VTI särtryck

Nr 251 • 1995

What is happening to the number of fatalities in road accidents? A model for forecasts and continuous monitoring of development up to the year 2000

Ulf Brüde

Reprint from Accident Analysis & Prevention, Vol. 27, No. 3, pp. 405–410, 1995

**Swedish National Road and
Transport Research Institute**

VTI särtryck

Nr 251 • 1995

What is happening to the number of fatalities in road accidents? A model for forecasts and continuous monitoring of development up to the year 2000

Ulf Brüde

Reprint from Accident Analysis & Prevention, Vol. 27,
No. 3, pp. 405–410, 1995

0001-4575(94)00062-X

BRIEF COMMUNICATIONS AND RESEARCH NOTES

WHAT IS HAPPENING TO THE NUMBER OF FATALITIES IN ROAD ACCIDENTS? A MODEL FOR FORECASTS AND CONTINUOUS MONITORING OF DEVELOPMENT UP TO THE YEAR 2000

ULF BRÜDE

Swedish Road and Transport Research Institute (VTI), S-581 95 Linköping, Sweden

(Received 11 March 1994; Accepted 21 June 1994)

Abstract—A model for successively forecasting and monitoring the development in the number of fatalities in traffic is presented. The model has been created through time series analysis covering the years 1977–1991. The model is simple, with the number of fatalities as the dependent variable and with time and traffic as the only predictors. The time factor describes the cumulative effect of changes such as better roads, vehicles, drivers, etc. The model is multiplicative and permits a nonproportional relationship with traffic volume. Taking into account the purely random fluctuations in the number of fatalities, the historical fit for the period 1977–1991 is very good. Also the forecasts for 1992 and 1993 have proved very accurate. The model will be revised as new annual data are received. At present, the model points to a favorable development in the reduction of the number of fatalities up to the year 2000, assuming a moderate increase in traffic.

INTRODUCTION

In accordance with the WHO guidelines, the Swedish Parliament and traffic safety authorities have declared that the number of killed and injured in traffic accidents must be reduced. In the case of fatalities, the goal is a maximum of 600 by the year 2000. To enable continuous monitoring of progress towards this goal, some sort of forecasting method is required. In particular, information is needed for determining whether further actions are necessary. The general responsibility for traffic safety work in Sweden is nowadays handled on both central and regional levels by the National Road Administration.

HISTORICAL DEVELOPMENT

As shown in Fig. 1, the number of fatalities in traffic in Sweden increased successively up to about 1,300 in 1970. A large decrease occurred temporarily during the Second World War and also in connection with the changeover to right-hand traffic in 1967.

During the 1970s and early 1980s, the number of fatalities decreased. Legislation on using seat belts in front seats was introduced in 1975 after a campaign lasting a number of years. However, the latter part of the 1980s saw an increase again. At this point, it was feared that the traffic safety goal

would prove difficult to achieve without extraordinary measures, even if traffic increased only moderately up to the year 2000.

Figure 1 also shows that, except during the Second World War, the number of cars on the roads increased successively up to the beginning of the 1990s. The pattern of development in fatalities and traffic in Sweden agrees closely with that in several other countries (Oppe 1989).

Table 1 shows the development in fatalities in Sweden for the 15-year period 1977–1991 and the development in index form for traffic mileage, estimated with the aid of petrol consumption data. The period 1977–1991 has been considered sufficiently recent and long enough to provide a basis for a forecasting model. In addition, the period has not been disturbed by individual major events affecting the number of fatalities.

Table 2 and Figure 2 show the development in the death risk, defined as the number of fatalities divided by the traffic index. The absolute decrease in the death risk diminishes with time. What is happening to the relative decrease? As can be seen with the aid of the table, this is on the order of -3.5% on the average. How long this will continue is of course unknown. At the time of writing, it can be seen that the change in 1991–1992 was $+0.2\%$ and in 1992–1993, -12.4% . It can also be seen that the

Fig. 1. Number of fatalities and number of cars during the 20th century. Source: Statistics Sweden.

decreased death risk appears to have had a more favorable development at times when traffic has been unchanged or has decreased compared to when it has increased.

MODEL

Several attempts have been made to use regression models in adaptation to historical development and also to use the models in forecasting the number

of fatalities. This is described in several articles in a special issue of *Accident Analysis & Prevention* (Haight 1991).

The model chosen in this study has the following structure:

$$Fatalities = a \times b^{Year} \times Traffic^c, \quad (1)$$

which can alternatively be written as

$$Fatalities = \exp(\alpha + \beta \times Year + \gamma \times \ln Traffic) \quad (2)$$

where $\alpha = \ln a$, $\beta = \ln b$ and $\gamma = c$.

Table 1. Number of fatalities and traffic index based on petrol consumption, 1977–1991

Year	Fatalities	Traffic index
1977	1,031	100
1978	1,034	103
1979	926	102
1980	848	99
1981	784	97
1982	758	98
1983	779	100
1984	801	104
1985	808	105
1986	844	111
1987	787	115
1988	813	119
1989	904	124
1990	772	117
1991	745	120

Source: Statistics Sweden.

Table 2. Death risk 1977–1991

Year	Death risk	Annual change (%)
1977	10.31	
1978	10.04	-2.6
1979	9.08	-9.6
1980	8.57	-5.6
1981	8.08	-5.7
1982	7.73	-4.3
1983	7.79	+0.8
1984	7.70	-1.2
1985	7.70	±0
1986	7.60	-1.3
1987	6.84	-10.0
1988	6.83	-0.1
1989	7.29	+6.7
1990	6.60	-9.5
1991	6.21	-5.9

Fig. 2. Death risk 1977–1991. (Solid line indicates actual development. The broken line indicates adjusted development for the three subperiods with different traffic development).

The regression coefficients a , b , and c are estimated with the aid of the GLIM computer package and assume that the number of fatalities follows a Poisson distribution (which naturally is not strictly applicable).

Model 1 (i) is simple and interpretable, (ii) shows directly the number of fatalities without needing to use the death risk, (iii) permits a nonproportional relationship to traffic volume for fatalities (as well as for the death risk).

Coefficient b indicates the average annual change in number of fatalities (in unchanged traffic) resulting from better roads, vehicles, drivers, etc. Coefficient b also pays attention to other changes over time, such as speed changes and changes in the age composition of the population.

Coefficient c indicates how the number of fatalities varies with traffic volume, which in turn is strongly correlated with business cycle fluctuations.

ESTIMATED MODEL COEFFICIENTS

For the 15-year period 1977–1991, the following regression model has been obtained.

$$\text{Fatalities} = 0.2091 \times 0.9562^{\text{Year}} \times \text{Traffic}^{1.851} \quad (3)$$

where $\text{Year} = 1$ for 1977, 2 for 1978 . . . and 15 for

1991. Traffic is equal to traffic (mileage) index with 1977 as base year (index = 100). The value of coefficient b ($= 0.9562^*$) means that in the event of unchanged traffic mileage, the number of fatalities is expected to decrease by just under 5% from one year the next. The value 1.851^+ for coefficient c means that the number of fatalities in a particular year is expected to increase (or decrease) just over 1.5% if the traffic increases (or decreases) by 1%. The relationship between fatalities and traffic is therefore somewhat stronger than purely proportional.

On the basis of equation (3), it is also understood that the death risk has a positive relationship to traffic. In a certain given year, the death risk may be expected to increase with increased traffic and decrease with decreased traffic. This was already evident in Figure 2.

Owing to the intercorrelations between estimates of coefficients b and c , the numerical value of the regression coefficients should be interpreted with caution. From the predictive aspect, the intercorrelations fortunately need not play an important part. In addition, the credibility of the regression coefficients is strengthened in a further inspection

* Approximately 95% confidence interval: 0.9487–0.9639.

† Approximately 95% confidence interval: 1.4159–2.2861.

Fig. 3. Observed and predicted numbers of fatalities and traffic index, 1977–1991.

of the fatality development from 1977 onwards (cf. Table 1).

—During 1977–1983, traffic was almost unchanged (increased on average by 0.02% per year). The number of fatalities decreased by an average of 4.4% per year.

—During 1983–1989, traffic increased on average by 3.7% per year. The number of fatalities increased by an average of 2.7% per year.

—During 1989–1991, traffic decreased on average by 1.5% per year. The number of fatalities decreased by an average of 9.1% per year.

The observed and predicted fatality numbers for the study period 1977–1991, i.e. the years used to estimate the regression model, are described together with traffic development in Fig. 3. It is striking how well the pattern of variation in traffic is reflected by variations in the number of fatalities.

HOW WELL IS THE HISTORICAL DEVELOPMENT EXPLAINED?

The agreement between the actual and predicted number of fatalities is naturally not completely perfect. This cannot be expected for a variable regarded as random.

The coefficient of determination, R^2 , i.e. the squared correlation coefficient between observed and predicted fatalities in 1977–1991, is calculated as 0.92. In other words, 92% of the variation in number of fatalities is explained by the factors, Year and Traffic.

On the assumption that the annual expected values agree fully with the predicted values and that the number of fatalities in a given year follows a Poisson distribution, the maximal coefficient of determination can be expected (Brüde and Larsson 1993):

$$R_{\max}^2 = V(\text{pred})/[E(\text{pred}) + V(\text{pred})] = 0.90, \text{ i.e. } 90\%$$

From this, it follows:

$$(1 - R_{\max}^2) \times 100\% = 10\% \text{ constitutes purely random variation.}$$

As can be seen, the coefficient of determination, R^2 , actually obtained exceeds $R_{\max}^2$, and therefore no remaining unexplained variation can be demonstrated. (This is also indicated by the fact that the value obtained for scaled deviance in the GLIM anal-

Fig. 4. Forecasts for 1992, 1993, and 2000. (Solid rectangles indicate observed fatalities. Open rectangles indicate predicted and forecast values.)

ysis, 11.3, closely agrees with the value of 12 degrees of freedom).

According to this approach, the model describes historical development very closely.

FORECASTS AND MONITORING UP TO THE YEAR 2000

Figure 4 shows the observed and predicted values for 1977–1991, as well as forecasts for 1992, 1993 and 2000.

In recent years, traffic has (temporarily?) stagnated and even decreased somewhat. However, it is probable that traffic will increase again in the coming years. According to the forecast of the Swedish National Road Administration, traffic will be about 10% greater in 2000 than in 1990.

It is important to emphasize that, with regard to future traffic, this figure can be reviewed if necessary. It is also important to emphasize that the model is intended to be continuously updated. Up to the year 2000, the time series will be extended annually.

For 1992 and 1993, the actual traffic index was 122 and 116, respectively. Given this traffic index, 743 and 647 fatalities, respectively, would have been forecast. The actual number of fatalities was 759 and 632, respectively. In passing, it should be noted that

in 1992 an accident with an empty runaway tram killed 13 people.

For 1994, 619 fatalities are forecast, given the same traffic as in 1993 (traffic index 116). If the traffic index were to increase by 5%, the forecast would be 679 fatalities, and if traffic decreased by 5%, 561 would be killed.

On the assumption of about 10% more traffic in 2000 compared with 1990, which would mean a traffic index of 129 in 2000, 576 fatalities are forecast for the year 2000. If so, this would mean that the traffic policy goal for fewer fatalities was achieved. *This type of monitoring will be performed annually up to the year 2000.*

For unchanged traffic in the year 2000 compared with 1990, 481 fatalities are forecast and if a 20% increase in traffic is assumed, 670 fatalities.

By using the covariance matrix, a confidence range can be calculated for the forecasts (the expected values). An approximately 95% confidence range for the predicted value in 2000 (given a traffic index of 129) of 537–618 is then obtained. A prediction interval for the observed outcome would of course be somewhat larger still.

UNCERTAINTY OF THE FORECASTS

The forecasts for the numbers of fatalities are uncertain in several respects. First, it is assumed

that history repeats itself and that future traffic safety work will be as extensive and successful as before.

The fact that the model has a very good coefficient of determination for the data used to estimate the model is no guarantee that the model will be as reliable in the future. Different models can be estimated with almost the same coefficient of determination, but which nevertheless produce different forecasts.

Making forecasts entails extrapolating the regression model outside the area where the observations were made. Making forecasts differs considerably from taking a random sample from a well-defined population and performing statistical inference, i.e. point and interval estimates of expected values within the range of observations. This requires calculated confidence intervals for forecasts and the prediction intervals for observed outcomes to be regarded with some caution.

As has already been pointed out, the model is intended to be updated successively as new annual data become available.

FURTHER NOTES

The model will probably be used also with a division based on fatalities among motorists and un-

protected road users and a division between children and adults. According to the goals, priority is to be given to unprotected road users and, in particular, to children.

The same type of model can also be used for injured persons, although forecasts will be comparatively uncertain compared with those for fatalities. However, a small part of the variation in the number of injured depends on random variation. On the other hand, the coefficient of determination is also small and the unexplained variation thereby fairly large. A contributory cause may be a varying degree of reporting of injured persons over time. Another cause may be that further predictors are required.

Acknowledgements—Valuable opinions and suggestions for processing have been received from Mats Wiklund, VTI. Rune Elvik, Institute of Transport Economics in Norway, has contributed useful information on earlier studies in this field.

REFERENCES

- Brüde, U.; Larsson, J. Models for predicting accidents at junctions where pedestrians and cyclists are involved. How well do they fit? *Accid. Anal. Prev.* 25:499–509; 1993.
- Haight, F. A., editor. Theoretical models for traffic safety. Special issue. *Accid. Anal. Prev.* 23(5); 1991.
- Oppe, S. Macroscopic models for traffic and traffic safety. *Accid. Anal. Prev.* 21:225–232; 1989.

***Swedish National Road and
Transport Research Institute***

RESEARCH FOR MOBILITY

***The Swedish Road and Transport Research Institute
has laboratories and know-how for advanced research
commissions in transport and welfare economics,
road safety, vehicles and the environment. It also has
research capabilities for the construction, operation
and maintenance of roads and railways.***

Address
S-581 95 Linköping
Sweden

Telephone
+ 46 13 20 40 00

Fax
+ 46 13 14 14 36