9th International Conference
ROAD SAFETY IN EUROPE
September 21–23, 1998
Bergisch Gladbach, Germany

SPEED-ACCIDENT RELATIONSHIPS ON EUROPEAN ROADS

A. Baruya
Transport Research Laboratory, Crowthorne, Berks, UK

ABSTRACT:

This paper summarises the results of our investigation on speed-accident relationships on different kinds of European roads. Available literature on the subject is rich, but further research was needed to meet the specific requirements of the MASTER project. To meet those requirements European speed and accident data were critically examined in order to gain a better understanding of speeds on European roads and their relationships to accident occurrences. This paper summarises the results of the investigation published in several reports under Work Package 1.1 of the MASTER Project. The paper gives a comprehensive description and analysis of flow and speed distributions on a sample of UK and European roads and examines critically the effects of extreme speeds, road environment and road geometry on traffic speed and accidents. An accident predictive model is proposed for wider application in the European Union. Despite the likelihood that the model will successfully predict accident frequency for a wide variety of roads it is recommended that, where possible, the model is tested and assessed prior to use in a particular situation.

1 INTRODUCTION

One of the important road safety issues in the European Union is speeding on rural highways and urban roads. While "Speed Kills" has been an important slogan for many years, speed still takes its toll in terms of life and serious injuries. According to a recent report by the United Nations about 44,000 deaths and 1.67 million injuries occurred in 1.24 million personal injury accidents in the 15 member states of the EU in 1995 alone. The annual loss is enormous and the total cost accruing from these accidents has been estimated to be ECU 162 billion in 1995, a significant proportion of which is attributable to speed. According to one estimate (ETSC, 1997), the impact of speed on this loss is such that a reduction of 5 km/h in the average speed would result in a saving of ECU 30-40 billion annually.

Against this background the MASTER (MAnaging Speeds of Traffic on European Roads) project set its goal to provide information for decision making concerning speed management on both Community and national levels. One of the objectives of this project is to construct a framework for the assessment of the effects of speed, with due regard to the effect on environment, so that it could be used by the commission and national authorities in the determination of appropriate and acceptable ranges of speed for different kinds of roads in the European Union.

The effect of speed on accidents has been an important issue for a long time. Much research has been done on this subject, but opinions still seem to differ amongst scholars as to whether the mean speed, or the speed variance, affects accidents and if so to what extent.

The most visible effect of speed on accidents is reported to be the effect of speed limit reductions on fatalities in various countries during the period of oil shortage in 1973. Some US reports summarised that soon after the introduction of reduced speed limit the overall fatality rate per driven mile on all US roads decreased by 15% in a single year (1974) and on inter-state highways the fatality rates fell by 32%. Similar studies in Europe of ‘before-after’ type, particularly in the Nordic countries, reported accidents falling as a result of reduction in mean traffic speeds caused by speed limit reductions (Finch et al, 1994).

Cross-sectional studies, on the other hand, produced widely varying results. Some researchers have reported a positive relationship between accident rates (per unit exposure of vehicle kilometres) and mean speed, others have failed to find any such relationship, but found a positive relationship with speed variance (Lave, 1985; Garber and Gadirau, 1988). In the early sixties Solomon (1964) reported a U-shaped relationship between *accident involvement rate* (number of drivers involved in accidents divided by the related vehicle miles of travel) and *travel speed.* Since the publication of the result the U-shaped relationship has been a subject of debate. Some researchers tried to find an explanation of the phenomenon (Hauer, 1971) and others tried to find evidence from cross-sectional studies. A review of the published literature on this subject is available in Baruya (1997) (see also Fields and Lee, 1993; Finch et. al, 1994).

Garber and Gadirau’s (1988) cross-sectional study reported a negative relationship between accident rate and the mean speed. They qualified this result by suggesting that this effect depicted the effects of different geometric characteristics rather than the effects of the mean speed. The latter explanation is plausible, if much of the effects of geometric characteristics and traffic environment, which can directly or indirectly influence the mean speed, is reflected through the mean speed in a seemingly perverse form. The true relationship in that case remains hidden often due to ‘masking’ or ‘confounding’ effect of those factors. An evidence of such phenomenon was found in the UK relationship on urban roads, where pedestrian activity was identified as such a factor (Baruya, 1995).

Related to this topic is the subject of extreme speeds, which can have strong influence on accidents, injuries and injury severity, as well as on mean speed and speed variance. Very little research has so far been done on the subject of *excess speed* (exceeding the speed limit) and *inappropriate speed* (driving too fast for the prevailing conditions). Some research on speed limit violation and its effect on road safety is just emerging (eg. Andersson and Nilsson, 1997).

One of the tasks under the MASTER Project is to investigate the speed-accident relationships on different kinds of rural single-carriageway roads in Europe. Some research has been carried out recently on this subject, the results of which have been published in four Working Papers by Baruya (1997, 1998a, 1998b and 1998c) under Work Package 1 of the MASTER Project. This paper summarises those results.

What follows is a brief description of the sources of data in Section 2 and a brief summary of the results of flow and speed data analyses in Section 3. The influence of extreme speeds on the mean and the standard deviation of speeds will also be discussed in this section. Accident data will be described in Section 4 and speed-accident relationship, as well as the influence of extreme speeds, will be investigated in Section 5. Despite diversity the European roads have some common features on the basis of which they can be classified into several homogeneous groups. This will be discussed in Section 6. The results of the investigation will be discussed in Section 7. Conclusions will be drawn and recommendations will be made in Section 8

2 DATA SOURCES

The European data were supplied by the Institute of Road Safety Research (SWOV) of The Netherlands, Swedish Road and Transport Research Institute (VTI) and the Instituto Superior Tecnico (TRANS-POR) of Portugal. The UK data were supplied by eleven local authorities mainly from central and southern England. The primary database consists of information regarding speed, road geometry and accidents on 28 links from The Netherlands, 73 from Sweden , 39 from Portugal and 63 links from 'A' and 'B' class roads of England - altogether 203 links.

The speed data used in this study were gathered during the ‘off-peak’ period (09:00 to 16:00) of a ‘normal’ weekday (Monday to Friday), with the exception of Portugal for which 24 hour speed data were used. (We were informed that on the sampled Portuguese roads no distinction could be made between peak and off-peak as far as speed was concerned.) The flow data used were 24 hour ADT flows, where available, or estimated 24 hour flows from several days’ vehicle counts.

An initial investigation of the speed data revealed that there were 5 outliers in the primary database. These 5 links had very low mean speeds - 39 to 47 km/h and high ‘congestion’. Full data on accidents and geometric variables (eg. link length, road width) were only available for a reduced set of 171 links. However, where possible all the available information was used in the subsequent analysis depending on the context, but the inevitable consequence of missing information was that some results had to be derived from a reduced set of data.

Speed data were available from five speed limit groups - 70, 80, 90, 100 and 110 km/h. The UK speed data were originally available in miles per hour (miles/h) units. These were converted to km/h units by using a multiplication factor of 1.609 (1.609 km/h = 1 miles/h). Hence, the UK speed limits of 50 and 60 miles/h were treated as 80 and 100 km/h respectively, even though 60 miles/h corresponds to 96 km/h.

3 FLOW AND SPEED DATA

High flows were observed on the Portuguese links and low flows on the Swedish links compared with UK and The Netherlands. Mean speed and the standard deviation of speed also varied between countries. These variations may not be due to variation between countries but may be due to disproportionate representation from different flow regimes and speed limit groups within each data sample. Mean speeds varied with different speed limits. On some 70 km/h roads the mean speed exceeded the speed limit by up to 12.3 km/h and on 80 and 90 km/h roads by up to 9 km/h. But on 100 and 110 km/h roads the mean speeds remained below the speed limits. This characteristic of the roads is complemented by the fact that, except for the roads with high speed limits (100 & 110 km/h), speed limit violations were widespread and in some cases 85% exceeded the speed limit particularly where a low speed limit was in force.

Extreme speeds will be represented by two tail areas of speed distributions - the proportion (**P**) exceeding the speed limit and the proportion (**Cg**) of vehicle speeds not exceeding half the speed limit, both expressed as percentages. The latter will be treated as an indicator of congestion. On average the speed limit violation was found to be high on the Dutch (54.5%) and Swedish (41.9 %) roads and congestion low (0.9 and 1.1% respectively).

The mean speed (**V**) and the standard deviation of speed (**SD**) were examined critically to investigate the reasons for variations in them. An initial investigation suggested that the relationship between the **V** and **Cg** was nearly hyperbolic (Baruya 1998a, 1998b), so was that between **V** and **P**, suggesting that a multiplicative model would be more appropriate for **V**. Hence a multiplicative model was fitted to **V** using the proportion of speeders(**P**), speed limit (**S**), congestion (**Cg**), link length (**L**) and flow (**Q**) as explanatory variables.

The following equation describes the relationship in log-linear form (using '**ln**' for natural logarithm).

$$\mathbf{\ln(V) = 3.4465 + 0.0088 * S + 0.0744 * \ln(P) - 0.0327 * \ln(Cg) - 0.0097 * \ln(Q) + 0.0139 * \ln(L)} \qquad (1)$$

This equation explained about 92.3% of the total variation in **ln(V)**. The standard error of residuals SE(Res), say, the standard deviation of the unexplained random error, was 0.0467. The coefficients of **ln(P), ln(Cg)** etc are the elasticities of speed with respect to the corresponding explanatory variables. They are interpretable as the percentage increase in the mean speed following a 1% change in the explanatory variable. The equation shows that there is a small, but significant, negative effect of flow (**Q**) on the mean speed, indicating that increasing flow can slow down traffic. The coefficient of **ln(P)** is **0.074**, which reflects the elasticity of the mean speed with respect to **P**.

Similarly a multiplicative model was considered appropriate for the standard deviation (**SD)** of speed. The log-linear form of the multiplicative model for **SD** was as follows (Baruya, 1998b):

$$\ln(SD) = 0.9256 + 0.0147 * S - 0.0219 * W + 0.1601 * \ln(P) + 0.1176 * \ln(Cg) \qquad (2)$$

This equation explained about 74.3% of the total variation in **ln(SD).** The standard error of 'residuals' SE(Res) was 0.111 with 163 degrees of freedom. It is interesting to note there is a significant negative effect of road width (**W**), suggesting that on wider roads **SD** is smaller. This result indicates that on wider roads most vehicle speeds are closer to the mean speed, hence more synchronised. The positive effects of speeders (**P**) and congestion (**Cg**) suggest that increasing proportion of extreme speeds will cause the standard deviation to increase in magnitude.

4 ACCIDENT DATA

The accident data for the European roads covered different lengths of time period. The Dutch accidents cover a period of 4 years, the Portuguese 3 years and the Swedish and the UK accidents 5 years. When these accidents are converted to accidents per year to obtain *accident frequencies* (**AF**), it is found that the accident frequencies varied widely from road to road in the sample, ranging from 0 to 115.3 per year. The Swedish roads had the lowest average (0.67 per year per link), followed by UK (1.59), The Netherlands (2.66) and Portugal (25.33), in that order. But because the link length varied widely, if the accident frequencies are normalised by the link length (**L**) to obtain accident frequency per kilometre, then it is found that, on average, the **AF** per km varied from 0 to 7.7 during the reported period. On average Sweden had the lowest average (0.18), followed by The Netherlands (0.49), UK (0.92) and Portugal (2.09). A comparison of accident rates, expressed as *accidents per 100 million vehicle kilometres*, shows that the Swedish sample had the lowest average accident rate per link (15.3), followed by The Netherlands (21.5), UK (33.3) and Portugal (53.3).

5 SPEED-ACCIDENT RELATIONSHIP

We shall treat the number of accidents occurring during a period of time as discrete Poisson variates and employ a multiplicative Poisson model to the accident data [See Jorgensen (1974), Maycock and Hall (1984), Baruya and Finch (1994)]. One of the most important characteristics of a multiplicative model is that it ensures positivity of predictions and satisfies other technical requirements. The multiplicative model is defined as follows (using subscript 'i' for the i-th link):

$$\textbf{Number of accidents } (\mathbf{Y_i}) \sim \textbf{Independent Poisson } (\boldsymbol{\mu_i})$$

or

$$\textbf{Probability } (\mathbf{Y_i} = \mathbf{y}) = \mathbf{p(y)} = \mathbf{exp}\ [-\boldsymbol{\mu_i}]\ .\ \boldsymbol{\mu_i}^{\mathbf{y}} / \mathbf{y!}$$

$$\textbf{where } \boldsymbol{\mu_i} = \mathbf{E}(\mathbf{Y_i}) = (\mathbf{YR})\ .\ \mathbf{exp}\ [\ \mathbf{a} + \Sigma\ \mathbf{b_j}\ \mathbf{x_{ij}}\ ]$$

Here, **E(.)** stands for 'expectation', **YR** stands for the number of years of accident data, **'a'** is a constant, $\mathbf{b_j}$'s are Poisson regression coefficients corresponding to the j-th explanatory variable and $\mathbf{x_{ij}}$ is the value of the explanatory variable for the i-th link and the j-th variable. The observed accident frequency per year (**AF**) is defined as ($\mathbf{Y_i}$ / **YR**).

The technique of Generalized Linear Modelling (GLM) was employed to estimate the parameters of the multiplicative Poisson model. The procedure employed here was as recommended by McCullagh and Nelder (1989). Explanatory variables are entered (or omitted) at every stage of the model building and inclusion of a variable is decided on the basis of its significance. The significance of the effect of an explanatory variable is determined by the reduction (increase) of the 'deviance' measure achieved as a result of its inclusion (omission).

Regression analysis is a powerful tool for identifying the variables that affect accidents, but the choice of independent variables must be guided by theory as well as by professional and engineering judgment rather than by curve fitting (Fridstrom et al, 1995; Maher and Summersgill, 1996). The independent explanatory variables are to be carefully chosen, particularly in a situation like speed-accident relationship where collinearity can be a real problem due to inter-correlation between the explanatory variables. Hence in developing the model and in selecting the explanatory variables several criteria were used (Maher and Summersgill (1996). They are - (1) significance level should be better than 5%, (2) the model must be stable, (3) the effects must be meaningful and comprehensible and (4) the size of the effect must be reasonable.

The generalised linear modelling technique was employed to make a preliminary study on the English data (Baruya, 1998a), which highlighted some explanatory variables that could have significant influence on rural speed-accident relationship. The results of the study suggest that accident frequencies are strongly related to flow (**Q**), link length (**L**), number of minor junctions (**NJ**), mean speed (**V**), proportion of speeders (**P**) and proportion of slow vehicles (**Cg**). These variables were examined with respect to their effects on the European accident data and the investigation was extended to include additional explanatory variables such as road width (**W**), speed limit (**S**). In doing so some UK links, for which road width information was not available, had to be excluded. Hence this study was based on 165 links - 28 from The Netherlands, 73 from Sweden, 38 from UK and 26 from Portugal.

Initial investigation suggests that the accident figures on the Portuguese roads were so high that it was almost impossible to derive anything meaningful from the combined data for the four countries (Baruya, 1998b). It was, therefore, decided to investigate the Portuguese data separately, and derive a model based on the data for the 139 links from The Netherlands, Sweden and UK.

The following is the accident predictive model, to be called **EURO** model, based on the 139 links -

$$\mathbf{AF = 5.663 * Q^{0.748} . L^{0.847} . V^{-2.492} . P^{0.114} . e^{0.038 * NJ - 0.056 * W + 0.023 * S}} \qquad (3)$$

The model explains about 75% of the variation in **AF** due to non-Poisson sources, which is a good fit. The residual deviance was 305.58 with 131 degrees of freedom, which gives a ‘variance factor’ (σ^2) of 2.333 in Wedderburn’s variance ($\mu\sigma^2$). As expected both flow (**Q**) and link length (**L**) have strong positive effects on accidents. The negative power of speed (**V**) indicates that slower roads are associated with higher accidents. The effect of speed limit violation (**P**) is significantly positive, so is that of the number of junctions (**NJ**) and speed limit (**S**). The effect of road width (**W**) is negative and significant, suggesting that wider roads have fewer accidents. One explanation may be that wider roads are of better design, which have probably been built for safer passage of traffic.

Using this model, predictions were made for the four countries. Figure 1 shows the plot of the observed accident frequencies against their respective predictions for The Netherlands, Sweden and UK and Figure 2 is the plot for the Portuguese data. The solid line through the scatter is the 45° line on which the observed values should equal the predicted values. The 95% confidence limits are shown by the dashed lines. Figure 2 clearly indicates that the observed **AF**'s on the Portuguese roads are much higher than their corresponding predicted values.

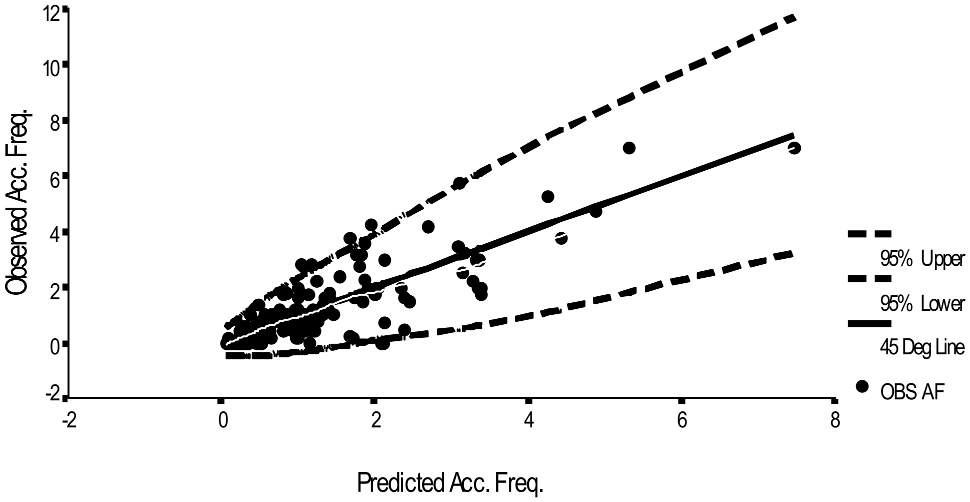

Figure 1. Observed vs predicted accident frequency fot the Netherlands, Sweden and UK (prediction by Euro Model).

To assess the goodness of fit the χ^2 (Chi-squared) statistic was used, given by

$$\chi^2 = \Sigma\ (\mathbf{AF} - \mu)^2 / \mathbf{Var(AF)}$$

where **Var(AF)** = $\sigma^2(\mu\ /\ \mathbf{YR})$, μ is the predicted value of **AF**, **YR** is the period (in years) for which the accident data were available, and σ^2 (= 2.333) is the variance factor mentioned earlier. The calculated values of the chi-squared for the individual countries are respectively 28.9 (df = 28, $P > 0.5$) for The Netherlands, 34.8 (df = 73, $P > 0.9$) for Sweden, 53.93 (df 38, $P > 0.05$) for UK and 1761.7 (df = 26, $P < 0.0001$) for Portugal. Except for Portugal the fits are good for the other three countries.

A comparison of the predicted totals with the observed totals of accident frequencies shows that the totals match very well for The Netherlands, Sweden and UK, but differ considerably for Portugal. The observed totals were 74.5 (predicted 72.89) for The Netherlands, 49.00 (53.17) for Sweden, 61.20 (58.56) for UK and 726.33 (265.20) for Portugal. The Portuguese accidents are under-predicted by a factor of 2.74.

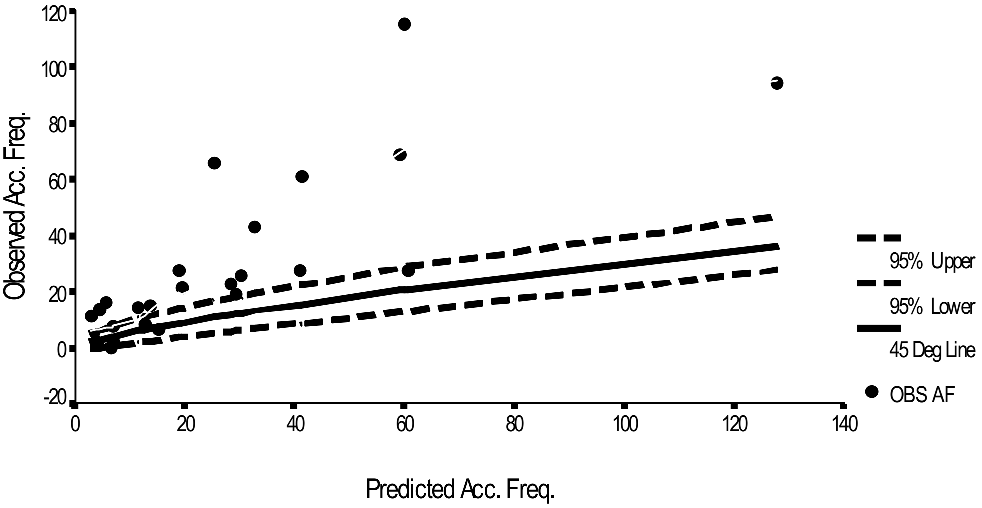

Figure 2. Observed against predicted accident frequency for Portugal (prediction by Euro Model).

Investigation of the Portuguese data

Purely on the basis of the Portuguese data a model can be derived. The model has been discussed in Baruya (1998b) requiring all the explanatory variables but speed limit (**S**), used in the EURO model. A comparison with the **EURO** model reveals that most of the variables have similar effects (indicated by the same signs), but different in magnitude. The effect of speed limit (**S**) was not detectable, because almost all the Portuguese links came from 90 km/h speed limit group. The model showed a stronger effect (47% higher) of flow (**Q**), number of minor junctions (**NJ)** and road width (**W**). An interesting feature of the accident data is that, if they are plotted against the predictions made by the **EURO** model in log-log plot they show a linear relationship. This empirical evidence suggests that a variant of the **EURO** model will be adequate for the Portuguese accidents. The variant is proposed to be $\mathbf{E(AF) = \alpha\ (EP)^{\beta}}$, where '**E**' stands for expectation, α and β are two parameters to define the relationship between accident frequency (**AF**) and the predicted values (**EP)** given by the **EURO** model. The parameter α is the proportionality factor, and β is the power parameter to which the predictions have to be raised so that they match the observed accident frequencies. Further investigation revealed that the proportionality factor α is close to unity (Baruya 1998b,c). Using $\alpha = 1$ it is found that the estimate of β is 1.3508 [**se(**β **)** = 0.0592], which is significantly greater than unity ($t = 5.92$, $df = 24$; $P < 0.0001$). We, therefore, can write as an approximation -

$$\mathbf{E[(AF)] = (EP)^{1.35}} \qquad (4)$$

The fit of the model is shown in Figure 3 in a log-log plot, with 95% confidence limits indicated by two parallel lines. This model will be called **'Powered EURO'** model for Portugal. The fit of the **Powered EURO** model is good in view of the fact that there is only one additional parameter (β) involved. The observed total AF for the 26 Portuguese links was 726.33, the **Powered EURO** model predicts 673.55, which means that it under-predicts by 8%, primarily due to the two links for which the model predictions are significantly lower than the observations. These two links (not included in the plot of Figure 3) have been identified to have the highest accident record (198 and 346 accidents) in the three year period reported as compared to the other Portuguese roads. The observed accident frequencies on these two roads are 66.0 and 115.3 as against the predicted values of 25.21 and 59.90 respectively. The speed limit is 90 km/h on both the links and the mean speeds are 72.6 and 57.9 km/h respectively.

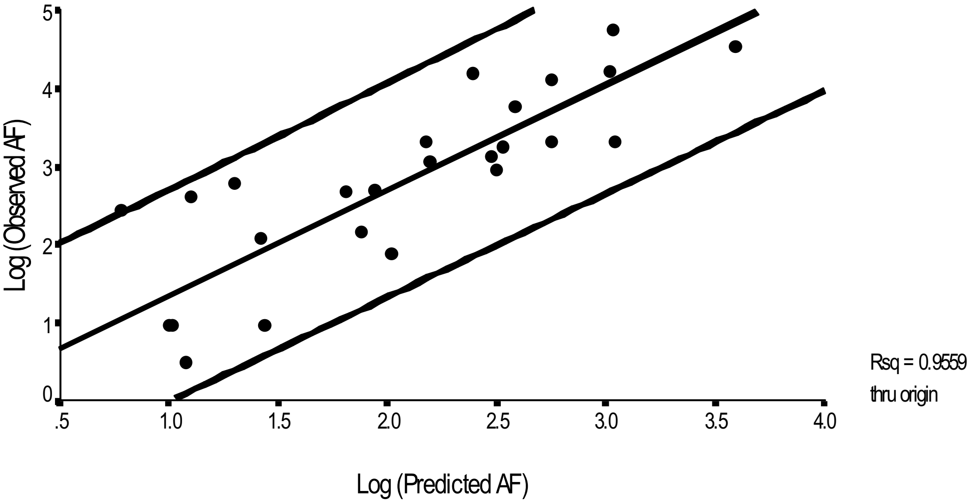

Figure 3. Observed vs predicted accident frequency for Portugal (prediction by powered Euro Model).

6 CLASSIFICATION OF THE EUROPEAN ROADS

Every country has a system of classification of their roads using some criteria, which may differ from country to country. For example, in UK the roads are classified as A, B, C class roads. In Portugal, for example, road types are indicated by letters eg. EN, IP, IC etc. Such classifications are administratively convenient for the nations concerned, but they do not necessarily reflect the statistical features of roads, such as traffic speed, traffic volume, degree of congestion, road environment and so on. Many of these statistical features can be found common across national boundaries. If one wants to identify similar roads in different countries on the basis of their statistical features then a system of statistical classification would be useful. Such a classification could provide a useful guide for speed management and control.

Here we shall attempt to classify the European roads by employing the multivariate technique of *Cluster Analysis* (Anderberg, 1973), whereby road groups can be formed based on the similarity of flow, speed and other characteristics. To obtain such road groups by the cluster analysis six feature variables were chosen, namely - **ln(Q),** road width **(W), ln(V),** standard deviation of speed **(SD), ln(P)** and **ln(Cg).** Link length (**L**) was not used as a feature variable, as in many cases the link length was chosen arbitrarily. Number of minor junctions **(NJ)** was also not used, as it can depend on the choice of the link length.

The number of groups was arbitrarily chosen to be four. (A preliminary investigation suggested that four was an acceptable number.) Non-hierarchical K-means cluster analysis was employed. The cluster analysis was performed by using a statistical computer package (SPSS V6.0). These groups are to be called **RGROUP** (to indicate Rural by '**R**'). Table 1 shows the cross-tabulation of the links by Country and **RGROUP**. The four groups contain 39, 24, 77 and 28 links. Nearly all the Dutch links belong to **RGROUP 3**, Portuguese links to **RGROUP**s **1** and **2**, Swedish links to RGROUPs 3 and 4 and UK links to **RGROUPs 1, 2** and **3**. Table 2 gives the summary statistics of the four groups.

In comparison with the grand means given in the penultimate row of this table, the group means can be categorised as 'below' or 'above' average. And on that basis the four groups can be described as in Table 3. Here the averages of the predicted **(Pred)** and observed (**Obs**) values of accident frequencies have been shown in the last column. The predicted values were obtained by averaging the predictions made by the **EURO** model for The Netherlands, Sweden, UK and by the Powered **EURO** model for Portugal. As Table 3 shows the predictions and the observations are reasonably close. The first two groups are highly congested, with average to high flow, narrow to medium width, and low to average traffic speed. These roads come from predominantly UK and Portugal. RGROUP 3 consists of roads with low flow, narrow to medium width and low congestion. The proportion of speeders in the group is high, perhaps, because the flow is low the temptation to violate the speed limit is high on these roads. RGROUP 4 is at the other end of the scale with low flow, very wide roads, low congestion and high traffic speed. However, the above table is a rough guide only based on group averages. It should be noted that even though their mean values are significantly different there will be a small overlap in the range of the 'discriminating' variables.

Table 4 shows the cross-tabulation of links by **RGROUP** and Speed Limit. It shows that most of the links in **RGROUP 1** and **2** come from roads where 90 and 100 km/h speed limits are in force and those in **RGROUP 4** are from 110 km/h speed limit roads. **RGROUP 3** is a good mixture of roads from 70 to 100 km/h speed limit groups.

Table 1: Distribution of links amongst Groups by Country

Country	RGROUPS				Total
	RGROUP 1	**RGROUP 2**	**RGROUP 3**	**RGROUP 4**	
Netherlands	0	1	26	1	28
Portugal	14	12	3	0	29
Sweden	4	5	37	27	73
UK	21	6	11	0	38
All	39	24	77	28	168

Note: Data excludes outliers and links that do not have the information on features used in the cluster analysis.

Table 2: Summary statistics of feature variables by RGROUP

RGROUP		ln(Q) [Q]	ln(V) [V]	SD	Width(W)	ln(P) [P]	ln(Cg) [Cg]
1		8.76317 *(6394)*	4.17393 *(64.97)*	11.13	7.01	0.15636 *(1.17)*	1.71786 *(5.57)*
2		8.52725 *(5050)*	4.39145 *(80.75)*	17.89	7.03	3.13015 *(22.88)*	1.13306 *(3.10)*
3		8.26277 *(3877)*	4.42546 *(83.55)*	12.50	7.32	3.65708 *(38.75)*	-0.54744 *(0.58)*
4		8.59077 *(5381)*	4.56246 *(95.81)*	13.84	12.85	3.28448 *(26.69)*	-0.66492 *(0.51)*
All	Av.	8.47138 *(4776)*	4.38504 *(80.24)*	13.18	8.13	2.70704 *(14.98)*	0.19892 *(1.22)*
	sd	0.87441 *(2.40)*	0.16597 *(1.18)*	2.80	2.38	1.65748 *(5.24)*	1.44101 *(4.22)*

Note: Anti-logs are shown within brackets in (italics).

Table 3: Qualitative description of road groups.

RG	Q	V(km/h)	Width(m)	P(%)	Cg(%)	AF Pred (Obs)
1	High	Low	(5 to 10)	Very Low	High	11.97 (14.31)
2	Av.	Av.	(6 to 8)	Above Av.	High	10.13 (8.96)
3.	Low	Av.	(5 to 10)	High	Low	1.89 (1.87)
4.	Av.	High	(10 to 14.5)	Above Av.	Low	1.15 (1.15)

Table 4.: Distribution of links by Group and Speed Limit.

Speed Limit (km/h)	RGROUP				Total
	RGROUP 1	**RGROUP 2**	**RGROUP 3**	**RGROUP 4**	
70	2	1	12	0	15
80	0	1	24	0	25
90	16	14	28	9	67
100	21	6	13	1	41
110	0	2	0	18	20
All	39	24	77	28	168

Multivariate stepwise Discriminant analysis (LDA) was carried out to determine the best discriminating variables between the four groups. LDA generates linear combinations of variables, called 'Discriminant Functions', which separate groups in such a way that the between group variation is maximised relative to their within group variation. Further investigation revealed that the best discriminators between the four groups were road width (**W**), mean speed (**V**), standard deviation (**SD**) of speed and proportion exceeding the speed limit (**P**). This means that the groups can best be described in terms of these four variables.

7 DISCUSSION OF RESULTS

The road classification in the previous section and the summary description of the road groups (Table 3) give some insight into the way the road groups differ in respect of their descriptive variables. Despite the fact that the description of each group is complicated there is one thing clear -that the average accident frequencies decrease from **RGROUP 1** to **RGROUP 4**, and this decrease is associated with increasing mean speeds (**V**) and decreasing proportion of slow vehicles (**Cg**). This negative association between the mean speed and congestion is supported by the European model for the mean speed [Equation (1)], where the mean speed (**V**) was found to be inversely related to **Cg**.

The **EURO** model describes how accident frequency (**AF**) is inversely related to the mean speed (**V**), which is possibly due to the effect of road environment and/or road geometry. A partial explanation will be found in the following example. To understand the effect of **P** on the mean speed (**V**) and accident frequency (**AF**), we shall hold the geometry and flow variables fixed at the following values - **S** = 90 km/h, **W** = 7m, **L** = 4 km, **NJ** = 3 and **Q** = 8000 per day. Consider four hypothetical roads - R1, R2, R3 and R4, of the same specification, but to differ in the value of **Cg** . Let **Cg** = 5, 10, 20 and

40 percent respectively. So these roads will differ only by the level of congestion, otherwise they are identical.

The mean speed (**V**) and the accident frequency (**AF**) at P = 1% can be calculated using Equations (1) and (3). These values of **V** will be treated as the 'baseline' values (V_0, say) of speed when the speed limit violation is minimal (**P** = 1%), and they are (61.57, 60.19, 58.84, 57.52) km/h respectively. The associated values of **AF**, calculated at V_0, using Equation (3), will also be treated as the 'baseline' values, (AF_0, say), of accident frequencies at **P** = 1%. They are 3.15, 3.34, 3.53 and 3.74 respectively. These calculated values show that for increasing values of **Cg** the values of V_0 are decreasing, at the same time the values of AF_0 are increasing.

The coefficient of **ln(P)** in the **EURO** model is 0.114, which suggests that for every 10 percent (absolute) change in **P** (ie Δ**P** = 10) accident frequency **(AF)** is changed by **114/P** percent. This does not immediately translate into accident reduction for a unit (1 km/h) change in speed. To convert the statement to one in terms of a unit change in speed we need to invoke the relationship between **V** and **P**. Equations (1) and (3) can be expressed in terms of **P**, by

$$\mathbf{V = V_0 \,.\, e^{0.07443 * \ln(P)}}$$

$$\mathbf{AF = AF_0 \,.\, e^{0.1143 * \ln(P)}}$$

assuming that everything else is constant. Using different values of **P**, the values of **V** and **AF** can be calculated to see how a *marginal* change in speed, as well as in accident frequency, is ‘effected’ simultaneously by a change in **P**. The marginal change in **AF** is associated with the marginal change in speed. Figure 4 shows the plot of the calculated values of AF against the corresponding calculated values of V for each road, for **P** = 1, 5, 10, 15, 20, ... ,80 percent.

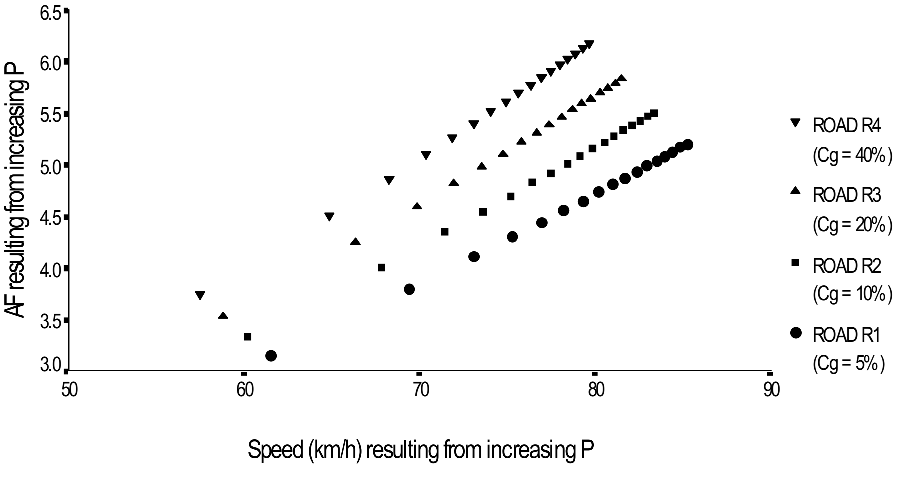

Figure 4. Plot showing accident frequency vs speed, resulting from increasing proportion of speeders (see text, W = 7 m, L = 4 km, NJ = 3, S = 90 km/h, Q = 8000).

In this plot the 'baseline' values are represented by the starting point (bottom left) for each road. The plot clearly shows increasing values of the baseline **AF** for decreasing values of **V** resulting from the increasing values of congestion. This phenomenon partly explains why a negative relationship is found when **AF** is regressed on **V**. In this simple example the roads differ only by a single variable - **Cg**. It is, therefore, reasonable to suggest that poor road environment, and perhaps poor road design, can cause high congestion, which can affect traffic speeds as well as road safety.

Figure 4 also shows points moving diagonally upwards to the right for each hypothetical road, for increasing values of **P**, indicating that increasing speed limit violation causes the mean speed to rise and accident frequencies to increase on the same road. This feature of the plot suggests that it will be possible to make a statement for changes in **AF** for a unit change in **V** resulting from the marginal effect of **P** on both.

The elasticity of **V** with respect to **P**, η_{VP} $[= \delta\ln(V) / \delta\ln(P)]$, say, is from Equation (1) 0.07443. The elasticity of **AF** with respect to **P**, η_{AP} $[= \delta\ln(AF) / \delta\ln(P)]$, say, is [from Equation (3)] 0.1143. This gives $[\, \delta\ln(AF) / \delta\ln(V) \,] = 0.1143 / \eta_{VP} = 1.536$. Hence, $\delta\ln(AF)/\delta V = (1.536 / V)$ which enables us to estimate the percentage change in accident frequency with respect to a unit change in speed for The Netherlands, Sweden and UK. Using this equation for V = 60, 70, 80 and 90 km/h, the percentage change in accidents for 1 km/h change in speed are respectively 2.56, 2.19, 1.92 and 1.71 percent. For the Portuguese links, the Powered **EURO** model gives $\delta\ln(AF) = (1.536*1.35 / V)\, \delta V = (2.074 / V)\, \delta V$, which suggests that the percentage changes are, respectively, 3.45, 2.95, 2.59 and 2.31 percent.

Allsop (1998) recently evaluated the net effect of speed on accident frequency, using the models for the mean speed [Equations (1)] and accident frequency [Equation (3)], and estimated that a reduction of about 8% can be achieved in the number of accidents corresponding to a 5 km/h reduction (from 90 to 85 km/h) in mean speed. This estimate compares well with the estimate of 8.55 % reduction (5 x 1.71 at 90 km/h) mentioned above.

Thus for the rural roads the following two equations can be used to calculate the percentage reduction of AF for a 1 km/h reduction in the mean speed.

$$\Delta \ln(AF) = [1.536 / V\,] \,.\, \Delta V \quad \text{for The Netherlands, Sweden \& UK} \qquad (5)$$

$$\Delta \ln(AF) = [2.074 / V\,] \,.\, \Delta V \quad \text{for Portugal} \qquad (6)$$

For the urban roads of UK Baruya and Finch (1994) found that accident frequency (**AF**) was proportional to $V^{1.573}\ e^{4.427.\ Cv}$, where **Cv** is the *coefficient of variation* (standard deviation to mean ratio) of speed. Since **Cv** is related to **V**, which can be approximated by $Cv = 0.448 - 0.00485*V$ (for **V** in km/h) the percentage change in **AF** for 1 km/h change in **V** can be calculated from

$$\Delta \ln(AF) = [1.573 / V - 0.02146] . \Delta V \tag{7}$$

There being no further published results available, to the best of our knowledge, for urban roads we shall use this equation for the European Union. This equation as well as Equations (5) and (6) for rural roads produce the plot of Figure 5, which shows the percentage change in **AF** for 1 km/h change in the mean speed (**V**). The mean speeds of the UK urban roads on which the urban model was based ranged from 18 miles/h (29 km/h) to 35 miles/h (56 km/h), and those of the European rural roads ranged from 50 km/h to 110 km/h. The plot shows a slight overlap in urban and rural speeds, and demonstrates the extent to which the different road designs which give rise to similar speeds in the two areas, result in a different potential for accident reductions through speed change.

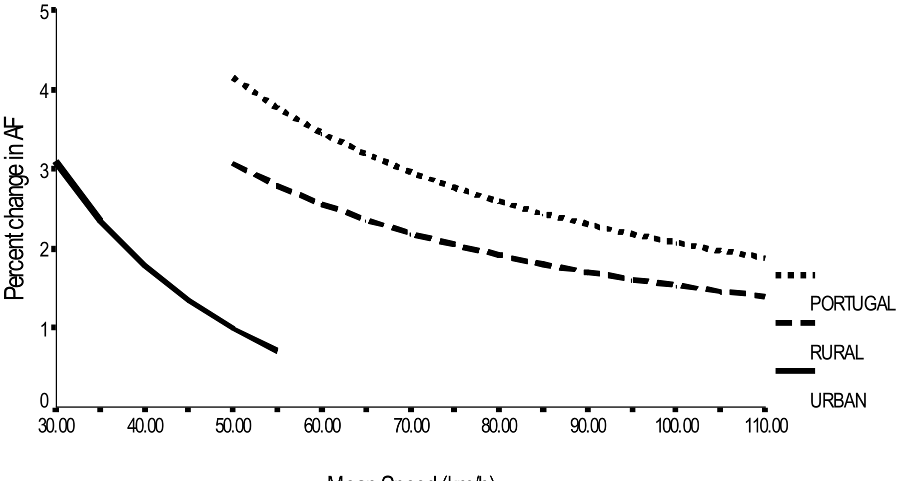

Figure 5. Percent change in accident frequency for 1 km/h change in mean speed.

8 CONCLUSIONS AND RECOMMENDATIONS

Despite diversity and individuality most European roads have common characteristics in respect of their range of flows, speeds and traffic environment, irrespective of their country of origin. A comparison of flow, speed and road geometry reveals that there is a good deal common amongst the European roads in respect of road environment. This enables us to classify them into homogeneous groups whose membership extends beyond national boundaries.

Speed parameters, such as mean speed and standard deviation of speed are strongly influenced by speed limit, flow, geometry and traffic variables (eg. proportion of speeders, congestion, etc). A speed-accident model (**EURO**) has been developed from multi-national data for application to a variety of European rural roads. The investigation also highlights the type of roads, particularly those with high accident rates, where a modified version **(Powered EURO**) of the model may need to be applied. The **EURO** model provides a reasonable estimate of accident frequency both for individual links and collectively for a number of links. It is less satisfactory in its original form for roads where accident rates are vary high, where the use of the **Powered EURO** is recommended.

Every effort has been made to make the **EURO** model suitable for application to different types of rural roads in the European Union. Despite the likelihood that the model will successfully predict accident frequency for a wide variety of roads it is recommended that, where possible, the model is tested and assessed prior to use in a particular situation. The scope of the model is limited to the European road environment.

The speed-accident relationship for urban roads, referred to in this report, was based on a UK study carried out by Baruya and Finch (1994). In absence of any similar study in other European countries it has been assumed that a similar relationship exists in other European countries. This assumption has not been verified. It is recommended that the urban relationship study is assessed on the urban roads of other European countries in order that the scope of the relationship or a modified version of it is made wider.

Similar studies are recommended for other European countries for even wider application. This study covers only single carriageway roads. Similar study is recommended for motorways and dual carriageways.

ACKNOWLEDGEMENTS

The author is grateful to Oei Hway-Liem of SWOV (The Netherlands), Gunnar Andersson of VTI (Sweden), Paulo Simoes of TRANS-POR (Portugal) and the local authorities of England who supplied the data for this research. The author is also grateful to Veli-Pekka Kallberg of VTI (Finland), Professor Richard Allsop of UCL (UK), Heather Ward of UCL (UK), David Lynam (TRL), Pat Wells (TRL) and Dave Finch (TRL) for their constructive comments, suggestions and support during various stages of this research. Thanks also to Mike Winnett, Elaine Woodgate and Fred James and others of TRL for their assistance in the compilation of the UK speed and accident data used in this report. The UK data were gathered with financial support from the Road Safety Division of the UK Department of Environment Transport and the Regions (DETR).

REFERENCES

ALLSOP, R E. (1998). Summary of Research Area 1: Basis for appraisal of effects of different levels of speed. MASTER Working Paper R 1.3.1.

ANDERBERG, M R (1973). *Cluster analysis for applications.* Academic Press, New York.

ANDERSSON, G and G NILSSON (1997). Speed management in Sweden: Speed, Speed limits and safety. Published by VTI.

ETSC - European Transport Safety Council (1997). *Transport accident costs and the value of safety.* Brussels. ISBN 90-801936-9-0

BARUYA, A and D J FINCH (1994). Investigation of traffic speeds and accidents on urban roads. *Proceedings of The 22nd European Transport Forum, PTRC.* Warwick University.

BARUYA, A (1995). An investigation of the effect of speed on accidents on UK urban roads. *TRL Annual Review, pp 55 - 61.*

BARUYA, A (1997). MASTER: A review of Speed-Accident relationship for European Roads. Working Paper R 1.1.1.

BARUYA, A (1998a). MASTER: Speed-Accident relationship on single-carriageway roads of UK, Working Paper R 1.1.2.

BARUYA, A (1998b). MASTER: Speed-Accident relationship on European roads. Working Paper R 1.1.3.

BARUYA, A (1988c). MASTER: Speed-Accident relationship on different kinds of European roads, Working Paper R 1.1.4, (Deliverable D7).

ETSC - European Transport Safety Council (1997). *Transport accident costs and the value of safety.* Brussels. ISBN 90-801936-9-0

FIELDS, B N and S J LEE (1993). The Speed Review: Road Environment, Behaviour, Speed Limits, Enforcement and Crashes. Monash University, Australia.

FINCH, DJ, P KOMPFNER, C R LOCKWOOD and G MAYCOCK (1994). Speed, speed limits and accidents. Department of Transport, *TRL Project Report PR 58*, Crowthorne: Transport Research Laboratory.

FRIDSTROM L., J IFVER, S. INGEBRIDGTSEN, R. KULMALA and L K THOMSEN (1995). Measuring the contribution of randomness, exposure, weather, and daylight to the variation in road accident counts. *Acc. Anal and Prev. , Vol 27, No 1, pp 1-20.*

GARBER, N J and R GADIRAU (1988). Speed variance and its influence on accidents. *AAA Foundation for Traffic Safety*, Washington, DC.

HAUER, E (1971) Accidents, overtaking and speed control. *Acc. Anal. and Prevention, Vol 3, pp 1-13.*

JORGENSEN, N O (1969). A model for forecasting traffic accidents. *Proceedings of the symposium on the use of statistical methods in the analysis of road accidents.* Held at Road Research Laboratory, Crowthorne, Berks.

LAVE, C A (1985). Speeding co-ordination and the 55-miles/h limit. *American Economic Review, Vol 75 (Dec), pp 1159-1164.*

MAYCOCK, G and R D HALL (1984). Accidents at 4-arm roundabouts. Department of Environment Department of Transport, *TRRL Laboratory Report 1120*. Crowthorne: Transport and Road Research Laboratory.

McCULLAGH P and J A NELDER (1989). *Generalised Linear Models,* 2nd Edition, Chapman and Hall, London.

MAHER, M J and I SUMMERSGILL, (1996). A comprehensive methodology for the fitting of predictive accident models, *Acc. Anal & Prev., Vol 28, No. 3, pp 281-296.*

McCULLAGH P and J A NELDER (1989). *Generalised Linear Models,* 2nd Edition, Chapman and Hall, London.

SOLOMON, D. (1964) *Accidents on main rural highways related to speed, driver and vehicle.* Bureau of Public Roads, Department of Commerce, Washington, USA.