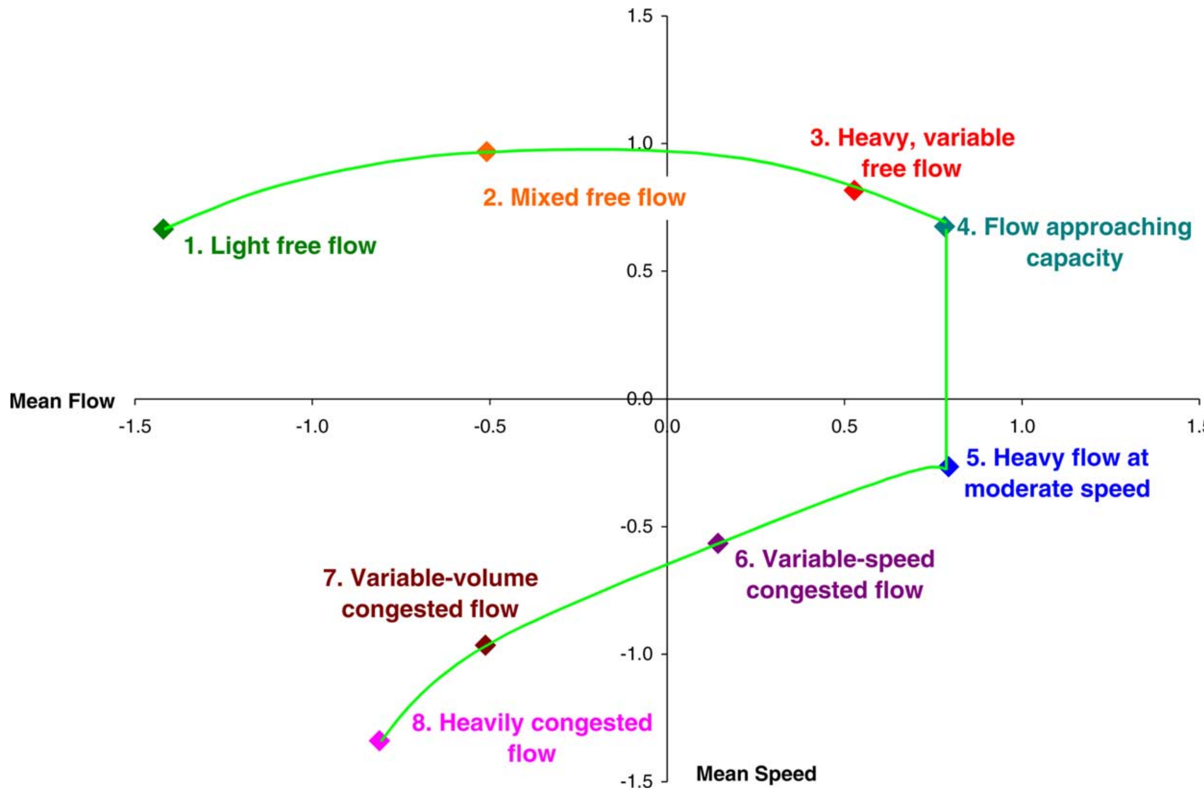

Fig. 3. Speed–flow curve implied by locations of the eight traffic flow regimes in standardized speed–flow space.

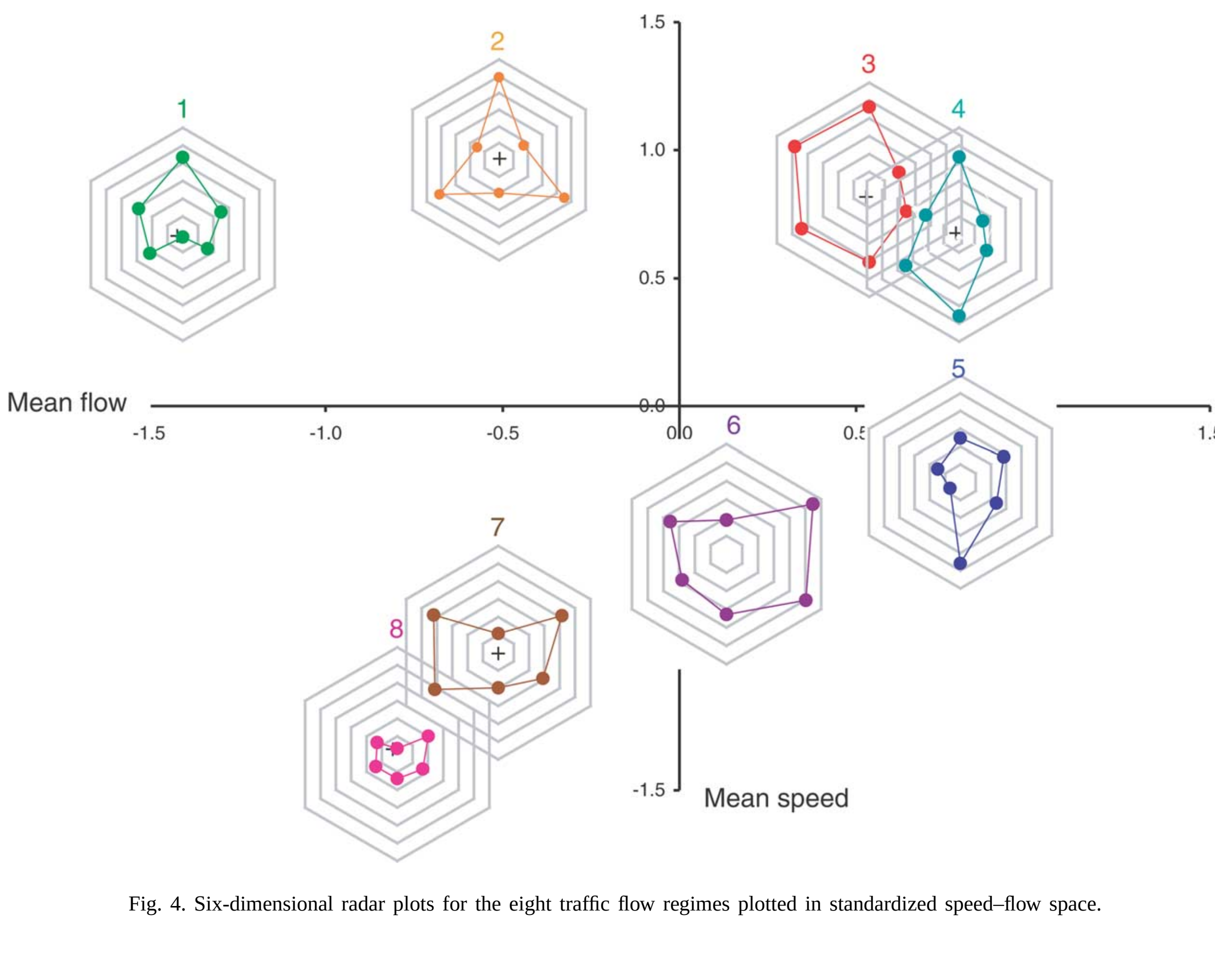

Fig. 4. Six-dimensional radar plots for the eight traffic flow regimes plotted in standardized speed–flow space.

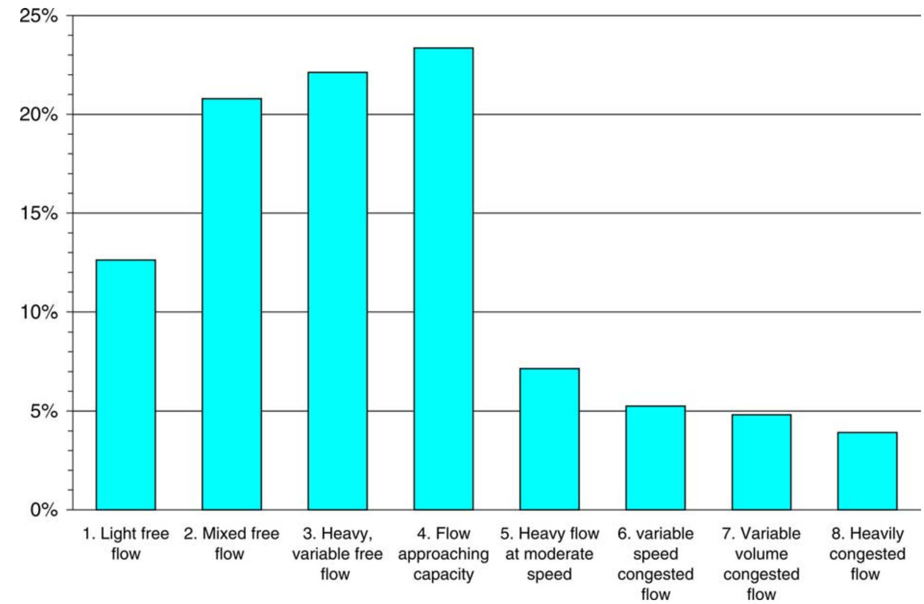

Fig. 5. Distribution of daylight, dry road regimes on six Orange County freeways for 1998 morning weekday peak hours.

5.4. FITS application to 1998 weekday morning peak period data

For purposes of testing 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 a.m. peak hours (6:00 a.m. to 9:00 a.m. inclusive) for all of calendar year 1998 (Golob et al., in press). Because of systematic biases introduced by non-reporting loop stations in 1998, the following is

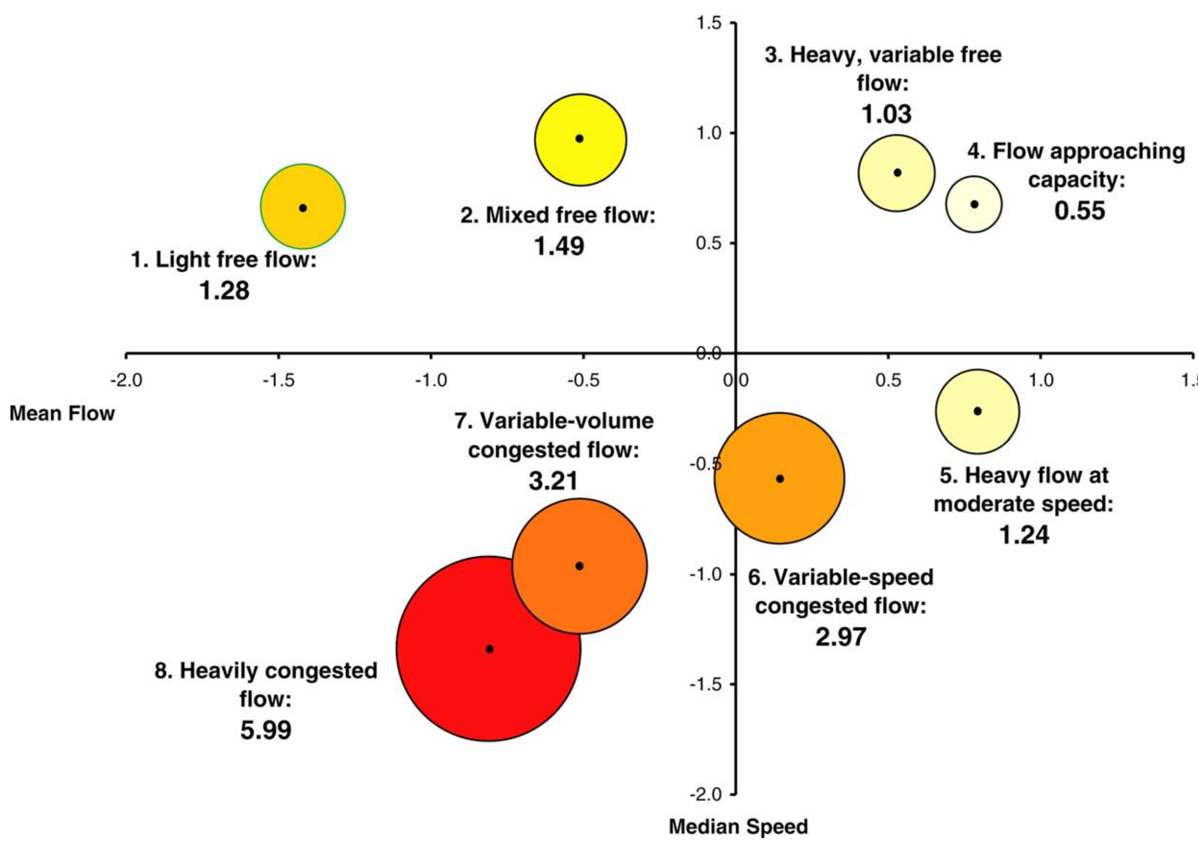

Fig. 6. Estimated total crashes per million vehicle miles of travel for the eight traffic flow regimes during a.m. peak hours, plotted in standardized speed–flow space.

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 a.m. peak period in 1998 is graphed in Fig. 5.

Regimes 1–4 are on the free-flow branch of the speed–flow curve (Figs. 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 a.m. peak hours. The remaining 20% 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 5–8, 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 a.m. peak period crashes across the eight regimes. Total volumes associated with each observed regime occurrence were calculated from total 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 Fig. 6.

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: 1.49 crashes per million vehicle miles for Regime 2 “mixed free flow” versus 3.21 for Regime 7 “variable volume congested flow;” and 0.55 for Regime 4 “flow

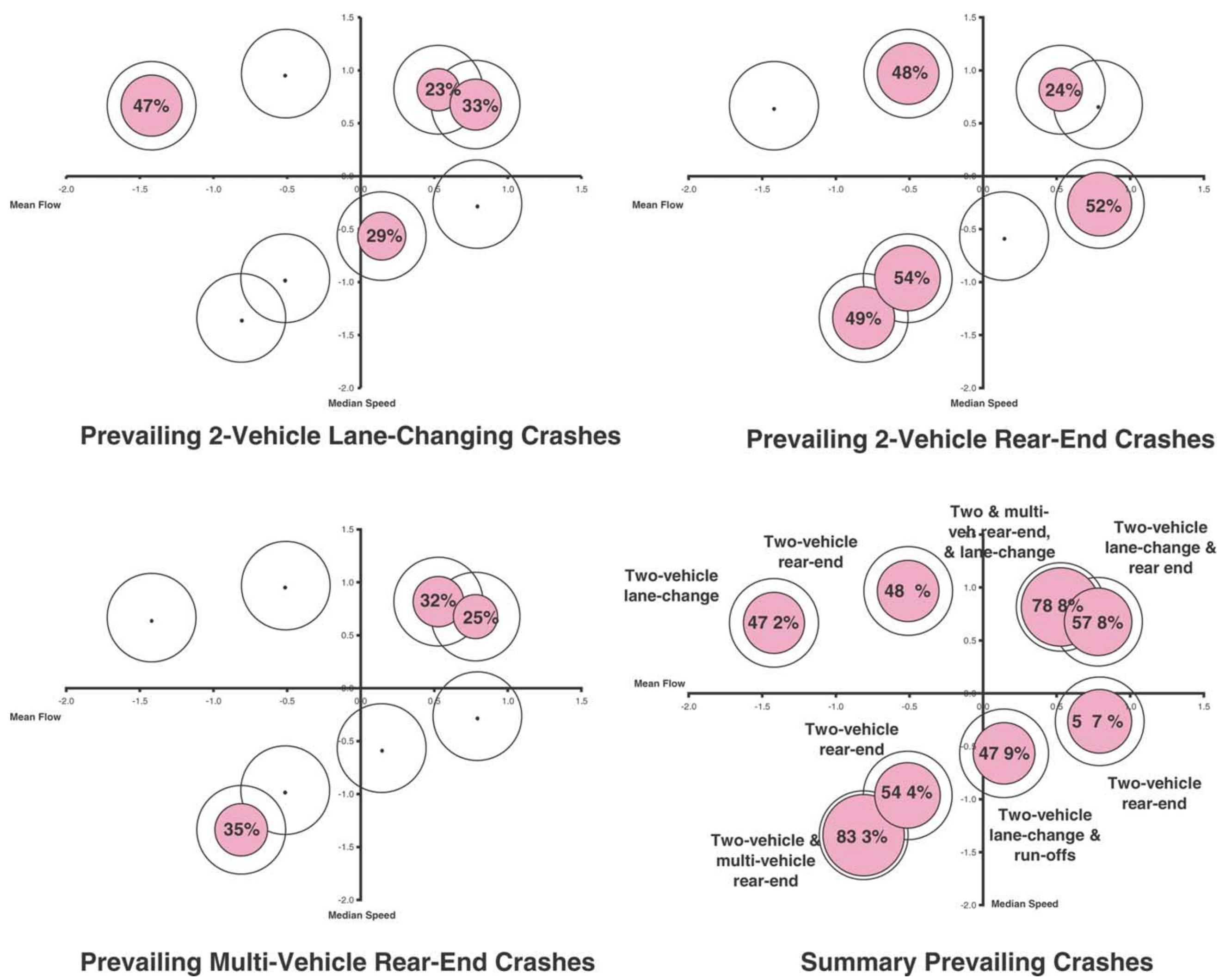

Fig. 7. Prevailing crash types for each regime as a percentage of crashes of all types for that regime during a.m. peak hours, plotted in standardized speed–flow space.

approaching capacity" versus 1.24 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.

5.5. 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 a.m. peak hour traffic revealed that the eight regimes for daylight and dry road conditions were characterized by different patterns of crash types. In Fig. 7, the prevailing crash types for each regime, arranged by mean speed and mean flow, are displayed as a percentage of all crashes within that particular regime; the percentages are displayed against a background (clear) circle representing 100%.

Two-vehicle weaving crashes make up 18.9% of all morning peak period crashes. The graph in the upper-left-hand quadrant of Fig. 7 shows that these two-vehicle weaving crashes are highly concentrated in Regimes 1 "light free flow," 4 "flow approaching capacity," 6 "variable-speed congested flow," and 3 "heavy, variable free flow," where they make up between 23 and 47% of all crashes during the morning peak hours. Two-vehicle rear end crashes make up 40.9% of crashes, but the graph in the upper-right-hand quadrant of Fig. 7 shows that such crashes are more likely to occur when traffic flow is operating under Regime 7 "variable volume congested flow;" Regime 5 "heavy flow at moderate speeds," Regime 8 "Heavily congested flow," or Regime 2 "mixed free flow." Finally, three-or-more-vehicle rear ends make up 25.6% of all morning peak period crashes, but these types of crashes are more prevalent when traffic is operating under Regimes 8 "heavily congested flow," 3 "heavy, variable free flow," and 4 "flow approaching capacity." To further understand these patterns, it is useful to look at the roles of variables that define flow turbulence.

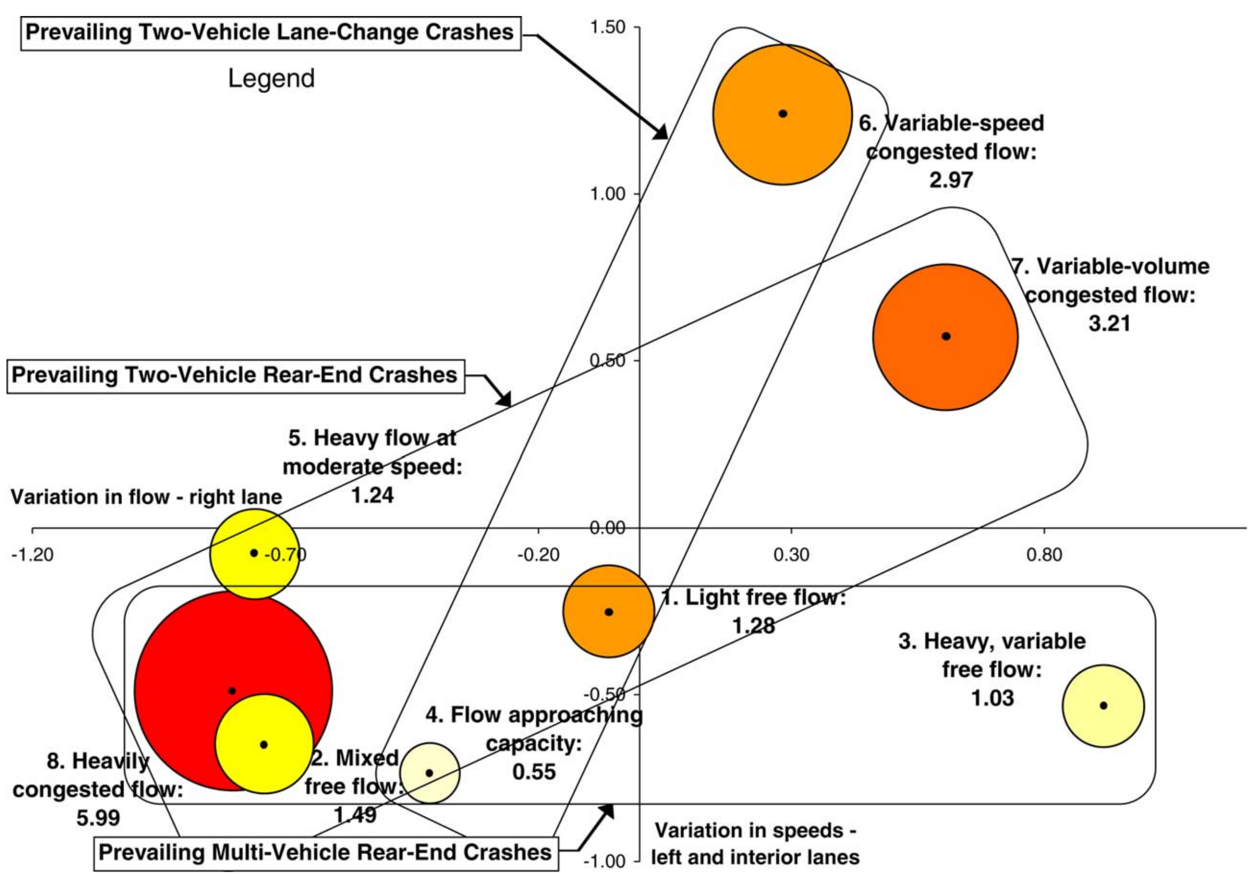

Fig. 8. Estimated total crashes per million vehicle miles of travel by traffic flow regimes plotted in standardized space of (x) variation in flow in right lane vs. (y) variation in speeds in left and interior lanes.

5.6. The influences 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. Results suggest that all of these variables are effective in explaining some aspects of safety. For example, the estimated number of morning peak period crashes per million vehicle miles is plotted in Fig. 8 as a function of variation in flow in the right lane versus variation in speeds in the left and interior lanes. The three rectangles superimposed on Fig. 8 capture the patterns of prevailing crash types for the regimes. These prevailing types were shown according to mean speed and flow in Figs. 7 and 8 represents a different two-dimensional plane passed through the six-dimensional space represented in the radar diagrams of Fig. 2.

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 as defined by these two dimensions. This demonstrates how reducing variations in speed and flow should lead to safer conditions. Fig. 8 also reveals that lane-changing crashes involving two vehicles tend to be distributed anywhere along the y-axis defining variation in speed in the left and interior lanes (that is, they are prevalent under both highly-variable and stable speeds in the left and interior lanes), but these types of crashes prevail only for a range of average variations in flow conditions in the rightmost lane. Conversely, multi-vehicle rear-end crashes tend to be distributed all along the x-axis defining variance in flow in the rightmost lane, but multi-vehicle rear-end crashes are more likely only when speed variations on the rest of the lanes are low. Finally, rear-end crashes involving only two vehicles tend to follow a pattern that is somewhere between these two extremes, occurring primarily under conditions of either relatively high or low variability in both of these two variables.

Viewed from a slightly different perspective—one of passing a plane through the median speed and left- and interior-lane speed variability facet of the radar diagrams (Fig. 9)—we see that the large cluster of crashes in the lower left quadrant of Fig. 8 (low variability in both flow in the rightmost lane and speed in the left and interior lanes) is more distinctly separated into a single cluster of crashes involving two- and multi-vehicle rear end crashes, which are associated with low, relatively stable speeds (and occurring with the highest crash rate), and a collection of different

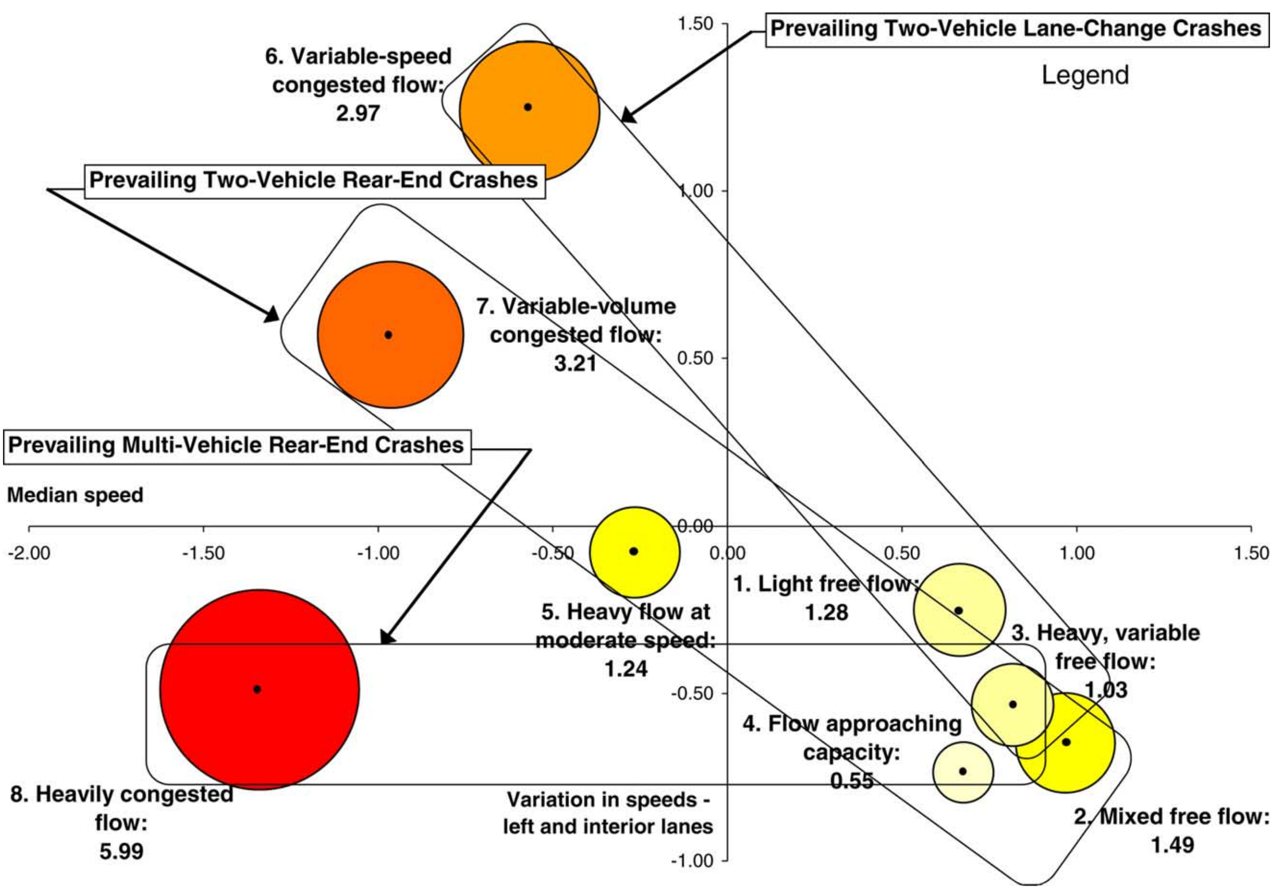

Fig. 9. Estimated total crashes per million vehicle miles of travel by traffic flow regimes plotted in standardized space of (x) median speed vs. (y) variation in speeds in left and interior lanes.

types of crashes, all of which are associated with relatively high, stable speeds. The relatively high crash rates associated with high turbulence in Fig. 8, is restricted primarily to conditions in which the speed is relatively low. Combining the trends shown in the two Figs. 8 and 9, we note that lane-change crashes tend to occur under conditions in which there is the highest variability in speeds (presumably, conditions in which switching lanes could prove advantageous to the driver), while rear-end crashes tend to cluster where there is both somewhat lower variation in speed and at somewhat lower speeds (presumably, stop-and-go conditions, where there is both volatility coupled with only marginal advantage to switching lanes).

6. Conclusions

Our object is to demonstrate the potential of implementing a tool for real-time assessment of the level of safety of any pattern of traffic flow on an urban freeway. Such a tool requires only a stream of 30 s (or similar interval) observations from ubiquitous single inductive loop detectors. This stream is processed to provide a continual assessment of safety, updated every interval, based on central tendencies of flow and speed, and variations in flow and speed for different lanes of the freeway. We are not purporting that the method described here is the only basis for such a tool. Future research is needed to compare the present method with the other approaches that have been recently developed (e.g., Oh et al., 2001; Lee et al., 2003), in order to implement a safety performance evaluation tool that incorporates the best features of each approach.

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 promising application is to forecast the safety implications of proposed projects by evaluating the levels of safety implied by traffic simulation model outputs.

Additional benefits of this course of research are the insights gained in relating accident and traffic flow typologies. Identifying the types of crashes that are most likely to occur under different traffic conditions should aid in identifying treatments aimed at enhancing safety. Identifying where and when on the freeway system these conditions occur should aid in efficiently directing these treatments. Treatments to reduce specific types of accidents might include traffic engineering improvements (signage, lighting, surface treatment, lane re-striping or realignment, barrier adjustments, ramp metering), implementation of intelligent transportation systems (variable message signs, highway advisory radio, information for in-vehicle navigation systems), and enhanced driver education.

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., Metcalfe, A.V., 1999. Speed. Accid. Anal. Prev. 31, 161–168.

Baruya, A., Finch, D.J., 1994. Investigation of traffic speeds and accidents on urban roads. In: Proceedings of Seminar J, PTRC European Transport Forum. PTRC, London, pp. 219–230.

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

Ceder, A., 1982. Relationship between road accidents and hourly traffic flow-II: probabilistic approach. Accid. Anal. Prev. 14, 19–34.

Ceder, A., Livneh, L., 1982. Relationship between road accidents and hourly traffic flow. Accid. Anal. Prev. 14, 19–44.

Chen, C., Petty, K.F., Skabardonis, A., Varaiya, P.P., Jia, Z., 2001. Freeway performance measurement system: mining loop detector data. Transport. Res. Rec. 1748, 96–102.

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

Daganzo, C.F., Cassidy, M.J., Bertini, R.L., 1999. Possible explanations of phase transitions in highway traffic. Transport. Res. Part A 33, 365–379.

Davis, G.A., 2002. Is the claim that 'variance kills' an ecological fallacy? Accid. Anal. Prev. 34, 343–346.

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

Garber, N.J., Ehrhart, A.A., 2000. Effects of speed, flow, and geometric characteristics on crash frequency for two-lane highways. Transport. Res. Rec. 1717, 76–83.

Garber, N.J., Gadiraju, R., 1990. Factors influencing speed variance and its influence on accidents. Transport. Res. Rec. 1213, 64–71.

Garber, N.J., Subramanyan, S., 2001. Incorporating crash risk in selecting congestion-mitigation strategies. Transport. Res. Rec. 1746, 1–5.

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

Golob, T.F., Recker, W.W., 2003. Relationships among urban freeway accidents, traffic flow, weather and lighting conditions. ASCE J. Transport. Eng. 129, 342–353.

Golob, T.F., Recker, W.W., 2004. A Method for relating type of cash to traffic flow characteristics on urban freeways. Transport. Res. Part A 38, 53–80.

Golob, T.F., Recker, W.W., 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., Alvarez, V.M. A tool to evaluate the safety effects of changes in freeway traffic flow. ASCE J. Transport. Eng., in press.

Hall, F.L., Hurdle, V.F., Banks, J.H., 1992. Synthesis of recent work on the nature of speed–flow and flow-occupancy (or density) relationships on freeways. Transport. Res. Rec. 1365, 12–18.

Helbing, D., Hennecke, A., Treiber, M., 1999. Phase diagram of traffic states in the presence of inhomogeneities. Phys. Rev. Lett. 82, 4360–4363.

Helbing, D., Huberman, B.A., 1998. Coherent moving states in highway traffic. Nature 396, 738–740.

Helbing, D., Schreckenberg, M., 1999. Cellular automata simulating experimental properties of traffic flow. Phys. Rev. E 59, R2505–R2508.

Holt, D., Steel, D.G., Tanmer, M., Wrigley, N., 1996. Aggregation and ecological effects in geographically based data. Geogr. Anal. 28, 244–261.

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

Jovanis, P.P., Chang, H., 1986. Modeling the relationship of accidents to miles traveled. Transport. Res. Rec. 567, 42–51.

Kerner, B.S., Rehborn, H., 1996a. Experimental properties of complexity in traffic flow. Phys. Rev. E 53, R4275-–R4278.

Kerner, B.S., Rehborn, H., 1996b. Experimental features and characteristics of traffic jams. Phys. Rev. E 53, R1297-–R1300.

Lee, C., Saccomanno, F., Hellinga, B., 2002. Analysis of crash precursors on instrumented freeways. Transport. Res. Rec. 1784, 1–8.

Lee, C., Hellinga, B., 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, 12–16 January 2003, Washington, DC.

Martin, J.-L., 2002. Relationship between crash rate and hourly traffic flow on interurban motorways. Accid. Anal. Prev. 34, 619–629.

Mensah, A., Hauer, E., 1998. Two problems of averaging arising from the estimation of the relationship between accidents and traffic flow. Transport. Res. Rec. 1635, 37–43.

Michailidis, G., de Leeuw, J., 1998. The GIFI system of descriptive multivariate analysis. Stat. Sci. 13, 307–336.

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

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

Persaud, B., Dzbik, L., 1992. Accident prediction models for freeways. Transport. Res. Rec. 1401, 55–60.

Pushkar, A., Hall, F.L., Acha-Daza, J.A., 1994. estimation of speeds from single-loop freeway flow and occupancy data using cusp catastrophe theory model. Transport. Res. Rec. 1457, 149–157.

Roess, R.P, McShane, W.R. Prassas, E.S., 1998. Traffic Engineering, second ed. Prentice-Hall, Upper Saddle River, NJ.

Sullivan, E.C., 1990. Estimating accident benefits of reduced freeway congestion. J. Transport. Eng. 116, 167–180.

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

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

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

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.