Accid. Anal. and Prev. Vol. 25, No. 5, pp. 499–509, 1993
Printed in the U.S.A.

0001-4575/93 $6.00 + .00

MODELS FOR PREDICTING ACCIDENTS AT JUNCTIONS WHERE PEDESTRIANS AND CYCLISTS ARE INVOLVED. HOW WELL DO THEY FIT?

ULF BRÜDE and JÖRGEN LARSSON

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

(*Received* 12 *May* 1992)

Abstract—The coefficient of determination, R^2, i.e. the squared correlation coefficient between observed and fitted values, is often used as a measure of how well a model predicts the number of accidents at road junctions, for instance. The purpose of this article is to show that the R^2 values obtained in different studies are rarely comparable with each other and that a prediction model can be "nearly perfect" even if the coefficient of determination is small. Another purpose of the article is to present some results of interest from a practical viewpoint in regard to accidents where pedestrians and cyclists are involved. Empirical R^2 values for models predicting accidents at junctions where pedestrians or cyclists are involved are compared with the maximal R^2 values that could possibly be obtained. The latter can be calculated both theoretically and with the aid of simulation. How the maximal R^2 value depends on the average accident level and the relative dispersion of the expected values for the studied junctions is also shown theoretically. The results obtained show how difficult it can be to determine whether and how far the number of accidents is influenced by additional factors, over and above the traffic flows, which describe the design in greater detail.

INTRODUCTION

Several attempts have been made over the years to construct models for predicting the number of accidents at road junctions or on sections of road. Common to all the models is that they include traffic flows in some form or other. Furthermore, although the traffic flows are often estimated very roughly, they are frequently the primary factor behind accidents, at least after a classification has been made according to the type of junction, i.e. ordinary (non-signalized) junctions, signalized junctions, roundabouts, and grade-separated junctions; or type of road, e.g. ordinary two-lane roads, expressways, and motorways; and the legal speed limit.

A frequently used measure describing how good the models are is the coefficient of determination, R^2, i.e. the squared correlation coefficient between observed and fitted values. Other existing goodness-of-fit measures will not be treated in this paper. Occasionally, it is stated that a particular model is better than another in a different study, because it has a larger R^2 value. Sometimes it is stated that a model is not particularly good because its R^2 value is low, perhaps "only" 0.30. In some studies, the R^2 value has been calculated with individual junctions as observation units, while groups of road sections make up the observation units in other studies.

The purpose of this article is to illustrate that a prediction model may be extremely good even if it has a small R^2 value. The article also attempts to show that the R^2 values obtained in different studies are not, as a rule, directly comparable with each other because of differences in the length of the observation periods, differences in the number of accidents, or differences in the dispersion of the expected number of accidents at the observation units studied. Only cases having natural *individual observation units* are covered here. A number of results that are of interest from a practical viewpoint are also presented.

PREDICTION MODELS, DATA AND METHOD

The prediction models described here relate to accidents at individual junctions where either pedestrians or cyclists are involved.

Data has been collected from some 30 Swedish municipalities ranging from those with about 25,000 inhabitants up to the largest in the country, i.e. Gothenburg and Stockholm. All junctions are part of the major road and street network in urban areas. Their design varies from simple, ordinary junctions to junctions with traffic signals or roundabouts. The accident data consist of accidents reported by the police from 1983 to 1988.

Only junctions with 100 or more pedestrians or 100 or more cyclists per annual average day have been included. Counts of unprotected road users have generally been made on a single weekday during the spring or autumn.

A pedestrian using two pedestrian crossings is counted twice. Similarly, a cyclist using two bicycle crossings is counted twice. On the other hand, a cyclist arriving on the carriageway is counted only once. This latter procedure is, of course, not wholly unobjectionable.

In the construction of prediction models, 285 junctions with 165 accidents have been used for accidents involving pedestrians, and 377 junctions with 432 accidents have been used for accidents involving cyclists. The statistical technique used was the least squares method with the aid of NLIN in the SAS statistics program and with the accident rate as the dependent variable. (We have also tried the GLIM procedure, with the number of accidents as the dependent variable, and obtained essentially the same results.)

EQUATIONS FOR THE PREDICTION MODELS

For accidents involving pedestrians, the number of accidents involving pedestrians per million passing pedestrians has been obtained as

$$PACCRATE = 0.0201 \times TOTINC^{0.50} \times TOTPED^{-0.28} \quad (1)$$

where *TOTINC* and *TOTPED* are equal to the number of incoming motor vehicles and passing pedestrians per average annual day.

The annual number of accidents involving pedestrians, *PACCPERYEAR*, is then obtained as

$$PACCRATE \times 365 \times TOTPED^{1}/10^{6},$$

i.e.

$$PACCPERYEAR = 0.00000734 \times TOTINC^{0.50} \times TOTPED^{0.72} \quad (2)$$

(Conversely, *PACCRATE* can be expressed as $PACCPERYEAR/365/TOTPED^{1} \times 10^{6}$.)

The corresponding equations have been obtained for accidents involving cyclists:

$$CACCRATE = 0.0494 \times TOTINC^{0.52} \times TOTCYC^{-0.35}, \quad (3)$$

and

$$CACCPERYEAR = 0.0000180 \times TOTINC^{0.52\ (4)} \times TOTCYC^{0.65}. \quad (4)$$

As expected, the exponents in the equations for *PACCPERYEAR* and *CACCPERYEAR* assume values between 0 and 1. In both cases, the exponents for the number of unprotected road users are larger than for the number of motor vehicles.

Note that in the equations for *PACCRATE* and *CACCRATE*, the exponents for *TOTPED* and *TOTCYC* are negative. The exponents for *TOTINC* are of course positive and are the same as for *PACCPERYEAR* and *CACCPERYEAR*.

Put into words, the relationships shown above mean that both *PACCPERYEAR* and *CACCPERYEAR* increase with increasing numbers of motor vehicles and unprotected road users (see Figs. 1 and 2). It is even more interesting that the risk, the accident rate for unprotected road users, i.e. *PACCRATE* and *CACCRATE*, increases with increasing numbers of motor vehicles, but decreases with increasing numbers of pedestrians and cyclists, respectively (see Figs. 3 and 4). To a large extent, the explanation for this could very well be that the more motor vehicles there are, then the larger is the number that can conflict with pedestrians or cyclists. And the more pedestrians and cyclists there are, then the larger is the number of "living warning signs" and the greater the consideration shown them by motorists. An additional explanation in the latter case may of course be that the more unprotected road users there are, then the more extensive are the measures applied in junction design with regard to pedestrians and cyclists. However, it should be pointed out that the number of cars also has a bearing on the inclination to apply such measures.

Another interesting observation is that the risk of an accident (with the given definition of risk) is roughly twice as high for cyclists as for pedestrians. However, it should be mentioned that pedestrian single accidents are not reported by the police, as they are not defined as road accidents. And that pedestrian-cyclist accidents (quite a small number) are counted only as cycle accidents. If the omissions due to deficiencies in police reporting were taken into account, the accident numbers for pedestrians and cyclists would be roughly twice as high.

Tables 1 and 2 show how the junctions in total are distributed according to numbers of motor vehicles on the one hand and numbers of pedestrians and cyclists on the other hand, and also how the subset of junctions with PC* signals is similarly dis-

* The abbreviation PC used here refers to pedestrians and cyclists.

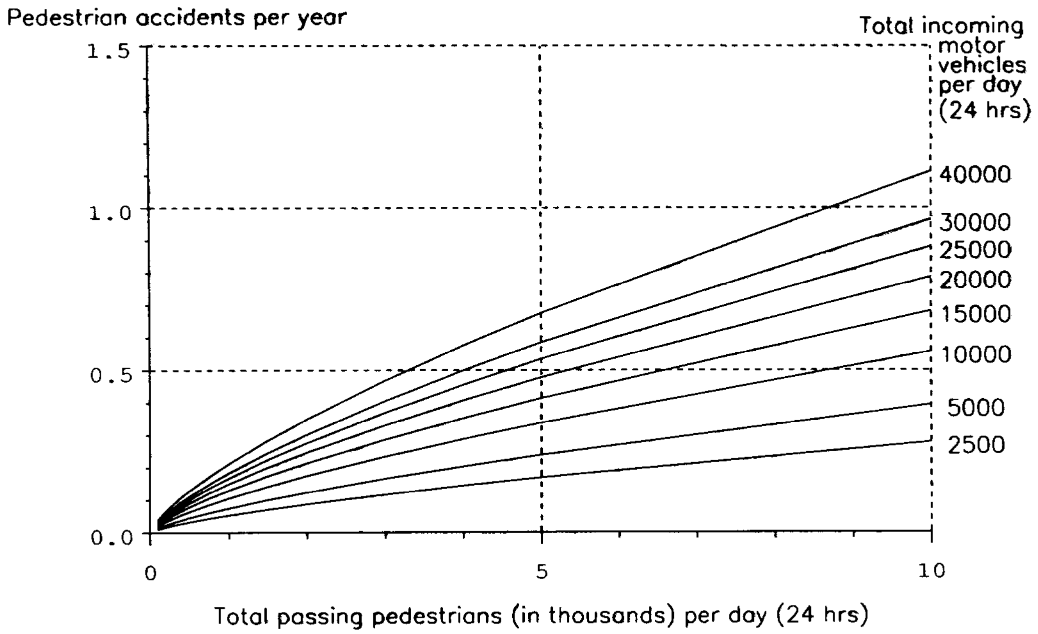

Fig. 1. $PACCPERYEAR = 0.00000734 \times TOTINC^{0.50} \times TOTPED^{0.72}$.

tributed. As can be seen, there is a strong relationship between the occurrence of PC signals and size of flow.

From a practical viewpoint, the results are interesting and at first perhaps somewhat hard to interpret or accept. The junctions having the lowest PC risk, i.e. the lowest *PACCRATE* or *CACCRATE* (given a certain number of motor vehicles), are those having the largest PC flows. At the same time, however, these junctions are the ones where numerous PC accidents occur precisely because of the high unprotected road-user exposure. Grouping pedestrians and cyclists together at a smaller number of junctions with a high PC standard (and not too much motor traffic) would therefore appear to be maximally safety-effective. From a practical point of view, however, this might be somewhat difficult to achieve.

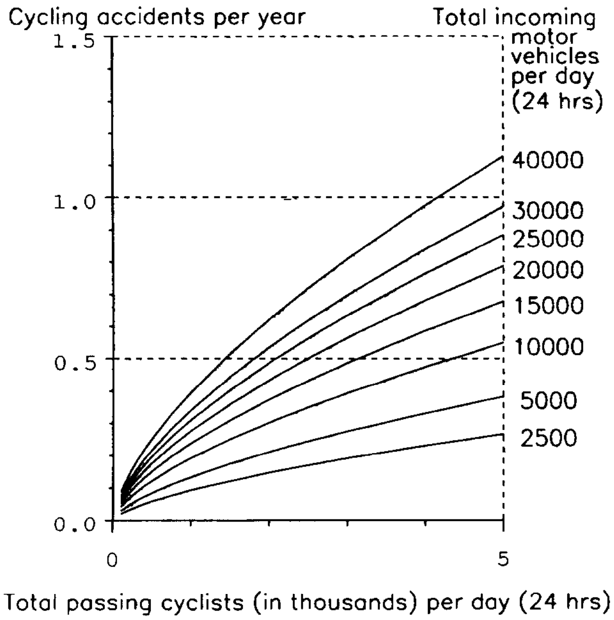

Fig. 2. $CACCPERYEAR = 0.0000180 \times TOTINC^{0.52} \times TOTCYC^{0.65}$.

OVERALL "RESIDUAL ANALYSIS"

Tables 3 and 4 contain an overall comparison of the number of observed and predicted accidents for groups of junctions, broken down by speed limit and type of junction. The deviations are alternately positive and negative and no systematic pattern is discernible.

It is unfortunate that the studied material contains such a small number of roundabouts. In point of fact, the question of how effective or ineffective roundabouts are from the viewpoint of the road safety of unprotected road users is a subject of much controversy.

DEGREE OF DETERMINATION

If *PACCPERYEAR* and *CACCPERYEAR* are used to predict the annual accident numbers for 1983, . . . , 1988, the following R^2 values (squared

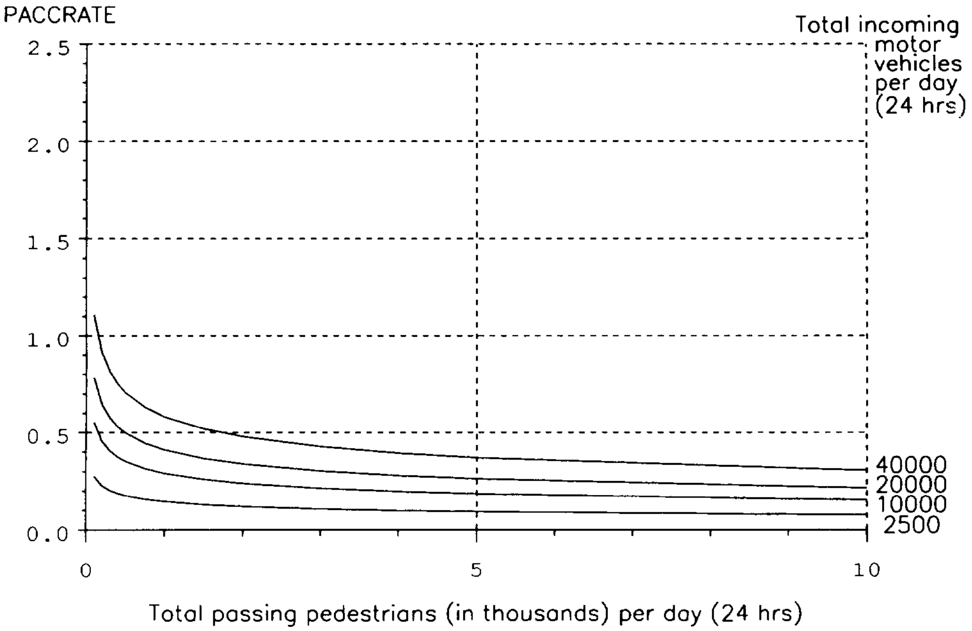

Fig. 3. $PACCRATE = 0.0201 \times TOTINC^{0.50} \times TOTPED^{-0.28}$.

Table 1. Junctions where pedestrian accidents have been analysed. Total number of junctions and number of junctions with PC signals

100	250	750	1250	3750	6250	13750	16250	Number of incoming motor vehicles per day (24 hrs)
		1 1	1 1					47500
1 1	4 3	1 1	2 2					32500
1 1	3 3	2 1		1 1	1 1	1 1		27500
3 3	7 5	6 6	6 6	1 1	1 1			22500
13 8	16 13	5 4	6 4	1 1	1 1			17500
16 6	27 15	6 3	9 6	1 1	1 1			12500
23 7	39 13	8 4	16 5	1 0				7500
24 1	22 4	2 1						2500
2 0	3 0							0

Number of passing pedestrians per day (24 hrs)

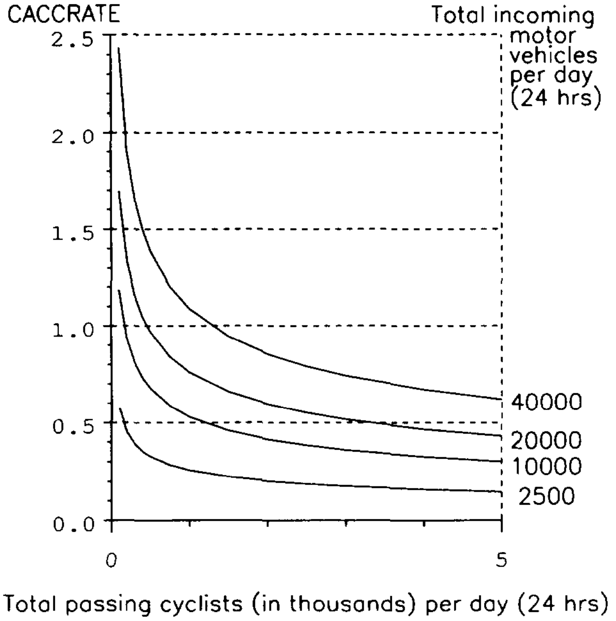

Fig. 4. $CACCRATE = 0.0494 \times TOTINC^{0.52} \times TOTCYC^{-0.35}$.

correlation coefficients) will be obtained (Tables 5 and 6):

On average, the coefficients of determination will be small, about 0.17 and 0.08, respectively.

If we calculate the R^2 values for the entire 1983–1988 period instead, i.e. the period used in constructing the prediction models, the coefficients of determination usually reported will be obtained. As will be evident from Table 7, the R^2 values then increase to 0.46 and 0.31, respectively.

If studies of the entire 1983–1988 period are continued but the R^2 values are calculated separately for the 50% of junctions having the lowest number of predicted accidents and the 50% having the highest number of predicted accidents, then the coefficients of determination will be smaller. This applies above all to the first-mentioned junction (Table 8).

Really high correlations are seldom to be expected when the observation units are individual junctions and when, as here, the study concerns random variables such as the occurrence of acci-

Table 2. Junctions where cycling accidents have been analysed. Total number of junctions and number of junctions with PC signals

Number of incoming motor vehicles per day (24 hrs) \ Number of passing cyclists per day (24 hrs)	100	250	750	1250	3750	6250	13750	16250
(above 47500)	1 1			1 1				
47500		1 1	2 2	4 4				
32500	1 1	1 1	3 3	5 4				
27500	2 2	12 11	6 6	7 6				
22500	6 2	13 9	16 12	12 10				
17500	10 3	29 16	16 7	17 9	1 1			
12500	10 2	43 13	34 11	27 4				
7500	18 0	37 1	17 5	14 1	1 0			
2500	1 0	6 1	2 0	1 0				
0								

Table 3. Accidents involving pedestrians

	Number of junctions	Acc. 83–88	6 × *PACCPERYEAR*
50 km/h			
3-way ordinary junctions	59	14	17.9
3-way signalized junctions	15	28	21.2
4-way ordinary junctions	83	37	36.9
4-way signalized junctions	89	75	81.9
4-way roundabouts	8	2	4.2
70 km/h			
3-way ordinary junctions	6	1	0.9
3-way signalized junctions	1	0	0.1
4-way ordinary junctions	7	2	1.6
4-way signalized junctions	16	6	4.7
4-way roundabouts	1	0	0.2
Total	285	165	169.6

Table 4. Accidents involving cyclists

	Number of junctions	Acc. 83–88	6 × *CACCPERYEAR*
50 km/h			
3-way ordinary junctions	93	61	75.3
3-way signalized junctions	18	17	23.2
4-way ordinary junctions	103	97	102.3
4-way signalized junctions	91	166	161.1
4-way roundabouts	8	15	11.2
70 km/h			
3-way ordinary junctions	16	16	10.0
3-way signalized junctions	2	0	0.9
4-way ordinary junctions	22	18	20.5
4-way signalized junctions	23	40	35.6
4-way roundabouts	1	2	1.7
Total	377	432	441.9

Table 5. Accidents involving pedestrians, year by year, R^2 values

	Acc 83	Acc 84	Acc 85	Acc 86	Acc 87	Acc 88
PACCPERYEAR	0.09	0.35	0.16	0.09	0.11	0.20

Table 6. Accidents involving cyclists, year by year, R^2 values

	Acc 83	Acc 84	Acc 85	Acc 86	Acc 87	Acc 88
CACCPERYEAR	0.05	0.12	0.11	0.04	0.05	0.09

Table 7. Accidents throughout the entire period from 1983 to 1988, R^2 values

	Acc 83–88
6 × *PACCPERYEAR*	0.46
6 × *CACCPERYEAR*	0.31

dents. In particular, they are not expected in cases of low accident levels when the relative uncertainty will be especially high.

Even if absolutely correct expected values have been successfully predicted, perfect agreement between observed and predicted (fitted) values can never be achieved on account of the randomness of the accidents.

Appendices 1 and 2 list the numbers of observed and predicted accidents, junction by junction, for the entire 1983–1988 period, broken down by accidents involving pedestrians and accidents involving cyclists. From these Appendices it is evident, for instance, that there are 40 junctions where the number of accidents involving pedestrians, rounded off to one decimal place, is predicted as 0.1. At 34 of these junctions, the observed number of accidents was zero and at six of them only one accident occurred. According to the theoretical Poisson distribution, with expected value m equal to 0.1, no accidents should have occurred at 36 of the junctions and one or more accidents at four of them. In general, similarly good agreement can be noted throughout.

RELATIONSHIP BETWEEN PREDICTED VALUES AND ESTIMATED EXPECTED VALUES

By weighting observed and predicted accidents together as in the following equation

$$\hat{m} = pred + (a \times pred/(1 + a \times pred)) \times (obs - pred) \quad (5)$$

an estimate of the "true" (expected) number of accidents m will be obtained (Brüde and Larsson 1988).

In eqn 5, a can be determined by means of cross validation (Junghard 1990). In Swedish junction studies covering all accidents reported by the police, a has usually ended up in the 0.10–0.30 interval. Furthermore, it has been found that the value is fairly "robust" in the sense that it makes little difference whether a is assumed to be 0.10, 0.20, or 0.30.

For the accident material we are concerned with here, which covers accidents involving pedestrians and accidents involving cyclists, a values of 0.29 and 0.08, respectively, have been obtained.

Apart from listing observed and predicted values, Appendices 1 and 2 also contain estimates of m for each junction. The latter are extremely close to the predicted values throughout. This also becomes evident when calculating the squared correlation coefficients between predicted numbers of accidents and estimates of m, which will be equal to 0.90 for accidents involving pedestrians and 0.95 for accidents involving cyclists (Table 9).

In actual fact, the main interest lies in the ability of the models to predict the (nonobservable) expected values. The results indicate that in this respect the prediction models are "nearly perfect". However, it is natural for the objection to be raised that the estimates of m constitute a function of the predicted values and that this is why the squared correlation coefficients lie so close to 1.

MAXIMAL COEFFICIENTS OF DETERMINATION, SIMULATED AND THEORETICAL

It is possible to examine what happens if it is assumed for each individual junction that the predicted value is equal to the actual expected value and a number of accidents (on the assumption of a Poisson distribution) is then simulated. Ten such simulations, each of accidents involving pedestrians and accidents involving cyclists during the 1983–1988 period, have been performed. The squared correlation coefficients between the assumed expected values and the simulated values show the maximal R^2 values that can be obtained for the prediction models used (Table 10).

Table 8. Subgroups of junctions, accidents throughout the entire period from 1983 to 1988, R^2 values

	Acc 83–88
Lowest number of predicted accidents, 6 × *PACCPERYEAR*	0.02
Highest number of predicted accidents, 6 × *PACCPERYEAR*	0.45
Lowest number of predicted accidents, 6 × *CACCPERYEAR*	0.04
Highest number of predicted accidents, 6 × *CACCPERYEAR*	0.20

Table 9. Squared correlation coefficients

	$\hat{m}$ 83–88
6 × *PACCPERYEAR*	0.90
6 × *CACCPERYEAR*	0.95

On average, the squared correlation coefficients will be 0.49 and 0.40, respectively. According to Table 7, squared correlation coefficients of 0.46 and 0.31, respectively, between predicted and observed values were obtained. The difference is surprisingly small, particularly in regard to accidents involving pedestrians. It is also interesting to note that the squared correlation coefficients can vary so widely in different simulations.

Furthermore, the maximal R^2 values are doubtless somewhat overestimated because the number of accidents occurring at individual junctions probably has a somewhat larger dispersion than is indicated by the Poisson distribution and also because omissions in police reporting vary between different municipalities.

On the other hand, it should of course be pointed out that the same data that have been used for constructing the prediction models have been used for comparing predicted and observed numbers of accidents.

The maximal degree of determination can also be derived and calculated theoretically.

Let X denote the predicted number of accidents and assume that the expected values for individual junctions are in full agreement with the predicted values.

Further, let Y denote the observed number of accidents.

Seen over all the junctions studied, X has a distribution with expected value $E(X)$ and variance $V(X)$.

For a certain given junction with $X = x$ it is assumed that Y has a Poisson distribution with expected value $E(Y|X = x) = x$ and variance $V(Y|X = x) = x$.

Seen over all junctions, then

$$E(Y) = E[E(Y|X)] = E(X)$$

$$E(XY) = E[E(XY|X)] = E[XE(Y|X)] = E(X^2)$$

$$V(Y) = E[V(Y|X)] + V[E(Y|X)] = E(X) + V(X)$$

$$\Rightarrow$$

$$Corr(X, Y) = \frac{Cov(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} = \frac{E(X^2) - E^2(X)}{\sqrt{V(X)[E(X) + V(X)]}} = \frac{V(X)}{\sqrt{V(X)[E(X) + V(X)]}}$$

$$\Rightarrow$$

$$Corr^2(X, Y) = \frac{V(X)}{E(X) + V(X)}, \qquad (6)$$

which means that the theoretical maximal coefficient of determination, $Corr^2(X, Y)$ increases with increased dispersion in the predicted number of accidents at the junctions studied. The measure will also be greater the larger the variance in the predicted values is in relation to the mean value of the predicted values.

Table 11 shows the mean value and variance for the predicted number of accidents at the junctions as well as $Corr^2(X, Y)$ for all junctions and for the subgroups with the lowest and highest number of predicted accidents.

Seen over all junctions, agreement with the simulated results in Table 10 is extremely good, as expected, but also with the results in Table 7 and Table 8.

It is easy to see what would happen with the maximal theoretical coefficient of determination $Corr^2(X, Y)$ if the length of the study period were to be halved, doubled, or extended by a factor of five or ten. If, in the equation for $Corr^2(X, Y)$, X is replaced by kX, where k denotes the length of the

Table 10. Squared correlation coefficients

	Sim. acc. 83–88									
	1	2	3	4	5	6	7	8	9	10
6 × *PACCPERYEAR*	0.49	0.49	0.49	0.36	0.56	0.53	0.50	0.39	0.56	0.48
6 × *CACCPERYEAR*	0.44	0.38	0.32	0.34	0.34	0.42	0.48	0.42	0.49	0.37

Table 11. $Corr^2(X, Y)$ for different groups

	Predicted number of accidents		
	Mean	Variance	$Corr^2(X, Y)$
Accidents involving pedestrians			
All junctions	0.5931	0.5890	0.50
Junctions with lowest prediction	0.2016	0.0062	0.03
Junctions with highest prediction	0.9818	0.8664	0.47
Accidents involving cyclists			
All junctions	1.1724	0.7009	0.37
Junctions with lowest prediction	0.5899	0.0514	0.08
Junctions with highest prediction	1.7519	0.6739	0.28

alternative period, then

$$Corr^2(kX, Y) = \frac{V(kX)}{E(kX) + V(kX)}$$

$$= \frac{k^2V(X)}{kE(X) + k^2V(X)} \quad (7)$$

is obtained, which will be greater or less than $Corr^2(X, Y)$, depending on whether k is greater or less than 1.

As will be seen, a change in the length of the study period, and with it the accident level and dispersion in the expected values, strongly affects the maximal degree of determination (Table 12).

STANDARDIZED COEFFICIENTS OF DETERMINATION?

Is there any good way of comparing more fairly the R^2 values of different prediction models? Something comparable with the coefficient of variation (standard deviation divided by the mean), for instance, which is sometimes used for comparing dispersion in different data materials. Would a conceivable means of standardizing the coefficient of determination be to divide it by the theoretical maximal coefficient of determination? These are some of the questions that come to mind.

Table 12. $Corr^2(kX, Y)$ for different values of k

	$Corr^2(kX, Y)$
Accidents involving pedestrians	
$k = 0.5$	0.33
$k = 1$	0.50
$k = 2$	0.67
$k = 5$	0.83
$k = 10$	0.91
Accidents involving cyclists	
$k = 0.5$	0.23
$k = 1$	0.37
$k = 2$	0.54
$k = 5$	0.75
$k = 10$	0.86

CONCLUSIONS

The results show that even with a simple model and simple estimates of the annual average daily numbers of motor vehicles and unprotected road users, it is possible to obtain "nearly perfect" models for predicting the number of accidents involving pedestrians or cyclists. This is in spite of the fact that the R^2 values obtained, the squared correlation coefficients between observed, and fitted values, may be extremely low.

Simulations and theoretical calculations can both be used for calculating the maximal R^2 value of a prediction model. The maximal R^2 value depends on the average level and the relative dispersion of the expected values for the junctions studied.

According to results obtained, the risk—the number of accidents involving unprotected road users per unprotected road user—increases with increasing numbers of motor vehicles but decreases with increasing numbers of pedestrians and cyclists. It would therefore appear to be favourable from a road safety viewpoint to "group together" pedestrians and cyclists, if possible, at a small number of road junctions with a high PC* standard and little motor traffic.

The risk of being involved in an accident is roughly twice as high for cyclists as for pedestrians in comparable traffic flows.

The results also indicate how hard it may be to decide whether additional factors that describe the design in greater detail influence the number of accidents, and in such a case by how much. This is

* The abbreviation PC used here refers to pedestrians and cyclists.

partly because traffic flows explain the systematic variation in accident frequency to such a large extent and partly because of the strong connection between traffic volume and the occurrence of PC signals, for instance.

Acknowledgements—The writing of this article has been possible through a project which the Swedish Road and Traffic Research Institute (VTI) has been carried out on behalf of the Swedish National Road Administration and also through funding from VTI's own resources. Greatly appreciated advice and suggestions on organizing and revising the contents of the article have been generously given by Mats Wiklund, VTI, and Stig Danielsson, University of Linköping.

REFERENCES

Brüde, U.; Larsson, J. The use of prediction models for eliminating effects due to regression-to-the-mean in road accident data. Accid. Anal. Prev. 20:299–310; 1988.

Junghard, O. Linear shrinkage in traffic accident models and their estimation by cross validation and bootstrap methods. Linköping Studies in Science and Technology. Thesis, No. 205. Linköping, Sweden: Linköping University; 1990.

APPENDIX A

Table A1. Accidents involving pedestrians. Junctions with 100 or more pedestrians per day

6 × *PACCPERYEAR*	Acc 83–88	$\hat{m}$ 83–88	Number of junctions
0.0	0	0.0	1
0.1	0	0.1	34
0.1	1	0.1	6
0.2	0	0.2	49
0.2	1	0.2	6
0.2	2	0.3	2
0.3	0	0.3	35
0.3	1	0.4	13
0.3	2	0.4	2
0.4	0	0.4	17
0.4	1	0.5	7
0.4	2	0.6	2
0.5	0	0.4	16
0.5	1	0.6	3
0.5	4	1.0	1
0.6	0	0.5	10
0.6	1	0.7	4
0.6	2	0.8	2
0.6	5	1.3	1
0.7	0	0.6	6
0.7	1	0.8	2
0.7	2	0.9	1
0.7	3	1.1	1
0.8	0	0.6	6
0.8	1	0.8	2
0.8	2	1.0	1
0.9	0	0.7	2
0.9	1	0.9	2
1.0	0	0.8	3
1.0	1	1.0	4
1.0	2	1.2	1
1.0	5	1.9	1
1.1	0	0.8	4

Table A1. (Continued)

6 × *PACCPERYEAR*	Acc 83–88	$\hat{m}$ 83–88	Number of junctions
1.1	1	1.1	1
1.2	0	0.9	1
1.2	1	1.1	1
1.2	2	1.4	4
1.3	0	0.9	2
1.3	1	1.2	1
1.3	2	1.5	1
1.3	3	1.8	2
1.3	4	2.1	1
1.4	0	1.0	1
1.4	1	1.3	1
1.4	2	1.6	1
1.4	3	1.9	1
1.4	4	2.2	1
1.5	0	1.0	1
1.5	1	1.3	1
1.6	0	1.1	1
1.7	0	1.1	2
1.7	1	1.5	2
1.8	1	1.5	1
2.0	2	2.0	2
2.5	0	1.4	1
2.7	5	3.7	1
2.8	0	1.5	1
3.1	5	4.0	1
3.3	2	2.7	1
4.3	2	3.0	1
4.6	4	4.3	1
4.6	5	4.8	1
7.3	13	11.2	1
Total 169.0	165	170.4	285

APPENDIX B

Table B1. Accidents involving cyclists. Junctions with 100 or more cyclists per day

6 × *CACCPERYEAR*	Acc 83–88	$\hat{m}$ 83–88	Number of junctions
0.1	0	0.1	1
0.1	2	0.1	1
0.2	0	0.2	5
0.2	1	0.2	2
0.3	0	0.3	18
0.3	1	0.3	4
0.4	0	0.4	14
0.4	1	0.4	6
0.4	2	0.5	1
0.4	3	0.5	1
0.5	0	0.5	27
0.5	1	0.5	7
0.5	2	0.6	4
0.6	0	0.6	15
0.6	1	0.6	10
0.6	2	0.7	3
0.6	3	0.7	2
0.7	0	0.7	6
0.7	1	0.7	5
0.7	2	0.8	1
0.7	4	0.9	1
0.8	0	0.7	7
0.8	1	0.8	7
0.8	2	0.9	3
0.9	0	0.8	16
0.9	1	0.9	7
0.9	2	1.0	4
0.9	3	1.1	1
1.0	0	0.9	12
1.0	1	1.0	6
1.0	2	1.1	4
1.0	3	1.2	2
1.0	4	1.3	1
1.1	0	1.0	7
1.1	1	1.1	13
1.1	2	1.2	2
1.2	0	1.1	9
1.2	1	1.2	7
1.2	2	1.3	2
1.2	3	1.4	4
1.3	0	1.2	1
1.3	1	1.3	6
1.3	2	1.4	5
1.3	3	1.5	3
1.4	0	1.2	5
1.4	1	1.4	2
1.4	2	1.5	3
1.4	3	1.6	4
1.4	4	1.7	2
1.4	5	1.8	1
1.5	0	1.3	6
1.5	1	1.4	4
1.5	2	1.6	2
1.5	3	1.7	1
1.6	1	1.5	1
1.6	2	1.7	4
1.6	3	1.8	2
1.7	1	1.6	2
1.7	2	1.7	3
1.7	3	1.9	1
1.7	4	2.0	3
1.8	0	1.5	5
1.8	1	1.7	2
1.8	2	1.8	1
1.8	3	2.0	1
1.8	4	2.1	1
1.9	0	1.6	1
1.9	1	1.8	2
1.9	3	2.1	1
2.0	0	1.7	3
2.0	1	1.8	1
2.0	2	2.0	3
2.0	4	2.3	1
2.0	5	2.5	1
2.1	1	1.9	1
2.1	2	2.1	1
2.1	4	2.4	1
2.1	5	2.6	1
2.2	0	1.8	1
2.2	1	2.0	2
2.2	2	2.2	1
2.2	3	2.3	1
2.2	4	2.5	1
2.3	1	2.1	2
2.3	2	2.2	1
2.3	4	2.6	1
2.4	2	2.3	1
2.4	3	2.5	1
2.4	5	2.9	1
2.5	1	2.2	1
2.5	3	2.6	1
2.6	2	2.5	1
2.6	5	3.1	1
2.7	2	2.6	3
2.8	2	2.6	2
2.8	3	2.8	2
2.9	0	2.2	1
2.9	1	2.5	1
3.0	5	3.5	1
3.0	7	3.9	1
3.1	2	2.8	2
3.2	4	3.4	1
3.4	8	4.6	1
3.5	9	4.9	1
3.7	0	2.7	1
4.0	6	4.6	1
4.1	3	3.8	1
4.4	3	4.0	1
4.4	4	4.3	1
4.5	4	4.3	1
5.0	1	3.7	1
5.9	4	5.2	1
Total 442.0	432	441.3	377