Title

An Analysis of the Salient Characteristics of Truck-Involved Freeway Accidents Using the Method of Log-Linear Modeling

Authors

Golob, Thomas F.
Recker, Will

Publication Date

1987-06-01

UCI-ITS-WP-87-2

An Analysis of the Salient Characteristics of Truck-Involved Freeway Accidents Using the Method of Log-Linear Modeling

UCI-ITS-WP-87-2

Thomas F. Golob [1]
Will Recker [2]

[1] Institute of Transportation Studies
University of California, Irvine

[2] Department of Civil Engineering and
Institute of Transportation Studies
University of California, Irvine

June 1987

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

Submitted to *Journal of Safety Research*. Not for citation or quotation.

ABSTRACT

The method of log-linear modeling is advanced as a procedure in identifying factors that underlie the relative frequency of occurrence of various accident characteristics, such as accident type, location, and severity. The method is centered on the estimation of saturated log-linear models for pairs of accident variables and determination of indices of association between categories of the variables.

Using data drawn from more than 9000 truck-involved accidents that occurred over a two-year period on freeways in three metropolitan counties in Southern California, the method is demonstrated by analyzing accident characteristics both by type as well as by freeway route segment. Specifically, accidents by collision type are analyzed relative to characteristics such as the primary collision factor, the location of the accident, the time period, road conditions, and weather conditions. Differences among 38 specific freeway segments in terms of accident characteristics are also analyzed.

The results of the analyses indicate that the method is a useful tool in uncovering underlying patterns in accident characteristics.

1. OVERVIEW AND SCOPE

The research reported here involves statistical analyses of the characteristics of truck-related accidents that occurred on freeways in three metropolitan counties in Southern California. It is part of a larger study involved with assessing congestion costs associated with truck-related freeway accidents. The analyses are based on data for more than 9000 truck-involved accidents that occurred during the 1983-1984 period. These data were drawn from the TASAS data base maintained by the California Department of Transportation, as described in Section 2.

The methodology, which is described in Section 3, is centered on the discrete multivariate method of log-linear modeling. The analysis involved the estimation of saturated log-linear models for pairs of accident variables, followed by calculation of indices of association between categories of the variables.

The analyses are divided into two categories: accident characteristics-teristics by type of collision, and accident characteristics by freeway route segment. In each category, the objective is to identify underlying patterns in accident characteristics. In Section 4, accidents by collision type are analyzed relative to characteristics such as the primary collision factor, the location of the accident, the time period, road conditions, and weather. In Section 5, statistical models are developed that identify differences among freeway segments in terms of accident characteristics. Thirty-eight specific freeway segments in Southern California are analyzed as an example of how the method can be used to identify roadways with varying accident characteristics. Conclusions are provided in Section 6.

2. THE DATA

The data source was the TASAS (Traffic Accident Surveillance and Analysis System) data base maintained by the California Department of Transportation (Caltrans, 1978). This data base contains all accidents on the State Highway System that involved police investigations at the scene of the accident. For 1983-1984, there were 9508 such accidents involving trucks larger than pickups or panel trucks on 22 freeway routes in Los Angeles, Orange, and Ventura Counties, three adjacent metropolitan counties in Southern California.

The analysis focused on the variables listed in Table 1. All variables are categorical in that there is no preconceived ordering of the categories. The category frequencies are included in Table 1. The overall sample size of 9508 truck-involved accidents over two years was sufficient to satisfy minimum cell size requirements in the cross-classifications of most variable pairs. A general rule for the accuracy of the statistical measures described in Section 3 is that all cells (category pairs) in a cross-classification have at least one observation, and 80 percent of the cells have at least five observations (Cochran, 1954; Haberman, 1978, Vol. I). This was satisfied except in a few situations, which are indicated in the descriptions of the results.

Freeway design, traffic levels, and many other factors that can be broadly defined as freeway conditions are expected to influence the characteristics of truck-involved accidents. The approach in the present study was to divide freeway routes in the case study region into segments within which conditions were relatively homogeneous when compared to differences in conditions between segments. The number of possible segments is limited by the necessity of having sufficient numbers of accidents to

TABLE 1

VARIABLES USED IN THE ANALYSIS

VARIABLE	CATEGORIES	FREQUENCY (n = 9508)
Collision Type	: 1. Sideswipe	: 4092
	: 2. Rear-end	: 2964
	: 3. Broadside	: 456
	: 4. Hit Object	: 1108
	: 5. Overturn	: 272
	: 6. All Other Types	: 616
Primary Collision Factor	: 1. Influence Alcohol	: 353
	: 2. Tailgating	: 263
	: 3. Failure to Yield	: 65
	: 4. Improper Turn	: 903
	: 5. Speeding	: 2786
	: 6. Other Violations (hazardous)	: 4276
	: 7. Other Improper Driving	: 189
	: 8. Not Driver	: 525
	: 9. Unknown	: 136
Generic Location	: 1. Highway	: 7889
	: 2. Ramp (includes connectors)	: 1619
Time Period	: 1. 00:00 - 05:59	: 669
	: 2. 06:00 - 08:59	: 1613
	: 3. 09:00 - 11:59	: 2039
	: 4. 12:00 - 14:59	: 2127
	: 5. 15:00 - 17:59	: 1871
	: 6. 18:00 - 20:59	: 728
	: 7. 21:00 - 23:59	: 438
Terrain	: 1. Flat	: 8057
	: 2. Rolling	: 904
	: 3. Mountainous	: 547
Road Conditions	: 1. No Unusual Conditions	: 9030
	: 2. Holes or Loose Material	: 76
	: 3. Construction	: 253
	: 4. Other Unusual Conditions	: 111
Weather	: 1. Clear	: 7415
	: 2. Cloudy	: 1327
	: 3. Rain or Fog	: 749
Road Surface Condition	: 1. Dry	: 8423
	: 2. Wet	: 987
	: 3. Icy or Otherwise Slippery	: 63
Ramp Direction (Ramp accidents only)	: 1. On-ramp	: 581
	: 2. Off-ramp	: 991
	: 3. Other (scales, etc.)	: 47
Ramp Location (Ramp accidents only)	: 1. Ramp Intersection (exit)	: 451
	: 2. Ramp	: 520
	: 3. Ramp Entry	: 229
	: 4. Intersecting Street	: 419

conduct reliable statistical analyses. Sixteen of the 22 freeway routes had sufficient numbers of accidents, and with the aid of the California Department of Transportation (Caltrans), 38 freeway segments were defined on these 16 routes. These 38 segments are mapped in Figure 1; the specification descriptions of the segments are provided in Table A.1 in the appendix.

3. METHODOLOGY

Log-linear models were used to determine relationships between the categorical variables measuring the characteristics and locations of truck-involved freeway accidents. The variables analyzed included the type of collision and the seven other accident characteristic variables listed in Table 1, plus the 38-category route segment variable depicted in Figure 1.

Log-linear models are designed to identify structure in the relationship between two or more categorical variables. In the following, the relationship between freeway route segment and collision type is used as an example in describing the modeling approach. The objective in this example is to determine whether or not there are differences among the types of collision that occur on specific routes. Given that a certain number of truck-involved accidents occur on a specific route, and given that there is a known distribution among types of collisions for all routes, is there a significant interaction between route and collision type that indicates that the distribution of collision types might be different for the route in question? The present approach to this question involves estimating a saturated log-linear model for the contingency table represented by the cross-tabulation of route segment by collision type, a 38 by 6 contingency table.

A test of independence between route and accident type involves whether or not the entries in the contingency table can be considered the result of

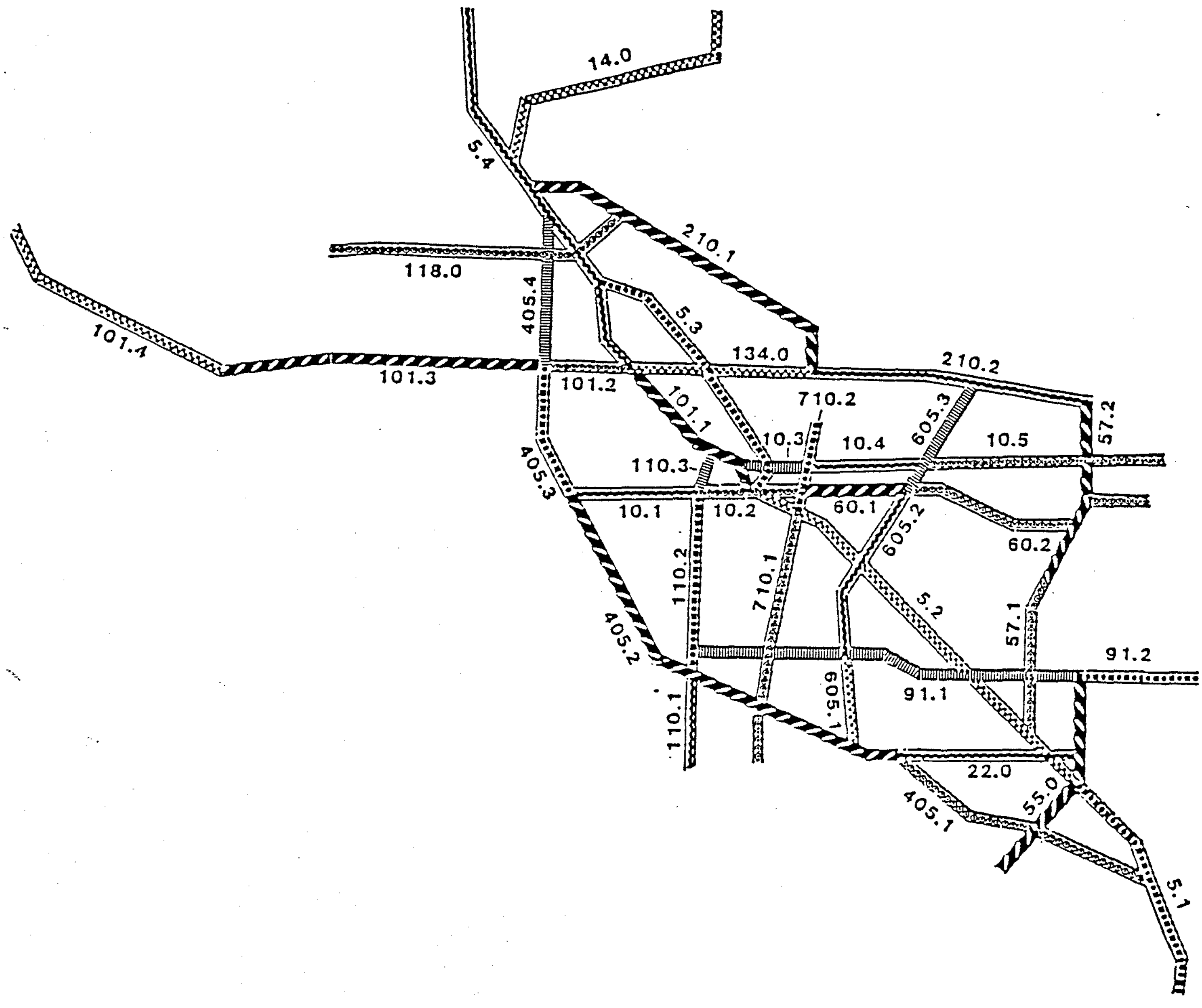

FIGURE 1

THIRTY-EIGHT FREEWAY SEGMENTS

a random process which depends only upon the expected number of accidents for each route (for all types) and the expected number of accidents by collision type (for all routes). Defining

n_{ij} = observed number of accidents of type j on route i,

the hypothesis of independence between route and type involves comparison of each n_{ij} with the randomly expected numbers, m_{ij}, given by the product of the sample size times the probability that an observation falls into the i^{th} row times the probability that the observation falls into the j^{th} column:

$$m_{ij} = N\,(n_{i.} / N)\,(n_{.j} / N) = n_{i.}\,n_{.j} / N \qquad (1)$$

where

$n_{i.}$ = total accidents of all types on route i,

$n_{.j}$ = total accidents of type j on all routes, and

N = total accidents (size of the sample)

The most common measure of association between n_{ij} and m_{ij} is given by:

$$\chi^2 = \sum_{ij} (n_{ij} - m_{ij})^2 / m_{ij} \qquad (2)$$

which has the known chi-square distribution for hypothesis testing under the usual assumption of multinomial distributions and sufficient expected cell frequencies.

Taking the natural logarithm of both sides of equation (1),

$$\ln m_{ij} = \ln N + \ln (n_{i.} / N) + \ln (n_{.j} / N) \qquad (3)$$

the test of independence for the (i,j) cell of the contingency table translates into a test of whether or not there is a statistically significant β_{ij} term in the log-linear equation

$$\ln n_{ij} = \alpha + \beta_i + \beta_j + \beta_{ij} \qquad (4)$$

where α accounts for the sample size (grand mean), β_i accounts for the route effect, β_j accounts for the accident type effect, and β_{ij} represents the interaction between route i and type j.

It is logical that the probability process underlying the accident counts is Poisson, the usual assumption for stochastic processes of Equation (4) is then assumed to include a Poisson error term and represents a saturated log-linear model (Plackett, 1962 and Birch, 1963). (Extensive overviews of general families of such models are provided in Haberman, 1974; Plackett, 1974; Bishop, et al., 1975; Goodman, 1978; Haberman, 1978; and McCullagh and Nelder, 1984.)

Estimation of the parameters of Equation (4) and their error terms is effectively accomplished using maximum likelihood methods (Nelder and Wedderburn, 1972; Bock and Yates, 1973; Haberman, 1973a). T-statistics, given by the ratios of the β_{ij} parameter estimates to the standard errors of the estimates, are used to determine which of the combinations of route (i) and accident type (j) have interaction terms that are significantly different from zero under the assumption of Poisson distributions.

The log-likelihood ratio statistic, given by

$$L^2 = 2 \sum_{ij} n_{ij} \log (n_{ij} / m_{ij}) \qquad (5)$$

has a distribution that is asymptotically chi-square (Cochran, 1954; Haberman, 1978), and can be used to test the hypothesis that the structure of the contingency table can be represented by a log-linear model with some coefficients set to zero.

A direct measure of the degree to which any route-accident type combination (in general, any cell i,j in a contingency table) varies from its expected value is given by the standardized chi-square residual for the cell:

$$r_{ij} = [\, n_{ij} - (n_{i.}\, n_{.j} / N)] / \sigma_{ij} \qquad (6)$$

where σ_{ij} is the standard error for the cell, given by

$$\sigma_{ij} = (n_{i.}\, n_{.j} / N)^{1/2} \qquad (7)$$

This residual is distributed as a standard normal variate under the probability assumptions and sufficient cell frequencies. The residuals are employed in the present analyses as indices of variation from expected values. They are listed for variable combinations (or interaction terms) that have significant coefficients in the log-linear models. They are not residuals associated with the fits of the log-linear models, which are exact because there are as many parameters as there are cells in the contingency tables ("saturated" models). The standardized residuals merely are one measure of the degree of variance between actual counts and counts expected under the assumption of independence between the variables.

The log-linear models for this example, as well as for the remaining associations tested in this analysis, were implemented using the GLIM (Generalized Linear Interactive Modeling) program (Nelder and Wedderburn, 1972; Baker and Nelder, 1978; McCullagh and Nelder, 1983). (Log-linear models are also available in most commonly used statistical analysis packages such as SAS, SPSS-X and BMDP.) The results of these analyses are described below.

4. ACCIDENT CHARACTERISTICS BY COLLISION TYPE

The relationship between collision type and primary collision factor was analyzed by estimating a log-linear model on the 54-cell cross-classification table for these two variables. The chi-square statistic (Equations (1) and (2)) for this cross-classification was 4925.4 with 40 degrees-of-freedom, indicating a very strong relationship between the variables (the critical chi-square value being 55.8 at the $p = .05$ level). The log-linear model (Equation (4)) for the table had 35 individual cell terms (β_{ij}) that were significantly different from zero at the $p = .05$ level. The standardized residuals (Equations (6) and (7)) are listed for these 35 cells in Table 2. (In Table 2 and all subsequent tables, standardized residuals are shown only for cells with significant log-linear model coefficients, all other cells are left blank.)

The residual values in Table 2 indicate relationships that are largely as expected. However, they do reveal some associations that can be useful in explaining accident causality. For instance, rear-end collisions have a strong relationship not only with tailgating driving behavior, but also with alcohol, speeding, and other improper driving. The strongest associations are

TABLE 2

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY PRIMARY COLLISION FACTOR

COLLISION TYPE	PRIMARY COLLISION FACTOR								
	INFLUENCE ALCOHOL	TAIL-GATING	FAILURE TO YIELD	IM-PROPER TURN	SPEEDING	OTHER VIOLA-TIONS	OTHER IMPROPER DRIVING	NOT DRIVER	UNKNOWN
SIDESWIPE	-4.2	-10.0		+9.2	-25.9	+24.2	-2.8	-10.1	
REAR-END	+6.2	+16.9	-4.3	-11.1	+29.8	-21.5	+2.8	-8.6	-2.5
BROADSIDE		-2.4	+19.2	+2.7					
HIT OBJECT		-3.0	-2.8	3.8	+3.2	-8.5		+13.9	
OVERTURN		-2.4		-2.3	+7.1	-5.9		+7.7	
OTHER TYPES	-2.9	-3.6		-5.4	-5.7			+22.2	
Sample Sizes:	353	263	65	903	2786	4276	189	525	136

for speeding (positively associated with rear-end collisions, negatively associated with sideswipes), other violations (conversely with the same collision types), and the not-driver factor (associated with other types of collisions).

There are substantial differences among collision types in terms of proportional occurrences at highway and ramp locations. The chi-square value of association for these two variables is 508.7 with 5 degrees-of-freedom, again indicating strong interactions (the critical value being 11.1). The residuals shown in Table 3 for significant terms in the log-linear model reveal that all collision types except the "other" category have varying

TABLE 3

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY LOCATION: HIGHWAY VERSUS RAMP

COLLISION TYPE	ACCIDENT LOCATION	
	HIGHWAY	RAMP
SIDESWIPE		-4.5
REAR-END		-7.3
BROADSIDE		+10.5
HIT OBJECT		+8.6
OVERTURN		+12.7
OTHER TYPES		

highway versus ramp splits, with rear-end and sideswipe collisions being located predominately at highway sites, and with overturns, broadsides, and hit-objects being located at ramp sites. The strongest associations are for overturns and broadside collisions at ramp locations.

Collision type and time period (in seven categories) are strongly related with a chi-square value of 186.1 with 30 degrees-of-freedom (compared to a critical value of 43.8). Significant time-of-day patterns for the collision types are shown in Table 4. Hit-object collisions tend to occur from midnight to 6:00 AM, while sideswipes do not. Rear-end collisions appear to be particularly a morning rush hour phenomenon, while overturns occur more frequently than expected by chance in the 9:00 PM to midnight period. The strongest association involves the occurrence of hit-object collisions during the midnight to 6:00 AM period. With regard to a related time of occurrence

TABLE 4

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY TIME PERIOD

COLLISION TYPE	TIME PERIOD						
	MIDNIGHT to 5:59 AM	6:00 AM to 8:59 AM	9:00 AM to 11:59 AM	NOON to 2:59 PM	3:00 PM to 5:59 PM	6:00 PM to 8:59 PM	9:00 PM to 11:59 PM
SIDESWIPE	-5.5						
REAR-END		+3.5					
BROADSIDE					-2.2		
HIT OBJECT	+7.3	-3.8			-3.5		
OVERTURN							+2.7
OTHER TYPES							
Sample Sizes:	669	1631	2039	2127	1871	728	438

variable, there were no significant differences among the collision types in terms of their occurrences over days of the week.

The relationships between collision types and roadway terrain are shown in Table 5. The chi-square is 101.5 with 10 degrees-of-freedom, indicating another highly significant overall relationship between the variables (the critical value being 18.3). Sections of Routes 14, 405, 118, and 5 are classified in the TASAS highway records as being "mountainous," and almost all routes have both "flat" and "rolling" sections. As shown in Table 5, only the "mountainous" category exhibits differences in the distribution of collision types: there are relatively more rear-end and overturn collisions on mountainous sections, and relatively fewer sideswipes.

TABLE 5

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY TERRAIN AT SITE

COLLISION TYPE	TERRAIN		
	FLAT	ROLLING	MOUNTAINOUS
SIDESWIPE			-6.5
REAR-END			+4.8
BROADSIDE			
HIT OBJECT			
OVERTURN			+4.1
OTHER TYPES			
Sample Sizes:	8057	904	547

With regard to collision type by road conditions, the significant chi-square is 75.9 with 15 degrees-of-freedom (compared to a critical value of 25.0). As shown in Table 6, hit-object collisions are more prevalent in areas of construction or "other unusual conditions." Collisions in the residual "other" category occur in areas being classified as having holes or loose material, and this is the strongest association in the table.

TABLE 6

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY ROAD CONDITIONS

COLLISION TYPE	ROAD CONDITIONS			
	NO UNUSUAL CONDITIONS	HOLES OR LOOSE MATERIAL	CONSTRUCTION	OTHER UNUSUAL CONDITIONS
SIDESWIPE		-2.2	-3.1	
REAR-END				
BROADSIDE				
HIT OBJECT			+3.2	+3.6
OVERTURN				
OTHER TYPES		+4.1		
Sample Sizes:	9030	76	253	111

With regard to weather conditions (Table 7), broadside, hit-object, and "other" types of collisions occur relatively more frequently in conditions of rain or fog, and these are the strongest associations in the table. Conversely, sideswipe collisions are less likely to occur during rainy or foggy conditions. The overall relationship between collision type and weather is measured by a chi-square value of 201.3 with 10 degrees-of-freedom, another highly significant value (the critical value being 18.3).

TABLE 7

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY WEATHER CONDITION

COLLISION TYPE	WEATHER		
	CLEAR	CLOUDY	RAIN OR FOG
SIDESWIPE			-6.2
REAR-END		-2.8	
BROADSIDE			+7.2
HIT OBJECT	-3.1		+7.6
OVERTURN			
OTHER TYPES			+3.8
Sample Sizes:	7415	1327	749

Following up the relationships involving weather, the relationship between collision type and the surface condition variable are listed in Table 8. The chi-square for this relationship is 248.0 with 10 degrees-of-freedom (compared to the critical value of 18.3). Hit-object and "other" collisions occur relatively more often under both wet and icy or slippery road surface conditions. However, broadsides are related to wet roads only, and overturns are related to icy or slippery conditions. The largest standardized deviations from randomly expected frequencies are associated with the occurrences of truck-involved hit-object and broadside collisions on wet freeways.

TABLE 8

STANDARDIZED RESIDUALS FOR COLLISION TYPES WITH SIGNIFICANTLY HIGH (+) OR LOW (-) FREQUENCIES BY ROAD SURFACE CONDITIONS

COLLISION TYPE	ROAD SURFACE CONDITION: DRY	WET	ICY OR OTHERWISE SLIPPERY
SIDESWIPE		-6.4	-3.1
REAR-END			
BROADSIDE		+7.7	
HIT OBJECT	-3.2	+8.2	+4.3
OVERTURN			+3.1
OTHER TYPES		+2.7	+2.4
Sample Sizes:	8423	987	63

5. ACCIDENT CHARACTERISTIC BY FREEWAY SEGMENT

Log-linear models can be used to identify roadway segments with varying accident characteristics. The following case study presents an example of the types of information that may be provided using this approach. The case study uses 38 freeway segments in Southern California. The segment locations are shown in Figure 1. The example focuses on several accident characteristics: collision type, relative concentration of ramp involvement, entry vs. exit incidents, location on the ramp, and time of occurrence. The analyses seek to identify freeway segments that tend to have either particularly severe or light association with various characteristics.

The relationship between freeway segment (38 categories) and collision type (6 categories) is indexed by a significant chi-square value of 558.8 with 185 degrees-of-freedom (compared to a critical value of 224.6). (Although there are 228 cells in the cross-classification of these two variables, only 37 (or 16.2 percent) of these cells had expected frequencies of fewer than five accidents, so the chi-square statistic is a fairly accurate indication of overall association.) There were 34 significant interaction terms in the log-linear model of the cross-classification, and the standardized residuals for the category combinations corresponding to these terms are given in Table 9. Because collision type captures an array of other accident characteristics, as described in the previous section, the results in Table 9 are depicted in the following figures on a collision-type basis (corresponding to a single column in Table 9).

The freeway segments with proportionally high or low concentrations of <u>sideswipe collisions</u> are segments 10.2 and 5.3 (Table 10). These two adjacent segments are highly congested and serve downtown Los Angeles. The segments with relatively low concentrations of sideswipes are 5.4, 14.0, 605.7, and 57.1. The first two of these segments are located at the northern edge of the metropolitan area, and overall congestion levels on the four segments are substantially lower than the average for all segments.

The indicated relationship between the proportion of sideswipe accidents and traffic congestion was further investigated through correlation analyses involving annual average daily traffic measures at locations along each of the thirty-eight freeway segments. The best indicator of the percentage of side-swipe accidents was found to be the maximum annual average daily traffic (AADT) per lane at all locations along a freeway segment; the correlation between this indicator variable and percent sideswipe collisions was 0.44,

TABLE 9

ADJUSTED RESIDUALS FOR FREEWAY-SEGMENT/ COLLISION TYPE COMBINATIONS WITH SIGNIFICANT CELL EFFECTS IN THE LOG-LINEAR MODEL

	COLLISION TYPE					
ROUTE	SIDESWIPE	REAR-END	BROADSIDE	HIT OBJECT	OVERTURN	OTHER
5.1				+2.3		
5.2					-2.8	
5.3	+2.6					
5.4	-5.8			+3.9	+4.5	
10.1						
10.2	+2.8			-2.2		
10.3					+2.2	
10.4						-2.4
10.5						
14.0	-4.5				+4.4	+4.2
22.0						
55.0						
57.1	-2.2		+2.1			
57.2						
60.1						
60.2						
91.1						+2.2
91.2						
101.1				-2.5		-3.5
101.2		+2.5				
101.3						
101.4		-2.3		+2.5		+3.0
110.1						
110.2				-2.6		
110.3		+2.8				
118.0			+2.6			
134.0						
210.1						
210.2						
405.1						
405.2			-2.5	-2.4		
405.3		+3.4	-2.3			
405.4.						
605.1	-2.3					
605.2						
605.3				+2.3		
710.1			-2.2	+4.5		
710.2		-2.2				

TABLE 10

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: SIDESWIPE COLLISIONS

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
10.2	54.1	14.0	13.0
5.3	49.6	5.4	25.4
		605.1	25.4
		57.1	33.3

(OVERALL AVERAGE = 43.2 PERCENT)

which is significant at the p = .003 level. The maximum AADT per lane for the two segments with high sideswipe incidences was 105,500 for the Route 10/ Route 60 segment and 91,500 for the Route 5 segment. The maximum AADT per lane for each of the three segments with low incidence of sideswipes was 18,000 for Route 14; 42,700 for the Route 5 segment; and 49,000 for the Route 57 segment. The median maximum AADT per lane for all 38 segments was approximately 54,300.

Rear-end collisions represent relatively high percentages of all truck-involved accidents on segment 110.3 and intersecting segments 405.3 and 101.2 (Table 11). These are three of the heaviest traveled freeway segments in the area. In contrast, such collisions represent relatively low percentages of accidents on segments 101.4 and 710.2, segments that are much less traveled.

TABLE 11

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: REAR-END COLLISIONS

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
110.3	46.4	710.2	15.6
405.3	45.2	101.4	22.9
101.2	45.1		

(OVERALL AVERAGE = 31.3 PERCENT)

The percentage of accidents that were rear-end collisions was found to be significantly related to an indicator of traffic level: the mean average annual daily traffic (AADT) at all locations along a freeway segment. The correlation between these two variables was 0.39 for the 38 segments, which is significant at the p = .008 level. Two of the three freeway segments with high incidences of rear-end collisions had the highest levels of mean AADT among all segments (206,300 for the Route 405 segment and 198,200 for the Route 101 segment); the third segment also had the very high mean AADT level of 163,800. Correlations with maximum AADT and maximum AADT per lane were not significant. Thus, relatively high percentages of rear-end collisions are associated with higher levels of overall traffic, while high percentages of sideswipe collisions are associated with high levels of traffic per lane at key locations.

With regard to statistically significant concentrations of <u>broadside collisions</u>, two segments were found to exhibit high concentrations and three segments had statistically significant low concentrations: segments 118.0 and 57.1 were high; and two adjacent segments of Route 405, 405.2 and 405.3, and segment 710.1 were low in terms of broadside collisions (Table 12).

Such collisions frequently occur on ramps (Table 3), and an investigation of the characteristics of the ramps for each freeway segment revealed that the percent of broadside collisions is a direct function of the percent of ramps that are components of diamond interchanges. Approximately 38 percent of all ramps in the study area on which truck-involved accidents occurred are diamond-interchange ramps, but 77 percent of the ramps on Route 118 and 64 percent of the ramps on the southern segment of Route 57 <u>are</u> diamond-interchange ramps. Conversely, only 26 percent, 10 percent, and 16

TABLE 12

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: <u>BROADSIDE COLLISIONS</u>

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
118.0	11.8	405.3	1.1
57.1	8.0	405.2	2.3
		710.1	2.9

(OVERALL AVERAGE = 4.8 PERCENT)

percent of the ramps on the three freeway segments with significantly low proportions of broadside collisions were diamond-interchange ramps. However, there are other freeway segments with very low or high percentages of diamond-interchange ramps, and further research is required to identify other possible causes of these collisions.

High concentrations of hit-object collisions are found on segments 710.1, 5.4, 101.4, 605.3, and 5.1 (Table 13). Low concentrations are found on intersecting segments 405.2, 110.2, 10.2, and 101.1. The percentage of hit-object collisions was found to be the complement of the percent of rear-end collisions in terms of a relationship with traffic intensity: the correlation between average annual daily traffic at all points along a freeway segment and percent hit-object collisions was -.60, which is highly

TABLE 13

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: HIT-OBJECT COLLISIONS

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
605.3	18.3	110.2	6.0
5.4	17.6	101.1	6.8
710.1	17.3	10.2	7.2
101.4	16.9	405.2	7.9
5.1	16.0		

(OVERALL AVERAGE = 11.5 PERCENT)

significant. It can be inferred that the total number of hit-object collisions on an urban freeway segment does not vary as much with traffic levels as do the numbers of rear-end and sideswipe collisions. Consequently, the percentage of hit-object collisions is an inverse function of traffic levels while rear-end and sideswipe collisions are direct functions.

Segments with outstanding concentrations of overturn accidents are listed in Table 14. Segments 14.0, 5.4, and 10.3 have a high concentration, while segment 5.2 has a significantly lower percentage of overturns. Two of the three segments with high percentages of overturns (Route 14 and Route 5/Route 170) are located in mountainous and rolling terrain. The third segment, Route 10 adjacent to downtown Los Angeles, is built primarily with roadways on separate structures with relatively steep ramps. Further investigation is required to isolate other potential causal effects of such types of truck-involved accidents.

TABLE 14

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: OVERTURNS

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
14.0	10.2	5.2	1.4
10.3	7.8		
5.4	6.2		

(OVERALL AVERAGE = 2.8 PERCENT)

Finally, high percentages of "other" types of collisions are found on segments 14.0, 101.4, and 91.1, while low percentages are found on segments 10.5 and 101.1 (Table 15). As in the case of hit-object collisions, there is an inverse relationship between the percentage of other types of collisions and average traffic levels on a segment, the correlation between the percentage variable and average annual daily traffic being -.65. However, the high incidence of other types of collisions on the Route 91 segment demonstrates that other factors are involved as well, because the Route 91 segment has a greater than median level of average AADT.

Freeway segments with varying splits between highway and ramp accident locations are listed in Table 16. Segments with relatively high concentrations of ramp accidents are intersecting segments 10.3 and 710.2,

TABLE 15

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: OTHER COLLISION TYPES

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
14.0	17.3	101.1	1.5
101.4	11.4	10.5	1.8
91.1	8.8		

(OVERALL AVERAGE = 6.5 PERCENT)

TABLE 16

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: RAMP ACCIDENTS

SIGNIFICANTLY HIGH CONCENTRATIONS		SIGNIFICANTLY LOW CONCENTRATIONS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
		101.1	8.4
10.3	41.2	101.3	9.2
710.2	34.4	60.1	9.3
22.0	31.0	405.3	9.6
605.3	29.0	5.4	9.7
57.1	26.3	110.2	10.8
10.1	25.9		
405.1	23.5	5.3	12.2
10.5	22.6		

(OVERALL AVERAGE = 16.8 PERCENT)

605.3 and 10.5, 22.0 and 405.1, and segments 10.1 and 57.1, the majority of which are east and south of downtown Los Angeles. Segments with relatively low concentrations of ramp accidents (or, high concentrations of highway accidents) are 101.1 and 101.3, 60.1, 5.3 and 5.4, 110.2, and 405.3, all of which are west or north of downtown Los Angeles.

Focusing further on the characteristics of ramp accidents, three freeway segments were found to have relatively high concentrations of on-ramp versus off-ramp accidents (Table 17). The overall split was 36 percent on-ramp versus 61 percent off-ramp (and 3 percent "other," such as truck scales and rest areas). However, these three segments, 605.2, 5.3, and 405.2, had from

TABLE 17

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: ON-RAMP VS. OFF-RAMP ACCIDENTS

SIGNIFICANTLY HIGH CONCENTRATIONS OF RAMP ACCIDENTS

HIGH CONCENTRATIONS OF ON-RAMP ACCIDENTS		HIGH CONCENTRATIONS OF OFF-RAMP ACCIDENTS	
Segment	Percent of All Collisions	Segment	Percent of All Collisions
605.2	63.1	101.3	90.9
5.3	51.2		
405.2	50.0		

(OVERALL SPLIT = 36.0 PERCENT ON RAMP / 61.0 PERCENT OFF RAMP / 3.0 PERCENT OTHER)

50 percent to 63 percent on-ramp accidents. In contrast, segment 101.3 had fewer than 10 percent on-ramp (over 90 percent off-ramp) accidents.

A completely different set of freeway segments were outstanding in terms of locations of ramp accidents, where four locations were distinguished: ramp entry, ramp itself, exit intersection, and intersecting street (Table 18). Two segments had relatively high percentages of accidents located at ramp entries, 10.4 and 10.2, both of which serve the immediate downtown Los Angeles area. One segment, 710.1, had a high percent of accidents on ramps themselves. Finally, four segments had high concentrations of accidents on intersecting streets: 91.2 and 91.1, 5.1, and 57.1, all of which are at least partially in Orange County in the southern portion of the metropolitan area.

TABLE 18

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: BY RAMP LOCATIONS

SIGNIFICANTLY HIGH CONCENTRATIONS OF RAMP ACCIDENTS

AT RAMP ENTRIES		ON RAMPS THEMSELVES		ON INTERSECTING STREETS	
Segment	Percent	Segment	Percent	Segment	Percent
10.4	37.5	710.1	54.3	91.2	70.0
10.2	28.3			5.1	53.3
				57.1	42.9
				91.1	40.8
OVERALL AVERAGE:	14.4	OVERALL AVERAGE:	32.0	OVERALL AVERAGE:	25.7

The final accident characteristic investigated by freeway segment was the time period during which an accident occurred. Seven time periods were distinguished, and freeway segments were found which were distinguished on five of these time periods (Table 19). Three adjacent segments to the northwest of downtown Los Angeles had relatively high concentrations of accidents in the early morning hours (midnight to 6:00 AM): 101.1, 5.4 and 14.0; all of which are major truck routes north from Los Angeles. The 5.4 segment also exhibits a high percent of accidents in the 9:00 PM to midnight period. Two segments, 57.1 and 10.5, have high percentages of accidents during the morning peak hours. Two segments, 405.3 and 5.2, have high percentages in the 9:00 AM to noon period. Finally, the three segments

TABLE 19

FREEWAY SEGMENTS WITH PROPORTIONS OF TRUCK ACCIDENTS SIGNIFICANTLY DIFFERENT FROM EXPECTED NUMBERS, AS DETERMINED BY LOG-LINEAR MODELS OF CONTINGENCY TABLES: TIME OF DAY

SIGNIFICANTLY HIGH CONCENTRATIONS OF RAMP ACCIDENTS

00:00-05:59 HOURS		06:00-08:59 HOURS		09:00-11:59 HOURS		12:00-14:59 HOURS		21:00-23-59 HOURS	
Segment	Percent	Segment	Percent	Segment	Percent	Segment	Percent	Segment	Percent
14.0	26.5	57.1	25.8	405.3	30.9	110.1	42.4	5.4	14.4
5.4	19.4	10.5	22.6	5.2	25.0	110.2	31.2		
101.1	10.8					110.3	32.7		
OVERALL AVERAGE:	7.0	OVERALL AVERAGE:	17.2	OVERALL AVERAGE:	21.4	OVERALL AVERAGE:	22.4	OVERALL AVERAGE:	4.6

comprising the entire length of Route 110 (110.1, 110.2, and 110.3) exhibit high concentrations of accidents in the noon to 3:00 PM period; this is the major harbor access route. No segments have significantly high or low concentrations of accidents during the afternoon peak hours or during the 6:00 PM to 9:00 PM period.

These and other similar analyses could be used in conjunction with data associated with corresponding freeway operational characteristics (e.g., traffic volume, congestion measures, geometric design) to identify potential causes of (and, possibly, solutions to) prevalent accident characteristics.

6. CONCLUSIONS

The method of log-linear modeling is potentially a powerful tool in identifying factors that underlie the relative frequency of occurrence of various accident characteristics, such as accident type, location, and severity. The method can be used to obtain a direct measure of the degree to which any accident characteristic represented in a standard contingency table varies from its expected value in the absence of interaction effects.

In the application presented in this paper, the method was used to associate accident characteristics with type of collision and to identify freeway segments on which various accident categories were more prevalent than expected. The results indicated substantial differences between the types of collisions that tend to occur at ramp locations and those more evident along the mainline freeway. The analysis was also able to uncover significant differences among the factors associated with the types of collision and to associate other characteristics, such as weather and road conditions, with particular collision types. The application of the method in the analysis of accident characteristics by freeway segment revealed several freeway segments that were particularly susceptible to certain types of accidents.

Some roadway characteristics, particularly overall traffic levels, were found to explain the pattern of freeway-segment results. However, the present research stops short of a thorough investigation of potential causal factors that distinguish freeway segments in terms of the characteristics of the accidents that occur on them.

Directions for further research include a multivariate analysis to relate the identified freeway-segment accident characteristics to roadway characteristics such as geometric design, shoulder provisions, and traffic patterns. The log-linear model residuals calculated in the present research could be used directly in future analysis. It would also be useful to contrast truck-involved accident characteristics with car-only accident characteristics as a control sample. The overall objectives of such future research would be to enhance truck-related safety considerations in freeway design.

ACKNOWLEDGEMENTS

The reported research was sponsored in part by the AAA Foundation for Traffic Safety and the California Department of Transportation. The authors wish to thank Messrs. Donald Juge and Goro Endo of Caltrans District Seven, Los Angeles, for their support and encouragement. Keith Rutherford prepared the graphics, and Cynthia Wenks prepared the manuscript. The authors alone are responsible for any errors of omission or commission.

REFERENCES

Baker, R.J. and J.A. Nelder. The GLIM System, Release 3. Royal Statistical Society, Numerical Algorithms Group, Oxford, 1978.

Birch, M.W. Maximum likelihood in three-way contingency tables. Journal of the Royal Statistical Society, Series B, 25: 220-233, 1963.

Bishop, Y.M.M., S.E. Fienberg, and P.W. Holland. Discrete Multivariate Analysis. M.I.T. Press, Cambridge, MA, 1975.

Bock, R.D. and G. Yates. MULTIQUAL: Log-Linear Analysis of Nominal or Ordinal Qualitative Data by the Method of Maximum Likelihood. International Education Services, Chicago, 1973.

Caltrans. Manual of Traffic Accident Surveillance and Analysis System. State of California, Business and Transportation Agency, Department of Transportation, Division of Operations, Office of Traffic Engineering, Sacramento, 1978.

Cochran, W.G. Some methods for strengthening the common ^{2} tests. Biometrics, 10: 417-451, 1954.

Goodman, L.A. Analyzing Qualitative/Categorical Data. Addison Wesley, London, 1978.

Goodman, L.A. Some multiplicative models for the analysis of cross-classification data. In L. LeCam, et al., eds., Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, Vol. I: 649-496. University of California Press, Berkeley, 1972.

Haberman, S. J. The analysis of residuals in cross-classification tables. Biometrics, 29: 205-220, 1973b.

Haberman, S.J. The Analysis of Frequency Data. University of Chicago Press, Chicago, 1974.

Haberman, S.J. Analysis of Qualitative Data (2 Volumes). Academic Press, New New York, 1978.

Haberman, S.J. Log-linear models for frequency data: sufficient statistics and likelihood equations. Annals of Statistics, 1: 617-632, 1973a.

Institute of Transportation Studies, University of California, Irvine. Analysis of Caltrans Major Incident Response Data on Freeway Incidents in Los Angeles, Orange, and Ventura Counties Involving Large Trucks, 1983-1985. Technical Memorandum TM-TAS-1, Irvine, CA, 1986.

References (cont'd)

McCullagh, P. and J.A. Nelder. Generalized Linear Models. Chapman and Hall, London, 1983.

Nelder, J.A. and R.W.M. Wedderburn. Generalised linear models. Journal of the Royal Statistical Society, Series A, 135: 370-384, 1972.

Plackett, R.L. The Analysis of Categorical Data. Griffin, London, 1974.

Plackett, R.L. A note on interactions in contingency tables. Journal of the Royal Statistical Society, Series B, 24: 162-166, 1962.

APPENDIX

CODE	DESCRIPTION OF FREEWAY SEGMENT
5.1	Santa Ana (I-5): Orange-San Diego Co. line to Jct. 55 (Costa Mesa Fwy.)
5.2	Santa Ana (I-5): Jct. 55 to Jct. 10/60 (Pomona Fwy.)
5.3	Santa Ana-Golden State (I-5) Jct. 10/60 to Jct. 170 (Hollywood Fwy.)
5.4	Golden State-Hollywood (SR 170) (I-5): Jct. 101/134 to Jct. 170/5 to Los Angeles-Kern Co. line
10.1	Santa Monica (I-10): Jct. 405 (San Diego Fwy.) to Jct. 110 (Harbor Fwy.)
10.2	Santa Monica (I-10)-Pomona (SR 60): Jct. 110 to Jct. 710 (Long Beach Fwy.)
10.3	San Bernardino (I-10): Jct. 101 to Jct. 710 (Long Beach Fwy.)
10.4	San Bernardino (I-10): Jct. 710 to Jct. 605 (San Gabriel R. Fwy.)
10.5	San Bernardino (I-10): Jct. 605 to Los Angeles-San Bernardino Co. line
14.0	Antelope Valley (SR 14): Begin Jct. 5 (Golden State Fwy.) to Los Angeles-Kern Co. line
22.0	Garden Grove (SR 22): Jct. 405 (San Diego Fwy.) to end, Jct. 55 (Costa Mesa Fwy.)
55.0	Costa Mesa (SR 55): Begin Fwy. southwest of 73 to end, Jct. 91 (Riverside Fwy.)
57.1	Orange (SR 57: Begin Jct. 5/22 to Orange-Los Angeles Co. line
57.2	Orange (SR 57)-Pomona (SR-60)-Foothill (I-210): Co. line to Jct. 30
60.1	Pomona (SR 60): Jct. 710 (Long Beach Fwy.) to Jct. 605 (San Gabriel R. Fwy.)
60.2	Pomona (SR 60): Jct. 605 to L.A.-San Bernardino Co. line (excluding overlap with Rte. 60)
91.1	Artesia-Redondo Beach-Riverside (SR 91): Begin Fwy. near Jct. 110 (Harbor Fwy.) to Jct. 55
91.2	Riverside (SR 91): Jct. 55 to Orange San Bernardino Co. line
101.1	Santa Ana-Hollywood (US 101): Begin, Jct. 5 (Golden State Fwy.) to Jct. 134/170
101.2	Ventura (US 101): Jct. 134/170 to Jct. 405 (San Diego Fwy.)
101.3	Ventura (US 101): Jct. 405 to Los Angeles-Ventura Co. line
101.4	Ventura (US 101): Los Angeles-Ventura Co. line to Ventura-Santa Barbara Co. line
110.1	Harbor (I-110): Begin Fwy. near Jct. 47 to Jct. 405 (San Diego Fwy.)
110.2	Harbor (I-110): Jct. 405 to Jct. 10 (Santa Monica Fwy.)
110.3	Harbor (I-110): Jct. 10 to Jct. 101 (Hollywood Fwy.)
118.0	Simi Valley-San Fernando Valley (SR 118): Begin Fwy. in Ventura Co. to Jct. Rte. 210
134.0	Ventura (SR 134): Jct. 101/170 (Hollywood Fwy.) to Jct. 210 Foothill Fwy.)
210.1	Foothill (I-210): Begin Jct. 5 (Golden State Fwy.) to Jct. 134 (Ventura Fwy.)
210.2	Foothill (I-210): Jct. 134 to end, Jct. 30
405.1	San Diego (I-405): Begin Jct. 5 (Santa Ana Fwy.) to Jct. 22 (Garden Grove Fwy.)
405.2	San Diego (I-405): Jct. 22 to Jct. 10 (Santa Monica Fwy.)
405.3	San Diego (I-405): Jct. 10 to Jct. 101 (Ventura Fwy.)
405.4	San Diego (I-405): Jct. 101 to end, Jct. 5 (Golden State Fwy.)
605.1	San Gabriel River (I-605): Begin Jct. 22 to Jct. 91 (Artesia Fwy.)
605.2	San Gabriel River (I-605): Jct. 91 to Jct. 60 (Pomona Fwy.)
605.3	San Gabriel River (I-605): Jct. 60 to end, Jct. 40 (Foothill Fwy.)
710.1	Long Beach (I-710): Begin Jct. 1 to Jct. 5 (Santa Ana Fwy.)
710.2	Long Beach (I-710): Jct. 5 to break in route, Valley Blvd., north of 10

TABLE A.1: DESCRIPTION OF FREEWAY SEGMENTS