CONCLUSION

The final product of this entire process is a list of projects ranked according to relative cost-effectiveness. By applying budget constraints to this listing of projects, a yearly or multiyear program is devised. The process explained presents a simple technique for facilitating the resource-allocation decision. It is designed to be applicable to all local highway organizations regardless of their size or sophistication.

ACKNOWLEDGMENT

The research presented in this paper was funded in part by the Federal Highway Administration. We also wish to express appreciation to the members of the Oakland County TSM Committee, particularly to Deborah Schutt, Jo Ann Soronen, and Steve Ron, for their assistance and encouragement in this research.

REFERENCES

1. J.I. Taylor and H.T. Thompson; Pennsylvania Transportation Institute. Identification of Hazardous Locations. FHWA, 1977. NTIS: PB-283 925/6ST.
2. Problem Identification Manual for Traffic Safety Programs. National Highway Traffic Safety Administration, 2 vols., HS-802 084 and HS-802 085, Dec. 1976.
3. D.L. Renshaw and E.C. Carter. Identification of High-Hazard Locations in the Baltimore County Road-Rating Project. TRB, Transportation Research Record 753, 1980, pp. 1-8.
4. Roy Jorgensen Associates. Methods for Evaluating Highway Safety Improvements. NCHRP, Rept. 162, 1975, 150 pp.
5. W.F. McFarland and others. Assessment of Techniques for Cost-Effectiveness of Highway Accident Countermeasures. Texas Transportation Institute, College Station, Jan. 1979.
6. A Procedure for the Analysis of High Accident Locations. Department of Civil Engineering, Wayne State Univ., Detroit, MI, Dec. 1976.
7. J.L. Graham and others. Identification, Analysis and Correction of High Accident Locations. Midwest Research Institute, April 1976.
8. Wayne County TSM Plan. Wayne County TSM Committee, Detroit, MI, May 1981.
9. Oakland County TSM Plan. Oakland County TSM Committee, Pontiac, MI, May 1981.
10. G. Smith and others. TSM Project Evaluation Process. Southeast Michigan Council of Governments, Detroit, MI, Unpublished Rept., Aug. 1981.
11. Highway Capacity Manual. HRB, Special Rept. 87, 1965, 411 pp.
12. Interim Materials on Highway Capacity. TRB, Circular 212, Jan. 1980, 276 pp.

Publication of this paper sponsored by Committee on Planning and Administration of Transportation Safety.

Analysis of Accidents in Traffic Situations By Means of Multiproportional Weighted Poisson Model

R. HAMERSLAG, J.P. ROOS, AND M. KWAKERNAAK

This article describes a model that enables traffic engineers to get insight into the factors that influence the occurrence of accidents. This model has a multiplicative form and describes how the expected number of accidents depends on road and traffic characteristics. Because of the input of observations where no accidents occurred, a logarithmic transformation to linearize the model was impossible without biasing the estimates considerably. By introducing the maximum likelihood estimation theory, a model was developed that also analyses situations where no accidents occur. This method was first applied successfully in 1974 for the analysis of accidents on Dutch polderroads. This article also describes the results obtained by the method from a study that tries to establish a relation between road and traffic characteristics on one hand and the safety of cyclists and moped riders on the other. Influencing factors are (a) motor car, moped, and cycle traffic flows; (b) width of cycle lane and median width; (c) access roads to houses; (d) type of road surface of the cycle lanes; and (e) parking bays and bus stops. A further application is given by the study of interurban car traffic. Daily traffic flows proved to be the most important variable, followed by the presence of obstacles and intersections and crossings of various kinds.

Traffic accidents are caused by errors of judgment on the part of road users or by defects in vehicles. The occurrence of accidents is related to the psychological characteristics of the traffic participants as well as to the physical characteristics under which they take part in traffic. These physical characteristics are, for instance, the weather conditions (e.g., fog or slipperiness), the light or dark period of the day, and the road characteristics. One of the tasks of the traffic engineer is to examine whether the accident rate can be lowered by improving the traffic situation.

The occurrence of accidents can be analyzed by means of mathematical models. Regression analysis is often used; sometimes analysis of variance and factor analysis are also used to ascertain the effect of road and traffic characteristics (1-3). Some have used linear regression. Often, a multiplicative model is made linear (4,5).

The use of multiple linear regression implicitly assumes that the observation results are distributed normally. This assumption is not very realistic since the analysis is specifically concerned with traffic situations in which few accidents occur. The probability that the number of accidents would become negative is not negligible in that case.

The drawback of an erroneous assumption with respect to the sampling distribution is even greater in the use of the multiplicative model linearized by a logarithmic transformation. The logarithm of zero is not defined, and a zero observation can therefore not be included in the investigation. The zero observations are sometimes omitted from the analysis. This seems undesirable because traffic situations where no accidents occur are of a very real importance. Other devices are sometimes used; for instance, a small number (e.g., 0.5) may be added to all observations (6-8). Such a pretreatment of observations can greatly affect the estimate and is therefore undesirable.

For the method we propose it is not required either. This contribution deals with the weighted multiproportional Poisson model and illustrates this method with some applications. The number of accidents is used as the dependent variable, whereas the accident rate is not. In fact, the rate depends on the dimensions used. The lengths of road segments where accidents have been observed lead to the introduction of weighted models. Accidents are related to road and traffic characteristics by means of a multiplicative or multiproportional model. The accidents are assumed to be Poisson distributed.

MULTIPROPORTIONAL POISSON MODEL

The multiproportional Poisson model is based on two assumptions. First, it is assumed that accidents are Poisson distributed with some expected value. Subsequent accidents are not correlated and the time interval between two subsequent accidents has a negative-exponential distribution. Second, the expected number of accidents (μ) is multiplicative (i.e., the product of the effects of independent variables). This model is based on the analysis of higher-order cross-classifications to test whether the factors (roadway and traffic characteristics) of influence are independent. In the use of the accident model, many roadway and traffic characteristics must be included simultaneously in the analysis. The multiplicative model introduced here is a logical continuation of the analysis of cross-classifications that contain one or two roadway features; thus, all detailed information available may be analyzed. Oppe (9) gives some theoretical and experimental justification for the use of a multiplicative model.

In addition, some road segments, which have a certain combination of factors, may differ considerably in length (L) from other segments, which have a different combination of factors. The experimental design is not balanced. As a consequence of the governmental road design policy, these are combinations of road and traffic characteristics that do not exist (e.g., roads that have a small lane width but a high car volume). Moreover, observations from long road segments are more reliable than those from short segments. The literature on this subject pays little attention to the analysis of such weighted cross-classifications (6,10). A computer package like BMDP does not contain software for the analysis of weighted cross-classifications.

The presence of weight factors is a vital difference between the method being proposed and the standard log-linear analysis of cross-classifications. Note that the ratio between the number of accidents and the weighting factor is not suitable for analysis since, in that case, the analysis will depend on the dimensions used (11).

The form of the model is

$$\mu_{klmn} = a_k \cdot b_l \cdot c_m \cdot d_n \ldots \; L_{klmn} \qquad (1)$$

where

μ_{klmn} = expected number of accidents in case the explanatory variables belong to the categories k,l,m, and n;
L_{klmn} = length of the segment that belongs to the categories k,l,m, and n (if necessary, weighted by period of observation);
$a_k, b_l, c_m, d_n \ldots$ = coefficients (estimate is indicated by ˆ);
a,b,c,d = factors (characteristics of the road and traffic situation); and
k,l,m, and n = classes with k = 1,2,3,4,...; l = 1,2,3,4,...; m = 1,2,3, 4,...; and n = 1,2,3,4,....

Interactions can also be taken into account. This means that the influence of several independent variables together differs from that of each separate independent variable.

Since it is possible to multiply the coefficients a_k by 100 and to divide the b_l coefficients by 100 without affecting the number of accidents, a normalization is used. The coefficients are not unique. The ratios between the coefficients of any factor are unique. In performing the computations this complication is taken into account. The influence of the traffic volume can be estimated by means of one of the traffic coefficients. It is also possible to include the traffic volume directly as an independent explanatory variable, if so required. In the latter case, L_{klmn} becomes equal to the product of volume, length, and observation period.

ESTIMATION EQUATIONS

The coefficients in the accident model are estimated on the basis of the maximum likelihood method. Maximization of the likelihood gives the estimation equations. As indicated above, the nature of the occurrence of an accident is a Poisson process. Consequently, the probability of y_{klmn} accidents at an expected value μ_{klmn} is given by the equation

$$\Pr[y_{klmn}] = [\exp(-\mu_{klmn}) \cdot \mu_{klmn}{}^{y_{klmn}}]/y_{klmn}! \qquad (2)$$

The numbers of accidents (y_{klmn}) are assumed to be independent for all combinations of k,l,m,n,.... As a result, the value of the log-likelihood function (λ) becomes

$$\lambda = \sum_k \sum_l \sum_m \sum_n \ln \Pr[y_{klmn}] \qquad (3)$$

In Equation 1 the coefficients should be chosen in such a way that the log-likelihood has a maximum value. Substitution of Equations 1 and 2 in Equation 3 gives the log-likelihood function:

$$\lambda = \sum_k \sum_l \sum_m \sum_n [-a_k \cdot b_l \cdot c_m \cdot d_n \ldots L_{klmn} + y_{klmn} \cdot \ln(a_k b_l c_m d_n \ldots L_{klmn}) - \ln(y_{klmn}!)] \qquad (4)$$

The maximum value of the log-likelihood is found by determining the first partial derivative for each of the coefficients and by equating it to zero:

$$\partial\lambda/\partial\hat{a}_k = \sum_l \sum_m \sum_n (-\hat{b}_l \cdot \hat{c}_m \cdot \hat{d}_n \ldots L_{klmn}) + \sum_l \sum_m \sum_n (y_{klmn}/\hat{a}_k) = 0; \quad \forall k \qquad (5a)$$

It is also (equivalently) true that

$$\partial\lambda/\partial\hat{b}_l = 0, \forall l \qquad (5b)$$

$$\partial\lambda/\partial\hat{c} = 0, \quad \forall m \qquad (5c)$$

$$\partial\lambda/\partial\hat{d}_n = 0; \quad \forall n \qquad (5d)$$

A set of nonlinear equations is developed, with which the coefficients are determined.

$$\hat{a}_k = y_{k\ldots}/\sum_l \sum_m \sum_n \hat{b}_l \hat{c}_m \hat{d}_n \ldots L_{klmn}; \quad \forall k \qquad (6)$$

$$\hat{b}_l = y_{.l..}/\sum_k \sum_m \sum_n \hat{a}_k \hat{c}_m \hat{d}_n \ldots L_{klmn}; \quad \forall l \qquad (7)$$

$$\hat{c}_m = y_{..m.}/\sum_k \sum_l \sum_n \hat{a}_k \hat{b}_l \hat{d}_n \ldots L_{klmn}; \quad \forall m \qquad (8)$$

$$\hat{d}_n = y_{...n} / \sum_k \sum_l \sum_m \hat{a}_k \hat{b}_l \hat{c}_m \ldots L_{klmn}; \quad \forall n \quad (9)$$

In these formulas,

$$\sum_l \sum_m \sum_n y_{klmn} = y_{k...}; \quad \forall k \quad (10a)$$

$$\sum_k \sum_m \sum_n y_{klmn} = y_{.l..}; \quad \forall l \quad (10b)$$

$$\sum_k \sum_l \sum_n y_{klmn} = y_{..m.}; \quad \forall m \quad (10c)$$

$$\sum_k \sum_l \sum_m y_{klmn} = y_{...n}; \quad \forall n \quad (10d)$$

$y_{k...}$, $y_{.l..}$, $y_{..m.}$, and $y_{...n}$ are the observed marginal frequency distributions of accidents. The coefficients are determined by an iterative method in accordance with the Gauss-Seidel principle.

The method being proposed can be modified, if necessary; for example, the Poisson distribution of the accidents could be replaced with some other distribution (gamma, Erlang) if the empirical data would indicate so.

STATISTICAL TESTING

The estimators of the coefficients are stochastic variables. Each of these stochastic variables has a probability distribution, a mean, and a standard deviation. The smaller the standard deviation of the estimator, the more reliable a coefficient is considered to be. It is examined by testing whether certain assumptions concerning the parameters of a distribution (comprised in the null hypothesis) can be rejected in favor of the alternative hypothesis. Since a multiproportional model is being used here, it should be examined whether the ratio between the coefficients of each set of classes per factor (a_k/a_1, c_m/c_1, $d_n/d_1, \ldots$) differs significantly from one.

The variation in the coefficients can be determined by means of the matrix of the second derivatives of the log-likelihood function (λ). The negative expectation of the inverse of this matrix gives (asymptotically) the variance-covariance matrix. The square root of the value of the diagonal elements of this variance-covariance matrix gives the estimated standard deviation in the coefficients. The probability distribution of the ratios (a_k/a_1, c_m/c_1, $d_n/d_1, \ldots$) is skew. As the values of the coefficients are positive integer numbers, values smaller than or equal to zero cannot occur. The natural assumption to make in testing is that the estimated coefficients are log-normally distributed. It has been ascertained by Monte Carlo simulation that this assumption is very useful (12,13). Because the procedure of drawing random numbers requires lengthy calculations, the method with the second derivative is used. The normalization is done in such a way that only the estimated values of the normalized coefficients greater than one occur.

SELECTION OF FACTORS

In some studies we estimated the effects of a large number of roadway characteristics. The results of the simultaneous estimation can then be supported by some simple strategies. Depending on the problem, a distinction can be made between the various roadway characteristics. These are selected on the basis of the hypotheses that are to be analyzed. As a tool in selecting roadway features, the likelihood-ratio test statistic (G^2) is used [see, for instance, Bishop and others (6)].

$$G^2 \overset{\text{def}}{=} 2[\lambda^* - \lambda(\mu)] = 2 \Sigma\, y_{klmn} \ln (y_{klmn}/\mu_{klmn})$$

(sum over all the observations)

where

$\lambda(\mu)$ = value of the log-likelihood function with estimated coefficients ($\hat{a},\hat{b},\hat{c},\hat{d},\ldots$);

λ^* = highest attainable value of the log-likelihood; this value of the log-likelihood function is attained if the model results become equal to the observation results; and

G^2 = chi-square distributed if the number of observations is sufficiently large (i.e., asymptotic).

The value of G^2 is calculated for each separate roadway characteristic. It can be examined by testing whether the roadway feature in question contributes significantly to the explanation. This method can also be continued for combinations of roadway characteristics. Table 1 presents an example of this methodology. Average daily traffic volume, obstacle distance, and pavement, for instance, are significant. The legally permitted maximum speed and the gradient are not.

In this way the roadway characteristics can be classified according to their explanatory strength. It would be incorrect to regard these (simple) analyses as definitive, since the effects of all other factors are ignored and the observations are classified solely on the basis of the one factor considered. Although these simple analyses indicate the relative importance of the various factors, the weighted multiproportional Poisson model, which takes account of many factors simultaneously, must be considered decisive.

The test statistic G^2 can also be used when several factors are considered. In broad outline the procedure is as follows: After two factors have been considered and analyzed separately, this is repeated for the two factors together. The effect of the two factors together is then compared with the sum of the effects of the two factors taken separately. If, for instance, the effect of the two factors together is found to be considerably smaller than the sum of the two separate effects, correlation between these factors is obvious--they explain, in an (almost) identical way, the occurrence of accidents.

COMPARISON OF OBSERVED AND ESTIMATED RESULTS

The value of the application of the model is illustrated by comparing the observation results with the estimation results, for the bicycle traffic study (one-sided bicycle lanes). The estimation results have been taken from Table 2.

In practice, two-dimensional tables are often used to search for independent variables. In the table below the relation between the accident rate per kilometer per year and the volumes of car traffic and (motorized) bicycle traffic is shown. To allow comparison with the model results, all observation results were divided by the number of accidents in the upper left cell.

	Annual Accident Rate per Kilometer			
Volume of Motorized Bicycles	<2000 Motor Vehicles	2000-4000 Motor Vehicles	4000-6000 Motor Vehicles	>6000 Motor Vehicles
<250	1.00	0.91	0.77	0.94
250-700	0.78	1.50	1.23	1.47
700-1000	0.00	2.63	3.33	4.88
>1000	0.52	0.45	5.83	6.27

The number of accidents is expected to increase with the increasing volumes of the motor vehicles and cycles. In the above table, however, the first

Table 1. Example of methodology.

Roadway Characteristic	G^2	df	Significance Level
Average daily traffic volume	266.90	6	>0.999
Points of conflict	94.32	3	>0.999
Type of obstacle	102.00	6	>0.999
Horizontal curve	63.81	2	>0.999
Obstacle distance	75.33	4	>0.999
Parallel facility	54.64	4	>0.999
Environment features	40.41	3	>0.999
Median width	49.33	4	>0.999
Sight distance	49.30	5	>0.999
Profile narrowings	9.93	1	0.997
Truck percentage	12.74	3	0.995
Shoulder width	15.96	5	0.994
Pavement width	10.27	7	∿0.80
Pavement	18.50	7	0.99
Lane width	10.13	6	0.90
Permitted speed	0.67	1	∿0.60
Gradient	0.29	1	∿0.40
Discontinuities	0.26	2	∿0.20

line and the first column might suggest that the accident rate decreases with an increase of the traffic volume. The second column and last line also present a rather illogical picture. The reason is that other roadway characteristics also affect the occurrence of accidents. This effect cannot be demonstrated in a two-dimensional table.

In the first table in the next column, the model results are shown. The table is the result of multiplying the estimated coefficients for motor vehicle volume by those for motorized bicycle volume (Table 2). The effect of other roadway features was incorporated in the other estimated coefficients but is omitted from this table. The model results are well in line with the expectations. An increase in traffic volume results in more accidents.

Volume of Motorized Bicycles	Annual Accident Rate per Kilometer: <2000 Motor Vehicles	2000-4000 Motor Vehicles	4000-6000 Motor Vehicles	>6000 Motor Vehicles
<250	1.00	1.23	1.56	1.74
250-700	1.27	1.56	1.98	2.21
700-1000	2.99	3.68	4.66	5.20
>1000	3.92	4.82	6.12	6.82

Observed results are presented in the table below. In this table the width of the bicycle lane and the width of the median between bicycle lane and roadway are included:

Median Width (m)	Accident Rate by Bicycle Lane Width: <2.7 m	> 2.7 m
<2.3	1.00	1.68
>2.3	0.81	0.73

The wider a bicycle lane with a narrow median width, the more dangerous it is. This table has led to the hypothesis of interaction. Consequently, coefficients were estimated for each cell of the matrix (so 4). These were included in model results in the following table.

Median Width (m)	Accident Rate by Bicycle Lane Width: <2.7 m	>2.7 m
<2.3	1.00	0.85
>2.3	0.65	0.54

Table 2. Estimation results for roads that have a bicycle lane on one side.

Factor	Class	L	Y	C	t-Value Between Classes: Class 1	Class 2	Class 3
1. Motor vehicle volume							
<2000	1	147	66	1.00	–	–	–
2000-4000	2	150	98	1.23	1.29	–	–
4000-6000	3	123	102	1.56	3.02	1.68	–
>6000	4	129	255	1.74	3.02	2.31	0.85
2. Bicycle volume							
<250	1	130	105	1.00	–	–	–
250-700	2	192	120	1.27	1.83	–	–
700-1000	3	48	83	2.99	8.66	5.35	–
>1000	4	78	213	3.92	9.69	9.35	1.94
3. Access points, bicycle lane side							
<225 m	1	135	199	1.00	–	–	–
≥225 m	2	414	232	0.89	0.64	–	–
4. Access points other side							
<225 m	1	130	211	1.00	–	–	–
≥225 m	2	419	310	0.97	0.17	–	–
5. Median bicycle lane							
<2.3 m	1	215	281	1.00	–	–	–
≥2.3 m	2	334	240	0.65	3.13	–	–
6. Width bicycle lane							
<2.7 m	1	235	230	1.00	–	–	–
≥2.7	2	314	291	0.85	0.96	–	–
7. Pavement bicycle lane							
asphalt + concrete	1	318	307	1.00	–	–	–
brick pavement	2	231	214	1.21	2.03	–	–
8. Sight distance							
100 percent	1	338	308	1.00	–	–	–
<100 percent	2	211	213	1.15	1.23	–	–
9. Obstacle							
No obstacle	1	411	290	1.00	–	–	–
Others	2	61	73	1.15	0.85	–	–
Trees, continuous	3	55	120	1.57	3.84	1.87	–
Trees, discrete + lights	4	22	38	1.66	2.85	1.51	0.33

Note: L = total segment length (km) weighted by the analysis period, Y = number of accidents per class, and C = coefficients estimated.

If the cells outside the diagonal are multiplied (0.85 x 0.65 = 0.55), this value differs little from the estimated value of 0.54. Consequently, the assumed interaction apparently does not exist. In the final estimation (presented in Table 2), therefore, interaction is not present.

The estimation results show that wide bicycle lanes are safer than narrow ones and that a wide median is safer than a narrow one, which is entirely in line with the expectations. The difference between model results and observation results must be attributed to the fact that apparently, in the table that shows observational results, other independent variables (e.g., volumes) come into play as well.

APPLICATION IN SPECIFIC SITUATIONS

In the Netherlands the weighted multiproportional Poisson model was used for polderroads in 1974 (14), for interurban bicycle traffic in 1978 (15), and for interurban car traffic in 1979 (16).

Figure 1. Bicycle lanes on both sides (T-roads).

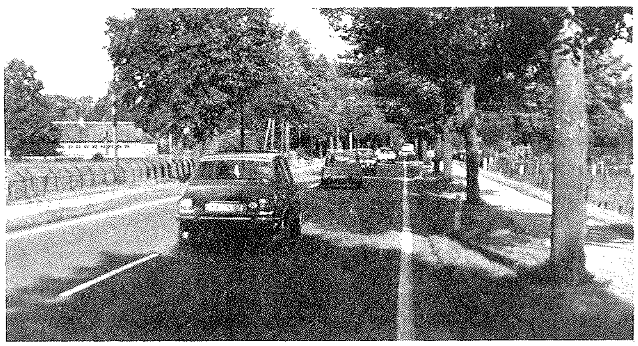


Figure 2. Bicycle lanes on one side (E-roads).

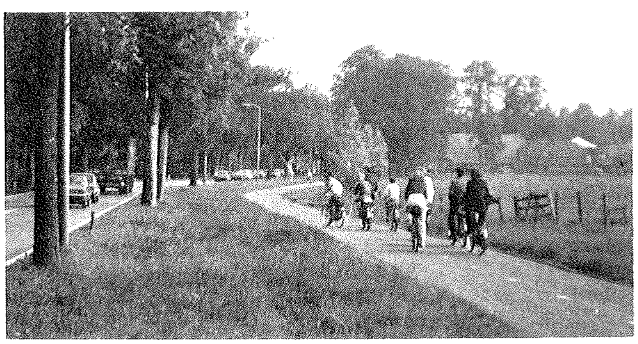
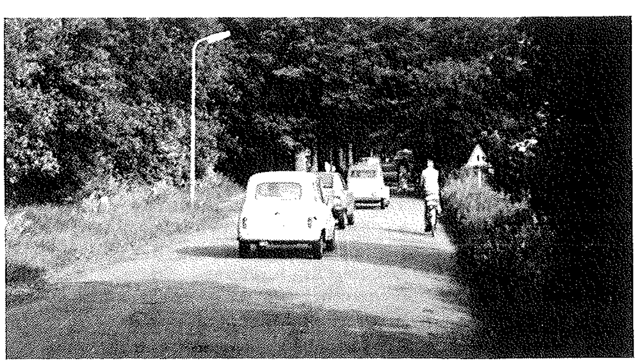


Figure 3. Road without bicycle lanes (Z-roads).



Interurban Bicycle Traffic

The study investigated 1774 accidents to motorized cyclists that resulted in severe injuries, some of which were fatal. The accidents have been taken from the national accidents survey of accidents per road segment. The roadway characteristics have been determined by a direct survey. The role of secondary and tertiary roads outside the built-up area (with a total length of 2439 km) in these accidents was studied. Some are roads with bicycle lanes on both sides (T-roads) (Figure 1), some are roads with bicycle lanes on one side (E-roads) (Figure 2), and some are roads without lanes (Z-roads) (Figure 3).

The inventory unit used was a road segment. A segment is a part of the road between intersections or junctions with public roads within which there are no changes in the most important characteristics of the road. It has a maximum length of 200 m. A distinction was made between roads without bicycle facilities, roads with a separate (i.e., reservation in between) bicycle lane on one side, and roads with separate bicycle lanes on both sides.

Some of the most important results are as follows:

1. The probability of accidents is greatly influenced by the motor vehicle volume. Average daily traffic volumes were used; the volume at the time of the accident could not be used because data were lacking. The influence of the volume on accidents is considerably greater for roads without bicycle facilities [factor 1 in Tables 2 and 3 (for illustration, only the results for E- and Z-roads are given)].
2. An increase in the bicycle volume greatly increases the probability of accidents (Tables 2 and 3, factor 2).
3. The probability of accidents is greater on roads with a wide bicycle lane and a narrow median than on roads with a less-wide bicycle lane and a wider median. This is particularly true for roads that have a bicycle lane on one side.
4. Roads that have many access points generate significantly more accidents than roads that have few or no access points (factor 3, Table 2, and factors 3 and 4, Table 3).
5. The influence of parking bays, bus stops, and so on is not significant on roads that have bicycle facilities on one side or on both sides as could be expected because there are no conflicts. On roads without bicycle facilities the probability of accidents is increased by more than 20 percent (see Table 3).
6. The influence of the presence or absence of an edge marking could not be proved (see Table 3, factor 6).

No significant influence could be demonstrated for other influence factors.

The model coefficients found are suitable for calculating the probability of accidents on a segment of a certain type. The method can also be used to determine whether no facilities, facilities on one side, or facilities on both sides are better.

Interurban Car Traffic

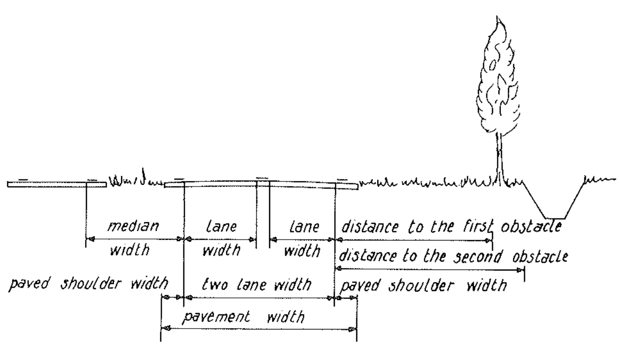
The following data are cited from Jager and Gijsbers (16). The study was concerned with 1545 accidents that caused severe injuries, some of which were fatal, on two-lane or equivalent roads maintained by the national or provincial governments (1300 km).

Table 3. Estimation results for roads that do not have bicycle lanes.

Factor	Class	L	Y	C	t-Value Between Classes: Class 1	Class 2	Class 3	Class 4
1. Motor vehicle volume								
<1000	1	323	74	1.00	–	–	–	–
1000-2000	2	469	180	1.34	2.13	–	–	–
2000-3000	3	318	227	2.05	5.72	4.28	–	–
>3000	4	199	242	3.21	7.55	7.48	4.69	–
2. Bicycle volume								
<150	1	357	82	1.00	–	–	–	–
150-250	2	266	110	1.61	3.48	–	–	–
250-500	3	363	185	1.62	3.71	0.03	–	–
500-700	4	161	125	1.95	5.02	1.38	1.79	–
>700	5	163	221	2.85	6.94	4.16	5.57	3.47
3. Access points								
>500 m	1	356	110	1.00	–	–	–	–
225-500 m	2	399	195	1.41	2.42	–	–	–
<225 m	3	553	418	1.69	4.60	1.69	–	–
4. Two lanes width								
<5.0 m	1	231	87	1.00	–	–	–	–
5.0-6.0 m	2	611	304	0.90	0.70	–	–	–
6.0-7.0 m	3	394	285	1.09	0.55	1.68	–	–
>7.0 m	4	72	47	0.91	0.44	0.07	0.96	–
5. Pavement of the road								
Asphalt + concrete	1	1234	657	1.00	–	–	–	–
Brick pavement	2	74	66	1.22	1.37	–	–	–
6. Parking bays, bus stops								
Yes	1	1054	583	1.00	–	–	–	–
No	2	255	140	0.78	2.64	–	–	–
7. Marking								
Yes	1	644	376	1.00	–	–	–	–
No	2	664	347	0.97	0.41	–	–	–
8. Sight distance								
100 percent	1	675	308	1.00	–	–	–	–
<100 percent	2	634	415	1.11	1.56	–	–	–
9. Obstacle								
No obstacle	1	693	301	1.00	–	–	–	–
Others	2	212	135	1.19	1.74	–	–	–
Trees, continuous	3	253	175	1.31	2.57	0.95	–	–
Trees, discrete + lights	4	150	112	1.72	4.68	2.80	1.69	–

Note: L = total segment length (km) weighted by the analysis period, Y = number of accidents per class, and C = coefficients estimated.

Figure 4. Road characteristics of cross section.



To obtain a unit of analysis inventory, the road network was subdivided into lengths of approximately 100 m. The analysis was carried out on segments of roads, most of which were 200 m in length. The used road characteristics in a cross section are given in Figure 4. Table 4 presents the estimation results.

As may be expected, the average daily traffic volume (factor 1 in Table 4) is by far the most-important explanatory variable. The accident density is approximately a factor 3 higher on roads that have an average daily traffic volume of more than 9000 motor vehicles than on roads that have a volume lower than 3000 motor vehicles/24 h (provided all other roadway features are equal). The accident density hardly increases with an increase of the intensity over 10 000 motor vehicles/24 h. These analysis results scarcely differ from the conclusions that can be drawn from West German and Danish studies (5,17-19).

If the regression coefficients are normalized by traffic performance (roughly speaking this means dividing by volume), those roads that have a high volume are safer than roads that have relatively low volume, if all other features are the same.

The two-lane width does not significantly influence the accident density, whereas the pavement width and shoulder width do (factor 3), although the pavement width less so than the shoulder width. The two-lane width, therefore, seems to influence the accident density much less than does the paved shoulder width. The effects found for the shoulder width are significant in all cases. A paved shoulder width smaller than 0.85 m produces a greater probability of accidents than does a wider paved shoulder. The accident density is not significantly changed by increasing the width of shoulders of 1.8-2.0 m wide. The accident density is as much for roads that have a shoulder width of 1.8-2.0 m as for roads that have considerably wider shoulders. Within the classes smaller than 0.9 m, the widths of 0.4-0.5 m appear to be significantly safer than the slightly wider (0.6-0.8 m) or the slightly narrower (≤0.3 m) widths. The results found for the lane width are not in line with the conclusion drawn from other studies; i.e., that the traffic safety increases with an increase of the lane width.

The various researchers, however, do not come to the same conclusion concerning the relation between lane width and accident density: Foody and Long (20) have found a linear relation; Dart and Mann (3)

Table 4. Estimation results for lane-width study.

Factor	Class	L	Y	C	t-Value Between Classes: Class 1	Class 2	Class 3	Class 4	Class 5
1. Motor vehicle volume									
<3000	1	256	172	1.00	–	–	–	–	–
3000-3999	2	177	156	1.23	1.83	–	–	–	–
4000-5999	3	282	329	1.55	4.48	2.37	–	–	–
6000-7999	4	238	329	2.07	7.59	5.30	3.56	–	–
8000-9499	5	123	236	2.72	9.83	7.56	6.32	3.18	–
≥9500	6	130	323	3.11	11.57	9.29	8.48	5.05	0.52
2. Truck percentage									
<20 percent	1	1003	1253	1.00	–	–	–	–	–
≥20 percent	2	204	292	1.14	1.89	–	–	–	–
3. Paved shoulder width									
≥0.85 m	1	131	136	1.00	–	–	–	–	–
<0.85 m	2	1077	1409	1.22	2.15	–	–	–	–
4. Obstacle distance									
Absent	1	39	37	1.00	–	–	–	–	–
2.5-3.5 m	2	1012	1197	1.21	1.13	–	–	–	–
<2.5 m	3	156	311	1.43	1.96	2.20	–	–	–
5. Median width									
Absent	1	653	703	1.00	–	–	–	–	–
≥4.0 m	2	412	612	1.15	2.29	–	–	–	–
<4.0 m	3	143	230	0.98	0.25	1.93	–	–	–
6. Type of obstacle									
Absent	1	764	829	1.00	–	–	–	–	–
Other	2	312	453	1.03	0.44	–	–	–	–
Open row of trees	3	87	145	1.26	2.41	2.04	–	–	–
Row of lighting columns	4	47	118	1.35	2.62	2.41	0.52	–	–
7. Sight distance									
≥900 m in both directions	1	263	258	1.00	–	–	–	–	–
Other	2	945	1287	1.14	1.91	–	–	–	–
8. Pavement									
Concrete	1	121	195	1.00	–	–	–	–	–
Asphalt	2	1086	1350	0.85	1.99	–	–	–	–
9. Horizontal curvature									
≥1500 m	1	982	1141	1.00	–	–	–	–	–
750-1499 m	2	128	198	1.25	2.29	–	–	–	–
≤749 m	3	96	206	1.64	6.29	2.70	–	–	–
10. Points of conflict									
Absent + crossings	1	1073	1261	1.00	–	–	–	–	–
Access points	2	111	206	1.32	3.45	–	–	–	–
Intersections	3	24	78	2.65	8.25	5.15	–	–	–
11. Profile narrowings									
Absent	1	1178	1487	1.00	–	–	–	–	–
Present	2	29	58	1.48	2.90	–	–	–	–

Note: L = total segment length (km) weighted by the analysis period, Y = number of accidents per class, and C = coefficients estimated.

and Nilsson (21) have found a parabolic relation; and Silyanov (22) has concluded a hyperbolic relation. On the whole the results for shoulder width and pavement width are in keeping with the conclusions drawn from the study of literature [Bitzl (17,18), Foody and Long (20), and Silyanov (22)].

The type of obstacle (factor 6 in Table 4) significantly influences the accident density. Open rows of trees and rows of lighting columns significantly increase the probability of accidents by 26 percent. With regard to the effects of the rows of lighting columns, a reservation should be made. In general, lighting is used along motorways only if the traffic volume in conjunction with the road situation calls for such a provision.

The obstacle distance (factor 4) significantly influences the accident density. The distance from the inner side of the edge marking of the lane to the first obstacle in the shoulder is used as a measure for the obstacle distance. The analysis makes obvious that the accident density decreases with an increase of the obstacle distance. The accident density for obstacle distances smaller than 2.5 m was found to differ significantly from that for obstacle distances greater than 2.5 m. Significant differences in accident density could be found neither for obstacle distances between 0 and 2.5 m nor for those greater than 2.5 m.

The presence of most types of conflict points (factor 10) significantly increases the accident density. After traffic volume, this road feature influences the accident density most. A distinction was made between the following points of conflict:

1. Pedestrian crossings and bicycle crossings, both with and without traffic lights, and crossings for mixed traffic;
2. Residential and agrarian access points and minor intersections; and
3. Type B intersections (intersections without road signs and without changes in their cross sections).

Access points and B-type intersections significantly influence the accident density. The presence of type B intersections increases the accident density considerably more than that of access points.

Horizontal curves (factor 9) that have a radius greater than 1500 m do not affect the accident density. With a decrease in the horizontal radius, the accident density increases significantly.

The sight distance (factor 7) significantly affects the accident density. A sight distance greater than 900 m is significantly safer than a more-restricted one. With a further decrease of sight distances smaller than 900 m, the increase of the probability of accidents was found to be insignificant. From other studies (5), it appeared that sight distances smaller than 400 m result in a higher accident density than do sight distances greater than this. The literature study showed that sight distances greater than 400 m hardly affect the accident density at all.

Beside the roadway features mentioned, the median width (factor 5), the type of pavement (factor 8), the truck percentage (factor 2), and the profile narrowings also affect the probability of accidents, though to a less extent. For some features no relation to the probability of accidents can be demonstrated. This applies to the legally permitted maximum speed and the presence of grades.

EVALUATION

The weighted multiproportional Poisson model presented here has yielded practical results when used in traffic situations with few accidents. Road and traffic characteristics affect the occurrence of accidents substantially. The estimated results enable the traffic engineer to design and evaluate safety measures.

ACKNOWLEDGMENT

M.C. Huisman was also closely connected with the development of the method. Tables and results were taken from the studies by Jager, Van der Wal, and Gijsbers, which were made for Rijkswaterstaat (Department of Public Works). We are grateful for having been allowed the use of these data.

REFERENCES

1. D.W. Schoppert. Predicting Traffic Accidents from Roadway Elements of Rural Two-Lane Highways with Gravel Shoulders. HRB, Bull. 158, 1957, pp. 4-26.
2. J.W. Hall, C.J. Burton, D.G. Coppage, and L.V. Dickinson. Roadside Hazards on Nonfreeway Facilities. TRB, Transportation Research Record 601, 1976, pp. 56-58.
3. O.K. Dart, Jr. and L. Mann, Jr. Relationship of Rural Highway Geometry to Accident Rates in Louisiana. HRB, Highway Research Record 312, 1970, pp. 1-16.
4. K.J. Kihlberg and J.K. Tharp; Cornell Aeronautical Laboratory. Accident Rates as Related to Design Elements of Rural Highways. NCHRP, Rept. 47, 1968, 173 pp.
5. J.H. Klöckner and H.G. Krebs. Untersuchungen über Unfallraten in Abhängigkeiten von Strassen--und Verkehrsbedingungen Ausserhalb Geschlossener Ortschaften. Universität Karlsruhe, Karlsruhe, Federal Republic of Germany, Oct. 1976.
6. Y.M.M. Bishop, S.E. Fienberg, and P.W. Holland. Discrete Multivariate Analysis. MIT Press, Cambridge, MA, 1975.
7. H.T. Reynolds. The Analysis of Cross-Classifications. Free Press, New York, 1977.
8. S.E. Fienberg. The Analysis of Cross-Classified Categorical Data. MIT Press, Cambridge, MA, 1977.
9. S. Oppe. The Use of Multiplicative Models for Analysis of Road Safety Data. Accident Analysis and Prevention, No. 11, 1979, pp. 101-115.
10. E.B. Anderson. Multiplicative Poisson Models with Unequal Cell Rates. Scandinavian Journal of Statistics, Vol. 4, 1977, pp. 153-158.
11. A.D. Ceder and O. Dressler. A Note on the Chi-Square Test with Applications to Road Accidents in Construction Zones. Accident Analysis and Prevention, No. 12, 1980, pp. 7-10.
12. R. Hamerslag and M.C. Huisman. Het Gebruik Van Het Multiproportioneel Schattingsmodel Bij Ongevallenanalyse. DHV Raadgevend Ingenieursbureau BV, Amersfoort, Netherlands, Internal Rept., 1978.
13. R. Hamerslag and J.P. Roos. Analyse van Ongevallen in Verkeerssituaties Met Een Multiproportioneel Poisson Model. Verkeerskunde, Vol. 11, No. 5, 1980, pp. 567-571.
14. M. Kwakernaak. Verkeersveiligheid Voor de Haarlemmermeer. Verkeerskunde No. 5, 1976, pp. 212-218.
15. M. Kwakernaak. Ongevallenkans Van Bromfietser op Wegvakken Buiten de Bebouwde Kom. Verkeerskunde, No. 5, 1980, pp. 561-565.
16. T.C.M. Jager and A.W.N. Gijsbers. Onderzoek Rijstrookbreedte. Verkeerskunde, Vol. 11, No. 5, 1980, pp. 584-589.
17. F. Bitzl. Der Sicherheitsgrad von Strassen. F.G., Schriftenreihe Strassenbau und Strassenverkehrstechnik, Vol. 28, 1964.
18. F. Bitzl. Der Einfluss der Strasseneigenschaften auf die Verkehrssicherheit. F.G., Schriftenreihe Strassenbau und Strassenverkehrstechnik, Vol. 55, 1976.
19. K. Pfundt. Vergleichende Unfalluntersuchungen auf Landstrassen. F.G., Schriftenreihe Strassenbau und Verkehrstechnik, Vol. 82, 1969, pp. 10-16, 42-45.
20. T.J. Foody and M.D. Long. The Identification of Relationships Between Safety and Roadway Obstructions. Bureau of Traffic, Ohio Department of Transportation, Columbus, Jan. 1974.
21. G. Nilsson; National Swedish Road and Traffic Research Institute. Methods for Determining Traffic Safety Standards in Relation to Geometric Road Design. In Geometric Road Design Standards, Organization for Economic Development and Cooperation, Paris, Road Research Rept., 1977, pp. 112-113.
22. V.V. Silyanov. Comparison of the Patterns of Accident Rates on Roads of Different Countries. Traffic Engineering and Control, Jan. 1973, Vol. 14, No. 9, pp. 432-435.
23. S.C. Carrol. Classification of Driving Exposure and Accident Rates for Highway Safety Analysis. Accident Analysis and Prevention, 1972, pp. 81-94.
24. RWS and DHV Consulting Engineers. Onderzoek Rijstrookbreedte. Rijkswaterstaat (Dienst Verkeerskunde)/DHV Raadgevend Ingenieursbureau, The Hague, Amersfoort, Netherlands, 1979.
25. RWS and DHV Consulting Engineers. Onderzoek Naar de Relatie Tussen Weg en Verkeerskenmerken en Ongevallenkans Van (Brom-) Fietsverkeer Langs Wegvakken Buiten de Bebouwde Kom. Rijkswaterstaat (Dienst Verkeerskunde)/DHV Raadgevend Ingenieursbureau, The Hague, Amersfoort, Netherlands, 1979.

Publication of this paper sponsored by Task Force on Highway Accident Statistics.