april 26 - 28, 1988
amsterdam
the netherlands

traffic safety theory & research methods

Session 5: Time dependent models

SWOV

This book was made possible through the support of 3M Europe

SESSION 5: TIME DEPENDENT MODELS

Summary of the paper presented by the additional speaker

Siem OPPE, Matthijs KOORNSTRA & Bob ROSZBACH, Institute for Road Safety Research SWOV, Leidschendam, The Netherlands
Macroscopic models for traffic and traffic safety

Full papers of other contributors

Angel MARIN & Pedro PABLOS, E.T.S. Ingenieros Aeronáuticos, Madrid, Spain
Flow optimization with traffic safety constraints

Frits BIJLEVELD, Siem OPPE & Frank POPPE, Institute for Road Safety Research SWOV, Leidschendam, The Netherlands
State space models in road safety research

J. WISMANS, P. DE COO & C. HUYSKENS TNO Road-Vehicles Research Institute, Delft, The Netherlands and T. HEIJER, Institute for Road Saety Research SWOV, Leidschendam, The Netherlands
Computer sinulation of crash dynamics

B.Ch. FÄRBER, B.A. FÄRBER & M. POPP, University of Tübingen, Federal Republic of Germany
Evaluation methods for traffic safety aspects of new technologies in vehicles

Helmut T. ZWAHLEN, Ohio University, Athens, Ohio, U.S.A.
Research methodology to assess the importance of peripheral visual detection at night

R.L. ERICKSON & H.L. WOLTMAN, 3M Traffic Control Materials Division, St. Paul, Minnesota, U.S.A.
Sign luminance as a methodology for matching driver needs, roadway variables and traffic signing materials

Marien G. BAKKER, Ministry of Transport and Public Works, The Hague, The Netherlands
Method floating car: A research method to study the speed-behaviour and routes of car drivers in residential areas

G. BLIKMAN, Delft University of Technology, The Netherlands
A new method for traffic safety research on driver distraction

Åke FORSSTRÖM, Chalmers University of Technology, Göteborg, Sweden
Simulation of casualties in person transport systems

Summary of the paper presented by the additional speaker

Siem OPPE, Matthijs KOORNSTRA & Bob ROSZBACH, Institute for Road Safety Research SWOV, Leidschendam, The Netherlands

Macroscopic models for traffic and traffic safety

MACROSCOPIC MODELS FOR TRAFFIC AND TRAFFIC SAFETY
These related approaches from SWOV

Siem Oppe, Matthijs Koornstra & Bob Roszbach
Institute for Road Safety Research SWOV, The Netherlands

FIRST APPROACH: MODELS FOR THE DEVELOPMENT OF TRAFFIC AND TRAFFIC SAFETY IN FOUR COUNTRIES (S. OPPE).

Introduction

Recently there has been an increased interest in the application of macroscopic models for the description of developments in traffic safety. At SWOV this new interest was initiated in the early eighties by the discussion on the causes of the sudden decrease in the numbers of fatal and injury accidents after 1974. Before that time these numbers had increased steadily over the years. A satisfactory explanation for this decrease could not be given.
Blokpoel [1982] presents data for the development of traffic volumes, accidents and accident rates in the Netherlands (see Figure 1a). Independently the same data was given by Appel [1982] for Germany (see Figure 1b).

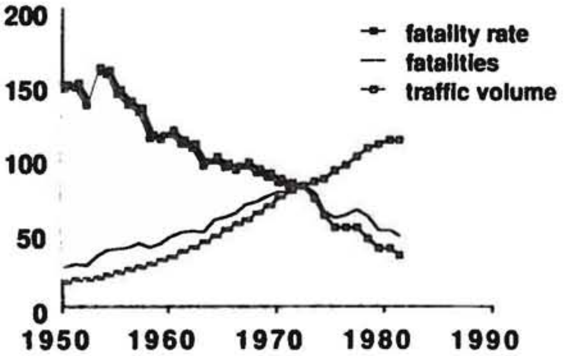


Fig. 1a. Traffic volume and traffic safety data for the Netherlands according to Blokpoel (1982).

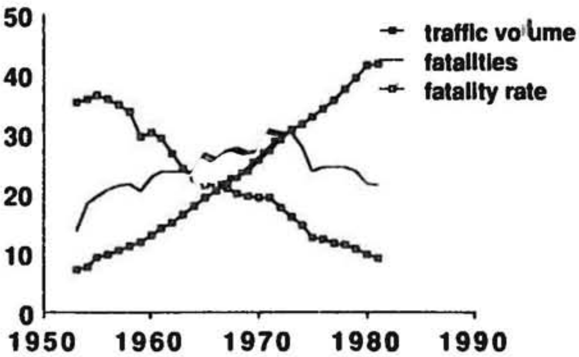


Fig. 1b. Traffic volume and traffic safety data for Germany according to Appel (1982).

Figures 1a and 1b support the assumption that the development of the accident numbers follows from the combination of two more basic processes, the development of the traffic volumes and of the accident rates. The first curve is monotonically increasing, the second monotonically decreasing, and the rise and fall of the accident curve is then supposed to result as the product of these two monotonic curves. The rise of the number of accidents up to 1974 is part of the same process as the fall after that year, and there is no specific explanation needed for the turning point in this curve. The combination of these basic curves may be used to predict developments in the number of accidents in the future. Several approaches start from one or both curves in order to describe or predict safety results. Oppe [1984] suggested two mathematical curves and estimated from this a total of 1080 fatal accidents in 1990 for the Netherlands.

This approach will be described and applied to the data of the Netherlands, the USA, West Germany and Great Britain. These data are collected from the various national sources. The US-data are from Accident facts 1974 and Traffic accident facts 1986. The data for West-Germany are from SBA Verkehrsunfälle 1986. The data for Great Britain are from Road Accidents G.B. 1985. The data for the Netherlands are from CBS, stat. verkeersongevallen op de O.W. (statistics of traffic accidents on public roads), and additional data from SWOV.

The model

The model is based on the above mentioned assumptions that:

- there is a monotonically increasing S-shaped saturation curve with regard to the development of the number of vehicle kilometers per year;
- there is a monotonically decreasing curve for the development of the fatality rates per year, to be called "the risk curve";
- as a consequence, the number of fatalities per year follows from these curves by multiplication of their respective values.

Two very simple mathematical functions turn out to fit the data rather well. A negative exponential according to model 1 is used for the fatality rates.

Model 1:

$$\log \left(\frac{f}{v}\right) = \alpha t + \beta \qquad (\alpha < 0) \qquad (1)$$

With f the total number of fatalities for a given year, v the total annual amount of vehicle kilometers, t the respective year and α and β the scale-parameters to fit.

This means that the decrease of the ratio between the number of accidents and the number of vehicle kilometers is proportional to time.
The decrease is supposed to be the combination of all efforts made to improve the traffic system, such as the improvement of the road system, vehicle design, crash measures, legislation, education and individual learning [SWOV, 1986]. The traffic density as such may also have had a direct effect on the decrease in the fatality rate.

For the description of the amount of traffic, a good fit was found from simple assumptions. First it was assumed that this development starts from zero and rises through time to a certain saturation level. A simple model of this kind is the S-shaped logistic curve. A generalization of the function for y-values between 0 and some arbitrary but positive value, instead of y-values between 0 and 1 results in

Model 2:

$$\log \left(\frac{v}{vmax - v} \right) = \alpha' t + \beta' \qquad (2)$$

The assumption is, that the rate between the traffic volume already realized at time t and the remaining traffic volume potential to be realized in the future increases proportionally to time. The value of vmax is not given in advance and will be chosen in such a way that the fit of model 2 is maximized.

Results

Both models fit the data rather well. As was already known before, the decrease in the log-rates for the fatalities per vehicle kilometer over the years, turns out to be fit indeed by a linear function for all four countries.
A maximum value for the annual amount of vehicle kilometers is found from the best fit of the linear function to the data according to model 2. Using this proportionality factor vmax, the fit for model 2 is, generally speaking, slightly better than the fit for model 1.
Furthermore an empirical relation has been found between the parameters of model 1 and 2, suggesting the combination of both models to.

Model 3:

$$f_t = c\sqrt{v_t.(vmax - v_t)} \qquad (3)$$

where f_t is the number of fatalities in year t, v_t the total amount of vehicle kilometers in that year and c is a given constant.

Koornstra (1988) noticed that this function is of a particular form. If we rewrite model 2 in its ordinary form as:

$$v_t = \frac{v_{max}}{1 + e^{\alpha t + \beta}} \qquad (4)$$

then it follows that the first derivative of this function with regard to t is:

$$v_t' = \frac{-\alpha}{v_{max}} v_t \ (vmax - v_t) \qquad (5)$$

(see also Mertens [1973])

This shows that the functional relationship between the number of fatalities and vehicle kilometers as suggested by the empirical data analysis can be written as $f_t = g(v_t') = c(v_t')^{\frac{1}{2}}$.

SECOND APPROACH: RISKREDUCTION AS A LEARNING PROCESS (M.J. Koornstra)

Minter (1987) interpretes the relation between traffic volume and the exponentional reduction of fatality-risk as a community learning process, based on the cumulative past experience with traffic.
He conjectured that this learning process is rather independent from actions of gouvernement, like legislative reforms.

Minter used Towill's model (Towill, 1973), in our notation written as.

$$\frac{V_n}{F_n} = d^\circ + e^{a^\circ - b^\circ \sum_{t=1}^{t=n} V_t} \qquad (1)$$

Rewritten as

$$\bar{P}_n = \frac{F_n}{cV_n} = \frac{1}{1 + e^{a' - b' \sum_{t=1}^{t=n} cV_t}} \qquad (2)$$

or as

$$\frac{1 - \bar{P}_n}{\bar{P}_n} = e^{a' - b' \sum_{t=1}^{t=n} cV_t} \qquad (3)$$

Where cV_t is interpretated as a measure for the number of learning events in time interval t and $\bar{P}_n$ as the mean probability of a fatal result of an event. Refering to Sternberg (1967) the postulated learning model is an agregrated Luce-Beta Model in which the learning probability P_{n+1} of the n+1 - event is related to P_n as follows:

$$\frac{1 - P_{n+1}}{P_{n+1}} = \frac{1}{a} \; \frac{1 - P_n}{P_n}$$

or

$$P_{n+1} = \frac{aP_n}{(1 - P_n) + aP_n} \qquad (4)$$

This is a one-parameter learning model characterized by
- path indepence of events
- commutativity of effects on events
- independence of irrelevent alternatives
  or arbitrariness of definition of classes of outcomes of events
- valid approximation by mean-values of parameters, assuming parameter distributions over individuals concentrated at its mean.

Although the learning-model interpretation of Minter is followed, we start with another assumption on the community learning process. We will assume a community learning process, that reduces the probability of negative outcomes of encounters by actions of institutions and that the amount of investments is such that a constant probability reduction for negative outcomes of encounters results. By this assumptions we obtain:

$$P_{t+1} = a P_t + (1 - a) \lambda \qquad (5)$$

Where a is the constant probability reduction factor and λ the level of the limit of the learning process.

$$P_t = \lambda \quad \text{for } t \dashrightarrow \infty \qquad (6)$$

Following Sternberg (1963) and using the explicit formula by repetitive application of (5) we obtain

$$P_t = a^{t-1} P_1 = a^t P_o \qquad (7)$$

and defining $P_o = e^{\beta}$ and $a = e^{\alpha}$ we obtain our first basic assumption:

<u>Assumption 1:</u> $$P_t - \lambda = e^{\alpha t + \beta} \qquad (8)$$

Formula (5) to (8) forms a generalisation of the so-called linear-operator learning model (see Sternberg) from Bush and Mosteller (1955) based on a society controlled learning process, where a is constant over time. As Sternberg noted linear-operator models and the Beta-models are hardly distinguishable on the basis of empirical data, since the difference is in the generally unobserved preliminary learning phase.

Let us now define a measure of encounters as the total exposure to traffic for a country in a year t as a function of traffic volume in year t

$$U_t = dV_t^{\,s} \qquad (9)$$

Let R_t be the total quantity of some well defined class of outcomes of encounters, then we identify P_t as

$$P_t = \frac{R_t}{U_t} \qquad (10)$$

Combining (10) and (8) we obtain

$$R_t = \lambda\, U_t + U_t\; e^{\alpha t+\beta} \qquad (11)$$

Identify R_t as F_t and assume that instutions are taking effective actions mainly on the basis of fatal accidents, then P_t for t --> ∞ approaches zero, than $\lambda = o$ for $R_t = F_t$, by which we obtain:

$$\frac{F_t}{U_t} = \frac{F_t}{dV_t^{\,s}} = e^{\alpha t\,+\,\beta} \qquad (12)$$

which for s = 1 gives a theoretical justification for the curve fitting of Oppe.

Taking (9) we generalize Oppe's logistic function for the growth of traffic volume as assumption 2:

<u>Assumption 2:</u> $$U_t = \frac{U_{max}}{1 + e^{\alpha' t\,+\,\beta'}} \text{, or } \frac{U_{max} - U_t}{U_t} = e^{\alpha' t\,+\,\beta'} \qquad (13)$$

which is only symetric for V_t if s=1, see Nelder (1961).
For 0<s<1 shows a curve for V_t that is initially growing fast and slowing down later, while the opposite is true for s>1.

It will be noted that the derivative of (13) with respect to t is.

$$U_t' = \frac{-\alpha'}{U_{max}} \cdot U_t \cdot (U_{max} - U_t) = w.U_t^2 . e^{\alpha' t+\beta'} \qquad (14)$$

$$\text{Where } w = \frac{-\alpha'}{U_{max}}$$

Oppe's results points to the finding that the number of fatalities is a simpel function of the derivative of U_t for s=1. We state this as our third basic assumption as:

<u>Assumption 3:</u> $F_t = p(U_t')^q$ (15)

From assumption 2 and 3 it follows that

$$F_t = p.w^q.e^{q\alpha' t+q\beta'} . U_t^{2q} \qquad (16)$$

Since p is a free parameter, we define $p=w^{-q}$ and combining (12) as the result of assumption 1 and (16) we obtain

$$e^{\alpha t+\beta} = e^{q\alpha' t + q\beta'} . U_t^{2q-1} \qquad (17)$$

This equivalence only holds if the last factor of (17) drops out by 2q-1=0. So we prove that if assumptions 1 to 3 are true, there only can exist a simpel relation between the logistic time parameters for U_t and exponential time parameter for F_t/U_t if $q = ½$, which for s = 1 must result in the finding of Oppe that

$$\alpha = ½\alpha' \text{ and } \beta = ½\beta' \qquad (18a)$$

and by which

$$F_t = c\,[V_t(V_{max} - V_t)]^{½} \qquad (18b)$$

From the functions of U_t and F_t it is easily proven that for s = 1

$$F_{max} = ½\, c\, V_{max}$$

so that $c = \frac{2\ F_{max}}{V_{max}}$ 18c)

It also follows that

$$F_t^2 = a'V_t - b'V_t^2 \tag{19}$$

from which we may estimate V_{max} and F_{max} as follows

$$F_{max} = \frac{a'}{b'} \tag{20a}$$

and

$$V_{max} = \frac{a'}{b'} \sqrt{b'} \tag{20b}$$

We take Oppe's results as an empirical justification for our assumptions.

Combining the results of (11) and assumption 3 as (15) we obtain a generalized assumption as

Generalized Assumption 1: $R_t = f\ (U_t) + g\ (U'_t)$ (21)

where by (9) and (15) for s=1 this reduces, for q = ½ as the solution for (17), to

$$R_t = \lambda.d.V_t + (1-\lambda).c.(V'_t)^{½} \tag{22}$$

Depending on λ, ranging from 1 to 0, we obtain R_t as the ordered quantities for exposure (λ=1), conflicts, less severe accidents, severe injuries and at last fatalities ($\lambda = 0$).

If we identify R_t as the quantity of severe injured people (S_t), we obtain from (22) for $o<\lambda<1$

$$S_t = \lambda.d.V_t + (1-\lambda).c.[V_t(V_{max}-V_t)]^{½} \tag{23}$$

or by (19) and (18)

$$S_t = \lambda.d.V_t + (1-\lambda).(a'V_t - b'V_t^2)^{½} \tag{24}$$

Fitting (19) and (23) directly on V_t, may give better fit for F_t and S_t, since departures of V_t from the logistic curve may explain variations in F_t and S_t.
Some results will be shown.

Moreover, transforming (23) as

$$\frac{S_t}{dV_t} = \lambda + b^\circ \frac{F_t}{V_t}$$

we obtain also as another example of assumption 1 for $1 > \lambda > 0$

$$\frac{S_t}{V_t} - a^\circ = e^{\frac{1}{2}\alpha' t + \phi} \qquad (25)$$

Where $\phi = \frac{1}{2}\beta' + \ln(b^\circ.d)$ and $a^\circ = \lambda.d$ while α' is the same parameter as in (13) and (12) where $\alpha = \frac{1}{2}\alpha'$

Comparable results to the curve of Oppe for F_t are shown for S_t and show a rather good fit as well.
It will be noted that curve fitting for V_t, F_t and S_t is extremely parsimonious by the assumptions of our theory.
Only 6 parameters are used to fit 3 observational independent times series by theoretical deduced shapes of curves.
The fit for the predicted shape of the curves and the empirical close identity of the α-parameter are taken as evidence for the justification of this mathematical theory.

We disagree with Minter's (1987) interpretation of the learning process as community learning based on cumulative experience, without effects of institutional actions, for two reasons:
Firstly: the difference in learning curves for F_t and S_t supposes discrimination between situations with fatal accidents and with less severe accidents, resulting in better learning for avoidance of fatal accidents than of less severe accidents. This cannot be explaind by individual cumulative experience.
Secondly: transforming Minter's Beta-model for learning to an equivalent linear-operator model would result in the expression as

$$\frac{F_n}{V_n} = e^{\bar{a} - \bar{b} \sum_{t=1}^{t=n} V_t}$$

which only reduces to the well fitted curves of Oppe if V_t is constant over time. This is evidently not true.

The observation that the curves do fit rather well for s=1 is somewhat puzzling, since this relates exposure to traffic volume by a ratio-scale factor only.

Taking encounters between two classes (say 1 as mopeds and 2 as passenger cars) and defining exposure as

$$U_{t,1,2} = d\,(V_{t,1})^{\frac{1}{2}} \cdot (V_{t,2})^{\frac{1}{2}} \qquad (26)$$

we would obtain for 1 = 2 according to s=1.

$$U_t = dV_t$$

Using (26) in the framework of our theory by substituting (26) into (11) we expect some well fitted results.

This leads to the conjecture that growth of traffic volume itselve reduces the number of independent encounters by a square root transformation of the relevant volumes. Roszbach (1988) points to this third approach of explaining for risk reduction by higher density of traffic.

THIRD APPROACH: DENSITY, EXPOSURE AND ACCIDENT RISK (R. Roszbach)

Processes proceed in time. In time related models, as presented in the preceding parts of this paper, basically some regularity in time is assumed with respect to some process. For explanatory or predictive purposes one would like to know more about these unknowns or, rather, eliminate time from the time dependent model and replace it by measures which are more directly related to the processes at hand.
At the same time, one should try to move in depth in the sense that, on the basis of the processes assumed, alternative predictions for subsets of the material are attempted. If not, one arrives at general formulae such as Smeed's, which pose comparable problems of interpretation. That such formulae are hard to interpret has been aptly demonstrated in last years issues of Traffic Engineering and Control (Adams, 1987; Minter, 1987; Andreassen, 1987 a,b), some forty (!) years after the birth of said formula.
(It is interesting to note, however, that some countries - among which the Netherlands - have already reached the safety level of about 3 fatalities per 10.000 motorvehicles per year - predicted by Smeed's formula at a saturation level of 1 vehicle per person - at much lower levels of vehicle penetration).
At the risk of incurring people's wrath (on stretching intended meanings or inducing unintended generalizations), I would hold that basic to such formulae are:

- a monotonic decrease in accident risks
- for conditions of growing motorization

With respect to the first part this leaves questions as to how these risks are defined. With respect to the second part we may wonder whether this is a fundamental constraint or an empirical (in the sense that we have as yet no data on conditions of non-growing motorization, with maybe the exception of the depression and wartime period in the USA. It would be hard, however, to generalize from such specific conditions).

Relating amount of travel to accidents may be relevant in cost-benefit considerations or in theories relating accidents to societies safety tolerances. It is not necessarily the best way, however, to define the set of potentially hazardous events from which accidents may or may not result (exposure). Also, we have to be very careful as to how we aggregate. I will give one example:

If, for the Dutch situation, we divide yearly fatalities by total amount of travel we get a neat exponential function, as Oppe shows. The decrease in fatality rate is about 50% per 10 years. If, however, we divide into vehicle categories and do the same per category (bicycle, moped, motorcycle, car, goods vehicle) we get fatality rates that are essentially constant for the period 1950-1970 and then rise and fall for the various categories in no easily interpretable manner. (A lovely result for those who cherish constant risk ideas). This is, of course, caused by the fact that car-kilometers predominate in total amount of travel and consequently e.g. bicycle victims are then divided by strongly increasing car-kilometers.
Constant risks for bicyclists over a period in which a tenfold increase in number of automobiles took place may of course be interpreted as a significant safety accomplishment, if one holds an exposure model which is multiplicative in nature or some other function of the various combined categories of vehicle involved. The exposure model Oppe uses, therefore, is more of a multiplicative model than it looks like, as a result of properties of the distribution of accidents and the distribution of growth in numbers over vehicle categories. (Taking the above results into account a variation on risk compensation theory may be offered in terms of compensation for increased mobility. We may be ordering our increasing traffic flows in such a way that we effectively control for any multiplicative exposure effects).

A basic exposure model defines single vehicle accidents as straightforwardly related to the amount of travel of that vehicle category and multiple vehicle accidents as related to the product of the amount of travel of the categories involved (Smeed, 1974). The second part of this model is not unlike a comparison of moving vehicles with randomly moving particles.
Vehicles, however, move in a network. If the number of vehicles increases and the network does not, vehicles tend to queue up. If we - loosely - introduce the term encounter for potentially hazardous events, it may be that encounters between single vehicles are replaced by encounters between such queue's or between single elements and such queue's. If such encounters are seen as units of exposure, it then follows that exposure is not related in any simple manner to the amount of traffic in the network at any one time. (A basic assumption in this model is that the continuous interactions within queue's are essentially without risk).

There does seem to be some plausibility in the idea that for e.g. a pedestrian who wants to cross a street it does not really matter - in terms of exposure to risk - whether he meets with one motorvehicle or with 5 or 10 vehicles in queue. If he acts on the first one, he is not likely to fail to act on the others. It does matter, however, if these vehicles are sufficiently spaced so as to lead to distinct encounters.
Conceived in this manner exposure would be relative to quite specific properties of the distribution of traffic over the road network and in time. An attractive side to such a conception is that, although derived from general considerations on the development of traffic safety, it is testable on the specific level of limited sets of locations.

Following this line of reasoning, two propositions can be made in relation to the development of traffic safety:

- descriptions of such developments may - from a process-oriented point of view - overestimate exposure and thereby overestimate risk reductions
- part of what may be conceived as risk reduction is inherent to the condition of growing traffic densities in the road network.

LITERATURE

- Adams, J.G.U., Smeed's law: Some further thoughts, Traffic Engineering and Control. pag. 70-, februari, 1987.

- Andreassen, D.C., Learning Curves of Casualties. Traffic Engineering and Control. pag. 411-, februari, 1987.

- Appel, H. Strategische Aspekten zur Erhöhung der Sicherheit im Straßenverkehr. Automobil-Industrie, 3, 1982.

- Blokpoel, A. Internal SWOV-note on the development of fatal accidents. SWOV, 1982.

- Bush, R.R. & F. Mosteller, Stochastic models for learning. New York, J. Wiley and Son Inc, 1955.

- Mertens, P. (1973). Mittel-und langfristige Absatz-prognose auf de Basis von Sättigungsmodellen. In: Prognoserechnung, Physica-Verlag, Würzburg, pag. 193, 1973.

- Minter, A.L. Road casualties - Improvement by learning processes. Traffic Engineering and Control, februari, 1987.

- Nelder, J.A. The fitting of a generalisation of the logistic curve. Biometrics, V17, pag. 289, 1961.

- Oppe, S. Internal note for the MacKinsey and Company projectgroup on traffic safety in the Netherlands, SWOV, 1984.

- Smeed, R.J. The frequency of road accidents. Zeitschrift für Verkehrssicherheit, 1974.

- Sternberg, S. (1967). Stochastic Learning Theory. In: Handbook of Mathematical Psychology Vol.II, Ch. 9, pag. 1 - 120, 1967, New York, J. Wiley and Son Inc.

- SWOV. SWOV voorspelt ca. 1200 verkeersdoden in 1990. (SWOV predicts approx. 1200 road traffic deaths in 1990). SWOV-schrift, 1986.

- Towill, D.R. A direct method for the determination of learning curve parameters from historical data. Int. J. Prod. Res. 11, pag. 97, 1973.

Full papers of other contributors

Angel MARIN & Pedro PABLOS, E.T.S. Ingenieros Aeronáuticos, Madrid, Spain
Flow optimization with traffic safety constraints

Frits BIJLEVELD, Siem OPPE & Frank POPPE, Institute for Road Safety Research SWOV, Leidschendam, The Netherlands
State space models in road safety research

J. WISMANS, P. DE COO & C. HUYSKENS TNO Road-Vehicles Research Institute, Delft, The Netherlands and T. HEIJER, Institute for Road Saety Research SWOV, Leidschendam, The Netherlands
Computer sinulation of crash dynamics

B.Ch. FÄRBER, B.A. FÄRBER & M. POPP, University of Tübingen, Federal Republic of Germany
Evaluation methods for traffic safety aspects of new technologies in vehicles

Helmut T. ZWAHLEN, Ohio University, Athens, Ohio, U.S.A.
Research methodology to assess the importance of peripheral visual detection at night

R.L. ERICKSON & H.L. WOLTMAN, 3M Traffic Control Materials Division, St. Paul, Minnesota, U.S.A.
Sign luminance as a methodology for matching driver needs, roadway variables and traffic signing materials

Marien G. BAKKER, Ministry of Transport and Public Works, The Hague, The Netherlands
Method floating car: A research method to study the speed-behaviour and routes of car drivers in residential areas

G. BLIKMAN, Delft University of Technology, The Netherlands
A new method for traffic safety research on driver distraction

Åke FORSSTRÖM, Chalmers University of Technology, Göteborg, Sweden
Simulation of casualties in person transport systems

FLOW OPTIMIZATION WITH TRAFFIC SAFETY CONSTRAINTS

by *Angel Marin and Pedro Pablos*
Departamento de Matemática Aplicada y Estadística
E. T. S. Ingenieros Aeronáuticos. MADRID.

ABSTRACT

The authors consider the introduction of safety criteria (depending on management) on the equilibrium traffic models (depending on users).

Two typical traffic situations have been considered where it is necessary to take into acount the safety. One consideration is safety distance between vehicles and the other is the acceleration limit used as vehicle enter the freeway.

To treat the non-linear bounds in the flow resulting in the consideration of safety situations, we adapt an ad hoc procedure of decomposition, where the management considerations relative to safety inform the flow bounds of the user's equilibrium model of the next iteration.

The use of the equilibrium models with upper and lower flow bounds implies an adaptation of the *Frank-Wolfe method* to include them.

INTRODUCTION

Traffic equilibrium models have ignored sistematically safety cosiderations, partly because they have not been valued enough and partly due to aditional mathematical complications. These are due to considerations of management decisions of responsible authorities besides user's characterizations.

On the other hand, whenever a study has been done, it has been restricted to a few links related to a crossroads. This work tries to extend techniques used in traffic equilibrium models of large dimension, including safety considerations.

If we want to include safety considerations it is necessary to separate on the one hand , classic equilibrium models and on the other hand, safety considerations. In spite of the reduced experience in this field and limited reach of the study, the results obtained and enormous possibilities that are beginning to be seen, allow us to say that these methods may be of great use in the considerations of reliable safety criteria for traffic systems.

USED TRAFFIC MODEL

This model tries to characterize on the one hand the user's route choice process and on the other hand, the authorities decision process, which controls the transport system.

For user's characterization the relations of the model of non-linear optimitation for convex and separable networks will be considered. This model results in stating the relations with a given fixed demand, route choice according with *Wardrop's first principle*; non-linear, separable, non decreasing monotone congestion functions and linear equilibrium relations in network nodes.

With all these hypothesis the following optimization model is obtained (*Florian, 1984*)

$$\left.\begin{aligned} &\text{Min} \quad \mathbb{F}(f_a) = \sum_a \int_0^{f_a} \hat{C}_a(x)\,dx \\ &f_a = \sum_k h_k\,\delta_{ka}\,, \quad \forall\, a \\ &d_\omega = \sum_{k\,\in\,K_\omega} h_k\,\delta_{k\omega}\,, \quad \forall\, \omega \\ &f_a \in [\,l_a\,,\,u_a\,]\,, \quad \forall\, a \end{aligned}\right\} \quad (1)$$

, where f_a is the flow along link a, $\hat{C}_a$ is the congestion function; h_k is the flow in path k; d_ω is the given demand for an origin-destination pair ω; δ_{ka} y δ_{kw} are indicators of wether the path k uses link a or belongs to origin-destination pair ω; l_a y u_a are fixed lower and upper bounds for link a.

Management models will consider a congestion function $\hat{C}_a$ that will depend on control variables t_a together with link flows f_a. They will determine the cost in the link a (c_a) and the capacity (q_a).

$$\left\{ c_a , q_a \right\} = \hat{C}_a (t_a , f_a) \quad , \forall a \qquad (2)$$

The management relation $\hat{m}$ represents decision making by traffic systems controllers, characterizes the control variable t depending on system's state $\{ f, c, q \}$ and management recourse parameters ($\bar{r}$).

$$t = \hat{m} (\bar{r}, f, c, q) \qquad (3)$$

The management model is complemented with feasability conditions from (1), where bounds are not fixed now:

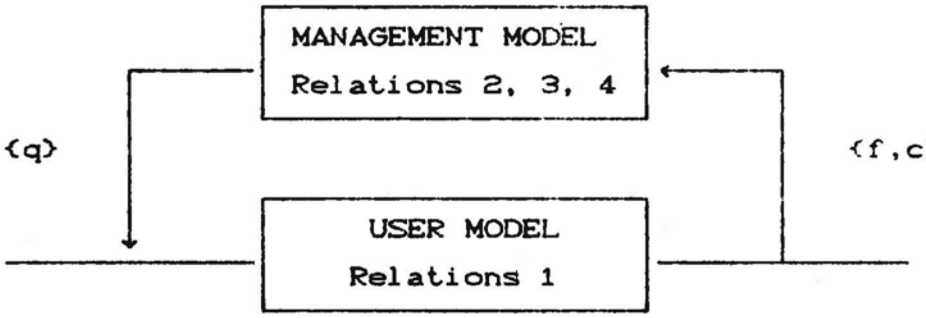
$$f_a \in [l_a , u_a] = q_a \qquad \forall a \qquad (4)$$

The resulting management model is so complicated that it requires the use of proper decomposition techniques which are fit for the kind of relations used in the model.

A simple and intuitive procedure of making decomposition is used by *Gartner, Gershwin and Little (1980)*, who consider a iterative scheme in which the traffic assignment is looped with a signal optimization program and with a mode split in relation with fuel consumption and reduced air polution.

We will consider a user's optimization cycle in which control variables and link capacities will be considered as parameters. The management cycle states new flow bounds { q } according to safety criteria and consideres as parameters the variables {f,c} related with user's equilibrium.

The global interaction procedure may be represented by the following diagram:



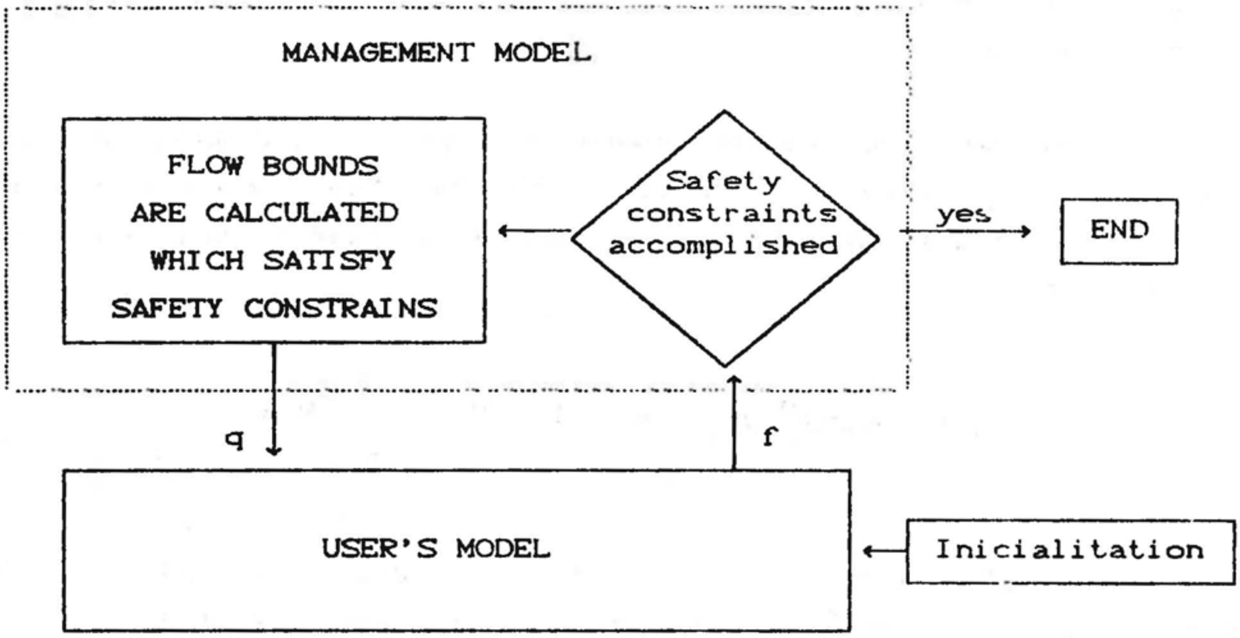
TRAFFIC MODEL SPECIFICATIONS

As congestion functions $\hat{C}_a$, the well-known BPR volume-delay curve has been used:

$$c_a = \hat{C}_a (f_a) = t_{oa} \left[1 + 0.15 \left(\frac{f_a}{k_a} \right)^4 \right], \ \forall a \qquad (5)$$

,where c_a is the average travel time in link a, t_{oa} is free flow average travel time (travel time without any congestion effect), f_a is the link flow and k_a is the practical capacity of the link a.

The management model can be reduced to the utilization of control variables (just as flow regulations, signal control in the influence area, etc) which improve flow bounds that guarantee the accomplishment of several safety constraints. In this way the above diagram may be represented as follows:



Now, let's see how safety constrains are obtained and bounds are assigned in the management model.

SAFETY CRITERIA

Minimum distance between vehicles

As a result of a flow along a link, an average distance between vehicles is obtained.

The average travel time c_a depens on the flow f_a (congestion function (5)). The average speed v_a can be expressed then as follows

$$v_a = \frac{\ell_a}{\hat{C}_a(f_a)} \qquad (6)$$

,where ℓ_a is the lenght of the link a. The average distance between vehicles will be

$$d_a = \frac{v_a}{f_a} = \frac{\ell_a}{f_a \, \hat{C}_a(f_a)} \qquad (7)$$

This is a decreasing function. This means that as the flow increases, distance becomes shorter. Therefore, an upper flow bound has to be imposed in order to satisfy a minimum distance between vehicles.

The safest distance between vehicles is that which allows a vehicle to stop without colliding into the leading vehicle when it suddenly stops. According to Papacostas (1987) the breaking distance bd_a is

$$bd_a = \frac{V_a^2}{2 \cdot g \cdot (\mu_a \pm tg\ \theta_a)} + 1.5 \cdot V_a \ , \qquad \forall\ a \qquad (8)$$

, where V_a is the speed in the moment the motorist notices he has to stop; g is 9.8 m/s^2; μ_a is the friction coefficient; $tg\ \theta_a$ is percent grade divided by 100 and 1.5 is the average motorist perception-reaction delay (in seconds).

We can establish a flow bound by imposing that the distance between vehicles must be greater than or equal to breaking distance

$$d_a \geq bd_a \ , \qquad \forall\ a \qquad (9)$$

Taking into account (6), (7), (8) and (9) we obtain the following inequality

$$\frac{1}{f_a\ \hat{C}_a(f_a)} \geq \frac{KT_a}{\hat{C}_a^2(f_a)} + \frac{1.5}{\hat{C}_a(f_a)} \ , \qquad \forall\ a \qquad (10)$$

, where

$$KT_a = \frac{\ell_a}{2g(\mu_a \pm tg\ \theta_a)} \ , \qquad \forall\ a \qquad (11)$$

We can obtain from (10) the following variable flow bound

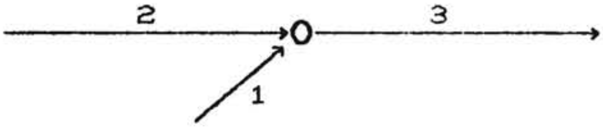
$$f_a \le \frac{2\ \hat{C}_a(f_a)}{2\ KT_a + 3\ \hat{C}_a(f_a)} , \qquad \forall\ a \qquad (12)$$

As a result of this upper flow bound we obtain flows so small (about 1000 veh/hr) that it is impossible to acomplish safety distance and satisfy ordinary demands in the networks studied. As d_a (7) is a decreasing function of flow, it will approach bd_a if we impose an upper bound to the flow as low as possible.

Nevertheless, another weaker minimum distance safety criteria should be developped.

Acceleration bounding

The entrance ramp to a freeway will be considered in order to bound the average acceleration that a vehicle needs to join the main stream.



If link 1 (see graph) represents the entrance ramp and link 2 and 3 are the freeway, the needed acceleration a_N is

$$a_N = \frac{\Delta v}{\Delta t} \qquad (13)$$

$$\Delta t = \frac{d_2}{av} = \frac{2\ d_2}{v_1 + v_3} \qquad (14)$$

$$\Delta v = v_3 - v_1 \qquad (15)$$

, where d_2 is vehicle spacing in link 2 ; v_1 and v_3 are the speeds in links 1 and 3 (initial and final speeds of acceleration maneuver); d_2 is the distance between vehicles in link 2, that is the maximum space in which the maneuver has to be finished; and av is the average velocity in the maneuver if acceleration is constant. As

$$v_1 = \frac{\ell_1}{\hat{C}_1(f_1)} \qquad \text{and} \qquad v_3 = \frac{\ell_3}{\hat{C}_3(f_3)} \qquad (16)$$

we use (7),(14),(15) and (16) in order to develop (13) as follows

$$a_N = \frac{1}{2\ell_2} \; f_2 \cdot \hat{C}_2(f_2) \left[\left(\frac{\ell_3}{\hat{C}_3(f_3)} \right)^2 - \left(\frac{\ell_1}{\hat{C}_1(f_1)} \right)^2 \right] \qquad (17)$$

If a_D is the maximum desired acceleration, it has to be greater or equal than a_N :

$$a_N \le a_D \qquad (18)$$

,where a_N is the function of f_1 (17). As $f_3 = f_1 + f_2$, if we take f_2 as a parameter, a_N depends only on f_1 so we can obtain a flow bound for this link.

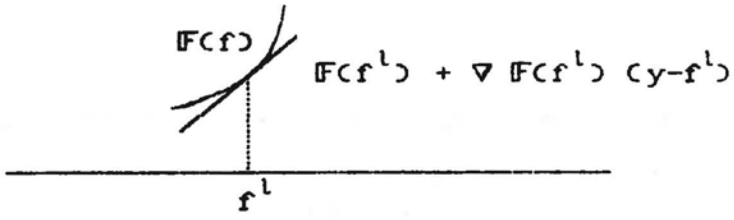
THE FRANK-WOLFE METHOD

The user's model (1) with bounds like

$$f_a \ge 0 \; , \quad \forall \; a \qquad (19)$$

will be refered to from now on as user's model (1-19). The Frank-Wolfe decomposition method is used to solve it. This method has been choosen because it is very efficient to resolve non-linear networks.

At first we look for a feasible descent direction, which is obtained by making the objective function linear and keeping the node equilibrium conditons. After we look for the optimum step along the above mentioned direction, taking into account the non linearity of the function.



a) Initial step:
Let f^l be a feasible solution of (1-19)

b) Descent direction search:

$$\left.\begin{array}{l} \text{Min} \quad \nabla \mathbb{F}(f^l)\, y \qquad\qquad , \text{ where } y \text{ is a feasible} \\ \quad \text{solution of } (1\text{-}19) \end{array}\right\} \quad (20)$$

Let y^l be the optimal solution of (20) in the step l.
If $\nabla \mathbb{F}(f^l)(y^l-f^l) < \varepsilon$, stop: y^l is an ε-optimal solution of (1-19)

c) Line search:

$$\text{Min} \quad \mathbb{F}[f^l + \alpha(y^l-f^l)] \ , \text{ with } \alpha \in [0,1] \qquad (21)$$

Let α^l be the optimal solution of (21) in the step l.
Let $f^{l+1} \longrightarrow f^l + \alpha^l(y^l-f^l)$ and come back to b)

The more important performance of the method is that it can solve the linear model (20) decomposing it by origin-destinations pairs. Then it is possible to solve it by using a shortest path method which provides the path to which we assign the known demand. (*Florian, 1984*).

The resulting method is a very efficient one that uses little memory. Therefore it is possible to use a Personal Computer with only 640k of internal memory, to solve network of 500 links in only a few minutes. For more information in relation to the use of this method to solve large networks the article of *A.Marin (1987)* may be used.

The introduction of a double flow bound as indicated in (1) is fundamental to solve the user's model taking into account the successive bounds associated to safety criteria, but then it is neccesary to include modifications in the indicated algorithm (20-21).

It is neccesary to maintain the shortest path algorithm with the typical bound (19) because it is very efficient in the resolution of the model (20), although it can not be adapted to treatment of double bounds. Then it is necessary to put between (20) and (21) a procedure to get the maximum step along the descent direction compatible with the mentioned double bound on the one hand, and a procedure to get a initial feasible solution with the double bounds on the other.

To get the maximum step (α_{max}) compatible with (1):

$$f_a^l + \alpha (y_a^l - f_a^l) \in [l_a , u_a] , \quad \forall a \qquad (22)$$

is enough to define it how:

$$\alpha_{max} = \min_a \left\{ \frac{f_a^l - l_a}{y_a^l - f_a^l} , \frac{u_a - f_a^l}{y_a^l - f_a^l} \right\} \qquad (23)$$

To get the initial feasible solution we begin in the lower bound (l) and from over there determine an extreme point (y^1), solution of the lineal model (20), from this point is localized a second extreme point (y^2), and then the segment from y^1 to y^2 is particioned looking for the initial solution of the model (1). If the search don't find the feasible point it is possible to try with other pair of extreme points.

With these modifications the Frank-Wolfe algorithm may be :

a) Initial step:

→ with the above method determine a feasible solution to model (1).

b) Descent direction search:

- → To determine the shortest possible path in order to join each origin with its correspoding destination.
- → Load the demand in the paths.
- → Verify the termination criteria.

c) Line search:

- → Determine the maximum step compatible with the double bounds as has been indicated.
- → Determine the optimum step, optimal solution of the unidimensional model (21).
- → Obtain the new flow and get back to solve the step b)

NETWORK DESCRIPTION AND RESULTS

A freeway corridor section used by *Gartner et al. (1980)* has been studied. A graphic representation of this network can be seen in the appendix.

The urban freeway and adjacent links of arterials is a specially interesting scenario for safety considerations. Shoulder-lanes have been modelled separately in order to better characterize entrance maneuver.

Freeway links have been characterized with a free flow speed of 120 km/hr. Arterials have been characterized with a free flow speed of 60 km/hr. The free flow speed together with the link length (of about 300-900m) allow us to obtain the free flow average travel time of the congestion function (5).

The practical capacity c_a that appears in (5) is about 2000 veh/hr. For an estimate of this capacity see Valdes (1978) and Highway Capacity Manual (1988).

The used vehicle demand is the same used by Gartner and others:

To	2	4	5	6	7	9	12	14	15
From 1					225	600			
3						560	375		
4						375	225		
5									225
6							300		
7	375								
8	330		150						225
10	3810			225				225	
11	375	187							
12				150					
13			225						
14					150	375			
15				225					
16						4510			

Not taking into account safety constraints, the needed acceleration a_N obtained by user's model (17) was 1.5 m/s^2 in the more critical entrance ramp: node 26 . The user's model (1) with safety consideration (18) has been solved for the mentioned method with smaller values of a_D at each time. In this way, the value of a_N was reduced to 0.7 m/s^2 in the mentioned node. For smaller values of a_D the lower bound in the ramp link is so large that it is impossible to find any feasible solution of (1).

However, when flow bounds are reduced in order to increase vehicle spacing, the upper flow in the shoulder-lanes (links 49 and 50) has proved specially critical. An upper flow bound is imposed in freeway links (as 49 and 50) in order to reduce vehicle spacing. Lower bounds are imposed in ramp links (as 31 and 39) according with (18) in order to minimize ramps acceleration.

We are trying to increase freeway entrance flow and reduce upper bound at the same time, so the conflict appears. If vehicle spacing in freeway links is trying to be increased, then ramps acceleration increases dangerously (until values of 2.8 m/s^2). Flow bounds in shoulder-lanes can be reduced to 3300 veh/hr, and 1700 veh/hr in the other links. However, these bounds are much greater than the safest one that we can obtain from (12), that is about 1000 veh/hr.

FURTHER RESEARCH DEVELOPMENT

It is necessary to acumulate more experience with real problems, and with large networks (1000 links or more).

It is important to work with other congestion functions that consider the signal control, not only because of the necessity to include it in an urban context, but their possibility to be used as a control element by the management in coordination with speed indicators (Smith, Van Vuren, Heydecker and Van Vliet, 1987).

We are working also in the characterization of other safety situations, for example, to study the lane-changing maneuver in the freeway. Other situations have to be also included.

Another possibility is the inclusion of an elastic demand in the user's equilibrium models, or using other less heuristic decomposition methods. The possibility of introducing stocasthic variables will be also consider.

REFERENCES

FLORIAN, M."An introduction to Network Models Used in Transportation Planning". M.Florian (editor), Elsevier Science Publishers, 1984.

GARTNER, N., S.B.GERSHWIN and J.D.LITTLE."Pilot Study of Computer-Based Urban Traffic Management". Transportation Research, V14B, pp 203-217, 1980

MARIN, A.G."Optimización no lineal de flujos en redes. Paquete de programas basado en el método de Frank-Wolfe". Informe interno del Departamento de Matemática Aplicada y Estadística. E.T.S.Ingenieros Aeronáuticos. Universidad Politécnica de Madrid, 1987.

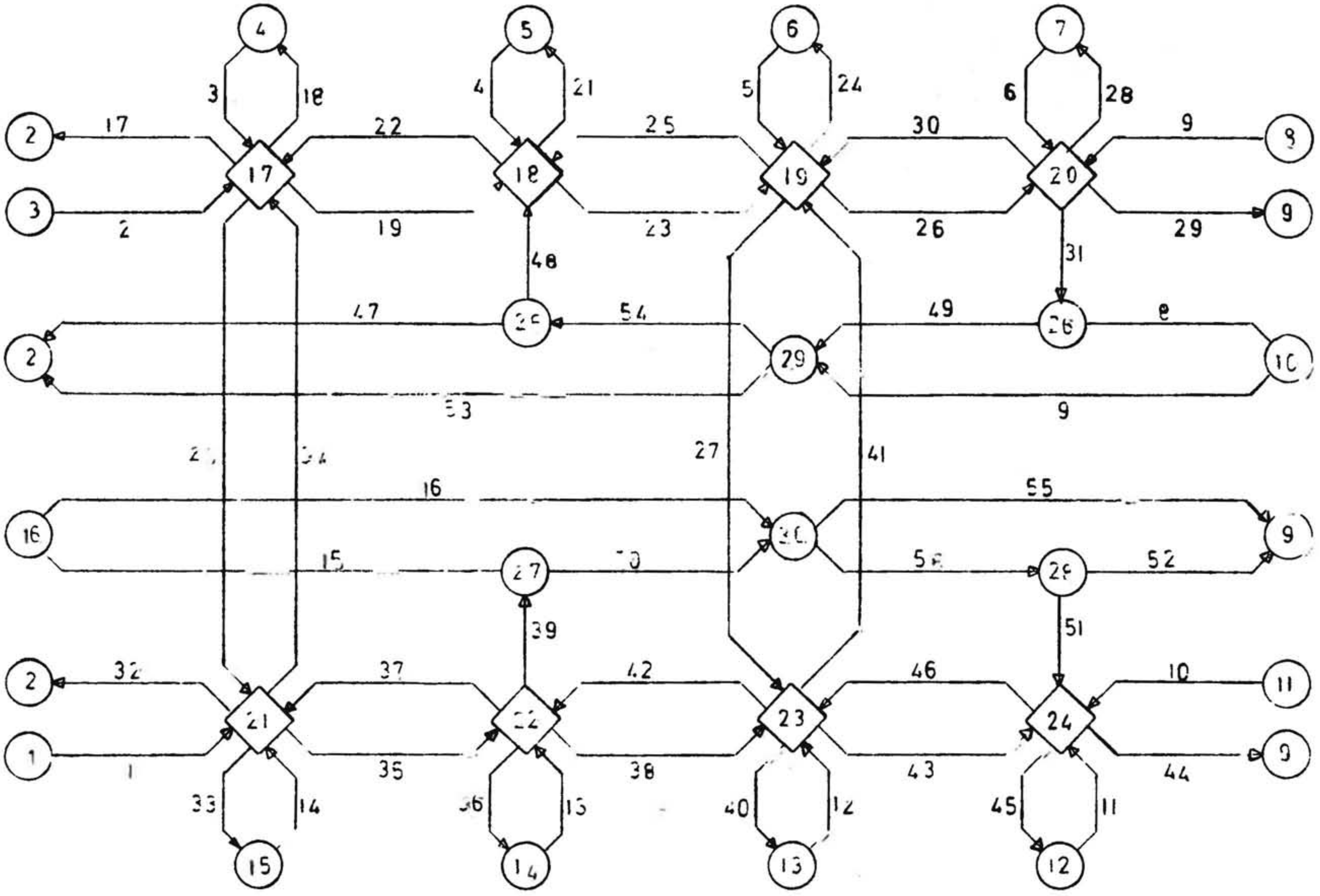
PAPACOSTAS, C.S. " Fundamentals of transportation engineering". Prentice-Hall, INC, Englewood Cliffs, 1987.

SMITH, M.J.,T.VAN VUREN, B.J. HEYDECKER and D.VAN VLIET. "The Interactions Between Signal Control Policies and Route Choice". Transport and Traffic Theory. Gartner & Wilson Editors. Elsevier Sciencie Publishers, 1987.

VALDES, A. "Ingenieria de tráfico". Editorial DOSSAT, S.A.,Madrid. 1978.

TRANSPORTATION RESEARCH BOARD. "Highway Capacity Manual". National Research Council, Washington D.C.

APPENDIX


4
5
6
7
3
18
4
21
5
24
6
28
2
17
22
25
30
9
8
17
18
19
20
3
2
19
23
26
29
9
48
31
47
54
49
2
29
9
10
27
41
16
16
55
9
15
27
52
39
51
2
32
37
42
46
10
11
21
22
23
24
1
1
35
38
43
44
9
33
14
36
13
40
12
45
11
15
14
13
12

STATE SPACE MODELS IN ROAD SAFETY RESEARCH

Frits Bijleveld, Siem Oppe & Frank Poppe
Institute for Road Safety Research SWOV, The Netherlands

INTRODUCTION

The OECD-S1-group developed in its report on Integrated Road Safety Programmes a conceptual framework to analyze the accident process (OECD, 1984). The three following elements characterize this framework.

1. The road user is the elementary unit of the system and must be viewed in his interaction with the surrounding system(s). The system is a dynamic one.
2. The processes are viewed as separated steps in succeeding order according to a phase model.
3. All (sub-)systems are governed by various levels of control. The control can be individual or collective, internal or external.

In this paper dynamic aspects of the model with regard to internal control are explored. In the S1-report this was seen as a risk control process. We will not bother with a predefined model of a control-mechanism. Instead we will show how the input and output of a system can be analyzed in such a way that the behaviour of the system under certain inputs can be defined and possibly interpreted. At SWOV the theory of state space models is used in order to investigate the different aspects of the dynamic (sub) systems in relation to traffic (un)safety.

THE DYNAMIC SYSTEMS IN TRAFFIC AND THE ROAD SAFETY ASPECT

The unsafety of road-traffic becomes clear from its end products. These final products can be material damage to cars, or physical or mental damage to people. The end products are the consequence of a series of events. This series of events (a process) can be divided into different phases, each phase having its own starting conditions, and its own context. The context defines the set of options that can be chosen

during each phase. The phases range from the conditions and events that lead to the planning and undertaking of a journey, through the different choices that have to be made during that journey, up to the events that lead to an accident, to damage and the consequences thereof. This way one views the system along a chronological line, which is at the same time the causal chain. The word "causal" does not mean that somewhere along this line the one-and-only "cause" of the accident can be found, but is used to reflect that the end of each phase determines the options in the next phase.
At the same time the well-known elements of the system must be considered, being a second viewpoint: the driver, the vehicle, the road and the surroundings (surroundings can be taken in a broad sense, including social surroundings).

The third way to describe the system is by means of the different control mechanisms. In this paper we want to explore the possibilities for such an analysis. If an analysis would prove to be possible it would have to take into account the following.
The system is governed by control mechanisms on different levels. The levels are hierarchical, going from the control model describing the individual driver, to different collective levels. E.g., an analysis of accidents of a particular type at a particular (type of) intersection can take into account the behaviour of the driver, but can be extended to the decisions and motivations of those who built the intersection or designed it, or the authorities that imposed regulations. The same goes for the car.

An analysis of the way a system is controlled by the operator can start from different points. In the OECD-report a general description is given of an operator who makes his decisions using an assessment of the risks related to the options he (thinks he) has. Research would then go on trying to build hypothetical models describing the perception of risk and the strategies handling it. A lot of work and a lot of discussion has been done in this direction.

An alternative would be to analyze the system itself, trying to find trends or cycles, relating known inputs to known outputs. Of course, the input one selects for analysis reflects some concept of what is relevant and what isn't. The information on how the system "behaves" and on the internal control necessary for such behaviour, could then be used to exercise an external control as efficiently as possible. The result of this kind of research (if any results are reached) would rather be a predictive model instead of an explanatory one.

A SYSTEM AND ITS STATE

Fundamental in the system approach is the idea that the relation between input and output of the system is not fixed. E.g. it is a well known fact from biological research that an organism adapts to a certain stimulus. A reaction to that stimulus therefore changes over time. The same effect is to be expected from the behaviour of road users. A road user having passed three intersections where he had right of way, will under the same conditions expect also to have right of way at the fourth intersection. This may change his reaction to a car coming from the right on this fourth intersection.

With other words, a system is supposed to have a history and a memory about this history that influences its reaction to input. This memory is incorporated in the state of the system, also being the link between input and output and to be taken into account in the model.

A "system" can be anything with some internal coherence. In this paragraph we introduce some general ideas about the systems we want to consider. Later on these ideas will be expressed in a more explicit and exact manner.

The systems will be time-dependent. As a consequence the ordering of the observations is important (in contrast, e.g., to a regression analysis where the ordering of the observations is totally irrelevant). At each moment in time the system can be characterized by the "state" it is in. The state can be seen as a memory of the system, in which all information necessary for a reaction on a particular input is contained in the state. The system is acted upon by input-variables which influence the state, and the system adapts and produces output. The input is supposed not to be influenced by the system.

A simple example that is often used is the bath-tub. Somebody has opened the tap but forgot to put in the plug. The input-variable is the rate of in-flowing water, the output is the rate of out-flowing water. The output is dependent on the level in the tub, which is the state of the tub-system. This state is only partly dependent on the input, it also depends on the previous state.

The assumption that the input is independent of the system is not always easily ensured. Sometimes the definition of the system-boundaries must be chosen carefully, regarding also the time-scale chosen. When one is interested in the choice of speed along a road one can define the individual driver in his car as the system, and regard the different speed-limits along the road as inputs. This is acceptable, although the behaviour of the driver can influence the speed-limit. However, this happens only in the long run (on a different time-scale) and through another level of the complex of systems.

INTRODUCTION TO THE THEORY OF THE LINEAR STATE SPACE MODEL.

As previously mentioned, road safety is one of the outputs of the traffic process we wish to study. A central assumption of state space models is that a process is assumed to have an (unobserved) internal state, which is essential in order to characterize the process. The output of the process is observed (at least what may be assumed to be output of that particular process). It is also possible that the process is influenced by input from the process environment. This influence can change the state of the process or even the process itself. This last phenomenon is assumed not to occur, the process itself is assumed to be invariant, at least for long enough to make this assumption practical. These processes are also called **stationary**. When we start monitoring a process, it is logical to assume that this process was in a particular state the instant just before we started monitoring. One might be interested what that state could have been, and whether, or, for how long this state influenced the successive states of the process. This last point can be useful comparing processes: two equal processes must end up in (almost) the same state when they are both kept under the same circumstances for a long enough time. This is only valid for a special, to be further specified, class of processes.
An example could be the state drivers on the highway are in at some point in time. Two distinct initial states could be whether they just came up the slip road or they have been on the high way for a long

time. It could be reasonable that there is hardly any difference to be seen (in state) if two drivers have been on the road for a long time.

One might ask whether the state of the process can be determined from it's output alone or not. This can be very useful if we want to know in what state the process must be to achieve something, or even which process states are to be avoided (e.g, some danger zone). Of related interest is whether and in what manner we can manipulate the state. Further it is assumed that the state of a process can be characterized by a finite number of real valued parameters, composing the state **vector**.

This state vector and the internal structure of the process are unknown, and this rises to the problem of its dimensionality. The selection of the number of dimensions of the state space is mostly done by trial and error.

These considerations led to the start of a research project concerned with the development of a mathematical model and the development of a field experiment based on the concept of traffic safety problems as part of the dynamical systems approach. This experiment was done using car drivers as individuals and measuring various (environmental) parameters, such as maneuvers of other road users, the speed and acceleration of the car, the driver's heart beat, position of the car on the road and information on the road the car is on at that moment, such as its type, the average traffic product and its accident ratio (Janssen, 1988). This experiment has not been completed yet. The other part of the project is the development of a program to analyze data in this manner. This is being done by the Department of Data Theory of the University of Leiden, (De Leeuw, J and Bijleveld,C.C.J.H, 1987 and 1988). This research led to an experimental version programmed in SAS (a package for statistical analysis), and a Fortran-77 version is in development. A special experimental version was developed at SWOV and used in this paper.

FORMULATION OF THE MODEL

First some assumptions must be made:

1. The process is invariant.
2. In the general case, a particular system is exposed to environmental influence, or input, and is producing output.

3. The output does not influence the input.
4. The input does not influence the output directly, the only effect being through the system itself.

This describes a system for which we will attempt to establish a mathematical model.Recapitulating, a state of the system is dependent on the system its previous state and the present input. This state 'produces' some output.

In more precise terms, for every given moment t (t=1,...,T) there is an input vector x_t (this maybe a zero dimension vector) with fixed dimension p and an output vector y_t (this may not be a zero dimension vector) with fixed dimension r. For each t (t=0,...,T) we get an unknown state vector z_t of dimension q.

This may be described as follows:

$$z_t = f(z_{t-1}, x_t) \qquad t=1,\dots,T$$
$$y_t = h(z_t) \qquad t=1,\dots,T$$

The first equation is called the **system equation**, the second equation is called the **measurement equation.**

Restricting ourselves to linear versions of these functions, this may be written as:

$$z_t = \mathbf{F}(z_{t-1}) + \mathbf{D}(x_t) \qquad t=1,\dots,T \qquad (1)$$
$$y_t = \mathbf{H}(z_t) \qquad t=1,\dots,T \qquad (2)$$

Where **F**, **D** and **H** are linear functions (transformations), often referred to by the matrices that symbolize them. **F** is called the transition matrix, symbolizing the transition function, **D** is called the control matrix and **H** is the measurement matrix. This approach is not new, descending from the linear control theory approach (Kwakernaak, 1972) the matrices **F**, **D** and **H** were assumed to be known, but in this case we wish to estimate them as well.

PARAMETER ESTIMATION

Generally all matrices and state vectors are unknown and must be estimated. It is assumed that neither the system equation nor the measurement equation are error free, therefore a particular loss function is minimized. In this case a sum of squares function denoted by first introducing some error terms in the formulae ($\mathbf{w}_t$ for the system equation, $\mathbf{v}_t$ for the measurement equation) in the following manner:

$$\mathbf{v}_t = z_t - \mathbf{F}(z_{t-1}) - \mathbf{D}(x_t) \qquad (1a)$$

$$\mathbf{w}_t = y_t - \mathbf{H}(z_t) \qquad (2a)$$

Both $\mathbf{v}_t$ and $\mathbf{w}_t$ are vectors of error terms whose components are assumed to be mutually independent random variables with zero mean normal distribution, implying unbiased estimation.
At this point, it can be useful to state that there is no unique solution to this minimization problem. Basically the freedom lies in variation in the state space. For example, all orthogonal transformations (i.e. choosing $\mathbf{R}(z_t)$ instead of z_t with $\mathbf{R}$ being an orthogonal transformation) on the state space result in equivalent solutions.

$$\begin{aligned}\mathbf{R}(\mathbf{v}_t) &= \mathbf{R}(z_t - \mathbf{F}(z_{t-1}) - \mathbf{D}(x_t)) \\ &= \mathbf{R}(z_t) - \mathbf{RF}(z_{t-1}) - \mathbf{RD}(x_t) \\ &= \mathbf{R}(z_t) - \mathbf{RFR}^{-1}(\mathbf{R}(z_{t-1})) - \mathbf{RD}(x_t)\end{aligned}$$

and

$$\mathbf{w}_t = y_t - \mathbf{HR}^{-1}(\mathbf{R}(z_t)) \qquad (2a)$$

This means that given a solution consisting of z_t ($t=0,\ldots,T$), $\mathbf{D}$, $\mathbf{F}$ and $\mathbf{H}$ one could use $\mathbf{R}(z_t)$ ($t=0,\ldots,T$), $\mathbf{RD}$, $\mathbf{RFR}^{-1}$ and $\mathbf{HR}^{-1}$ as well. In other words one could use another basis for the state space, for instance one for wich $\mathbf{RFR}^{-1}$ is a diagonal matrix or the components of $\mathbf{R}(z_t)$ are uncorrelated.
Another variation is choosing $a * z_t$ (a real) instead of z_t for all t resulting in an equivalent solution in (1) and (2) but not in (1a)

and (2a). This last remark leads to the necessity of some standardization in the state space or the measurement transformation.
A special case is choosing H fixed, which is very effective if the dimensionality of the state space is equal to the dimensionality of the measurement space.

THE MEASUREMENT FUNCTION , THE MATRIX H

The matrix **H**, or the measurement matrix, is of fundamental importance. It defines how the individual aspects of the output space are related to the state space, and if there's a kernel present. It also defines a subspace in the state space that is of no direct relevance to the output. This does not mean that it is not relevant at all, because an element of the kernel of **H** can be mapped outside the kernel by the transition function **F**. In this manner a periodic phenomena can be fitted in the model. A simple example could be **F** a rotation and **H** only mapping one dimension of the rotation surface onto the output space. The manner in which a variable of the output space is dependent on the state space , say the i^{th} variable , is defined by the i^{th} row of the matrix **H**. The inner product of this vector and the state at that moment offers an estimate of the i^{th} variable of the output space due to the model. If one is interested in comparing the dependency of the i^{th} and the j^{th} variable of the output space one could use the inner product of the i^{th} and the j^{th} row of the matrix **H** as a measure. Unfortunately this is not generally sufficient due to, for instance, over dimensionality of the state space or less than complete fit of the model. One could proceed as follows. First an estimate of every component of y_t, the output vector, is defined by the corresponding component of **H**(z_t). One first studies the correlations between the components of those vectors or the explained variance. This offers an impression of the extend the model explains the output. The correlations between the estimated values of the i^{th} and j^{th} component could then offer a more sensible measure of mutual dependence (within the system) of different components of the output.

THE CONTROL FUNCTION , THE MATRIX D

The matrix **D**, or the control matrix, is important only when input to the system is assumed. It defines how the input alters the state of the system. If a kernel is present, this subspace is the space in which the input can vary without changing the manner in which it alters the state. This can be clarified by stating that, defining **d** an element of kernel (**D**), then $D(x + \mathbf{d}) = D(x)$ for any x element of the input space (or set of the available input vectors). So:

$$z_t = F(z_{t-1}) + D(x) = F(z_{t-1}) + D(x + \mathbf{d})$$

This means that it makes no difference to the system whether it is exposed to the signal x or x + **d**. In other words, if the difference between two vectors of input lies within the kernel of **D**, then both vectors have the same effect on the system. Unfortunately, this is not generally useful while one will rarely get perfect fit of the model (no stress at all). This results in varying effects of each single input vector on the system. One could define the mean effect of each single input vector on the system.

Assuming all error in the system equation due to the control of the system, one derives while rearranging (1a):

$$D(x_t) + v_t = z_t - F(z_{t-1})$$

By computing correlations between components of $D(x_t)$ and $z_t - F(z_{t-1})$, one gets an impression how well the alteration of the state can be predicted by the input.

THE TRANSITION FUNCTION , THE MATRIX F

The interpretation of the matrix F in relation to the state space is simplified by the theory of invariant spaces. For instance, if the largest absolute value of the eigenvalues of F is sufficiently smaller than 1 than the effect of older states on later states is decreasing. This can easily be seen by stating: (using r as the largest absolute value)

$$| F(z) | \leq r * | z |$$
$$| F(F(z))| = |F^2(z)| \leq r * |F(z)| \leq r^2 * |z|$$
$$| F^n(z) | \leq r^n * | z |$$

Here | x | means the norm (length) of element x, in this case a vector. In this paper the euclidian norm is used , although this is not necessary.

In case there is no input one obtains:

$$z_t = F(z_{t-1})$$
$$z_n = F^n(z_0)$$

Hence:

$$| z_n | \leq r^n * | z_0 |.$$

This means that the state always converges to zero if we don't apply any input to the system. If, on the other hand, this value is greater than 1, then it is possible that the system is getting out of control, i.e, the norm on the state vector is increasing. This is the case if the initial state contains a non-zero element of a subspace that has a ratio larger than one. If we apply the system to one constant input vector, say x, the state vector develops as:

$$z_1 = F(z_0) + D(x).$$
$$z_2 = F(z_1) + D(x) =$$
$$= F(F(z_0) + D(x)) + D(x) =$$

$$\begin{aligned} &= F^2(z_0) + F(D(x)) + D(x) \\ z_n &= F^n(z_0) + \{ F^{n-1}(D(x)) + \dots + F(D(x)) + D(x) \} \\ z_n &= F^n(z_0) + M_n(x) \end{aligned}$$

For each n M_n is a linear transformation ($M_n = \Sigma_{i=0}^{n-1} F^i (D (x))$ and $F^0 = I$)
If the largest ratio r is less than one M_n even exists if n moves to infinity:

$$M_n \quad ---> \quad (I - F)^{-1} B \ (n --> \infty , \ | F | < 1).$$

(I is the identical transformation).
A general result is that the state vector converges to one state if the input remains constant. This can be of use if one is interested in controlling the system. If one is interested in keeping the state in one position, this can be extended to the case were the ratio is larger than one. Assuming the state is z, a suitable input vector x can be derived from:

$$\begin{aligned} z &= F(z) + D(x) \\ z - F(z) &= D(x) \end{aligned}$$

This results in the following equation:

$$(I - F) (z) = D(x)$$

Of course, if D has a kernel this equation cannot be solved uniquely, but in that case one has more optional input vectors to consolidate the state of the system.

A THEORETICAL EXAMPLE

Consider a particle moving in one direction while a force is acting upon it. While moving it encounters air resistance, proportional to the square of its velocity. This seems an example simple enough to understand and nonlinear enough to experiment with. To keep things simple, a force in one direction only is assumed and a particle with mass 1 kg is used in this simulation. We assume the air resistance to be $v_t^2 / 10$, acting in opposite direction as the movement. The force is taken as a function of time: F=SIN(w * t) + SIN(5 * w * t) / 3. This precise function was not chosen for any particular reason other than getting reasonable fluctuation in the input. The elapsed time between two time points in the model is denoted by δt. In this manner we end up with a nonlinear state space model defined as:

$$v_t = v_{t-1} + a_{t-1} * \delta t$$

$$a_t = (f_t - \beta * v_t * | v_t |) / M$$

Here v_t denotes the velocity at time t = $time_0$ + t * δt

a_t denotes the acceleration at time t,

ft denotes the external force at time t,

M denotes the mass of the particle, 1 kg,

β denotes the air resistance parameter.

Including the velocity being $v_t = v_{t-1} + a_{t-1} * \delta t$ in the second equation, while $v_t > 0$, gives:

$$a_t = (f_t - \beta * (v_{t-1} + a_{t-1} * \delta t)^2) / M$$

$$a_t = (f_t - \beta * (v_{t-1}^2 + 2*\delta t*a_{t-1}.v_{t-1} + (\delta t*a_{t-1})^2))/ M$$

Clearly this is a nonlinear system but if δt is taken small then this should not differ too much from a linear model.
As an experiment data was generated for five seconds of time sampling with time laps of 0.05 sec, 0,1 sec and 0.2 seconds. In this case, the model is slightly simplified using a fixed value for the measurement matrix **H**. This can be done without loss of generality and reduces possible rotations in the state space to the identical. It also normalizes the state space such that multiplication is not a valid transformation. Another advantage is that the interpretation is far simpler this way, the state space being the expected output space.

RESULTS

Notable is first a low correlation (less than 0.126 in absolute value) between the two output variables. This means they are not dependent and this will result in a proper distinction between the two dimensions in the state space. Another aspect is the low correlation (less than 0.075) between the input and the first output variable, the force and the velocity, in contrast with the high correlation (greater than 0.99) between the force and the acceleration. This last correlation is not equal to one, Therefore the acceleration is not totally explained by the force. At this point, one can only explain the acceleration from the input.

The results from the runs relevant to this aspect are:

First run:

Time step : 0.05

D -5.283669452E-05
9.962226748E-01

	row	mean	std dev	corr
Z_t-FZ_{t-1}	1	0.0002993	0.0008931	
DX_t	1	-7.309E-06	0.0000410	0.10729
Z_t-FZ_{t-1}	2	0.1442387	0.7713191	
DX_t	2	0.1378167	0.7723025	0.99977

Second run:

Time step : 0.1

D -7.550053997E-04
1.015115142E+00

	row	mean	std dev	corr
Z_t-FZ_{t-1}	1	0.0005538	0.0019106	
DX_t	1	-0.0001006	0.0005912	0.36769
Z_t-FZ_{t-1}	2	0.1416896	0.7939345	
DX_t	2	0.1352468	0.7948840	0.99979

Third run:

Time step : 0.2

D 1.572910580E-03
1.035519719E+00

	row	mean	std dev	corr
Z_t-FZ_{t-1}	1	0.0012209	0.0034674	
DX_t	1	0.0001932	0.0012555	0.31475
Z_t-FZ_{t-1}	2	0.1330888	0.8257702	
DX_t	2	0.1271998	0.8265505	0.99980

Printed here are the mean and standard deviations of the first and second rows of the matrices $DX_t = (DX_1,\dots,DX_T)$ and $Z_t - F(Z_{t-1}) = (Z_1 - F(Z_0),\dots,Z_T - F(Z_{T-1}))$.
One should note the low value of the first element of **D** together with the low standard deviations of and the low correlation between the first rows of DX_t and $Z_t - F(Z_{t-1})$. From this the conclusion is made that the input hardly influences the first component of the state space. The information about the second component of the state space suggests a reasonable influence of the input on the second component.

Interpreting the results, listed below, about the transition matrix **F** one can clearly see that the first row delivers excellent estimation. The first component, as suggested above, not dependent on the input, is perfectly estimated by the previous state. This can be seen using the correlation between the first rows of FZ_{t-1} and $Z_t - DX_t$:

First run:

Time step = 0.05

$$F \quad \begin{matrix} 1.000008106E+00 & 5.016898364E-02 \\ -1.427406520E-01 & -4.843976814E-03 \end{matrix}$$

	row	mean	std dev	corr
$Z_t - DX_t$	1	0.9830164	0.5617061	
FZ_{t-1}	1	0.9827098	0.5622465	1.00000
$Z_t - DX_t$	2	-0.1337881	0.0924366	
FZ_{t-1}	2	-0.1402101	0.0803185	0.99150

Second run:

Time step = 0.1

$$F \quad \begin{matrix} 1.000073552E+00 & 1.010929644E-01 \\ -1.450375617E-01 & -3.097852506E-02 \end{matrix}$$

	row	mean	std dev	corr
$Z_t - DX_t$	1	0.9837848	0.5670522	
FZ_{t-1}	1	0.9831304	0.5682030	1.00000
$Z_t - DX_t$	2	-0.1368614	0.0951249	
FZ_{t-1}	2	-0.1433041	0.0830380	0.99220

Third run:
Time step = 0.2

$$F \quad \begin{matrix} 9.999271035E\text{-}01 & 1.990711838E\text{-}01 \\ -1.459238529E\text{-}01 & -6.734193861E\text{-}02 \end{matrix}$$

	row	mean	std dev	corr
$Z_t - DX_t$	1	0.9827600	0.5782808	
FZ_{t-1}	1	0.9817323	0.5800890	0.99999
$Z_t - DX_t$	2	-0.1391711	0.1013326	
FZ_{t-1}	2	-0.1450600	0.0906569	0.99150

Applying the information due to F leads for the first dimension of the state space to:

$$Z_{t,1} = Z_{t-1,1} + \delta t * Z_{t-1,2}$$

Hence , due to H = I which means $Z_{t,1} = v_t$,

$$v_t = v_{t-1} + \delta t * a_{t-1}$$

The second dimension is somewhat less clear:

$$\begin{aligned} Z_{t,2} &= f_t - \Phi * Z_{t-1,1} - \Theta * Z_{t-1,2} \\ a_t &= f_t - \Phi * v_{t-1} - \Theta * a_{t-1} \end{aligned}$$

Hence:

$$\begin{aligned} v_t &= v_{t-1} + \delta t * a_{t-1} \\ a_t &= f_t - \Phi * v_{t-1} - \Theta * a_{t-1} \end{aligned}$$

The first equation could be substituted into the second, but for now it can be seen that raising the force at one time results in a raise in acceleration, this results in extra increase of the velocity and both result in a decrease in acceleration.

A REAL LIFE EXAMPLE

Oppe (1987) describes an analysis of the relation between fatalities and traffic volume in four countries. This analysis shows that the number of fatalities is a function of the derivative of the traffic volume v with respect to t.
This results in the following functional relation:

$$F_t^2 = a * v_t + b * v_t^2$$

Although this function fits the data quite good, there are still some systematic deviances between the model and the data. Furthermore this relation seems to be independent of time; at least time cancels out as an underlying variable. However if the development of the mass transportation and the side effects of it in terms of accidents are regarded as the development of a system, and system improvement over time in terms of accident reduction as part of this development, then time is an important variable. We therefore reanalyzed the data with the state space model to find out whether an improvement in fit was possible by the introduction of a system component. This seemed likely because there is a high correlation between the number of fatalities in one year and the next year (auto correlation). The auto correlation for number of fatalities is 0.957, for the squared number of fatalities it is 0.951.

Three analyses have been carried out:

1. State space model with input v_t and v_t^2 and output d_t.

2. Ordinary regression modelling $d_t^2=b*v_t+c*v_t^2$ (without intercept).

3. Ordinary regression modelling $d_t^2=a+b*v_t+c*v_t^2$ (with intercept).

The second analysis can be seen as an state space model eliminating the transition of the state component, assuming $F = 0$. This was done to find out to what extend the addition of the state space component improves the fit. In the third analysis the intercept was added to find out whether this parameter could account for the same improvement.

The results were as follows.

The first analysis, the state space model, resulted in:

$$z_t = 0.837 * z_{t-1} + 0.0830 * v_t - 0.00110 * v_t^2$$
$$d_t = z_t$$

The equation can be rearranged into:

$$z_t = 0.836 * z_{t-1} + 0.00110 * v_t * (75.5 - v_t)$$

The second analysis, using v_t and v_t^2 without intercept, resulted in:

$$d_t = 0.385 * v_t - 0.00410 * v_t^2$$

The third analysis, using v_t and v_t^2 with intercept, resulted in:

$$d_t = -2.12 + 0.476 * v_t - 0.00533 * v_t^2$$

Using v_t and v_t^2 with intercept.

Residuals of:	Mean	STD
State space model	0.0	0.767
Regression v_t , v_t^2	5.0	1.300
Regression v_t , v_t^2 with intercept	0.0	1.108

The correlation between z_t and $F z_{t-1} + D x_t$ was 0.996
.. $D x_t$ and $z_t - F z_{t-1}$ was 0.935
.. y_t and $H z_t$ was 0.975

The state is well explained from its previous state and the input and better quality estimates of the number of fatalities were acquired then using the regression models.
From figures 1 through 3 it can be seen that the state space model (fig 1) tends to explain the peak better than the simple regression models (fig 2 and fig 3). This phenomenon will be due to the property of the state space model that it is adapting itself to the output. The solution mechanism of the state space model that the model searches for a

vector z_t balanced between $F\ z_{t-1} + D\ x_t$ and $H^{-1}\ y_t$. This means that the prediction of the state as being wrong does not always mean that the next state of the system is also wrong. This could explain the fact that the one step ahead predictions of the model are better than the findings of Oppe.

CONCLUSION

In both examples the linear state space model proved to be able to supply good results. The model is able to improve the results of standard regression techniques, because it uses time dependent information to explain the variation in the data. The model has to be tested on a wider scale to find its usefulness in practice, furthermore some extensions are to be made; for instance a nonlinear transformation on the input and output data could supply a better fit of the model in various cases. Research in this direction is done at the moment by the Department of Datatheory of the University of Leiden in order to apply the technique to the data from SWOV projects mentioned above.

REFERENCES

Akaike, H. (1974). Markovian representation of stochastic processes and its applications to the analysis of autoregressive moving average processes: Annals of the institute of statistical mathematics, Vol 26 (1974) p 363-387.

Akaike, H. (1976). Canonical correlation analysis of time series and the use of an information criterion. In: Advances and case studies in system identification, edited by R. Mehra and D.G. Lainiotis. New York Academic Press, 1976.

Anderson, T.W. (1951). Estimating linear restrictions on regression coefficients for multivariate normal distributions: Annals of Mathematical Statistics, Vol 22 (1951), p 327-351.

Janssen, S.T.M.C. (1988). De verkeersonveiligheid van wegtypen in 1986 en 2010; Resultaten van berekeningen voor een beleidsscenario uit het Structuurschema Verkeer en Vervoer. R-88-03. SWOV, Leidschendam, 1988.

Jøreskøg, K.G. and Wold, H.O.A. (1982). Systems under indirect observation: causality, structure, prediction. North Holland Publishing Comp., Amsterdam 1982.

de Leeuw, J. and Bijleveld, C.C.J.H. (1987). Fitting reduced rank regression by alternating least squares, RR-87-05. Department of Data Theory of the University of Leiden, 1987.

de Leeuw, J. and Bijleveld, C.C.J.H. (1988). Fitting linear dynamical systems by alternating least squares, RR-88-03. Department of Data Theory of the University of Leiden, 1988.

Kwakernaak, H. and Sivan, R. (1972). Linear optimal control systems. Wiley, New York, 1972.

OECD (1984). Integrated Road Safety Programmes. OECD, Paris, 1984.

Oppe, S. (1987). Macroscopic models for traffic and traffic safety. SWOV, Leidschendam, November 1987.

Wood, E.F.(1984). Handbook of applicable mathematics, Ledermann, W (ed.),VOL VI, Statistics , Part B, Chapter 20. Wiley , New York 1984.

ANNUAL SQR(FATALITIES) BY V SQR(V)

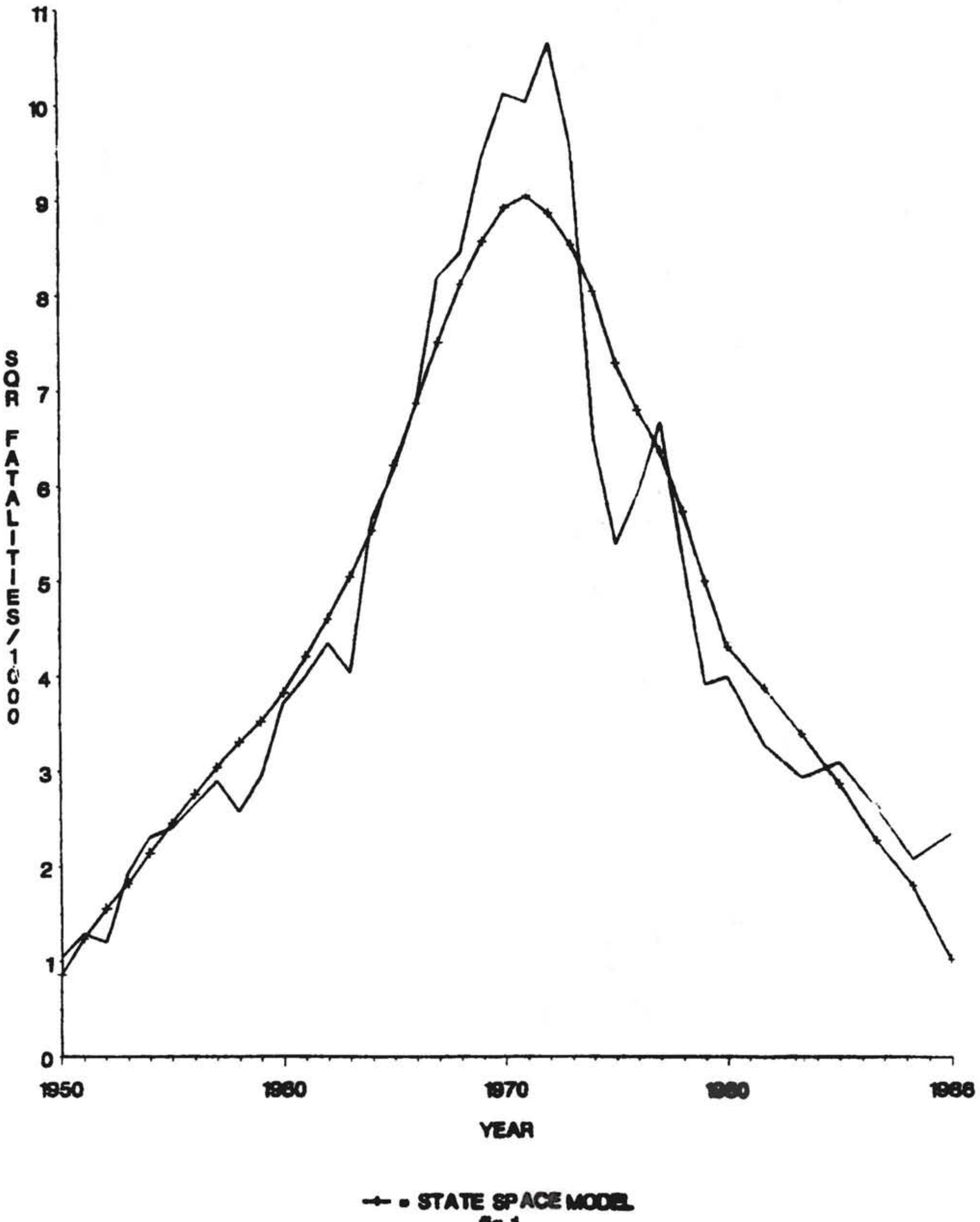
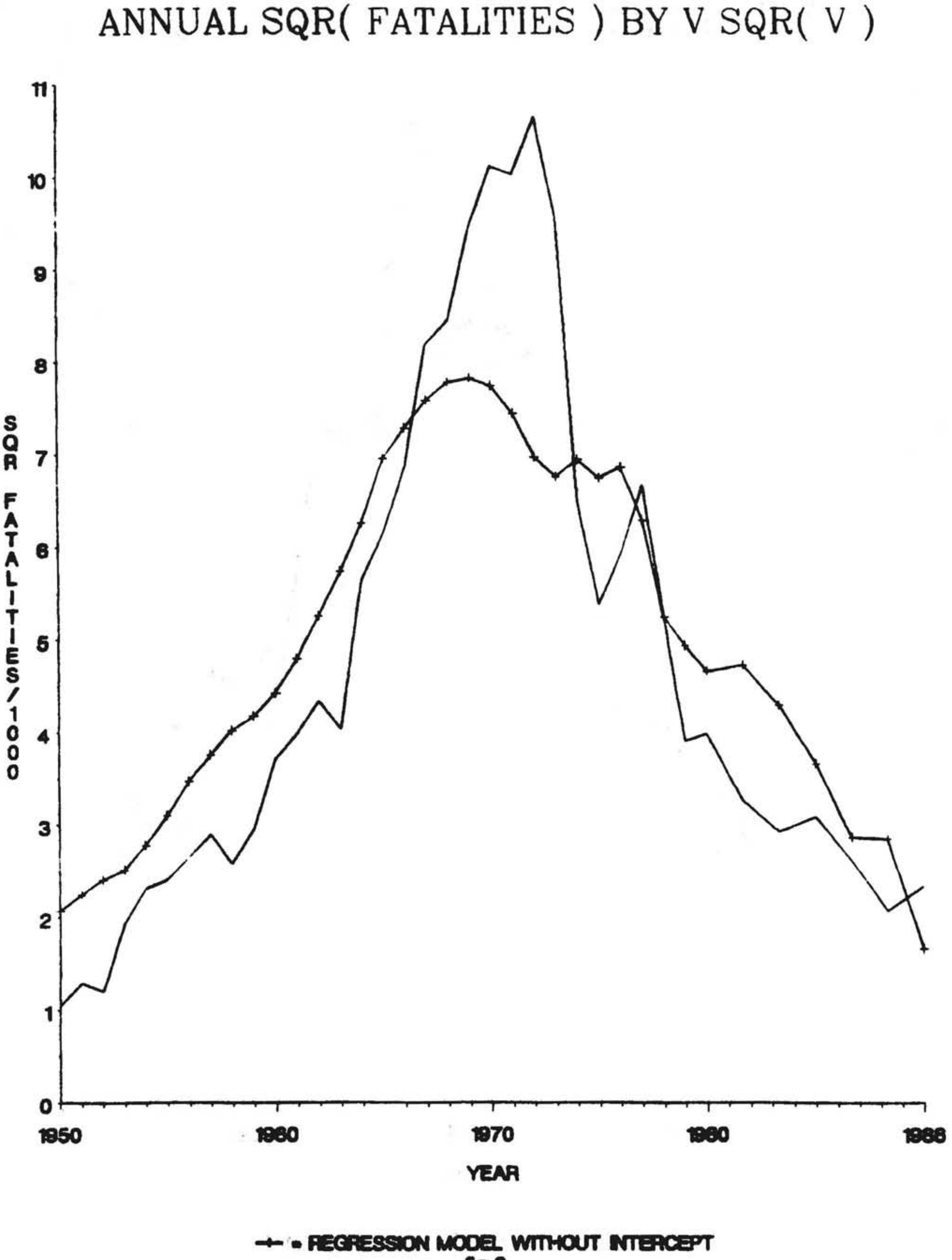


fig 1



ANNUAL SQR(FATALITIES) BY V SQR(V)

fig 2

ANNUAL SQR(FATALITIES) BY V SQR(V)

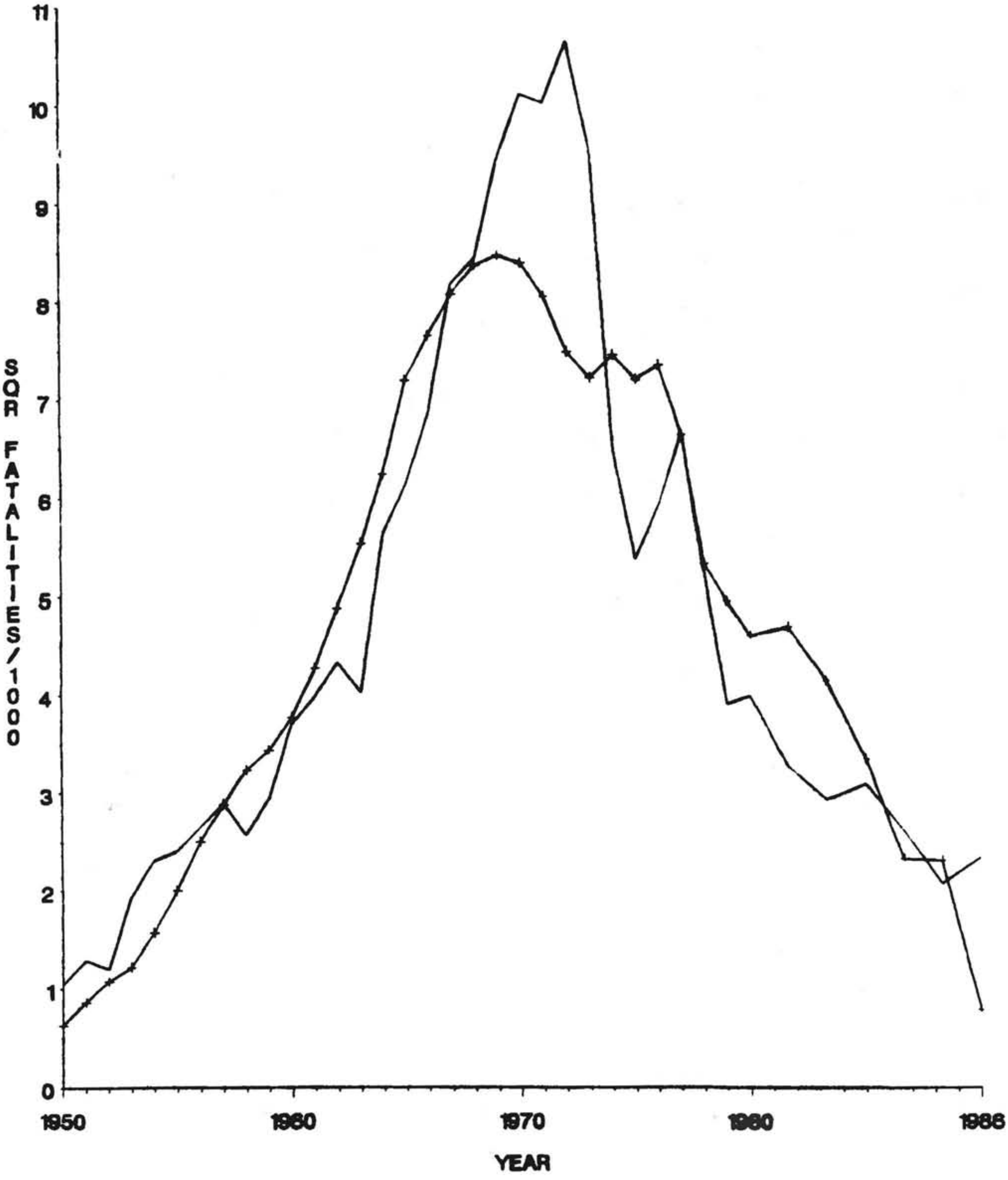


fig 3

COMPUTER SIMULATION OF CRASH DYNAMICS

J. Wismans, P. de Coo and C. Huyskens
TNO Road-Vehicles Research Institute,
Delft, the Netherlands

T. Heijer
Institute for Road Safety Research SWOV,
Leidschendam, the Netherlands

Abstract

The objective of this presentation is to give an overview of the 'state-of-the-art' of crash models in automotive safety. Crash models can be subdivided roughly into four categories:

- accident reconstruction models
- vehicle and obstacle structural models
- crash victim simulation models
- body segment injury models

The most important features of the various categories will be summarized and illustrated by means of typical representative examples. An outline of future developments in this field will be given.

Introduction

"Accidents will happen", an old adage and one of the many examples of proverbs produced by mankind to depict its fallibility in spite of its best efforts. Although accident rates in practically all western countries are steadily decreasing, the absolute number of accidents is still considerable; a pattern that will probably not change dramatically in the foreseeable future. In the Netherlands alone, almost 4 people are killed and 16 people are injured in traffic daily. So, as our attempts at accident prevention occasionally fails, we must resort to ways of controlling the outcome of these accidents, ways to increase the" passive safety" of vehicles and roadside appurtenances. In the past decades this field of research yielded some rather spectacular improvements in the safety of car occupants with the introduction of seat belts and crush zones while at the roadside, vastly improved guard rail and effective impact attenuators save lives every day. But also nowadays developments in injury prevention continue and, although results may not be as spectacular as before, much is still to be done. Traditionally, full scale crash testing and real world accident studies are the "tools of the trade" but in recent years, along with the expansion of digital computing, additional and powerful methods have been developed in the form of computer programs that emulate or reconstruct accident reality; this paper reviews some of the most important developments, their use and their prime characteristics.

Computer models for road accidents

Why computer models

As knowledge of the mechanism of accidents increased, the need for more detailed measuring and testing increased proportionally. This meant, that more and more sophisticated (and vulnerable) measuring and testing equipment was introduced in full scale testing and this, along with the increase in labour costs, has gradually driven the costs of the tests to a point where the development of alternative methods can pay off. Development of computer models in itself is neither cheap nor simple but the costs per simulated accident are only a tiny fraction of what a similar full scale test would cost. Apart from lower operational costs, these models offer additional advantages to the researcher like unconditional repeatability and the possibility to vary important parameters separately in order to establish their influence. The models can be applied as a substitute for or in addition to full scale tests in practically all types of research like:

- reconstruction of actual accidents,
- design of (crash characteristics of) vehicles and roadside facilities,
- biomechanics of impact to the human body and related safety measures,
- development of new and better safety regulations.

Furthermore, application and required results of a certain model may be part of safety regulations, much like standard crash tests.

Reliability of the models

Although the models are generally quite complex and usually employ a large number of parameters to describe the modelled objects and circumstances, "reality" is always even more complex. As a result, many models include generalized correction factors to account for influences that are not included in the causal chains of the model proper or contain some parameters that are impossible to determine exactly and therefore must be approximated. This is why calibration or verification of the models on "real life" data will always be necessary and why we will always need at least some full scale tests. Moreover, as correction factors or estimates of parameters will generally change as modelled circumstances differ signifficantly, new calibration tests are necessary whenever the models are to be used in thus far "unknown territory". Without these tests, models may still produce correct results but the user will not be certain as to when these results are correct and when not.

Types of models

Depending upon the nature of research, several types of computer programs have been developed, each with their own, but often overlapping, area of applicability. Most of the models are of the **predictive** type, that is, based upon (measured or estimated) parameter values representing characteristics of vehicle and surroundings and using the laws of common mechanics the most probable outcome of the modelled accident is calculated. Although the models may differ in many aspects, they are all dynamic models; they all account for inertial affects by somehow deriving equations of morion for all movable parts and solve these equations by some iterative method. Some models, with a slightly different setup, **reconstruct** accidents that actually took place, on the basis of "ex post facto" measurement of post impact positions, damages to vehicles, tyre marks etc.

Thus the following categories*) may be distinguished:
1- simulation of vehicle accidents (HVOSM, VEDYAC, CRASH3, EES)
2- simulation of structural deformations (MARC, ABAQUS, PAMCRASH, PIS CES/ELK 3D)
3- simulation of human whole body response in a crash (MADYMO, CALSPAN/CVS)
4- simulation of human body segments
Some of these models, with the application of some of them the authors have experience, will now be reviewed in more detail.

---

*) Program names between parentheses need not necessarily to be the only applicable programs in this field.

Simulation of vehicle accidents

Crash Reconstruction Modelling

Generally, these programs are used for investigation and interpretation of physical evidence from vehicle accidents. The rest and impact position data provide bases for estimating the speeds of the vehicles at the time of collision contact. On top of that, extra information can be provided by inserting the geometry of the structural deformation. From this, the deformation energy can be approximated, comparing the vehicle's structure with data base vehicle stiffnesses and deformation parameters. They are mostly based on work-energy relationships for the spinout trajectories and the principle of conservation of linear momentum for the collision or they make use of the locations and extends of structural deformations on the structures involved and are based on energy approximations.

One of the programs used at TNO for this purpose is CRASH3 [5]. As an example a reconstruction of a real world accident will be shown. From this accident, which happened in France, most of the parameters were known. As part of a comprehensive European research program on the biofidelity of prototype side-impact dummies, at the TNO Crash Facility this accident has been simulated by a number of full scale tests (see Figure 1) High speed filming was used to analyse the vehicle motions during the tests and to evaluate the CRASH3 reconstruction.

Figure 1: Snapshot of both vehicles during collision.

Some data from the real accident are:
- At the moment of collision the striking vehicle's velocity was 75 km/h and the struck vehicle was assumed to be still in a rest position.
- The collision angle was 70 degrees (see Figure 2); the steering wheels of vehicle 2 were 20 degrees left.
- The mass of vehicle 1 (including belted occupants) was 1385 kg.
  The mass of vehicle 2 (one unbelted occupant) was 1290 kg.

Data from the experimental reconstruction are:
- The initial velocities and positioning were as stated above.
- After impact, from both vehicles the wheels were blocked after 20 m; both vehicles sticked together for some time, where the rear end of vehicle 2 touched the right door of vehicle 1.
- Vehicle 1 slightly turned left. Vehicle 2 spinned to the right due to impact in front of the centre of gravity.

The results from the CRASH3 reconstruction appeared to be in good agreement with the experimental reconstruction. The vehicle motion simulation shows the same trajectory as was observed in the film (see Figure 3). Even the sticking between both vehicles is achieved in this simulation.
It can be concluded that programs like CRASH3 are of great importance to reconstruct accidents for instance for litigation purposes.

The VEDYAC model

This model has been originally set up as a general purpose computer program to simulate both vehicle manoeuvres and vehicle crashes. The basic tools, that the program offers the user to define and describe the vehicles and moveable or fixed obstacles to collide with, are so versatile however, that the simulations are not confined to simulation of road vehicles only: the program can be used to simulate dynamics of man-like structures, trains, helicopters, lighting poles, guard rail etc. Moreover, the number of simultaneously moving and interacting objects is not specifically limited and so it is possible to simulate a crashing car having occupants colliding with its interior at the same time. The model has been used extensively in the development of new or improved types of guard rail, research of slope accidents using various types of passenger cars and trucks, development of crashworthy helicopter parts and may be used to investigate train crashes. The model has proven a cheap and versatile replacement of often too costly full-scale tests. However, these high costs have also prevented an extensive validation of the model on "real world" data, which sometimes makes the results of the simulations uncertain. Still, verifications with the aid of smaller scale mechanical models in laboratory tests have been carried out successfully.

Characteristics of the program in brief

The basics of the model are simply the Newtonian laws of reaction and motion, applied to an unlimited amount of freely moveable rigid bodies. To these bodies, deformable shapes (cylinders or plane elements) can be attached in specifyable places to describe the outward shape and to enable forces to be generated upon contact with other (moveable or fixed) bodies. In order to model vehicles, several types of suspension are available (independant wheel movement, rigid axle, swing axle, steering gear) that can also be attached in any place to moveable bodies. Bodies may be coupled in one or more places by means of deformable joints, the characteristics of which (elastic-plastic-frangible deformation, damping) can be specified for each point separately.

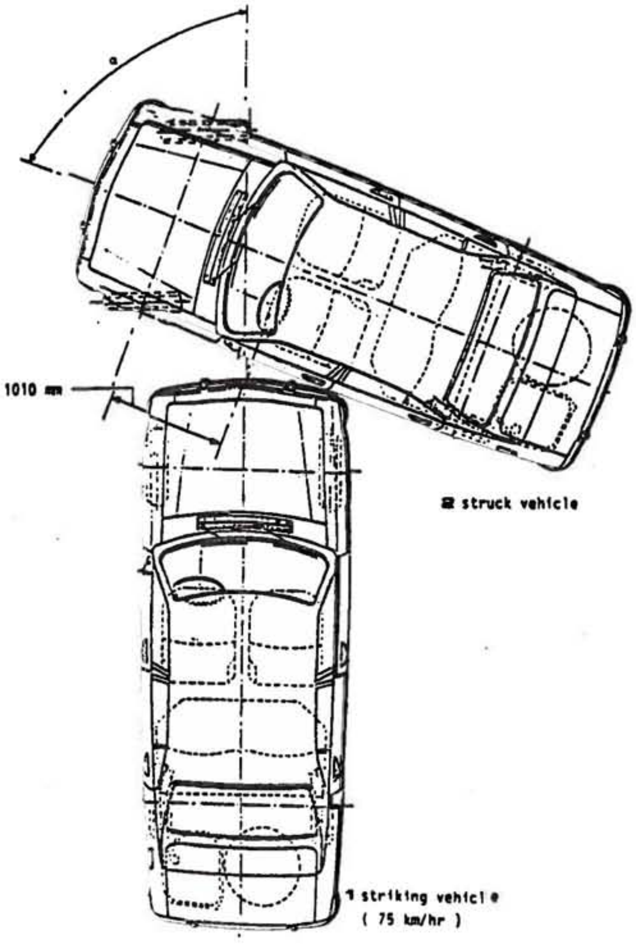


Figure 2: Impact position of both vehicles

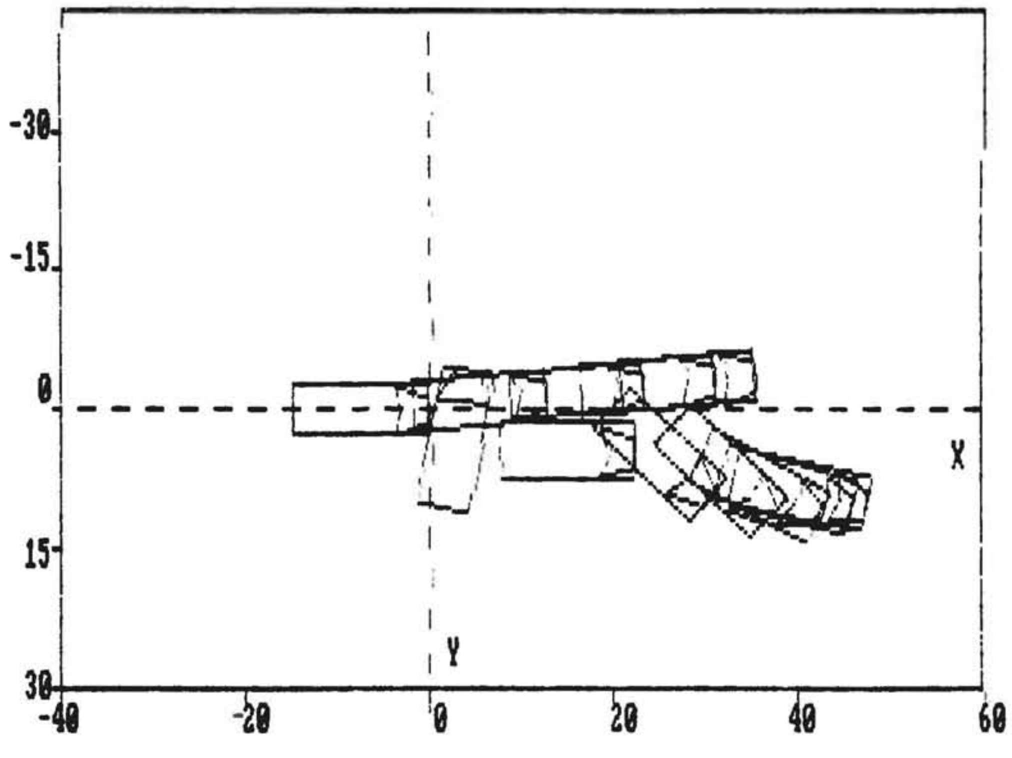


Figure 3: Vehicle trajectory as derived from computer simulation

Thus articulated vehicles (truck-trailer) can be modelled, but also more continuously deformable objects like guard rail beams can be described by dividing the beam into a large number of coupled elements of finite length: in such fashion, the program allows the use of "finite elements" to provide for deformation of complex constructions. Simple finite element shells can also be attached to moving bodies to enable modelling of more complex deformation properties than planes and cylinders would allow. The program solves the equations of motion of each body separately in a large number of small timesteps with the aid of a simple predictor- corrector algorithm. Results can be presented in the form of tables (time series indicating position, rotation, speed and acceleration of all bodies), graphs or 3 dimensional drawings.

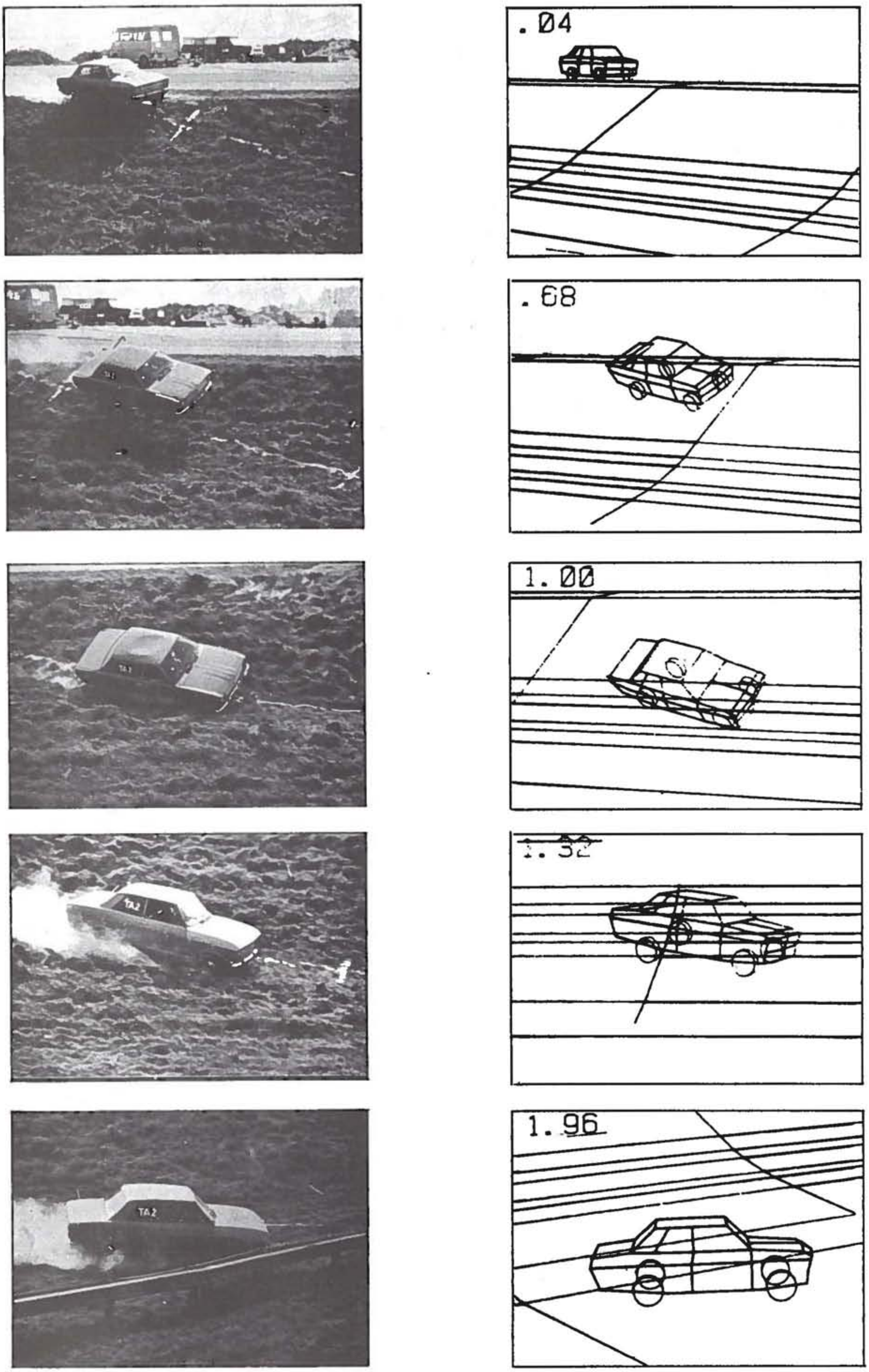
Applications
As already mentioned, the program is extremely flexible and applicable in a great range of dynamic problems ranging from simple manoeuvring on hard or soft soil via cars impacting pedestrians to impact of complexly structured train fronts against heavy trucks. The program has also shown some potential in dynamic analysis of assembled steel structures (vibration of networks of girders and beams). Some examples of recent applications of the program are illustrated in Figures 4, 5 and 6.
Figure 4: Comparison of full-scale test results and model output in case of a slope accident on a 1:2 slope at a speed of 100 km/h
Figure 5: Simulation of a truck, loaded with a 5 ton steel roll, colliding at 80 km/h with a concrete median barrier
Figure 6: An example of a finite element structure: an aluminum train body colliding at 70 km/h with a 30 ton truck.

Simulation of structural deformations

With these programs the vehicle's structure can be optimized and the structural behaviour can be predicted. They are based on finite element and finite difference techniques. With the current general purpose finite element and finite difference computer programs problems such as static crushing and low and medium velocity impact can be dealt with. By the appearance of supercomputer systems these problems can be of rather extent. However, to perform a simulation of a full vehicle crash, a very complicated and cpu consuming model is needed. Specialized computer programs for crash analysis have been derived from existing advanced packages by unifying their respective virtues and by adding new features. In this way, special purpose crash analysis packages have been developed. Until now, application of such programs in the automotive industry is not common use. Composing the model and defining the input parameters require specialized knowledge on engineering and computational mechanics [8]. As an example, in Figure 7 the difference is shown between an overall structural analysis mesh and a special front-end crash analysis mesh.
For type approval a road vehicle must undergo a series of tests. One of these tests is a frontal collision with a velocity of 50 km/h with a rigid wall. During this collision certain parameters must be measured which should not exceed prescribed values. A simulation of this test is presented here.

During a real crash, the kinetic energy (1/2.m.v2) is dissipated by the structure and transformed into deformation energy (F.s). Normally, the deformation length is of order 0.5 m, which results in an average deceleration of 20 g. The total impact has a duration of about 100 ms.

Figure 4: Comparison of full-scale test and model output



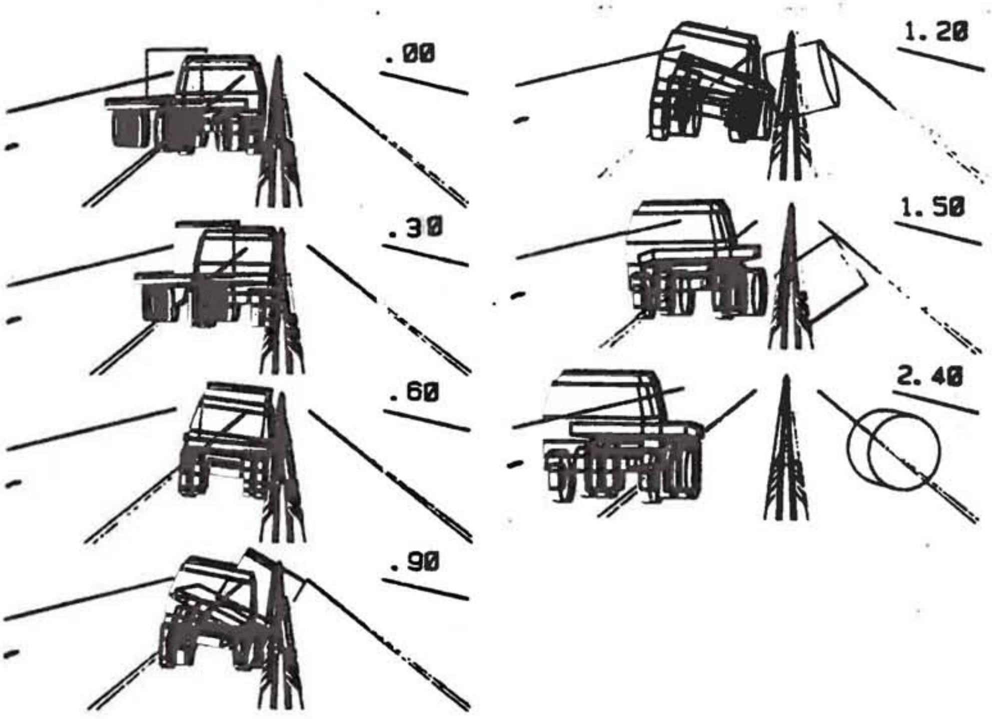


Figure 5: Truck, loaded with a 5 ton steel roll, colliding with a concrete median barrier

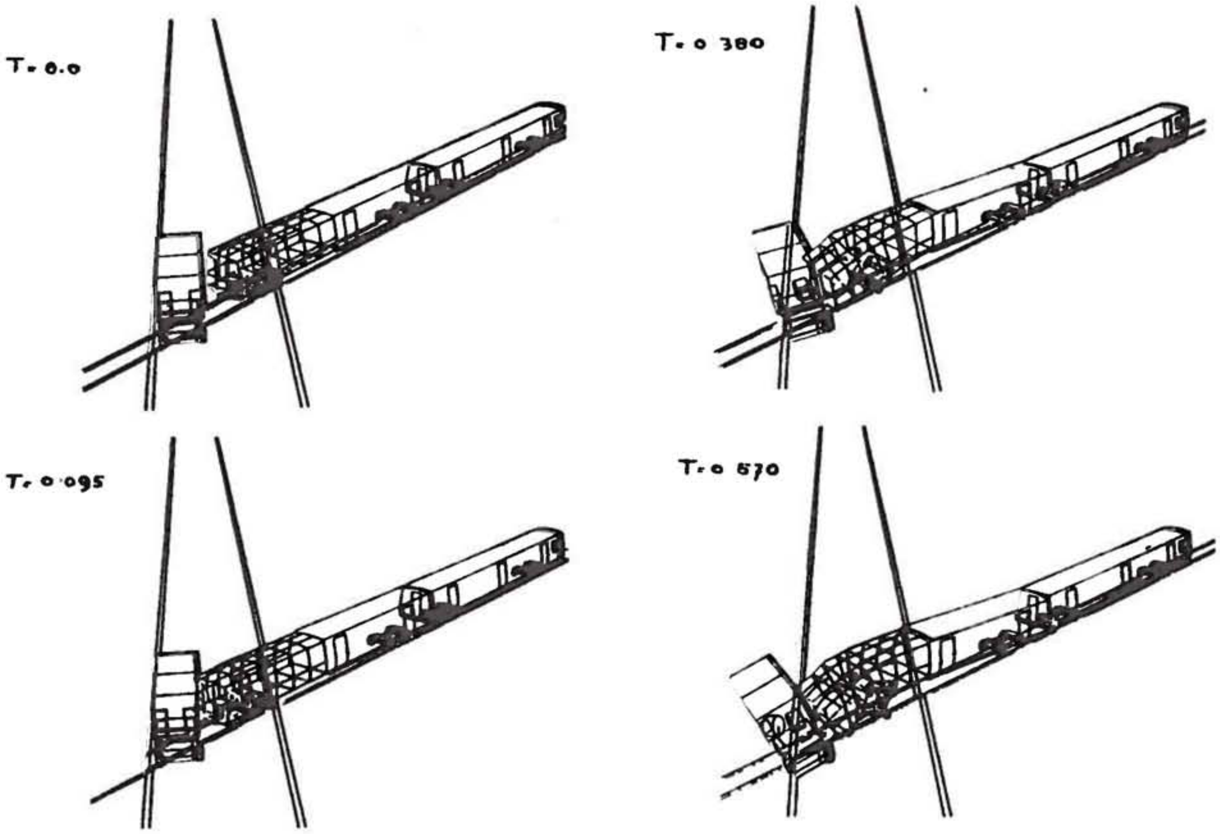
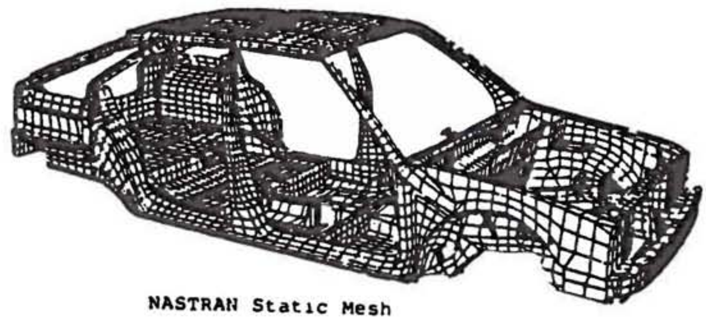
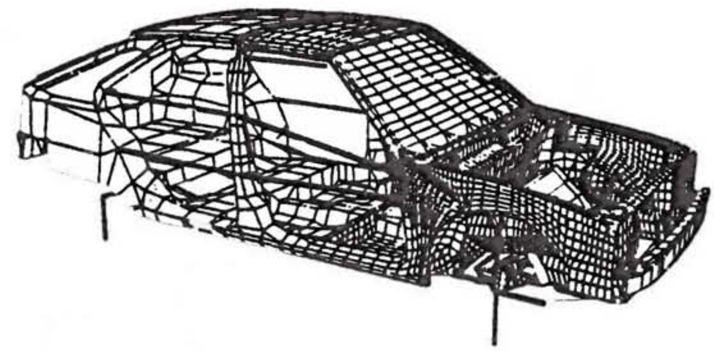


Figure 6: Aluminum train body colliding with 30 ton truck



NASTRAN Static Mesh



PAM-CRASH Dynamic Mesh

Figure 7: Static structural mesh and dynamic local impact mesh

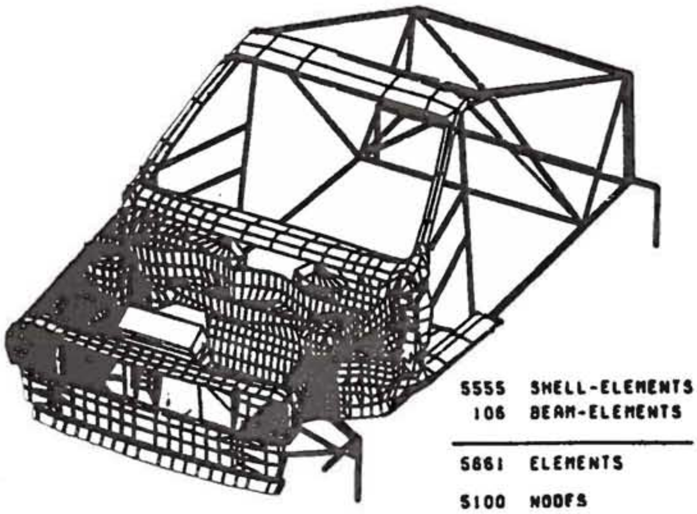


Figure 8: Finite element model

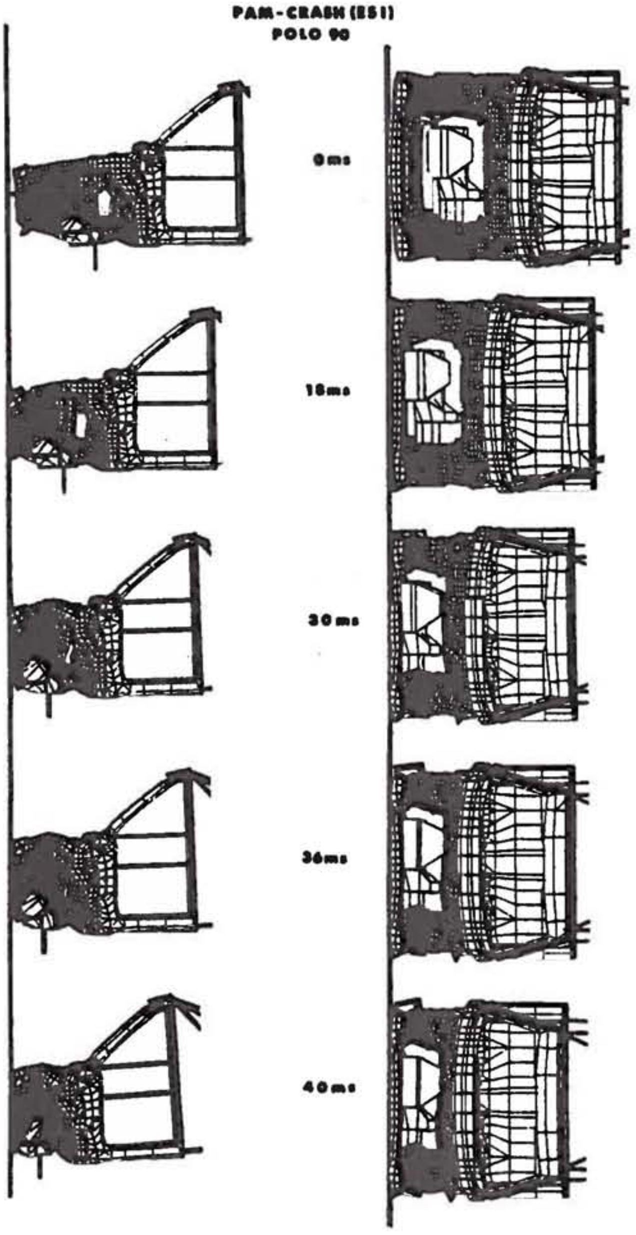


Figure 9: Finite element crash simulation

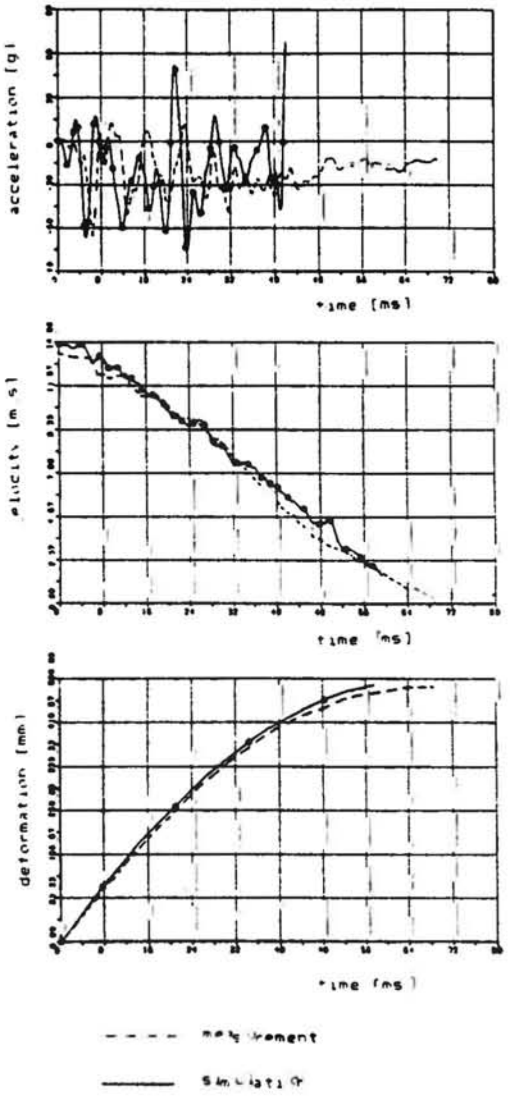


Figure 10: Acceleration-, velocity- and deformation-time relationship of a rear vehicle point

A finite element model of the front structure of a Volkswagen Polo is shown in Figure 8. It consists of 5555 shell elements and 106 beam elements, representing the structural and physical properties of the vehicle.
Only 60 ms of crash time have been simulated. This simulation took 4 hours of cpu time on a CRAY1 supercomputer. The results are shown in Figure 9.
In Figure 10 some comparisons are shown from simulation and a full-scale test. The results are in good agreement.

Simulation of human whole body response

This type of models has been developed to describe the dynamic response of a vehicle occupant involved in a collision event. However, they also can be used to study for instance the motions of a pedestrian or a cyclist if impacted by a vehicle [1]. The human body in this type of models is described by a number of rigid elements connected by hinge or ball and socket joints. The dimensions and the mass of the body usually can be changed to represent an actual accident victim. The earlier programs were two-dimensional models developed to study frontal impacts. Later models, however, have fully three-dimensional capabilities. A review of various models developed in the past is given by King and Chou [2].

One of the most recent programs available now is the MADYMO crash victim simulation program [3]. This program has been developed at the TNO Road-Vehicles Research Institute in the Netherlands in co-operation with the Institute for Road Safety Research SWOV. The program is widely used by car companies and research institutes for instance in computer aided design applications and for biomechanical analyses. Advanced graphics possibilities are available which allow visualization and animation of the motion of the body during a crash event. Special submodules allow for the simulation of the human body with the vehicle interior and crash safety devices like an employing airbag. The MADYMO program has been validated in the past in numerous studies by comparing the predictions with results of experimental crash simulations. Illustrations of the possibilities offered by MADYMO are given in Figure 11. Results are presented for the reponse of an occupant in a frontal collision, a side impact and the motions in a pedestrian impact.

Simulation of human body segments

The final type of crash models to be discussed here are the body segment models where the model representation is restricted to a specific organ or structure. They particularly have been developed to study injury mechanisms in the most frequently injured body regions like the head, neck, spine and thorax. For a review of this type of models see Ward and Nagendra [4].

In these more detailed models the distribution of the forces inside the tissue and bone is studied and internal stresses and strains are calculated.
Due to the complex geometry involved and the non-linear and elastic properties of biological tissue in mostly finite element or finite difference representations are used. However, many of the significant parameters of biological material are not or only partially known and consequently the realibility of the predictions of such models appears to be still rather limited if correlated with experimental test data. Figure 12 illustrates a typical example of this category of model namely a finite element representation of the human brain developed by Ward [4].

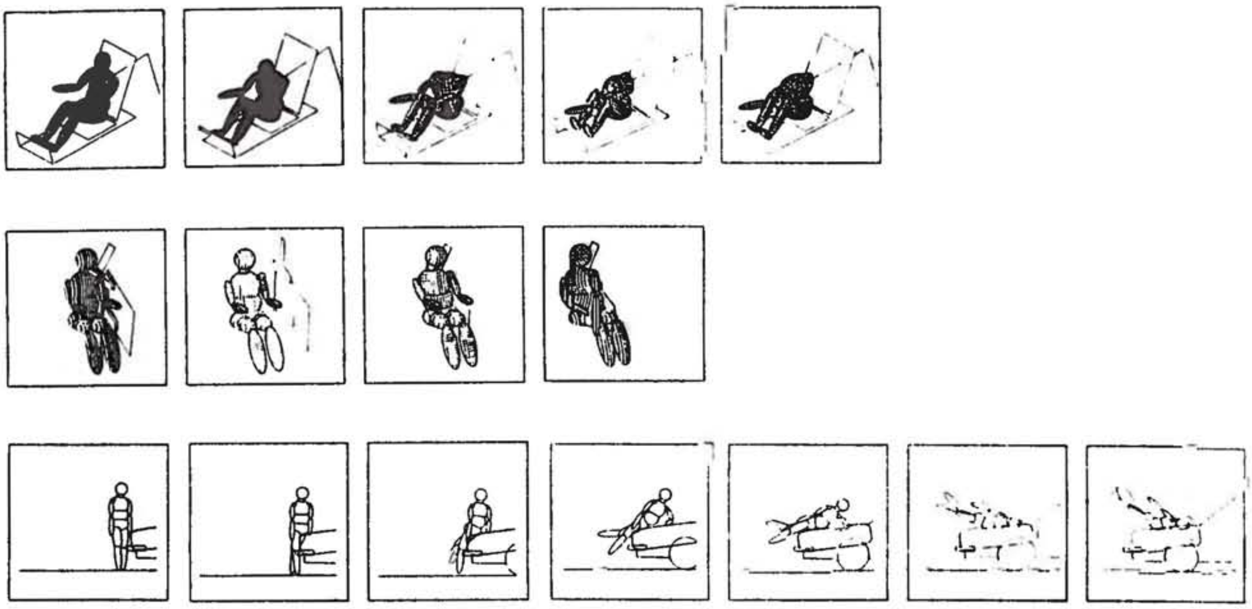


Figure 11: Examples of model simulations with the MADYMO Crash Victim Simulation program.

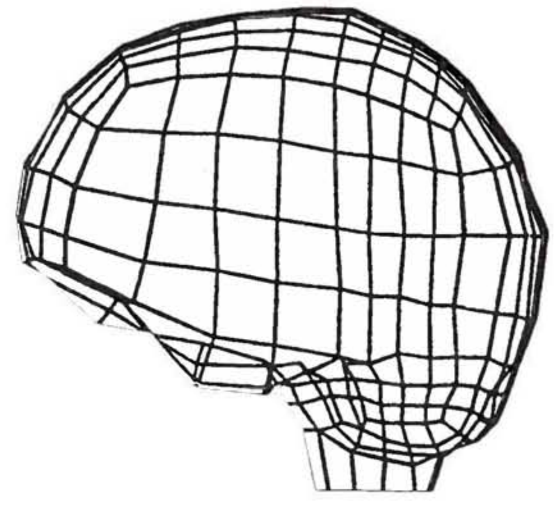


Figure 12: Finite element model of the human brain used for head impact simulation.

Discussion

In the past years application of crash models in traffic safety research has shown a rapid growth. To a large extend this is due to a strong increase in features offered by software packages. But also more powerful computer hardware systems have become available which have contributed significantly to this trend.

In this paper the most important categories and applications of crash models have been reviewed. They vary from relatively simple lumped mass models to study vehicle trajectories during a crash to very complicated non-linear finite element programs to simulate the vehicle deformation in a crash or the response and injuries in a specific body segment like the human brain. Practical application of this last type of models (i.e. human segment models) is still rather limited mainly due to the lack of information on material properties of biological tissue.

In the field of accident reconstructions, particularly for litigation purposes, models have been applied in the past years extensively. Also in vehicle design more and more use is made of crash models (computer aided design). Though these analyses are not common practice yet in automotive industries, it is expected that simulations will be of increasing importance during the design period of a vehicle. However, these simulations will never superfluit the need of testing, because it is not possible to take account for or to model all phenomena which may be of influence during a real crash test. By applying simulation techniques in an early stage of the design, a considerable decrease in the number of prototype crash tests can be expected. As a result, in a shorter time and a more effective way a vehicle can be designed with optimal safety features.
A final development which is expected to contribute significantly to the acceptance of simulation programs for crash analysis is the increasing use of technical graphical workstations. Most recent developments in this field allow real-time animation of the dynamics during a crash which appears to be a real advancement in understanding the complex mechanisms involved in a crash.

References

1. Janssen, E.G., Wismans, J.
   Experimental and mathematical simulation of pedestrian-vehicle and cyclist-vehicle accidents.
   10th Int. Conf. on Exp. Safety Vehicles. July 1985, Oxford, England.

2. King, A.I., Chou, C.C.
   Mathematical modelling, simulation and experimental testing of biomechanical system crash response.
   J. Biomechanics, 1976, Vol. 9, pp 301-317.

3. Wismans, J., Griffioen, J.A.
   MADYMO Crash Victim Simulations.
   AGARD 66th Meeting of the structures and material panels.
   Specialists meeting on energy absorption of aircraft structures as an aspect of crashworthiness, May 1988, Luxembourg.

4. Ward, C., Nagendra, K.
Mathematical models: Animals and Human models.
The biomechanics of trauma.
Nahum and Melvin (eds), Chapter 5, pp 77-100.
Appleton-Century-Crofts/Norwalk, Connecticut.

5. U.S. Department of Transportation.
National Highway Traffic Safety Administration.
CRASH3 User's Guide and Technical Manual.
September 1982, Washington, D.C.

6. Jones, N., Wierzbicki, T.
Structural Crashworthiness.
Butterworths, London, 1983.

7. Haug, E., Scharnhorst, T., Dubois, P.
FEM-Crash, Berechnung eines Fahrzeugfrontalaufpralls.
Berechnung im Automobilbau: Tagung Würzburg, November 1986.

8. Chedmail, J.F., Haug, E.
Numerical Techniques, Experimental Validation and Industrial Applications of Structural Impact and Crashworthiness Analysis with Super Com puters for the Automotive Industries.
Proc. Int. Conf. on Supercomputer Applications in the Automotive Industry, Zürich, October 1986.

EVALUATION METHODS FOR TRAFFIC SAFETY ASPECTS OF NEW TECHNOLOGIES IN VEHICLES

B. Ch. Färber, B.A. Färber & M. Popp
University of Tuebingen, FRG.

INTRODUCTION

Electronic systems in cars are becoming more and more important and the project PROMETHEUS is only the logic consequence of a long development in car electronics. Electronics gives us the possibility to measure and display a huge amount of information, which is more or less helpful to the driver. The problem for traffic safety researchers is to evaluate the value of these additional systems on a rational way, excluding esthetic and emotional points of view.

Some of us remember the "good old times", where car instrumentation consisted only of a speedometer and two control lamps for oil preasure and high beams. They may have the feeling, that these indicators have been sufficient, and are sufficient even in nower days. There is another group of persons who may be termed as "high-tech freaks". These drivers want to have as much information as possible and like the feeling of steering an aeroplane. Which group of drivers causes more risks or more accidents?
It is evident that this is an unanswerable question. The question should only demonstrate a problem:
we are able to compare cars with respect to their horse power, fuel consumption or crushable bin, but we don't have any safety marginal to compare different displays or controls. Until now, the designer of automobil dashboards may have some ergonomic rules in mind, all the rest is based on

principles like customers fancy or expected salery rates. In spite of the lack of such a measure, the design of displays and controls can be seen as an important part of the closed-loop "driver-car-environment", and therefore of save driving.

PREREQUISITES FOR EVALUATION PROCEDURES

The development of evaluation procedures for different dashboards must take into account several factors, which belong to the closed-loop driver-car-environment. These factors are:
- the optical, acoustical and mental processing capacity of the driver.
- the variability and constancy of driving situations.

Perceptual chanels and divided attention

Mental load: Several experiments from basic research (e.g. SPELKE et al. 1976) or from traffic safety research (FÄRBER, 1987) show mens ability to divide attention without lack of accuracy, if different input chanels are used. Other investigations (MORAY, 1967) demonstrate, that seperation between different input chanels is no sufficient condition for unrestricted information processing (compared to undivided attention).
There seems to be a second variable, which is responsable for the performance differences in single or multiple chanel processing. This second variable concerns the similarity between the task and the subjects internal model of that task. Therefore we have to look for appropriate criteria to predict, which information is best suited to the internal model of a driver. Unfortunately we are only partly able to

predict the accordance of physical lay-outs and cognitive structures of men. As a consequence we must find a general measure for mental load which allows the comparison of different realisations in representative driving situations.

Motor load: With respect to controls, motor load is mainly important for arms and hands. Motor load seems to be less important than visual distraction, if two restrictions are fulfilled:

1. handling of controls must be possible without visual feedback,
2. the driver rests in his normal position and one hand remains on the steering wheel.

These conditions are satisfied, if controls are positioned in the optimal reach and are haptically separable from each other.

Visual distraction is undoubtedly the biggest problem in car driving. Not only the huge amount of information in the road environment, but also inside the car tends to overload the driver. This general accepted statement leads to the question: how to measure visual distraction with sufficient accuracy and external (i.e ecological) validity?

Control of situational parameters

Another precondition for the comparison of traffic safety effects of two ore more realisations of dashboards is the systematic control of situational variables. On one hand it is important that the driving situation is held constant - otherwise the effects of the driving task and the secondary task are confounded. On the other hand, the repeated presentation of the same situation is impossible. Especially critical events change their characteristics from the first to the second presentation. What we need, is a model to construct parallel, equally loading traffic situations.

A cybernetic model of driving situations

The proposed model has three dimensions:

- The driven speed,
- the predictability of the situation and
- the manoeuvring space.

If each dimension has only two realisations, the eight situational variations in figure 1 result:

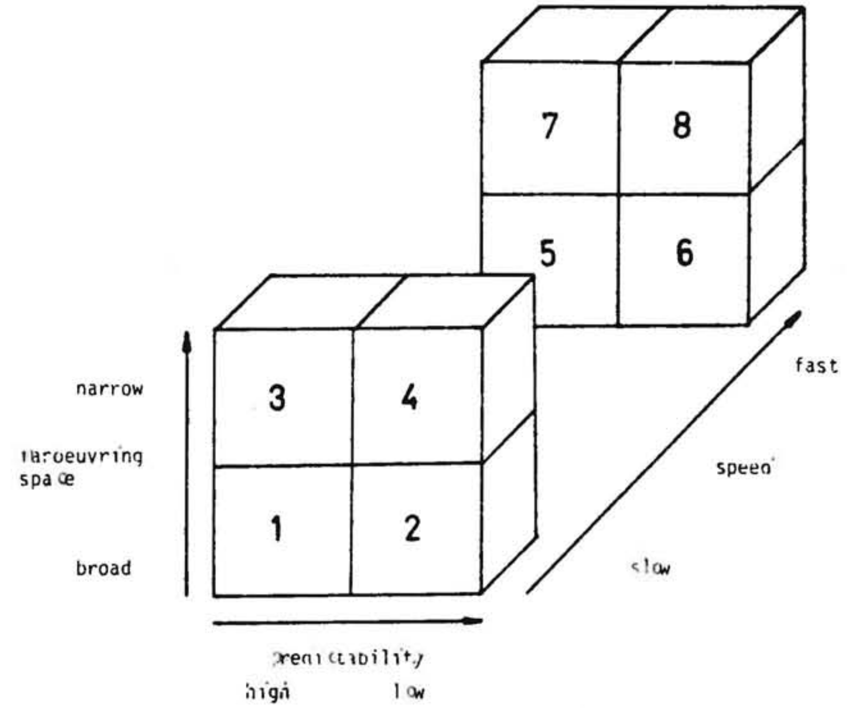


Fig. 1: Three-dimensional model of driving situations

The situation with the lowest operator load is characterized by low speed, high predictability and ampleous manoeuvring space (e.g. driving alone on a wide airfield with low speed). In the most loading situation speed is fast, predictability low and the street narrow (e.g. driving in a busy narrow urban road, where pedestrians can cross the road).

Following this model, several experimental realisations where definied and tested in two different driving simulators. The static simulator uses a film projection and has no motion simulation. The other one is a fully dynamic simulator with computer generated interactive video display of traffic scenes and realistic motion simulation. Heart rate, tapping (= producing a constant rhythm, MICHON, 1965) and subjective ratings were taken as dependent measures. For most experimental realisations the predicted change of subjective and objective measures could be observed. Generally speaking the lowest values were obtained in cell 1 of figure 1 and the highest values in cell 8.

Figure 2 exemplifies the change of the heart rate due to increasing speed.

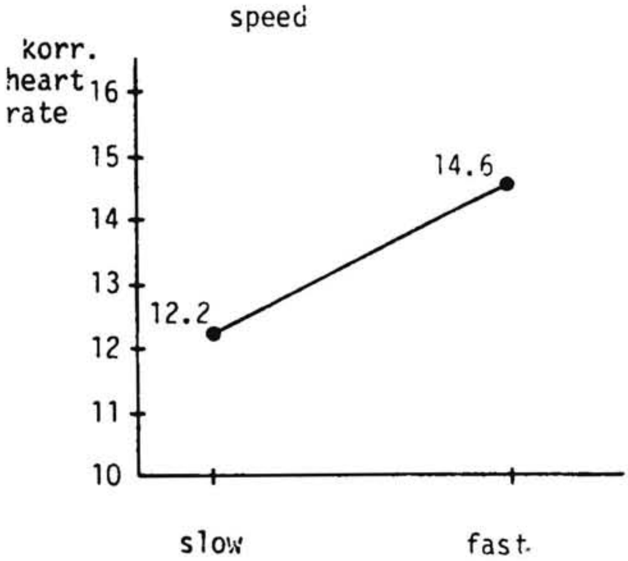


Fig. 2: Heart rate for "slow" and "fast" situations.
slow = 80 Km/h; fast = 130 Km/h.

Figure 3 and 4 show the increase of tapping irregularity for the variables "manoeuvring space" and "predictability".

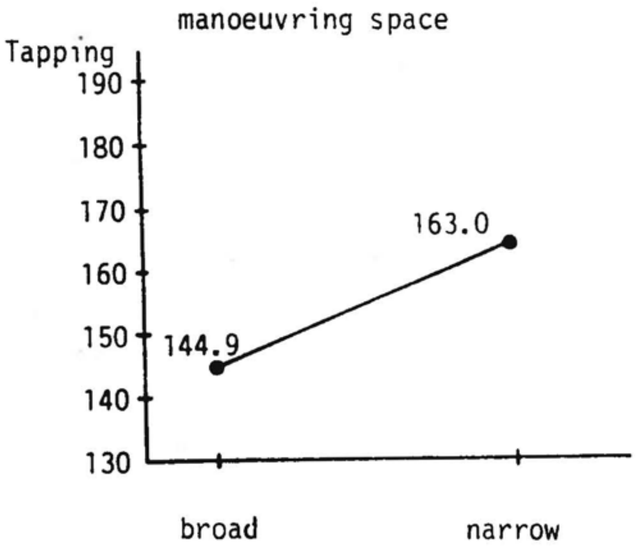


Fig. 3: Tapping irregularity for narrow and wide manoeuvring space

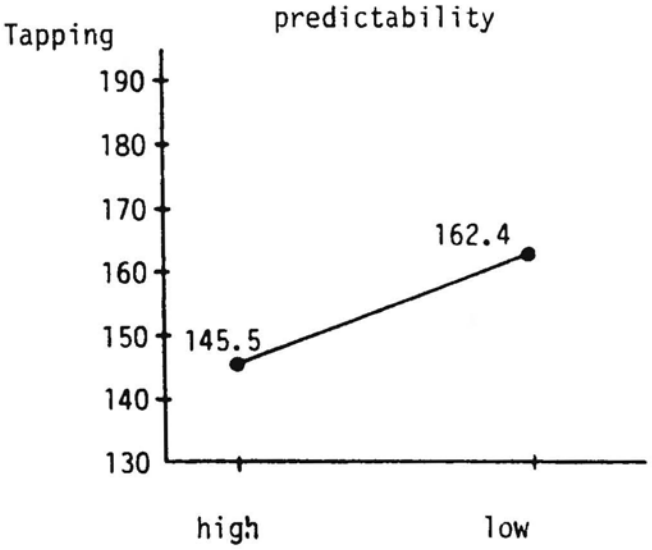


Fig. 4: Tapping irregularity for predictable and unpredictable situations

With this classification in mind we are now able to define traffic situations with high or low visual and mental operator load.

EVALUATION MEASURES

In analogy with other diagnostic processes (e.g. the diagnosis of managers), multiple measures are treated as more reliabel than single measures. Besides we must accept that a single safety measure for displays and controls will never be found. Too many aspects have to be taken into account. According to the above mentioned distractions of visual and motor channel and to mental load, lots of measures have been tested within the last years.

Parameters of driving accuracy:

Experiments in driving simulators provide a lot of measures like acceleration and decelaration, braking behavior or variance of side distance. Of course it is also possible to get all these measures in field experiments with instrumented cars - only the technical problems are somewhat higher. These measures can reflect driving accuracy, but they must not do so. For one reason the participation in any kind of experiment is an unnatural situation and people will not behave as they do in normal life. From a theretical point of view it would be best to test displays and controls in high loading and critical situations. The rational behind this experimental paradigma is the opinion that a system that works in critical situations will also function in normal situations. But, if the driving task in the experiment is highly difficult, subjects will never watch their instruments or try to change the radio station. They tell you: "in the moment I'm not able to do that". On the other hand we know from accident analysis and daily life, that people manipulate their controls even in busy traffic situations.

The second reason, why parameters of driving accuracy must not be overestimated, is a question of validity. WIERWILLE et al. (1968) found that experienced drivers have bigger variances in side distance than beginners. As a consequence they believe steering corrections to be the more valid measure. But, steering corrections themselves depend on the difficulty of the driving task and the amount of physical obsticles. As we can see, the objective driving parameters can only be taken as one indicator of divert from traffic environment.

The measurement of visual distraction

Visual distraction from traffic environment can be measured in several ways:

- The most simple way is a mirror at the dashboard and a video camera that watches the subjects eye-movements;
- the more sophisticated method uses Electro-Okulo-Gramm (EOG) - a measure that analyses the muscle potential near the eye;
- the most sophisticated way is eye-movement recording by an eye-movement camera.

Our own experiences show, that the measurement of visual distraction is best done by EOG. This measure is precise enough to decide, wether the subject looks at the traffic or watches the dashboard. Of course it is not possible to define the exact position of the dashboard where the driver looks on. But for most questions of traffic safety this is not necessary. The only relevant question is: how often and how long the gaze of the driver is distracted by the secondary task.

Compared to eye-movement cameras the main advantage of EOG is the minimal impairment of the subjects.

Performance measures

Performance measures in the secondary task are useful for the evaluation of displays and controls as well. The best parameter is solution time for a specified task. If a task is self-explaining, no long mental operations for the solution are necessary. Performance measures of the secondary task alone are useless, because the capacity distribution between primary and secondary task can change. Performance differences in the secondary task must coincide with constant performance in the primary task.

Some illustrative results.

The above mentioned measures haved been applied for the evaluation of several new display and control technologies. One example is the comparison of analog and digital devices in a driving simulator (see figure 5 and 6). Digital devices attracted subjects glances more often than analog ones (see also FÄRBER & FÄRBER, 1987). This is because the periphery of mens eye is sensitive for brightness differences. With respect to the task of the driver these looks are useless, because they are not necessary for speed control. Under this point of view analog devices are preferrable. On the other hand one must admit that no negative influence on driving behavior could be observed. Following our results, the discussion wether analog or digital speedometer are better, is more a question of esthetics and not of traffic safety.

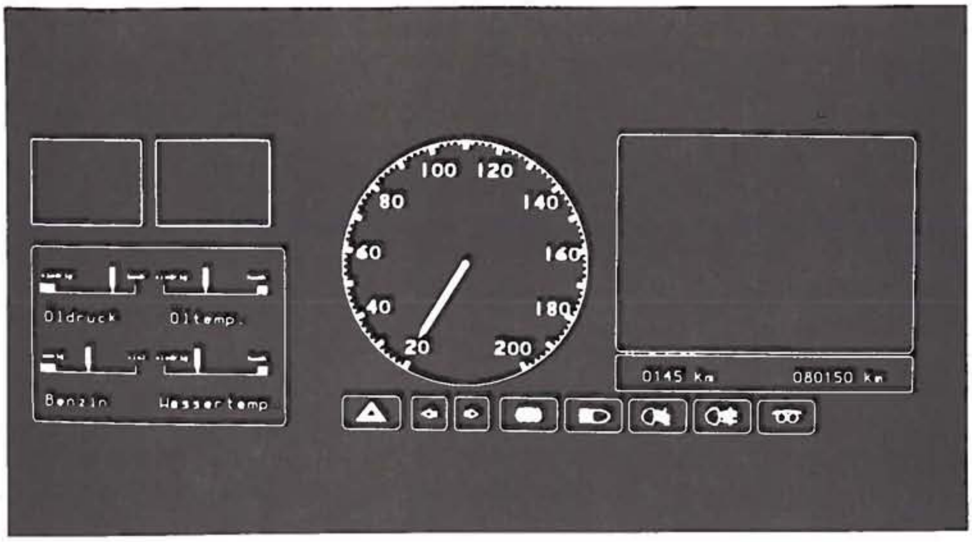


Fig. 5: Example of the investigated analog display

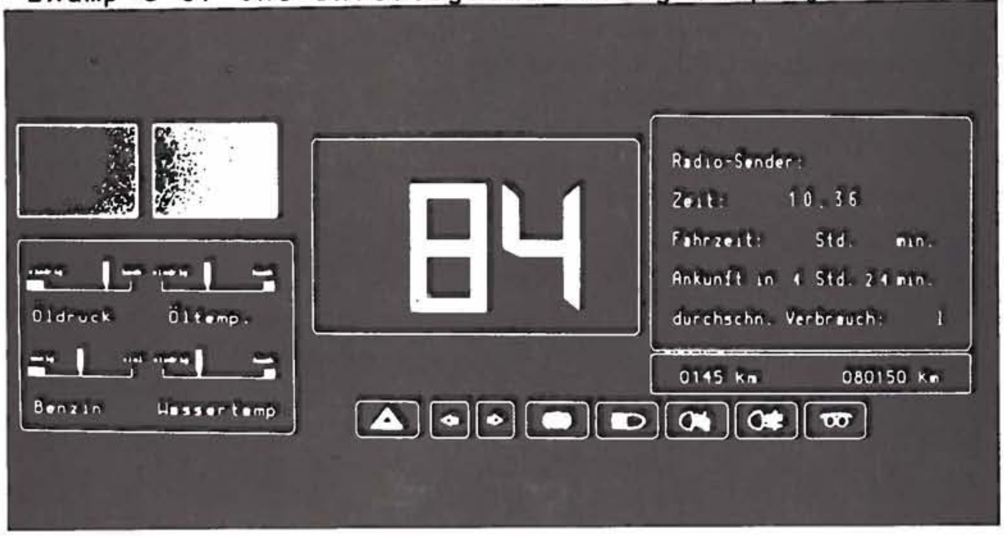


Fig. 6: Example of the digital display

A second example demonstrates the possibility, to evaluate new controls. Four different realisations of modern controls were tested:

- Softkeys (the key-function can be definded and changed by the operator)
- Hardkeys (controls with fixed function) in combination with cursor control

- Cursor control
- Voice control.

Figure 7 and 8 show the differences between these four experimental conditions in the EOG-measure.

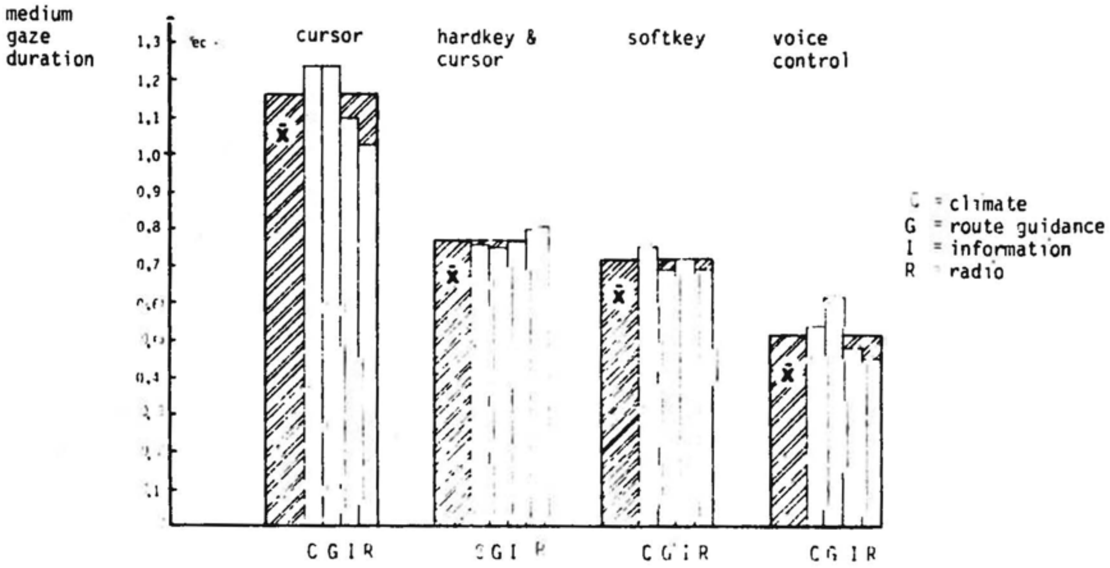


Fig. 7: Mean gaze duration of four different controls

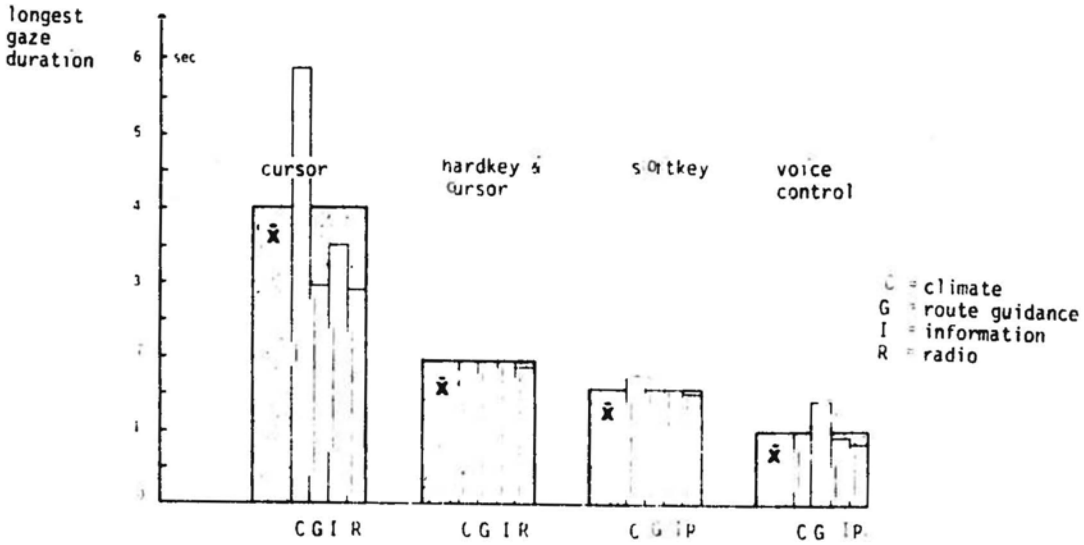


Fig. 8: Longest gaze duration of four different controls

In contrast to the monitoring of displays, the handling of controls has some influence on steering and braking behavior. Here, manual and visual distraction coincide. Brake reaction times increase significantly from voice control to softkey, hardkey and cursor control. The same negative effect can be observed for tracking behavior in critical sitations.

CONCLUSIONS

The adequate presentation of information in the cars has an important influence on traffic safety. Undoubtless many other factors like experience, risk perception and calculation, fatigue or motivation are also important for safe driving. But as 'conditio sine qua non' right information acquisition and processing can be seen. As evaluation procedure for dashboards with respect to traffic safety, a combination of:

- time and frequency of visual distraction using EOG-measure (visual load),
- driving performance in a simulator (motor load), and
- performance of control operations (mental load)

seems to be best. Prerequisite to the application of these measures is the definition of traffic situations with specified operator load. The proposed three parameter model using "speed, manoeuvring space and predictability" can generate parallel traffic situations that fulfill this requirement.

Even if we do not have a single safety marginal for dashboards, we are able to measure this part of traffic safety beyond the momentary taste of designers or customers.

References

FÄRBER, B. (1987). Geteilte Aufmerksamkeit: Grundlagen und Anwendung im motorisierten Straßenverkehr. Köln: TÜV-Rheinland.

FÄRBER, B. CH. & FÄRBER, B. A. (1987). Sicherheitsorientierte Bewertung von Anzeige- und Bedienelementen in Kraftfahrzeugen. FAT-Schriftenreihe Nr. 64. Frankfurt/M.

MICHON, J.A. (1965). Studies of subjective duration. II Subjective time measurement during tasks with different information content. Acta psychologica, 24, 205-219.

MORAY, N. (1967). Where is capacity limited? A survey and a model. Acta Psychologica, 17, 84-92.

SPELKE, E., HIRST, W. & NEISSER, U. (1976). Skills of divided attention. Cognition, 4, 215-230.

WIERWILLE, W.W., GAGNE, G.A. & KNIGHT, J.R. (1967). An experimental study of human operator models and closed-loop analysis methods for high-speed automobile driving. IEEE Transactions on Human Factors in Electronics, Vol. HFE-8, No 3.

RESEARCH METHODOLOGY TO ASSESS THE IMPORTANCE OF PERIPHERAL VISUAL DETECTION AT NIGHT

Helmut T. Zwahlen

Department of Industrial and Systems Engineering
Ohio University, Athens, Ohio 45701-2979

ABSTRACT

Past investigations and experimental studies dealing with the visual detection of either non-reflectorized or reflectorized objects or targets in the driving environment at night have been primarily limited to foveal or line of sight visual detection. This study presents a research methodology which includes a geometric analysis of reflectorized targets located ahead of the car at different locations along a tangent-curve and curve-tangent section of a highway, typical driver eye scanning data, and demonstrates that in many cases unknown or unexpected reflectorized targets such as a reflectorized license plate or an advance warning sign will initially appear at moderately large peripheral angles up to 20 or more degrees away from a driver's foveal eye fixation point or line of sight. A methodology to obtain experimental data from a field study involving the foveal and peripheral detection of a reflectorized target is presented and has been used to demonstrate that the peripheral visual detection distances decrease considerably as the peripheral visual angle away from the fovea or line of sight

increases. A 10 degree peripheral visual detection angle results in an average visual detection distance which is approximately one half of the average foveal detection distance. Based upon the geometric analysis, the eye scanning data and the experimentally obtained foveal and peripheral visual detection distances, it is concluded that in a situation where drivers approach or negotiate a curve at night and where reflectorized objects or targets will become visible for the first time, most likely in the periphery of the visual field, appropriate increases in the reflectivity of the target should be made to assure early detection in order for timely recognition, information processing, decision making and appropriate control actions.

INTRODUCTION

Past investigations and experimental studies dealing with the visual detection of either non-reflectorized or reflectorized targets in the driving environment at night have been primarily limited to foveal or line of sight visual detection. One exception which may be found in the current literature is a field study by Zwahlen (1986) which investigated human subjects' ability to detect an approaching reflectorized target in the field foveally, as well as at peripheral visual angles of 10, 20 and 30 degrees at night. The experimental methodology used and the results of this study are discussed in more detail later in this paper, however, it should be noted that Zwahlen (1986) found that at a 10 degree peripheral angle the average detection distance was 47 to 59 percent of the average foveal detection distance and at a 30 degree peripheral angle this distance declined to 25 to 33 percent of the average foveal detection distance. Like the investigations of human detection capabilities in the driving environment, the past investigations of a driver's recognition capabilities have also been primarily limited to foveal or line of sight recognition of symbols or shapes of targets. One exception to this is a laboratory study by Karttunen and Hakkinen (1981) which investigated subjects' ability to recognize various symbolic road signs commonly used in Finland at peripheral angles of 10, 20, 30, 40, and 50 degrees. Karttunen and Hakkinen (1981) found that when signs, which subtended a visual angle of 4 degrees of visual arc (from bottom to top), were projected on a screen in a laboratory for 125 milliseconds, the subjects' ability to correctly identify such a commonly used road sign decreased from 100 percent for foveal presentation to 92.4 percent for a peripheral angle of 10 degrees and to 32.5 percent for a peripheral angle of 50 degrees.

Since Zwahlen (1986) showed that a driver's ability to detect a reflectorized target at night decreases considerably as the peripheral angle at which the target is first presented increases and Karttunen and Hakkinen (1981) showed similar results for peripheral recognition accuracy of commonly used symbolic road signs, it would seem that the use of data, based solely upon human foveal visual detection capabilities, in the design of reflectorized targets in the highway environment may be inadequate if such a target is likely to first appear in the periphery of a driver's visual field. Therefore, the objective of this paper is to present a methodology to assess the importance of peripheral visual detection in the driving environment at night.

METHODOLOGY TO ASSESS THE MAGNITUDE OF PERIPHERAL VISUAL DETECTION ANGLES IN THE HIGHWAY ENVIRONMENT

Knowledge of human peripheral visual detection capabilities would not be of great practical value in traffic safety if it were not possible to demonstrate that driving situations exist where targets are most likely to first appear at relatively large peripheral visual detection angles in a driver's visual field. Since highways are designed based upon geometric principles, it should be possible to develop a geometric model to accurately determine the peripheral visual detection angles which exist for particular driving situations and targets in the driving environment and to employ a computer to quickly analyze multiple situations as an alternative to conducting much more time consuming and costly field investigations for situations where large peripheral visual detection angles might be present.

A geometric model was developed which is based on the assumptions that: 1) the driver looks ahead of the car in a direction which is parallel to the longitudinal center of the car, 2) the driver is driving on a two-lane highway, 3) the driver is driving on a level and flat road surface as opposed to a road with vertical curves and 4) there are no physical barriers along the highway which might obstruct a driver's direct view of the target of interest. The analytical model was developed to evaluate situations which involved tangent sections combined with either left or right curves, where large peripheral visual detection angles were expected to occur. These situations include tangent-curve sections of a highway, where a driver is on a tangent section of a highway approaching a target which is located along the right edge of a curve section of the highway, and curve-tangent sections of a highway, where a driver is negotiating a curve while approaching a target which is located along the right edge of the tangent section on the highway beyond the end of the curve. Large peripheral visual detection angles may also occur when a driver is negotiating a long curve and approaching a target which is located along the right edge of the same curve ahead of the driver. This case was not separately analyzed since for certain short distances between the driver and target in the tangent-curve or curve-tangent sections of a highway both the driver and the target are very close to being in the same curve. Tangent-curve and curve-tangent sections occur frequently in the driving environment since each curve is preceded and succeeded by either a tangent section or another curve section of a highway and curves along highways are common, especially in locations where there are hills or when highways follow natural rivers. For example, according to Zwahlen (1983) in the state of Ohio there are over 18,000 curves along the two-lane rural state highways. Since the state of Ohio has a total of about 19,000 miles of highways of which about 1200 miles are interstate highways and several hundred additional miles of highway are four-lane highways, one can see that about one curve exists for every mile of two-lane rural state highway in Ohio.

To explain how one defines and calculates the peripheral visual detection angle and develops the computer model, a single tangent section of a highway will be used. Looking at Figure 1 one can see the variables which must be considered in such an analysis and their relation to the eye position of a driver. These variables include the lane width (LW), the lateral distance from the driver's sagittal plane to the center of the driver's lane (DSP), the lateral distance from the edge of the highway to the center of the target of interest (D1) and the longitudinal distance from the driver's eyes to the target of interest (S1). Based upon geometrical calculations, one can see that the euclidean distance from the driver to the target of interest (L) and the peripheral visual detection angle (ALPHA) can be calculated from the equations shown in Figure 1. However, peripheral visual detection angles calculated for targets along a tangent section of highway are normally relatively small when the target is viewed at distances of 400 to 1000 feet.

Based upon the approach and equations for calculating the peripheral visual detection angles for tangent sections of a highway as shown in Figure 1, similar equations can be developed for tangent-curve and curve-tangent sections of a

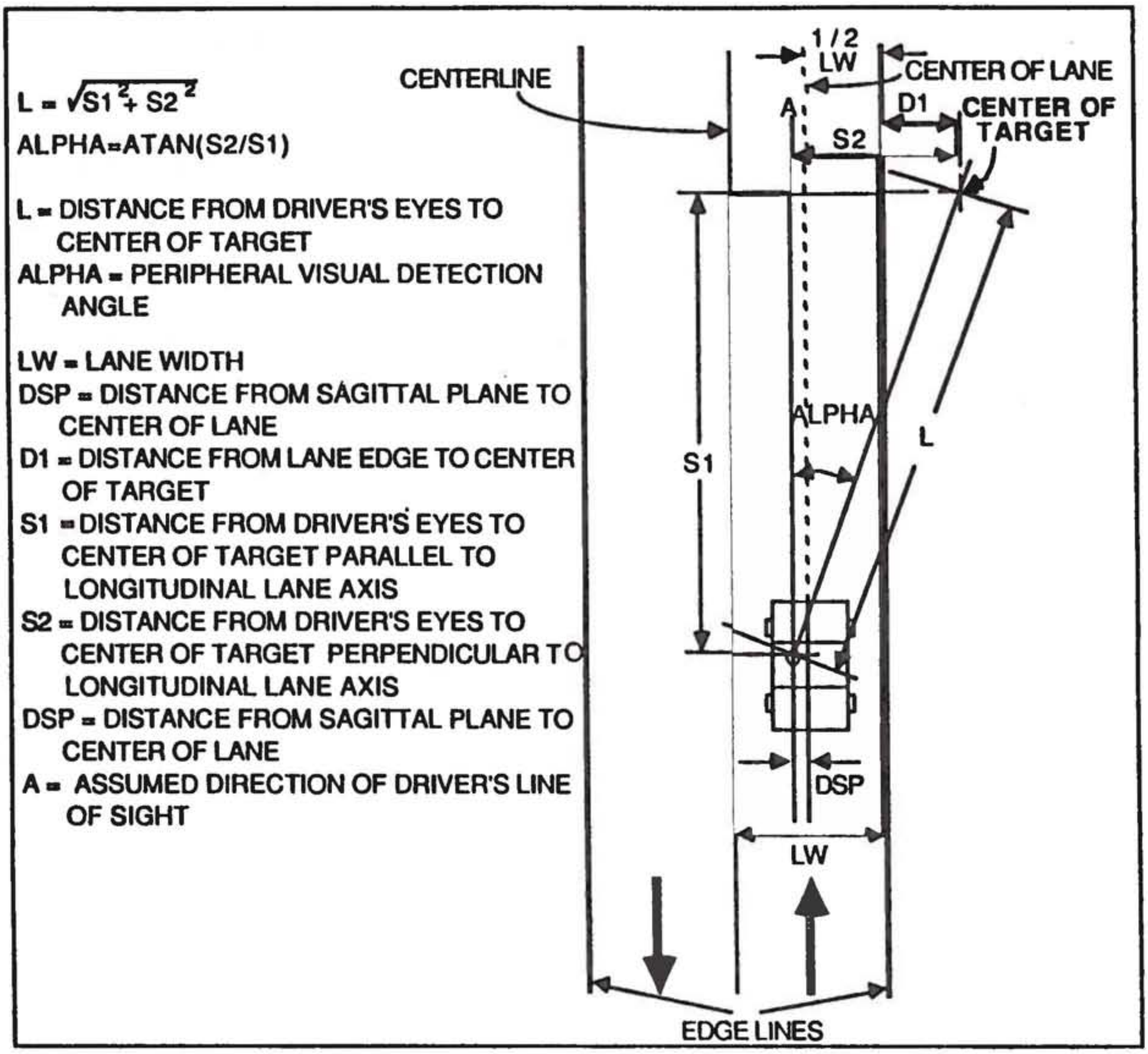
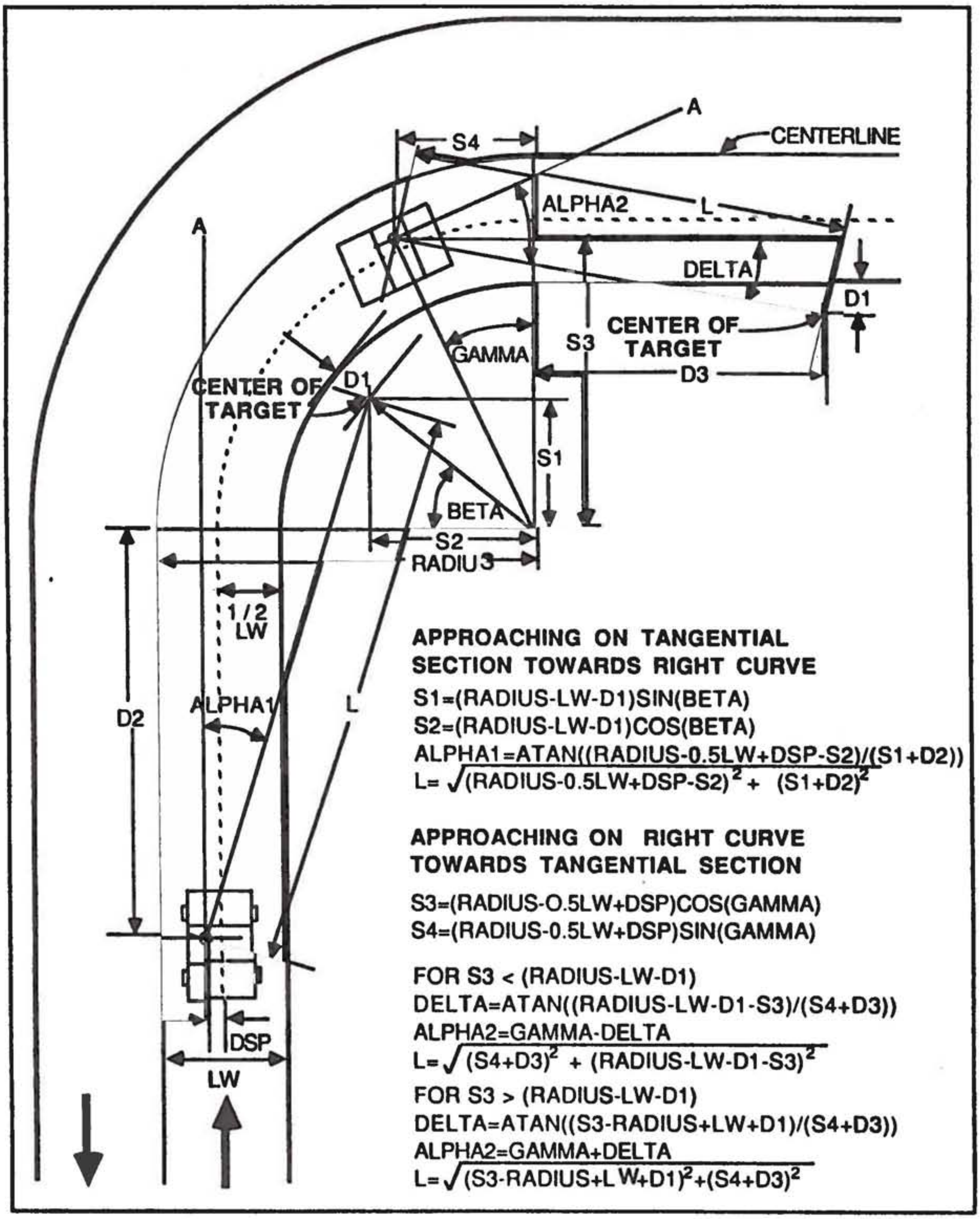


Figure 1 Geometric Configuration and Equations for Calculation of the Peripheral Visual Detection Angle for a Driver on a Tangent Section of a Two-Lane Rural Highway.

highway. As one can see from Figure 2, which shows the geometric conditions and equations for calculating the peripheral visual detection angles for tangent-curve and curve-tangent sections of a highway for right curves, in addition to the variables of lane width (LW), lateral distance from a driver's sagittal plane to the center of the highway (DSP) and the lateral distance from the right edge of the highway to the target of interest (D1), which were considered for the tangent section, one must also consider the radius of the curve (RADIUS), the distance from the driver's eyes to the beginning of the curve (D2) and the angle subtended by the radial line at the beginning of the curve and the radial line to the target of interest which will be called the curve position angle (BETA). It should be noted that the equations have been established assuming that the curve radius is measured from the center or the centerline of the two-lane highway.

Similarly, the peripheral visual detection angle can be calculated for the curve-tangent section of a highway, as shown in Figure 2. From Figure 2 one can see that the variables include the lane width (LW), the lateral distance from a driver's sagittal plane to the center of the highway (DSP), the lateral distance from the edge of the highway to the target of interest (D1), the radius of the curve (RADIUS), the rectilinear distance parallel to the direction of the tangent section of the highway from the end of the curve to the target of interest (D3), and the angle subtended by the radial line at the end of the curve and the



APPROACHING ON TANGENTIAL SECTION TOWARDS RIGHT CURVE

$$S1=(RADIUS-LW-D1)SIN(BETA)$$

$$S2=(RADIUS-LW-D1)COS(BETA)$$

$$ALPHA1=ATAN((RADIUS-0.5LW+DSP-S2)/(S1+D2))$$

$$L=\sqrt{(RADIUS-0.5LW+DSP-S2)^2+(S1+D2)^2}$$

APPROACHING ON RIGHT CURVE TOWARDS TANGENTIAL SECTION

$$S3=(RADIUS-O.5LW+DSP)COS(GAMMA)$$

$$S4=(RADIUS-0.5LW+DSP)SIN(GAMMA)$$

FOR $S3 < (RADIUS-LW-D1)$

$$DELTA=ATAN((RADIUS-LW-D1-S3)/(S4+D3))$$

$$ALPHA2=GAMMA-DELTA$$

$$L=\sqrt{(S4+D3)^2+(RADIUS-LW-D1-S3)^2}$$

FOR $S3 > (RADIUS-LW-D1)$

$$DELTA=ATAN((S3-RADIUS+LW+D1)/(S4+D3))$$

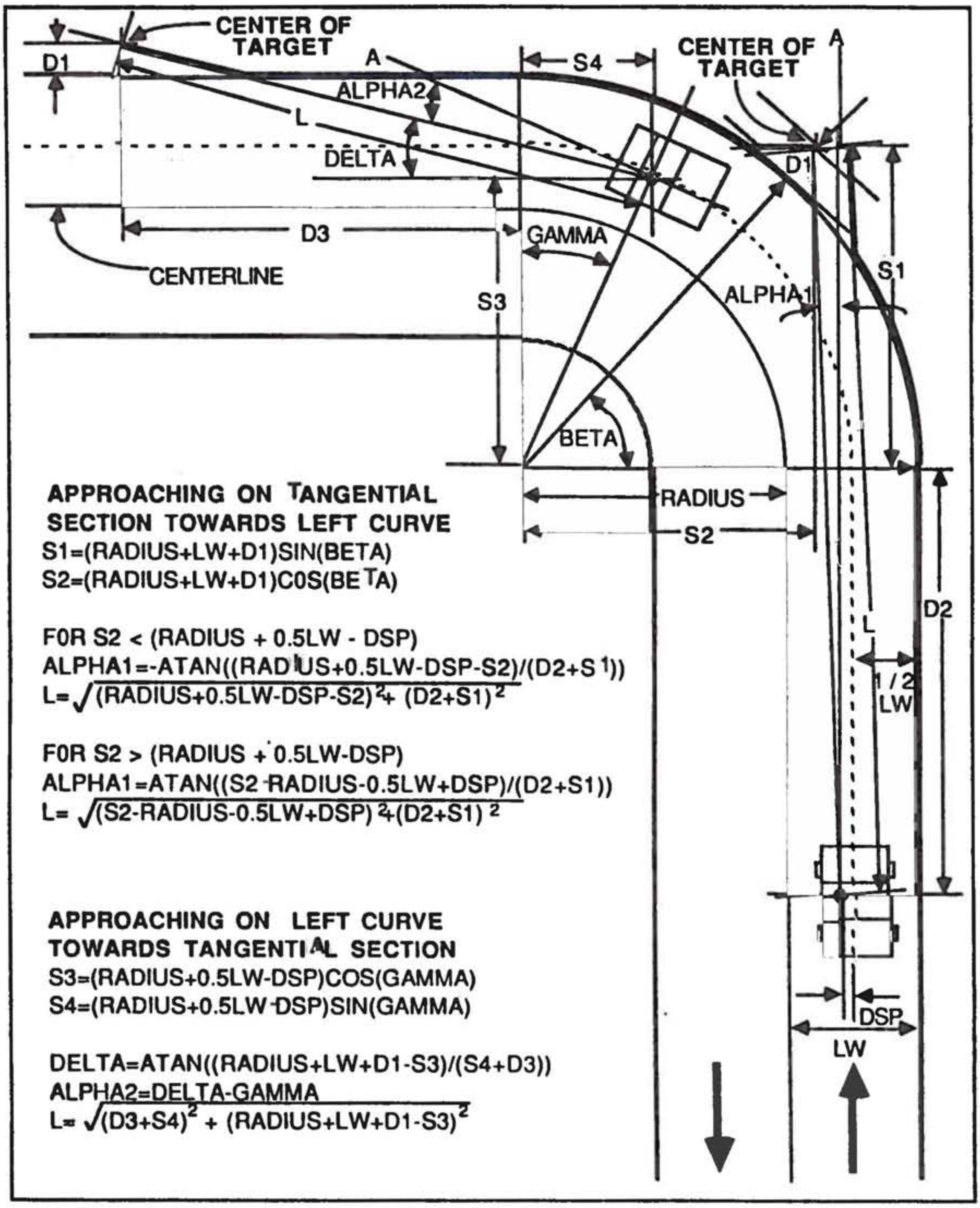
$$ALPHA2=GAMMA+DELTA$$

$$L=\sqrt{(S3-RADIUS+LW+D1)^2+(S4+D3)^2}$$

Figure 2 Geometric Configuration and Equations for Calculation of the Peripheral Visual Detection Angles for Tangent-Curve and Curve-Tangent Sections for a Right Curve of a Two-Lane Rural Highway.

radial line to the driver's eyes which will be called the curve completion angle (GAMMA). For the curve-tangent section for a right curve two equations had to be developed in order to calculate ALPHA2. The appropriate equation may be chosen by first calculating the distance S3 as shown in Figure 2. This distance may then be used to determine the appropriate equation for calculating the peripheral visual detection angle (ALPHA2). If one wishes to calculate peripheral visual detection angles to aid in the determination of what ranges of magnitudes could be considered as common in the driving environment, BETA and GAMMA should probably not exceed 40 degrees due to the optical properties of most retroreflective materials and the reduced projected areas of the targets.

Figure 3 shows the geometry and equations which can be used to calculate the peripheral visual detection angles (ALPHA1 and ALPHA2) and distances (L) for the tangent-curve and curve-tangent sections for a left curve. When calculating the peripheral visual detection angle (ALPHA1) and distance (L) for the tangent-curve section it was necessary to develop two equations from which one equation must be chosen based upon the position of the target in the curve. The distance S2 must be calculated before the appropriate set of equations can be chosen when investigating the tangent-curve section for left curves and the appropriate formula for the peripheral visual detection angle must be chosen based upon the magnitude of S2 as shown in Figure 3. It should be noted that calculated peripheral visual detection angles to



APPROACHING ON TANGENTIAL SECTION TOWARDS LEFT CURVE

$S1=(\mathrm{RADIUS}+LW+D1)\,\mathrm{SIN}(\mathrm{BETA})$

$S2=(\mathrm{RADIUS}+LW+D1)\,\mathrm{COS}(\mathrm{BETA})$

FOR $S2 < (\mathrm{RADIUS} + 0.5LW - DSP)$

$\mathrm{ALPHA1}=-\mathrm{ATAN}((\mathrm{RADIUS}+0.5LW-DSP-S2)/(D2+S1))$

$L=\sqrt{(\mathrm{RADIUS}+0.5LW-DSP-S2)^2+(D2+S1)^2}$

FOR $S2 > (\mathrm{RADIUS} + 0.5LW-DSP)$

$\mathrm{ALPHA1}=\mathrm{ATAN}((S2-\mathrm{RADIUS}-0.5LW+DSP)/(D2+S1))$

$L=\sqrt{(S2-\mathrm{RADIUS}-0.5LW+DSP)^2+(D2+S1)^2}$

APPROACHING ON LEFT CURVE TOWARDS TANGENTIAL SECTION

$S3=(\mathrm{RADIUS}+0.5LW-DSP)\,\mathrm{COS}(\mathrm{GAMMA})$

$S4=(\mathrm{RADIUS}+0.5LW-DSP)\,\mathrm{SIN}(\mathrm{GAMMA})$

$\mathrm{DELTA}=\mathrm{ATAN}((\mathrm{RADIUS}+LW+D1-S3)/(S4+D3))$

$\mathrm{ALPHA2}=\mathrm{DELTA}-\mathrm{GAMMA}$

$L=\sqrt{(D3+S4)^2+(\mathrm{RADIUS}+LW+D1-S3)^2}$

Figure 3 Geometric Configuration and Equations for Calculation of the Peripheral Visual Detection Angles for Tangent-Curve and Curve-Tangent Sections for a Left Curve on a Two-Lane Rural Highway.

the left of the driver's sagittal plane are represented by negative values, while calculated peripheral visual detection angles to the right of the driver's sagittal plane are represented by positive values. Figure 3 also shows an equation which may be used to calculate the peripheral visual detection angle (ALPHA2) for the curve-tangent section of a left curve. It should be noted that all peripheral visual detection angles for tangent-curve and curve-tangent sections for right curves are measured to the right of the driver's sagittal plane and all angles for tangent-curve and curve-tangent sections for left curves are measured to the left of the driver's sagittal plane except when S2>(RADIUS+0.5LW-DSP) for the tangent-curve section for left curves, where the peripheral visual detection angle is measured to the right of the driver's sagittal plane. A common spreadsheet package (Microsoft Excel) and a graphics package (Cricket Graph) for the Macintosh computer were utilized to perform the calculations and to graphically display the results for the selected combinations of the variables present in the model. The use of a spreadsheet package combined with a compatible graphics package enables one to calculate the peripheral visual detection angles and to display them as a function of the distance from the driver's eyes to the reflectorized target graphically in a quick and efficient manner without developing special software for this particular application.

In order to make this small set of calculations representative of conditions which might exist in the highway environment it was necessary to use conditions which commonly occur in the normal driving environment. According to Zwahlen (1983) there are 18,093 curves on two-lane state highways in Ohio. It was found that the vast majority of these curves have a curvature of between 3 and 28 degrees with an average of 12 degrees of curvature. Therefore, in performing these representative calculations the effect of curves with a curvature of 3, 12 and 28 degrees (radii of 1906, 477 and 204 feet respectively) upon the peripheral visual detection angles were investigated.

Two lateral offset values on the right hand side of the driving lane were chosen to represent two typical reflectorized targets which might appear in a driver's peripheral visual field. These targets include a reflectorized license plate of a disabled or abandoned vehicle and a reflectorized roadside warning sign. Once again, to achieve a representative sample the analysis assumed that the disabled vehicle would be situated such that the longitudinal center of the vehicle and the reflectorized license plate would be positioned above the right edge line of the highway. It might be noted that the same conditions would be applicable to delineation devices such as raised reflective pavement markers which could be located along the right edge line of the highway. It was further assumed that the reflectorized highway warning sign was positioned 12 feet to the right of the edge line (measured from the edge line to the inside edge of the sign) as specified by the Manual on Uniform Traffic Control Devices (1978). Therefore, the center of a 24 x 24 inches roadside warning sign would be 1.4 feet to the right of this mark and the distance measured from the edge of the highway to the center of the sign would be 13.4 feet. In order to further reduce the number of calculations it was assumed that: 1) the driver is driving in the center of his or her lane, 2) the driver is driving in a 12 feet wide lane and 3) the driver's sagittal plane is located 1.25 feet to the left of the vehicle's longitudinal center.

Figures 4 and 5 show the peripheral visual detection angle (ALPHA) as a function of the distance from the driver's eyes to the reflectorized target (L), the curve position angle (BETA), the radius of the curve (RAD), and the horizontal distance from the edge of the road to the target of interest (D1) for the tangent-curve conditions for right and left curves. Figures 4 and 5 show that as a driver gets closer to the target the peripheral visual detection angles increase for the tangent-curve sections of highway. From Figure 4 one can see that for the right curves when the distance from a driver's eyes to the target is within a range of 400 to 1000 feet the peripheral visual detection angles range from about 7 to about 20 degrees for curves with a radius of 477 feet (12 degrees of curvature) with a curve position angle of 40 degrees and for curves with a radius of 1906 feet (3 degrees of curvature) with a curve position angle of 20 degrees. There is very little difference in the peripheral visual detection angles whether or not the target is located on the right edge line or 13.4 feet to the right of the right edge line. Within a range of

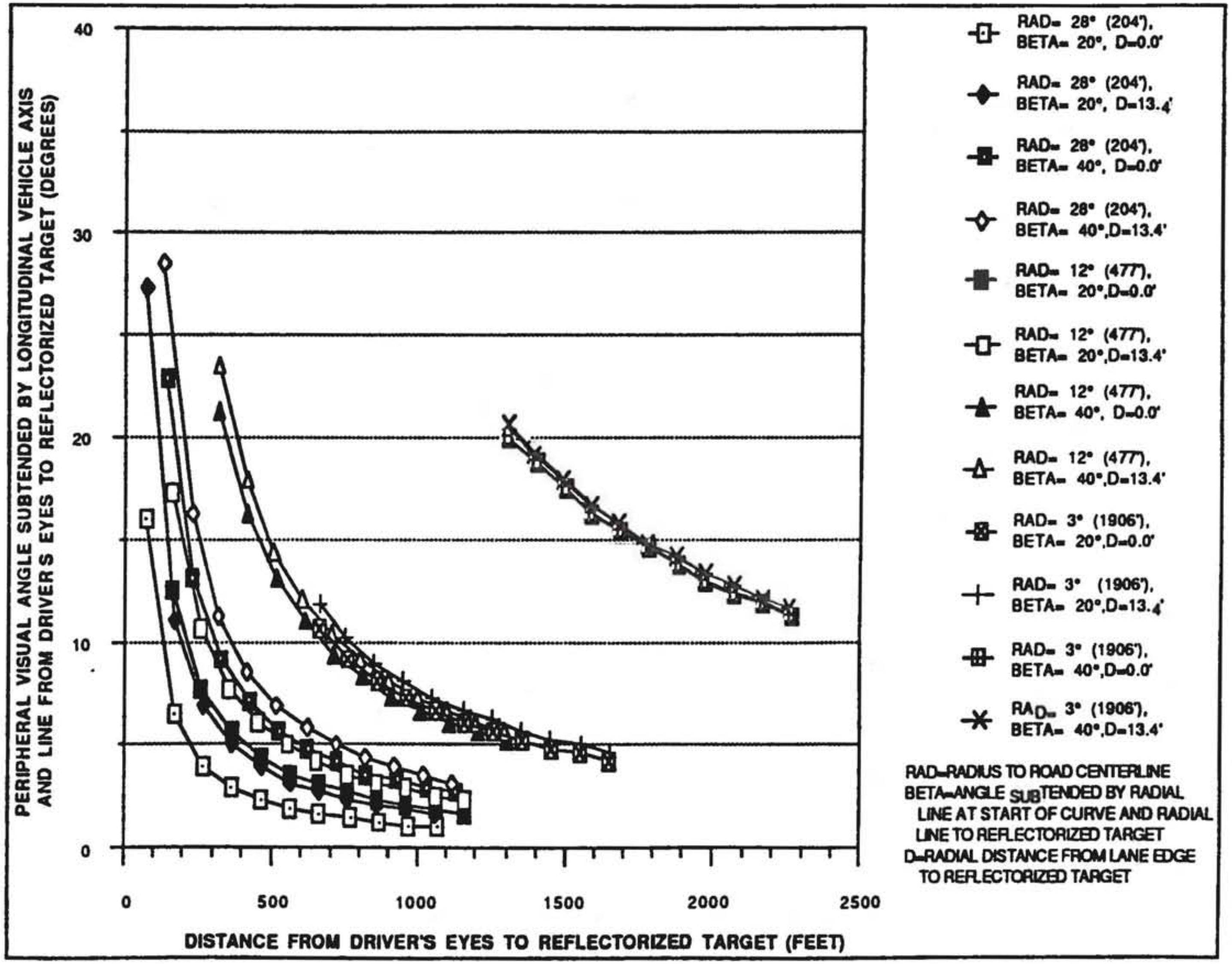


Figure 4 Peripheral Detection Angles for the Tangent-Curve Section for a Right Curve of a Highway for Various Distances, Curve Radii, Curve Position Angles and Lateral Distances.

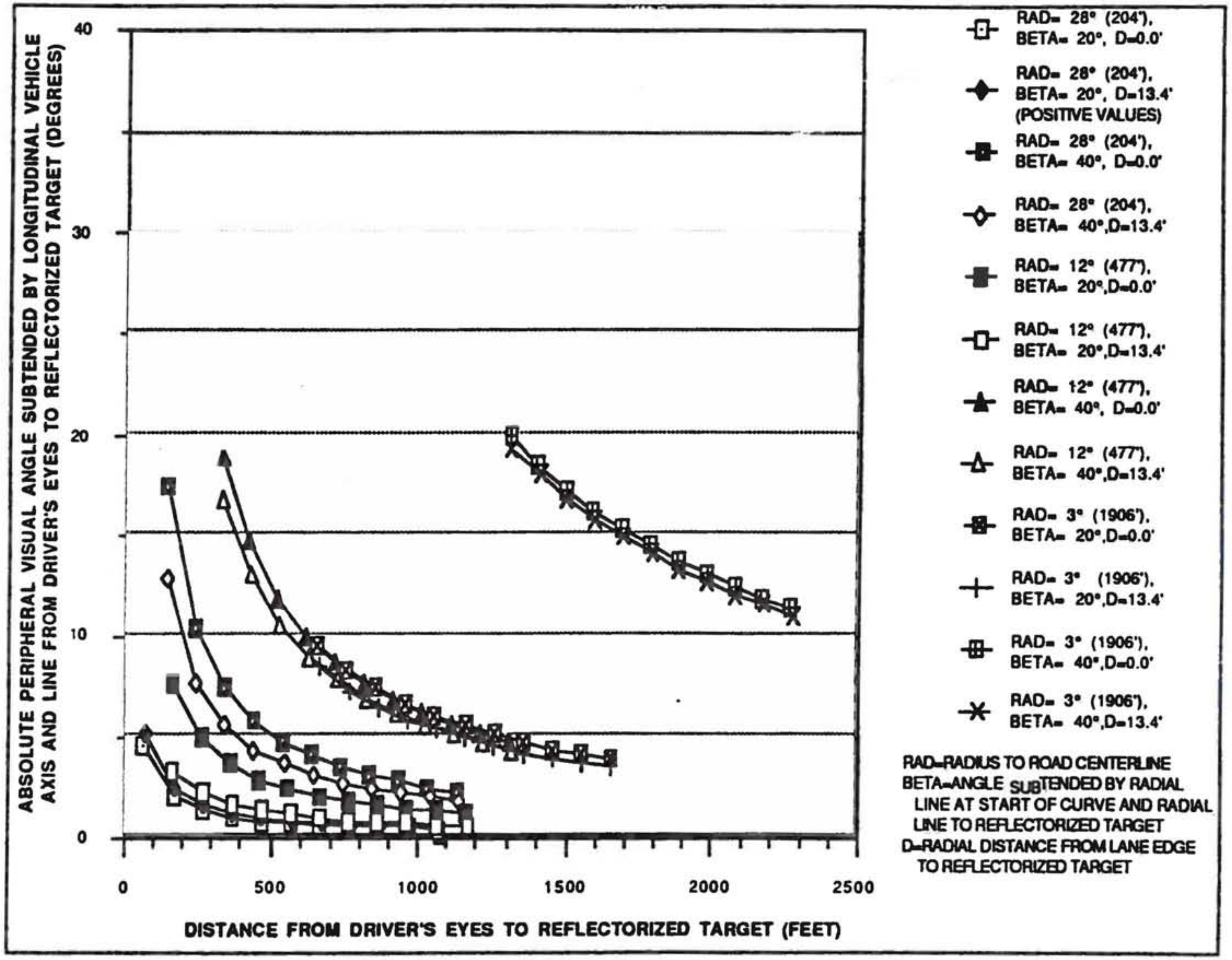


Figure 5 Peripheral Detection Angles for the Tangent-Curve Section for a Left Curve of a Highway for Various Distances, Curve Radii, Curve Position Angles and Lateral Distances.

1300 to 2200 feet the peripheral visual detection angles range from about 12 to about 21 degrees for curves with a radius of 1906 feet (3 degrees of curvature) and a curve position angle (BETA) of 40 degrees, however, it should be noted that these distances are rather large and therefore most likely not that relevant for this investigation. All other investigated conditions for tangent-curve sections of right curves have peripheral visual detection angles of 1 to 9 degrees at distances in the range of 400 to 1000 feet.

Figure 5 shows absolute peripheral visual detection angles since the values for a curve radius of 204 feet (28 degrees of curvature) with a curve position angle (BETA) of 20 degrees and a distance from the right edge of the road to the target of 13.4 is positive while the values for all other conditions are negative. From Figure 5 one can see that for the tangent-curve sections for left curves, peripheral visual detection angles of about -6 to -17 degrees exist within a range of 400 and 1000 feet for a radius of 477 feet (12 degrees curvature) with a curve position angle (BETA) of 40 degrees and for a radius of 1906 feet (3 degrees curvature) with a curve position angle (BETA) of 20 degrees. Again, there is very little difference in the peripheral visual detection angles whether or not the target is located on the right edge line or 13.4 feet to the right of the right edge line. Within a range of 1300 feet to 2300 feet the peripheral visual detection angles range from about -11 to about -20 degrees for a curve radius of 1906 feet (3 degrees of curvature) and a curve position angle (BETA) of

40 degrees, however, it should be noted that these distances are rather large and therefore most likely not relevant for this investigation. Further Figure 5 shows that for the distance range of 400 to 1000 feet the peripheral visual detection angles for all other conditions range from about -1 to about -7 degrees.

Figures 6 and 7 show peripheral visual angles as a function of the distance from the driver's eyes to the target of interest, the radius of the curve (RAD), the curve completion angle (GAMMA) and the horizontal distance from the edge of the highway to the target of interest (D1), for the curve-tangent condition for a right and a left curves respectively. From Figure 6 one can see that in general the peripheral visual detection angles decrease as the distance from the driver's eyes to the target decreases for the investigated conditions. The one exception occurs for a curve radius of 204 feet (28 degrees of curvature) with a curve completion angle (GAMMA) of 20 degrees when the target is located 13.4 feet to the right of the right edge of the highway. For this condition the peripheral visual detection angle increases as the distance from the driver's eyes to the target decreases. Looking at Figure 6 one can see that for a distance range of 400 to 1000 feet peripheral visual detection angles of about 34 to 38 degrees were observed for a curve radius of 204 feet (28 degrees of curvature) with a curve completion angle of 40 degrees (for 0 and 13.4 feet to the right of edge line target positions). In this same distance range, peripheral visual detection angles of about 24

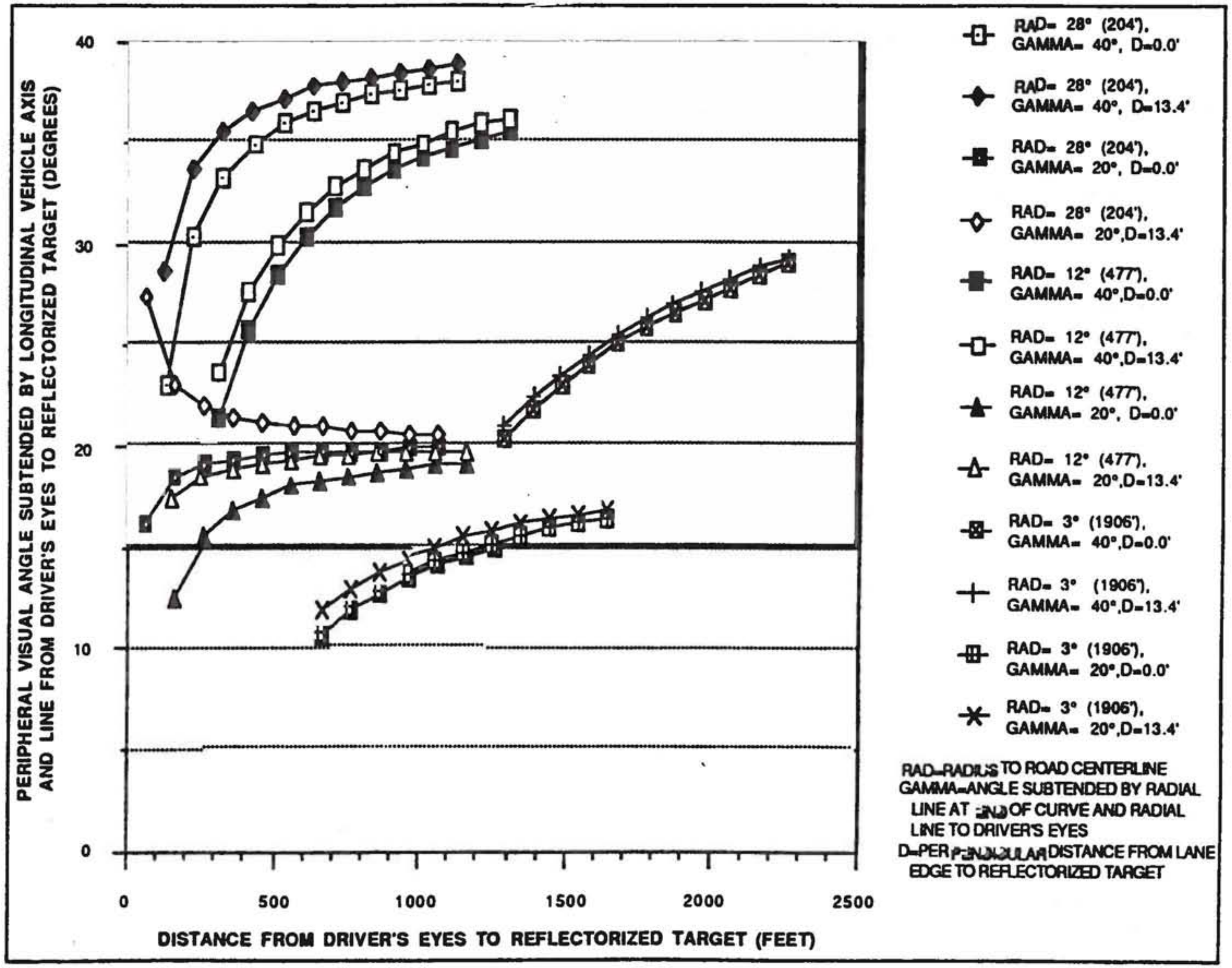


Figure 6 Peripheral Detection Angles for the Curve-Tangent Section for a Right Curve of a Highway, for Various Distances, Curve Radii, Curve Position Angles and Lateral Distances.

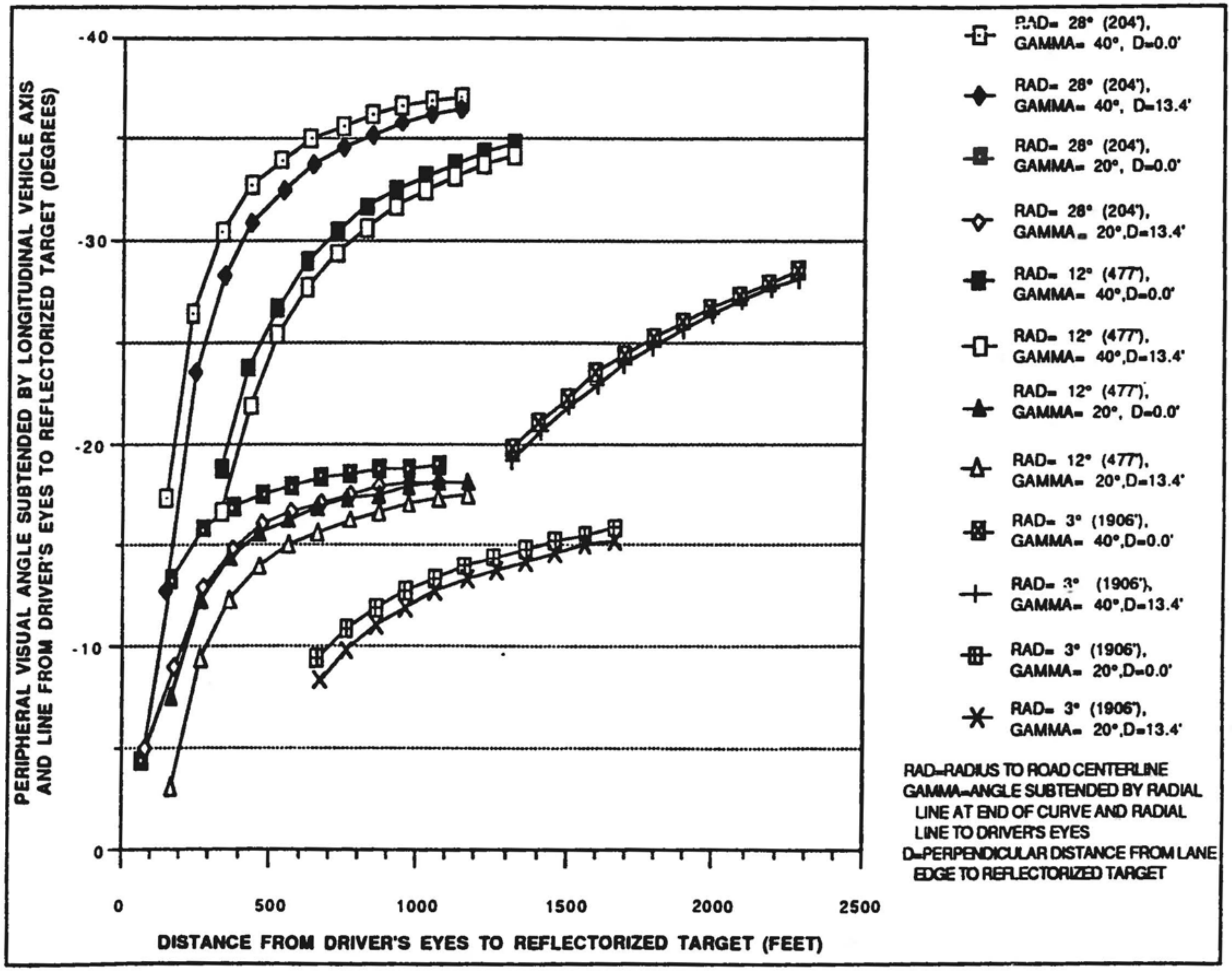


Figure 7 Peripheral Detection Angles for the Curve-Tangent Section for a Left Curve of a Highway, for Various Distances, Curve Radii, Curve Position Angles and Lateral Distances.

to 35 degrees were observed for a curve radius of 477 feet (12 degrees of curvature) with a curve completion angle of 40 degrees. Peripheral visual detection angles of about 18 to 21 degrees were observed for curve radii of 204 and 477 feet (28 and 12 degrees of curvature respectively) with a curve completion angle of 20 degrees for the 400 to 1000 feet range.

From Figure 7 one can see that the selected conditions for the curve-tangent section of a left curve of a highway all peripheral visual detection angles decrease in magnitude as the distance from the driver's eyes to the target decreases. Figure 7 also shows that between 400 and 1000 feet peripheral visual detection angles of -30 to -37 degrees are obtained for a curve radius of 204 feet (28 degrees of curvature) with a curve completion angle (GAMMA) of 40 degrees. In this same distance range peripheral visual detection angles of about -20 to -33 degrees were obtained for a curve radius of 477 feet (12 degrees of curvature) with a curve completion angle (GAMMA) of 40 degrees. Peripheral visual detection angles of about -13 to -19 degrees are obtained for curve radii of 204 and 477 feet (12 and 28 degrees of curvature respectively) with a curve completion angle (GAMMA) of 20 degrees within the 400 to 1000 feet range. It may also be observed that for distances of between 1300 and 2300 feet, peripheral visual detection angles of about -19 to -28 degrees may be observed for a curve radius of 1906 feet (3 degrees of curvature) with a curve completion angle (GAMMA) of 20 degrees, however, again it should be noted that these distances are too large and probably not that relevant for this investigation.

From the data presented in Figures 4 through 7 it would appear that relatively large peripheral visual detection angles may exist for targets which are located along or just beyond a curve. However, reviewing the assumptions which were made in developing the geometric model it should be noted that it was assumed that a driver's direction of his or her foveal fixation or line of sight is along a line which is parallel to the longitudinal center axis of the car. This assumption may not be valid since a driver fixates upon various targets located ahead of the car in the driving environment. Therefore, it might be necessary to adjust the obtained calculated peripheral visual detection angles according to the experimentally obtained spatial driver eye fixation densities. It should also be noted that only flat and level highways with horizontal curves were considered and vertical curves or combinations of horizontal and vertical curves, which could further increase the magnitude of the peripheral visual detection angles, were not considered.

In a prior study Zwahlen (1985) investigated driver eye scanning behavior as the subjects drove on four unlighted 1 mile long tangent sections of a four-lane interstate highway (Interstate 70 between Ohio State Routes 37 and 79) with a lane width of 12 feet and on four unlighted right 240 feet radius clover leaf type entrance/exit ramps (at the intersection of Interstate 70 and Ohio State Route 79) with a lane width of 16 feet at night under dry and light rain conditions. Eleven young licensed test drivers who were in good health, had about 20/20 uncorrected vision and were paid participated in this night driving study. The eye scanning behavior of eight of these test drivers was recorded during light rain or while the pavement was wet and the eye scanning behavior of four of the test drivers was recorded while they drove on dry pavement (one of the test drivers was tested for both conditions).

During the experiment the subjects drove an instrumented car (VW 412, automatic transmission, type 4000 low beams) which was equipped with an in car television eye scanning recording system and other electronic equipment. This equipment allowed the experimenter to monitor and record a driver's eye movements and the driving scene ahead with a maximum visual field of 18 by 18 degrees, as well as, a number of vehicle measures. For a more detailed description of the experimental apparatus see Zwahlen (1983). During the experiment each driver served as his or her own control and was asked to follow a selected route such that every subject drove the ramps and tangent sections in the same order two times. In order for the drivers to complete the loop it was necessary for them to drive two of the four ramps twice in each of the two experimental loops.

The results of this study show that there are no statistically significant differences in the eye scanning measures between the dry and light rain conditions. Therefore the data for these two conditions were combined to give larger sample sizes. Figure 8 shows the relative number of eye

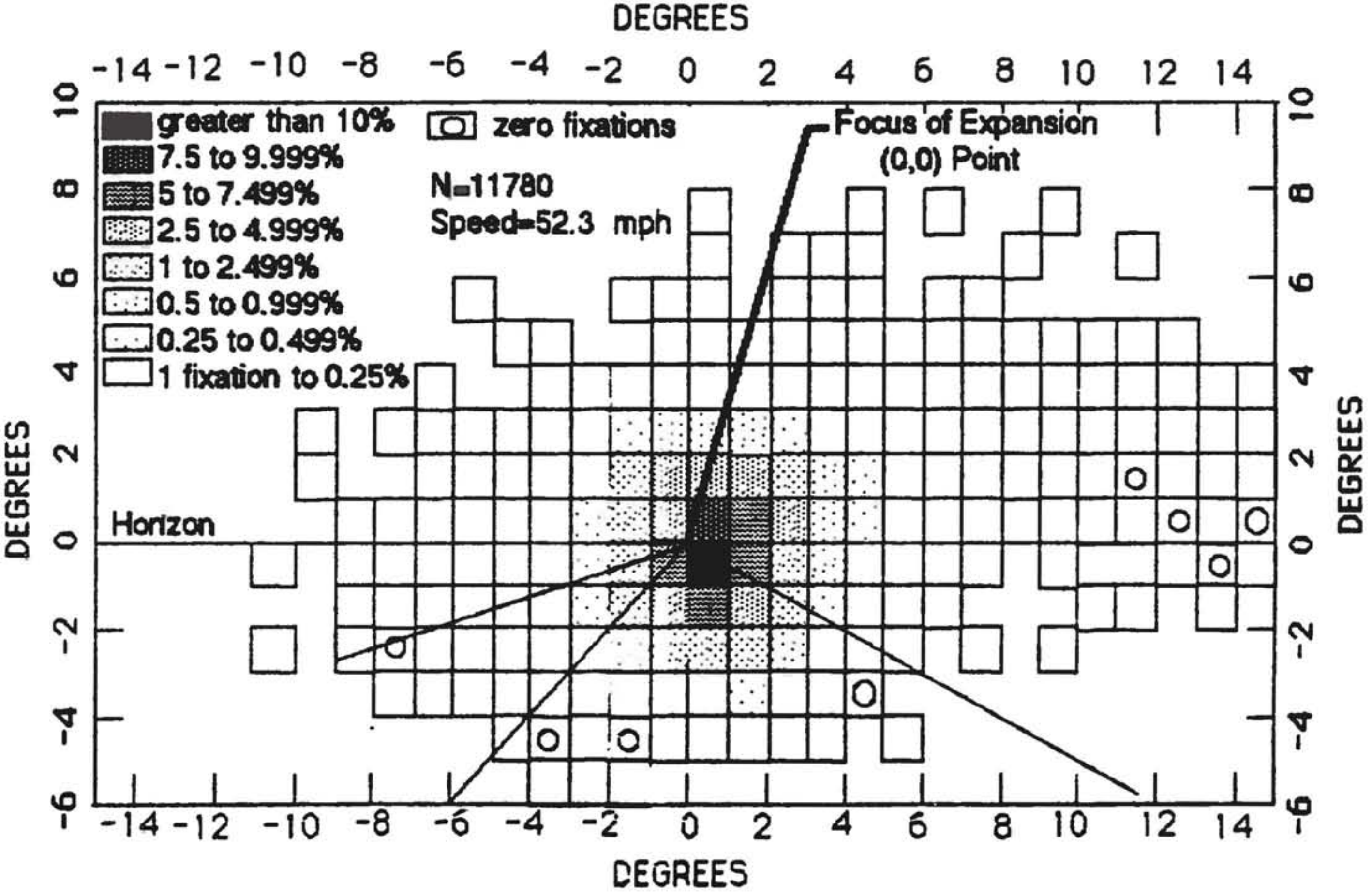


Figure 8 Eye Fixation Pattern for Tangent Sections of a Four-Lane Interstate Highway at Night.

fixations which occur in each one degree by one degree block within the viewing area for the tangent sections of a four-lane highway. From Figure 8 one can see that more fixations are focused in the one degree by one degree block which is centered 0.5 degrees to the right of the focus of expansion and 0.5 degrees below the horizon and focus of expansion than in any other block. In fact, this block contains 13.5 percent of all the eye fixations (Total N=11780 fixations) which were made by the test drivers. It should also be noted that the average of the horizontal eye fixation distribution for the sample size of 11780 eye fixations is about 0.84 degrees to the right of the focus of expansion with a standard deviation of 1.92 degrees.

Figure 9 shows the relative number of eye fixations which occur in each one degree by one degree block within the viewing area for the 240 feet radius right curves (16 feet lane width). Comparing Figure 9 with Figure 8 for the tangent sections one can see that the eye fixations are much more dispersed in the spatial distribution of the eye fixations for the curves. In fact, the block centered 5.5 degrees right of the imaginary focus of expansion and 0.5 degrees below the horizon or imaginary focus of expansion, which contains the most fixations, contains only 3.9 percent of the fixations made by the drivers as they negotiated the 240 feet radius right curves. Again, it may be noted that the average of the horizontal eye fixation distribution for a sample size of 8884 eye fixations is about 3.64 degrees right of the imaginary focus of expansion with a standard

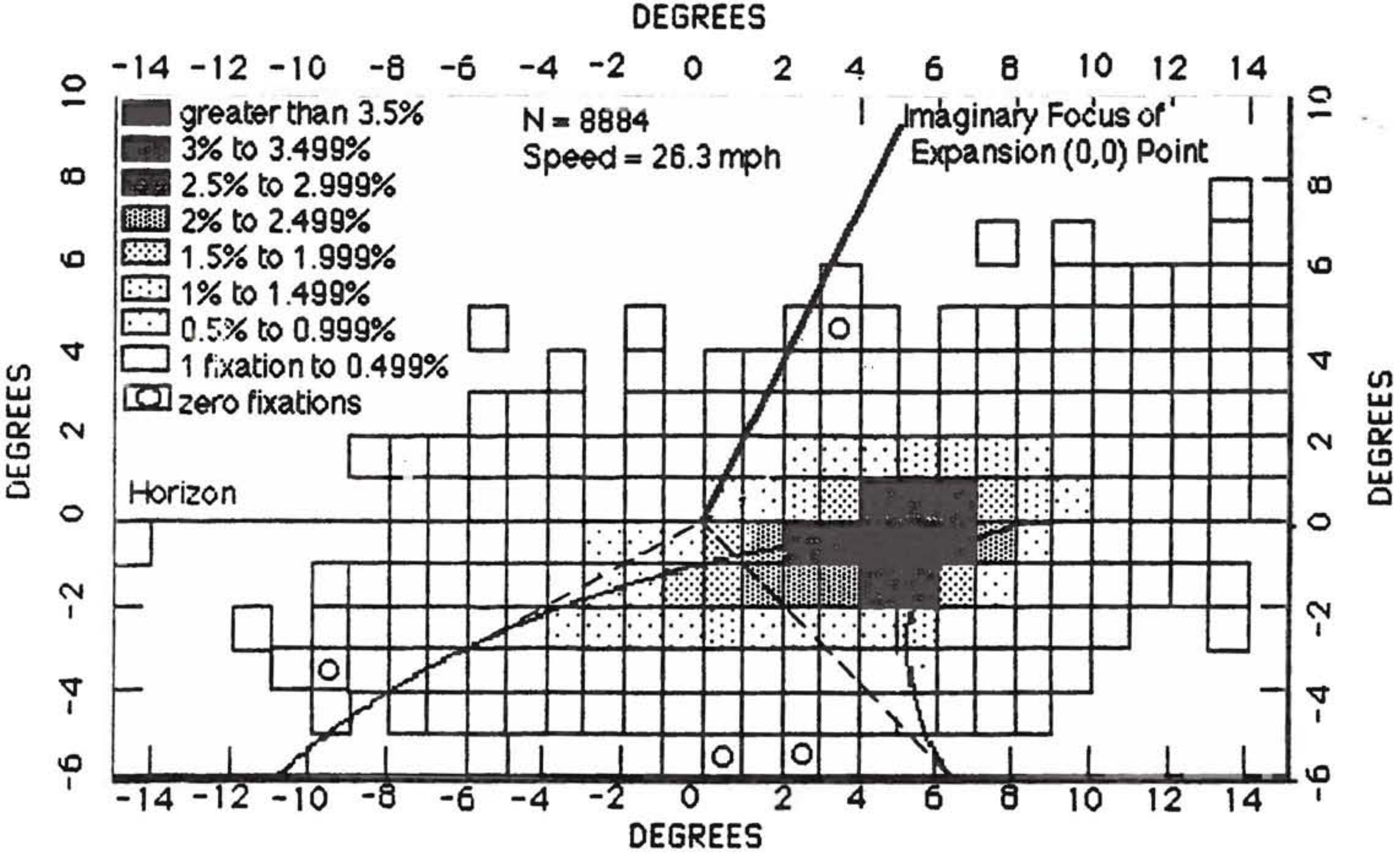


Figure 9 Eye Fixation Pattern for 240 Feet Radius Right Curved Entrance/Exit Ramps at Night.

deviation of 3.79 degrees.

Using the same experimental apparatus an exploratory study conducted at Ohio University investigated driver eye scanning behavior of two subjects on tangent and left and right curved sections (6 to 10 degrees of curvature, 954 to 572 feet radii) of a two-lane rural highway during daytime conditions and for one subject during nighttime conditions. The results of this study indicated that for six 800 feet sections of straight highway during the day the average of the horizontal eye fixation distribution was 0.81 degrees right of the focus of expansion with a standard deviation of 3.90 degrees (N=213 fixations) and for seven 800 feet sections of straight highway at night the average of the horizontal eye fixation distribution was 0.17 degrees left of the focus of expansion with a standard deviation of 2.63 degrees (N=46 fixations). The average of the horizontal eye fixation distribution was 2.16 degrees left of the imaginary focus of expansion for four left curves during the day with a standard deviation of 3.82 degrees (N=114 fixations). For right curves during the day the average of the horizontal eye fixation distribution was 4.59 degrees right of the imaginary focus of expansion with a standard deviation of 4.39 degrees (N=115 fixations). For seven left curves at night the average of the horizontal eye fixation distribution was 2.69 degrees left of the imaginary focus of expansion with a standard deviation of 4.06 degrees (N=46 fixations) and for seven right curves at night the average of the horizontal eye fixation distribution was 1.38 degrees right of the imaginary focus of expansion with a standard deviation of 3.59 degrees (N=25 fixations).

Shinar, McDowell and Rockwell (1977) used five drivers to investigate driver eye scanning behavior as they negotiated 22 curves with a central curvature of 5 to 19 degrees on two-lane hilly rural highways during the day. The results of this study show that the average of the horizontal eye fixation distribution was 1.6 degrees right of the focus of expansion for straight sections, 3.6 degrees right of the imaginary focus of expansion for right curves and 0.3 degrees left of the imaginary focus of expansion for left curves.

Since a driver's eye scanning behavior consists of a continuous string of discrete eye fixations there is no way of predicting exactly where a driver will look at any instant in time. There may be a very remote possibility that a driver will, by chance, look directly at an appearing target, however, looking at the spatial distribution of eye fixations it is very unlikely that this will occur, especially for a target that is located some 10 or 20 degrees away from the focus of expansion. Figure 8 indicates that on tangent sections more than 80 percent of all eye fixations are within a relatively small rectangle extending from -2 degrees left to 3 degrees right of the focus of expansion and from 2 degrees below to 2 degrees above the focus of expansion or an area of 20 degrees squared. Although this would indicate that a large number of a driver's eye fixations are concentrated fairly close around the focus of expansion it is still not reasonable to assume that a driver will be fixating in the close vicinity of the focus of expansion every time a target appears in the driving environment. In order to get some idea about how much the calculated peripheral visual angles should be adjusted to account for the observed horizontal eye fixation distributions the average of the horizontal eye fixation distribution and the average of the horizontal eye fixation distribution plus and minus one standard deviation were selected as representative horizontal eye fixation directions for the instant when a target becomes first visible in a driver's field of view. Table 1 provides information about peripheral visual detection angles calculated for a target distance of 500 feet and adjusted values based upon the driver eye scanning results for a few selected curve-tangent and tangent-curve situations of a two-lane rural highway.

Table 1 was developed for a target distance (the euclidean distance from the driver's eyes to the target of interest) of 500 feet since this distance would roughly correspond to the average peripheral visual detection distance minus one standard deviation for a 10 degree peripheral detection angle and near maximum low beam output (-3 degrees car heading angle) as it was presented by Zwahlen (1986). When adjusting the calculated peripheral angles for all tangent-curve sections an average foveal eye fixation position of .84 degrees to the right of the focus of expansion and a standard deviation of 1.92 degrees were used which are equivalent to the average and the standard deviation of the hor-

Table 1 Calculated and Adjusted Peripheral Visual Angles for Tangent-Curve and Curve-Tangent Sections for Left and Right Curves for a 500 Feet Target Viewing Distance.

CAR ON	CURVE DIRECT	CURVE RAD	BETA/ GAMMA	TARG POS. (D1)	CAL PER ANG	ADJ FOR X	ADJ FOR X-S	ADJ FOR X+S
TANGENT	RIGHT	12	40	0.0	13.5	12.7	14.6	10.7
TANGENT	RIGHT	12	40	13.4	14.7	13.9	15.8	11.9
TANGENT	RIGHT	28	40	0.0	6.0	5.2	7.1	3.2
TANGENT	RIGHT	28	40	13.4	7.2	6.4	8.3	4.4
CURVE	RIGHT	12	20	0.0	17.5	13.9	17.7	10.1
CURVE	RIGHT	12	20	13.4	19.1	15.5	19.3	11.7
CURVE	RIGHT	28	20	0.0	19.4	15.8	19.6	12.0
CURVE	RIGHT	28	20	13.4	21.0	17.4	21.2	13.6
TANGENT	LEFT	12	40	0.0	-12.5	-13.3	-11.4	-15.3
TANGENT	LEFT	12	40	13.4	-11.3	-12.1	-10.2	-14.1
TANGENT	LEFT	28	40	0.0	-5.0	-5.8	-3.9	-7.8
TANGENT	LEFT	28	40	13.4	-3.8	-4.6	-2.7	-6.6
CURVE	LEFT	12	20	0.0	-15.8	-13.1	-9.0	-17.2
CURVE	LEFT	12	20	13.4	-14.2	-11.5	-7.4	-15.6
CURVE	LEFT	28	20	0.0	-17.7	-15.0	-10.9	-19.1
CURVE	LEFT	28	20	13.4	-16.2	-13.5	-9.4	-17.6

CURVE DIRECT - CURVE DIRECTION
CURVE RAD - CURVE RADIUS (DEGREES)
BETA/GAMMA - CURVE POSITION OR COMPLETION ANGLE (DEGREES)
TARG POS. - TARGET POSITION FROM RIGHT EDGE OF HIGHWAY (FEET)
CAL PER ANG - CALCULATED PERIPHERAL VISUAL DETECTION ANGLE (DEGREES)
ADJ FOR X - PERIPHERAL VISUAL DETECTION ANGLE ADJUSTED USING THE AVERAGE OF THE HORIZONTAL EYE FIXATION DISTRIBUTION (DEGREES)
ADJ FOR X-S - PERIPHERAL VISUAL DETECTION ANGLE ADJUSTED USING THE AVERAGE MINUS ONE STANDARD DEVIATION OF THE HORIZONTAL EYE FIXATION DISTRIBUTION (DEGREES)
ADJ FOR X+S - PERIPHERAL VISUAL DETECTION ANGLE ADJUSTED USING THE AVERAGE PLUS ONE STANDARD DEVIATION OF THE HORIZONTAL EYE FIXATION DISTRIBUTION (DEGREES)

izontal eye fixation distribution presented earlier in this paper for tangent sections of highway. The curve-tangent sections for right curves were adjusted for an average horizontal foveal eye fixation position of 3.64 degrees to the right of the imaginary focus of expansion and a standard deviation of 3.79 degrees as was discussed for the 240 feet radius right curves at night and the curve-tangent sections for left curves were adjusted for an average horizontal foveal eye fixation position of 2.7 degrees to the left of the imaginary focus of expansion and a standard deviation of 4.06 degrees as found for the left curves at night in the exploratory experimental eye scanning behavior study conducted at Ohio University. Curves with 12 and 28 degrees of curvature (radii of 204 and 477 feet respectively) were selected since peripheral visual detection angles were not available for curves with 3 degrees of curvature (1906 feet radius) at a distance of 500 feet.

From Table 1, one can see that adjusting the calculated peripheral visual detection angle based upon the average of the horizontal eye fixation distribution does decrease the magnitude of the peripheral visual detection angle somewhat for the tangent-curve sections for the right curve and for the curve-tangent sections for the left and right curves, however, the magnitude of the peripheral visual detection angles actually increase slightly for the tangent-curve section for the left curve. One can also see that adjusting the calculated peripheral visual detection angles, based upon the average minus one standard deviation and the average plus one standard deviation of the horizontal eye fixation distribution data, has a fairly small effect upon the magnitude of the calculated peripheral visual detection angles. This is especially true for the larger peripheral visual detection angles since the magnitude of these angles remain relatively large even after adjustment.

METHODOLOGY TO OBTAIN PERIPHERAL VISUAL DETECTION CAPABILITIES

The objective of this section is to present an experimental methodology which would allow one to investigate a typical alerted driver's ability to detect a reflectorized target which may first appear in his or her peripheral visual field at a given peripheral angle and under selected experimental and environmental conditions. The research methodology which is presented in this section was used by Zwahlen (1986) and was designed to fulfill a number of requirements including: 1) the data which is collected should be of a form to allow one to directly apply it in the design of reflectorized targets for the highway environment, therefore, the dependent variable should be detection distance, rather than reaction time or accuracy, 2) one should be able to employ the methodology in the field at night as opposed to a laboratory setting, so that the effects of luminaires, advertising signs, pavement reflectance and other nonessential stimuli, which could produce a certain level of visual background noise from which the driver must extract the relevant information, are present as they are under normal driving conditions, 3) a real automobile and its headlamps should be used such that realistic night beam illumination conditions would exist, further, the heading direction of the car and the headlamps should be alterable such that one may investigate the effect of near maximum beam output and reduced beam output, 4) one should be able to keep a subject's information processing load controlled and close to a minimum, therefore, the subject should be sitting in a stationary car with a clean windshield in order to obtain near ideal or maximum performance data, 5) healthy, young subjects with good vision should be used in order to obtain near ideal or maximum performance data, 6) the experiment should be conducted under good weather conditions with clear visibility in order to obtain near ideal or maximum performance data, and 7) the methodology should be efficient and not require subjects to sit in the experimental car for much longer than one hour to perform the experiment.

In order to study the effects of two different beam output conditions (high candlepower values and moderate candlepower values in the direction of the reflectorized target) and the relative stability and reliability of the experimental results, two separate groups of subjects were used in this experiment. The first group had 7 subjects (5 males and 2 females) with an average age of 21.1 years (standard deviation of .9 years). This group of subjects had an average of 5.6 years driving experience during which they drove an average of 5000 miles per year, with respective standard deviations of 1.9 years and 3000 miles per year. The second group had 7 subjects (3 males and 4 females) with an average age of 23.5 years (standard deviation of 1.7 years). The second group had an average of 7.1 years driving experience during which they drove an average of 8700 miles per year, with respective standard deviations of 2.2 years and 3300 miles per year. All the subjects had normal visual acuity and volunteered their time as subjects. Each subject served as his or her own control.

A 1979 Ford Fairmont was used as the experimental car for the first group of subjects. The headlamps (H4656) were 24 inches above the ground and had a horizontal center to center distance of 48 inches. The actual established location of the hottest spot for the left low beam was 2 degrees to the right and 2 degrees down and the actual established location of the hottest spot for the right low beam was 1.5 degrees to the right and 1.7 degrees down. The electrical system of the car operated at an average of 13.3 volts. The average distance from the longitudinal vertical center plane of the car to the subject's sagittal plane while in the driver position was 14 inches. The average horizontal distance from the headlamps to the subject's eyes was 89 inches and the average subject eye height was 45 inches above the ground.

A 1979 Ford LTD II served as the experimental car for the second group of subjects. Its headlamps (GE 4562) were

29 inches above the ground with a vertical center to center distance of 46 inches. The actual established location of the hottest spot for the left low beam was 2 degrees to the right and 2 degrees down and the actual established location of the hottest spot for the right low beam was 1.5 degrees to the right and 1.7 degrees down. The electrical system of the car operated at an average of 14.1 volts. The distance from the headlamps to the subject's eyes was 97 inches and the average subject eye height was 43 inches above the ground.

A black bicycle was used as the target vehicle. A white license plate was mounted on the front of the bicycle such that the horizontal center of the license plate was 26.8 inches above the paved surface and its reflecting surface made an angle of -10 degrees with the transverse axis of the bicycle to simulate a vehicle parked at a slight angle along the highway. The license plate (size: 6 inches x 12 inches) had a reflectivity of 24 CIL (measured 23.5 cd./fc. at .2 degrees observation angle and -4 degrees entrance angle).

A 75 feet wide, 2000 feet long section of an abandoned concrete runway, which is located at the edge of the city of Athens, Ohio, was used as the experimental site. A two-lane state highway with moderate traffic was located parallel (about 200 feet away) to the runway. A number of luminaires, a few advertising signs, and other light sources were in the subject's field of view (mainly in the left peripheral field).

There were three approach paths parallel to the runway axis as shown in Figure 10. The front centers of the experimental cars were placed at the zero distance line, vertically above the center line of the runway. Looking forward from the car, path 1 was 12.5 feet to the left of the runway center line, path 2 was 6.25 feet to the right of the runway center line and path 3 was 25 feet to the right of the runway center line. The inclusion of three paths in the experiment was intended to introduce some uncertainty to the subject about the lateral location of the approaching target. The car was then positioned on the runway such that it was heading 3 degrees to the left of the center of the runway (-3 degree car heading angle) for group 1 or 10 degrees to the right of the center of the runway (10 degree car heading angle) for group 2. The -3 degree car heading angle produced close to maximum low beam candlepower values in the direction of the reflectorized target, whereas the 10 degree car heading angle produced considerably lower low beam candlepower values in the direction of the reflectorized target. Stakes were placed radially 500 feet away from the car at angles of -30, -20, -10, 0, 10, 20 and 30 degrees to indicate where one movable red dim light, (3 feet above the ground) which was to be used as a fixation point by the subjects, should be positioned.

During the experiment, a group of dark clothed experimenters positioned themselves at various locations along the side of the runway and signaled to the experimenter, who was sitting in the passenger seat of the stationary experimental car, the beginning of each trial by way of a flashlight.

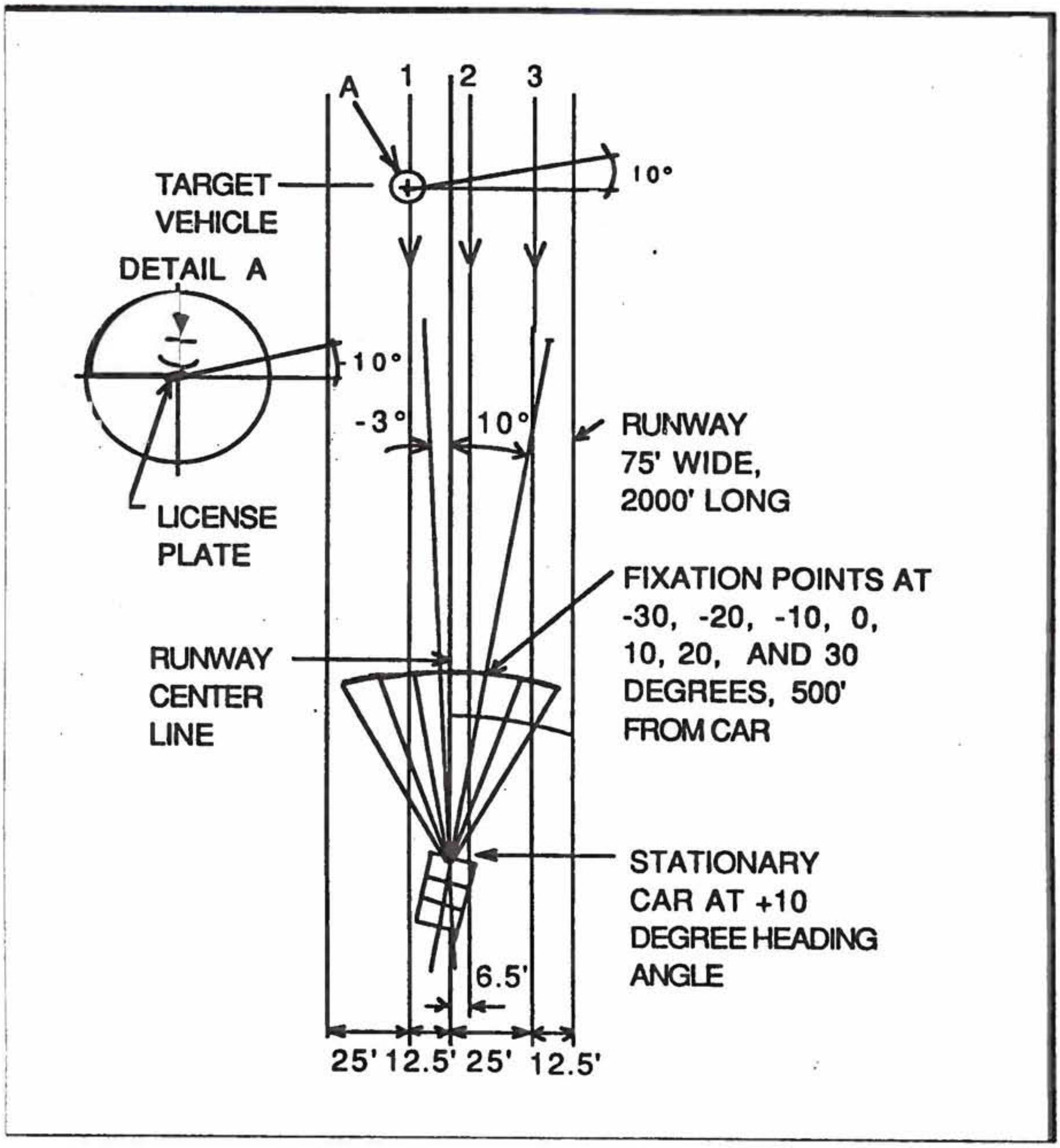


Figure 10 Layout of Experimental Site and Arrangements for the Peripheral Detection Experiment.

Another experimenter sitting in the car recorded the time of each trial, car voltage, weather conditions and the subject's responses. The stationary experimental car engine was kept idling throughout the experiment. When the experimenter in the car received the signal that the bicycle rider was ready for an approach and the measurement crew was off the runway the subject was asked to fixate on the dim red light which was positioned 500 feet ahead of the car at one of the seven selected detection angles. The subject was then instructed to turn on the low beams and be prepared to detect the approaching license plate while continuously fixating his or her eyes on the dim red light. The bicycle rider would approach the stationary car along one of the three approach paths at a constant speed of about 10 mph. As soon as the subject had the initial sensation of detection of the target he or she would immediately switch from low to high beams and keep them on for a few seconds. When the bicycle rider perceived the high beams, he or she would drop a small sandbag on the runway to indicate the detection distance. The measurement crew would then measure the detection distance, pick up the sandbag and return it to the bicycle rider.

After everyone had cleared the runway, the bicycle rider had moved back to the end of the runway and the bicycle was positioned perpendicular to the runway center line on the correct approach path for the next trial, the measurement crew would signal the experimenter in the car indicating the beginning of the next trial. The correct approach path of the bicycle with the target and fixation point position were checked by the experimenter in the car. Six practice trials were carried out for each subject to be sure they understood the procedure. This was then followed by the 63 actual trials (7 detection angles x 3 paths x 3 approaches). The experiment required about 1 hour and 20 minutes to complete for each subject.

The independent variables for this experiment were the seven detection angles (-30, -20, -10, 0, 10, 20 and 30 degrees with respect to the runway center line) and two car heading angles (-3 degrees to the left and 10 degrees to the right). The dependent variable was the detection distance measured in feet. The order of presentation of the peripheral detection angles was according to a latin square design (7 angles, 7 subjects) the 9 observations for each angle (3 paths x 3 replications) was blocked (3 blocks, each path assigned in random order within each block).

The detection distances obtained for each of the three approach paths were combined since paired t-tests indicated that almost all differences between the three paths were statistically not significant at the .05 level. The combined results are shown in Figure 11, which shows averages and standard deviations for the detection distances as a function of the peripheral visual detection angle. From the average detection distances shown in Figure 11, one can see that the average detection distances decrease considerably as the peripheral visual detection angle increases. At a peripheral visual detection angle of 10 degrees the average

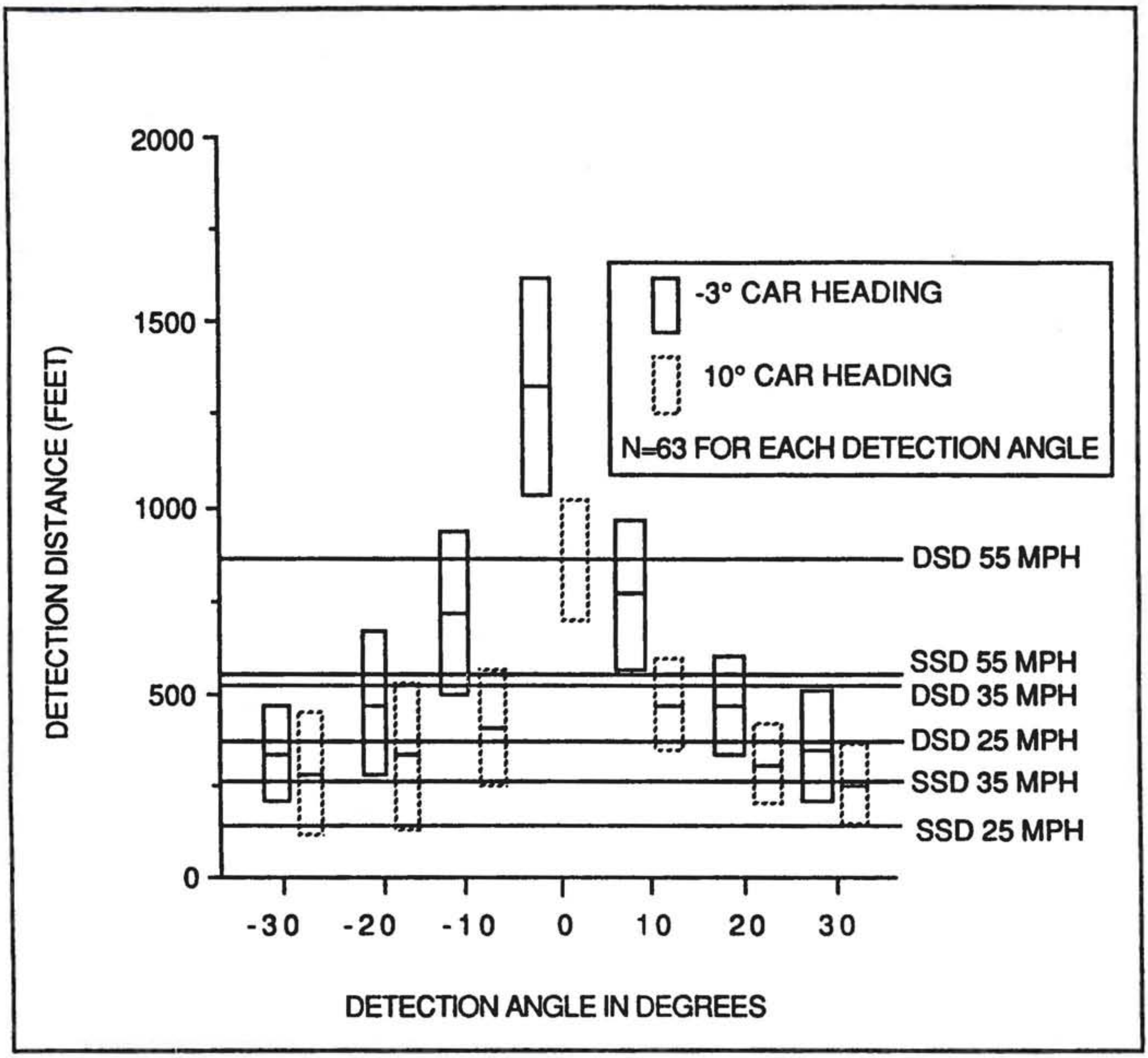


Figure 11 Averages with Plus and Minus one Standard Deviation of the Peripheral Visual Detection Distances for a -3 and 10 Degree Car Heading Angle Compared with Selected Stopping Sight Distances and Decision Sight Distances

detection distance was between 54.2 and 59.2 percent of the average foveal detection distance for the -3 degree car heading angle and between 47.3 and 55.6 percent of the average foveal detection distance for the 10 degree car heading angle. At a peripheral visual detection angle of 20 degrees the average peripheral detection distances were 35.7 and 36.0 percent of the average foveal detection distance for the -3 degree car heading angle and between 35.6 and 37.9 percent of the average foveal detection distance for the 10 degree car heading angle. At a peripheral angle of 30 degrees the average peripheral detection distances were 25.1 and 27.0 percent of the average foveal detection distance for the -3 degree car heading angle and 28.2 and 32.6 percent of the average foveal detection distance for the 10 degree car heading angle. As expected, one can also see in Figure 11 that the average detection distances obtained for the 10 degree car heading angle or the much lower candlepower values of the lowbeams are considerably shorter than the average detection distances obtained for the -3 degree car heading angle.

The average detection distances which were obtained in this study and are shown in Figure 11 can be further evaluated from a safety point of view. Comparing the peripheral detection distances obtained in this study with a recommended stopping sight distance of 563 feet for a speed of 55 miles per hour, one can see that for the -3 degree car heading angle only the average detection distances for the foveal and the 10 degree peripheral detection angles exceed the recommended stopping sight distance. If one compares the recommended stopping sight distance of 563 feet for a speed of 55 miles per hour to the average detection distance for the 10 degree car heading angle, one can see that the recommended stopping sight distance is larger than all of the average detection distances except the average detection distance for the 0 degree (foveal) detection angle. If one reduces the idealized average detection distances obtained in this study by 50 percent to adjust for driver alertness and expectancy, older drivers, information acquisition and information processing load while driving, somewhat degraded environmental visual conditions, dirty windshield, dirty headlamps, etc. then only the 50 percent reduced average detection distance for the 0 degree peripheral angle (foveal detection) for the -3 degree car heading angle exceeds the stopping sight distance for a speed of 55 mph. Comparing the average detection distances acquired for the -3 degree car heading angle after they are reduced by 50 percent with a stopping sight distance of 263 feet for a speed of 35 mph one can see that reduced average detection distances are larger than the stopping sight distance for only the foveal detection and the 10 degree peripheral visual detection angles. Further, comparing the average detection distances acquired for the 10 degree car heading angle one can see that only the reduced average detection distances for the 0 degree peripheral visual detection angle are larger than the stopping sight distance. In fact, once the average detection distances are reduced by 50 percent they are so small that for the 10 degree car heading angle the reduced average detection distances for the 30 degree peripheral angles are approximately equal to or slightly smaller than the stopping sight distance of 137 feet for a speed of 25 miles per hour.

McGee et al. (1978) recommended decision sight distances, the distances a driver needs to perceive a potentially hazardous situation and react to the impending danger efficiently, of 375 to 525 feet for a speed of 25 mph, 525 to 725 feet for a speed of 35 mph and 875 to 1150 feet for a speed of 55 mph. Comparing the smaller of each of these distances with the detection distances obtained in this study, one can see that the average detection distances obtained in this study are greater than the minimum decision sight distance for a design speed of 55 mph for only the foveal detection angle with the -3 degree car heading angle (near optimal low beam candlepower conditions). As the peripheral visual detection angle is increased the average detection distances decline so rapidly that for the relatively small peripheral detection angle of 10 degrees the average detection distances for 10 degree car heading angle are less than the decision sight distance for a speed of 35 mph and the average detection distances for a peripheral detection angle of 30 degrees are as much as 130 feet less than the decision sight distance for a speed of 25 mph. If the decision sight distances are compared to the detection distances reduced by 50 percent then the decision sight distance for a speed of 55 mph is larger than all of the

50 percent reduced average detection distances for both the -3 and 10 degree car heading angles. Comparing the decision sight distances for a speed of 25 mph to the average detection distances reduced by 50 percent, only the detection distance for the 0 degree car heading angle for the 10 degree car heading angle is larger than the decision sight distance for a speed of 25 mph. Similarly only the average detection distances obtained for a 0 degree and 10 degree peripheral visual detection angles for the -3 degree car heading angle are equal to or larger than the minimum decision sight distance for a speed of 25 mph. It should be noted that the much shorter 10 degree car heading detection distance results might be more applicable to the peripheral visual detection of a reflectorized target in the highway environment than the -3 degree car heading detection distance results, since the longitudinal direction of the car and its beams is such that the candlepower values of the beams in the direction of the reflectorized target are most likely considerably reduced in a situation where a reflectorized target first appears in a driver's peripheral visual field.

Zwahlen (1981) has shown that the multiples of threshold that a driver needs to detect a reflectorized target, such as a bicycle pedal, increases very rapidly as the peripheral visual detection angle is increased. The multiples of threshold are proportional to the specific intensity of a reflectorized target. Therefore, if for a given peripheral visual detection angle an average detection distance which is equal to the average foveal detection distance is desired one would have to appropriately increase the specific intensity or the reflectivity of the retroreflective target, assuming the environmental and beam conditions would remain the same.

Paired t-tests were performed to determine whether or not the average peripheral detection distances for the left side (peripheral visual detection angles of -30, -20 and -10 degrees) could be assumed to be equal to the corresponding average peripheral detection distances for the right side (peripheral visual detection angles of 10, 20 and 30 degrees). For both the -3 degree and the 10 degree car heading angles the average peripheral detection distances for the 10 degree peripheral visual detection angle for the left side were about 9.2 to 17.6 percent shorter (statistically significant at the 0.05 level) than the average peripheral detection distances for the right side. This might be partially explained by noting that there was a highway with moderate traffic located on the left parallel (about 200 feet away) to the airport runway and therefore there were more light sources (luminaires, advertising signs, etc.) in the left peripheral field of view (less uniform dark background). The average peripheral detection distances for the 20 and 30 degree peripheral visual detection angles were of about equal magnitude and were not statistically different.

CONCLUSIONS

The methodology to calculate peripheral visual detection angles for curve-tangent and tangent-curve sections of two-lane highways using a spreadsheet and graphics package provided magnitude ranges for peripheral visual detection angles in an efficient manner. These angles were further adjusted according to driver eye fixation density data which was collected for certain geometric, environmental and driving conditions (not related to this study). However, the magnitudes of these adjustments, when compared to the magnitude of the calculated peripheral visual detection angles, were for the most part rather small. Using the state of Ohio as an example, it has been shown that curve-tangent or tangent-curve sections occur fairly frequently along two-lane rural highways, especially in hilly regions, and therefore, one may conclude that relatively large peripheral visual detection angles (in the range of 10 to 20 degrees and in some cases up to 40 degrees) for reflectorized targets which will become visible for the first time in the periphery of a driver's visual field could be quite common.

The methodology to assess the peripheral visual detection capability of drivers for reflectorized targets at night has proved to be fairly efficient and reliable and has produced visual detection distances which were collected in the field, with nearly ideal subjects under fairly ideal and well controlled conditions. The results show that the peripheral visual detection ability or the detection distances for the detection of reflectorized targets decreases considerably as the peripheral visual detection angle increases. The decrease of the visual detection distances in the periphery can, however, be offset by increasing the reflectivity or specific intensity of the retroreflective target. It is, therefore, recommended that in cases where a target, such as a reflectorized warning sign, will become visible for the first time, most likely in the periphery of a driver's visual field, appropriate increases in the reflectivity of the target should be made to assure early detection in order for timely recognition, information processing, decision making and appropriate control actions.

REFERENCES

Karttunen, R. and Hakkinen, S., "Discrimination of Traffic Signs in the Peripheral Areas of the Field of Vision", Report for the Central Organization for Traffic Safety, Research Department, Iso Roobertinkatu 20, 00120 Helsinki 12, Finland, 1981, 42 pages.

McGee, M.W., et al. Decision Sight Distance For Highway Traffic Control Requirements, Report FHWA-RD-78-78. Federal Highway Administration, U.S. DOT, 1978.

Shinar, D., McDowell, E. and Rockwell, T., "Eye Movements in Curve Negotiation", Human Factors, Vol. 19, No. 1, February 1977, pp. 63-71.

U.S. Department of Transportation and the Federal Highway Administration, Manual on Uniform Traffic Control Devices, 1978, pp. 2A-8 to 2A-14.

Zwahlen, H.T., "Nighttime Detection of Bicycles", Transportation Research Circular, Number 229, May 1981, pp. 38-49.

Zwahlen, H.T., "Warning Signs and Advisory Speed Signs - Reevaluation of Practice", Prepared for Ohio Department of Transportation in Cooperation with US DOT, FHWA, State Job No. 14327(O), Final Report, Report No. FHWA/OH-84/003, June 1983, 185 pages.

Zwahlen, H.T., "Peripheral Detection of Reflectorized License Plates", Proceedings of the Human Factors Society 30th Annual Meeting, Vol. 1, 1986, pp. 408-412.

SIGN LUMINANCE AS A METHODOLOGY FOR MATCHING DRIVER NEEDS, ROADWAY VARIABLES AND TRAFFIC SIGNING MATERIALS

R. L. ERICKSON and H. L. WOLTMAN

The use of retroreflective materials to enhance the nighttime visibility of traffic control signs and other devices is very general. Various official standards possess the requirement that signs which must be seen by the motorist at night must be either retroreflecting or illuminated. Retroreflectorization alone is quite sufficient for sign visibility under reasonably optimum conditions. These conditions include satisfactory alignment of the vehicle with the sign, an uncluttered surround to permit timely discovery, headlights in satisfactory alignment and the use of retroreflective materials identified in such specifications as U.S. DOT FHWA FP-85[1], CIE Publication 39-2[2], ISO 3864[3], DIN 67520[4], BS 873[5] and others. However, sign perception is essentially dependent upon the development of an adequate level of luminance, and there are numerous factors which alter both the luminance capability of the retroreflective material, and the level of adequacy required by the driver. An adequate level of luminance is the satisfactory merging of supply with demand.

It is advantageous to describe the principle factors that influence both the luminance capability of retroreflective materials separately from those factors which influence the level of luminance deemed adequate for the driver. The first are highly mechanical: they have to do with the headlamp output, the choice of retroreflective material, the alignment of the sign with respect to the road, etc. The latter factors must deal exclusively with the driver's perceptual process and state of mind. The proper tailoring of the two is expected of the highway engineer for virtually all nighttime situations and drivers. A research methodology which explores a variety of scenarios representative of actual use is apt to be a more satisfactory descriptor of the retroreflector than is Specific Luminance, or Retroreflectance, its traditional descriptor.

The use of luminance as a criteria for evaluating performance of signs instead of retroreflectance, provides a means of direct matching to driver needs. Estimates of luminance to satisfy driver needs may be obtained from a number of investigations which may be illustrative: Driver needs from a review of nighttime sign legibility studies by Sivak[6] is provided in Table 1. Other studies dealing with Stop sign conspicuity by Morales[7], sign conspicuity for various background complexities and driver expectancy by Olson[8], estimates for sign priority by Perchonok[9]; all reveal a dependence on sign luminance for satisfactory performance.

The supply of luminance is dependent on relative sign position, roadway approach, headlamp quality; and is correlated with retroreflectance, the weathered state of the sign, presence of soil, etc. The direct inspection of either lumninance tables predicting performance, or the signs themselves from the driver's position is the most satisfactory method of assuring this necessary quality.

Table 1.

OPTIMAL AND REPLACEMENT
LUMINANCES OF TRAFFIC SIGNS:
A REVIEW OF APPLIED LEGIBILITY RESEARCH

Level	Sign Luminance
Optimal	75 cd/m2
Replacement	
85th Percentile	16.8 cd/m2
75th Percentile	7.2 cd/m2
50th Percentile	2.4 cd/m2

To determine the effects of vehicle and roadway variables on sign luminance, the authors have employed their previous findings of sign luminance for U. S. guide sign legends and backgrounds as well as the luminance enhancement from stream traffic and rainfall. These assessments employed a mix of U. S. cars operating on both low and high beams while carefully photometring a variety of retroreflective materials, from the driver's eye position in standard size passenger vehicles. Measurements were taken of samples in typical sign positions, from distances corresponding to the longest of decision sight distance models, to relatively short sign reading distances. Headlamps used were either typical of new vehicle equipment or were supplied by equipment manufacturers following photometric testing. Aim was adjusted to correspond to SAE recommendations usually employing the aiming screen method of SAE J 599.[10] Level tangent sections of roadway were employed. A full description of the methods is contained in three separate papers by Youngblood and Woltman.[11,12,13]

The findings are well suited for adaptation to the problem at hand: retroreflective materials will provide the same response curve given similar vehicle dimensions. Luminance values are proportional to illuminance, so that an accurate comparison may be made between headlamps. Beyond this relationship, careful characterization of the retroreflective materials is required, as is the headlamp/driver-eye relationship. Angularity of the signs with respect to the approach should be considered. Allowance for dirt on signs is the same as for dirt on headlamps; both are treated as a diminution of illuminance.

The procedure employed in this study is the modeling of sign luminance using a procedure first detailed by Elstad,[14] with further refinement by Szczech.[15] The model employs carefully detailed headlamp outputs in a matrix encompassing all directions of interest for sign positions at any distance. The values derived for sign luminance involve complex geometric and retroreflective response relationships: nevertheless, correspondence with the previously cited field studies is within the extreme measurements and permits comparisons of resultant sign luminances by headlamp type, angularity, and a number of common alignment conditions for three types of retroreflective materials.

SIGNING MATERIALS

The signing materials studied are representative of new white retroreflective materials employed for traffic control signs. Luminances for other colors and their ratios to white may be expected to fall within the following limits:

COLOR	LUMINANCE RATIO TO WHITE (Percent)
Yellow	61 to 76
Orange	33 to 42
Red	17 to 30
Green	13 to 19
Blue	7 to 10

The materials studied are described in Table 2. The Coefficients of Retroreflection, which are essential for sign luminance computations, are are determined according to ASTM method E810.[16]

TABLE 2.

COEFFICIENT OF RETROREFLECTION

FOR

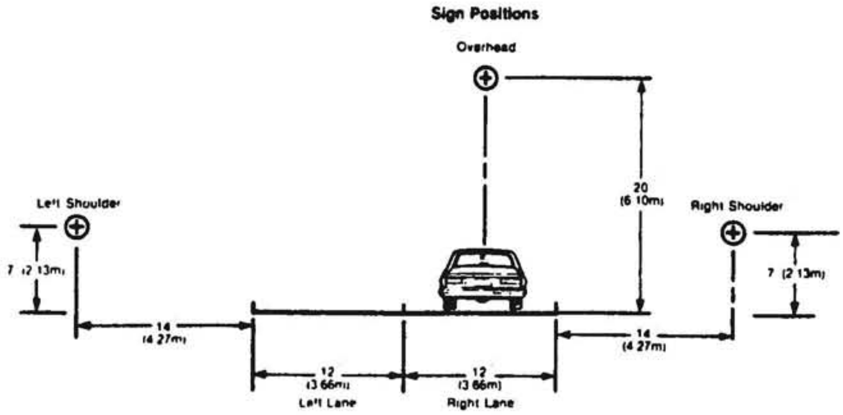
WHITE RETROREFLECTIVE SIGNING MATERIALS

Reflective Sheeting			Coefficient of Retroreflection (candelas/lux/sq meter) @ 0.2° Observation -4° Entrance
Material	Type	Color	
"A"	Enclosed Lens	White	120
"B"	Encapsulated Lens	White	310
"C"	Micro-prism	White	1100

SIGN POSITIONS

Sign positions are illustrated in Figure 1. The positions are typical of regulatory and warning signs commonly displayed on the right shoulder, overhead signs over the driver's lane of travel, and signs on the left side such as No Passing Zone pennants, bridge end barricades, other warning and informational signs. The authors employed off-sets and elevations specified in the U.S. MUTCD.[17]

Figure 1.



ROADWAY

The approach to the sign is not always a straight, level, tangent section. It is frequently a horizontal or vertical curve. The authors have chosen five cases which may be considered an incomplete senario, but which are hoped to be relatively representative and illustrative of a variety of approach conditions. These roadway geometries are diagrammed in Figure 2 and include; straight tangent approach, right and left curves of 2000 feet radius, sag and hill of 6000 feet radius.

Figure 2.

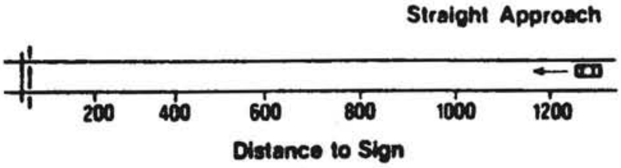
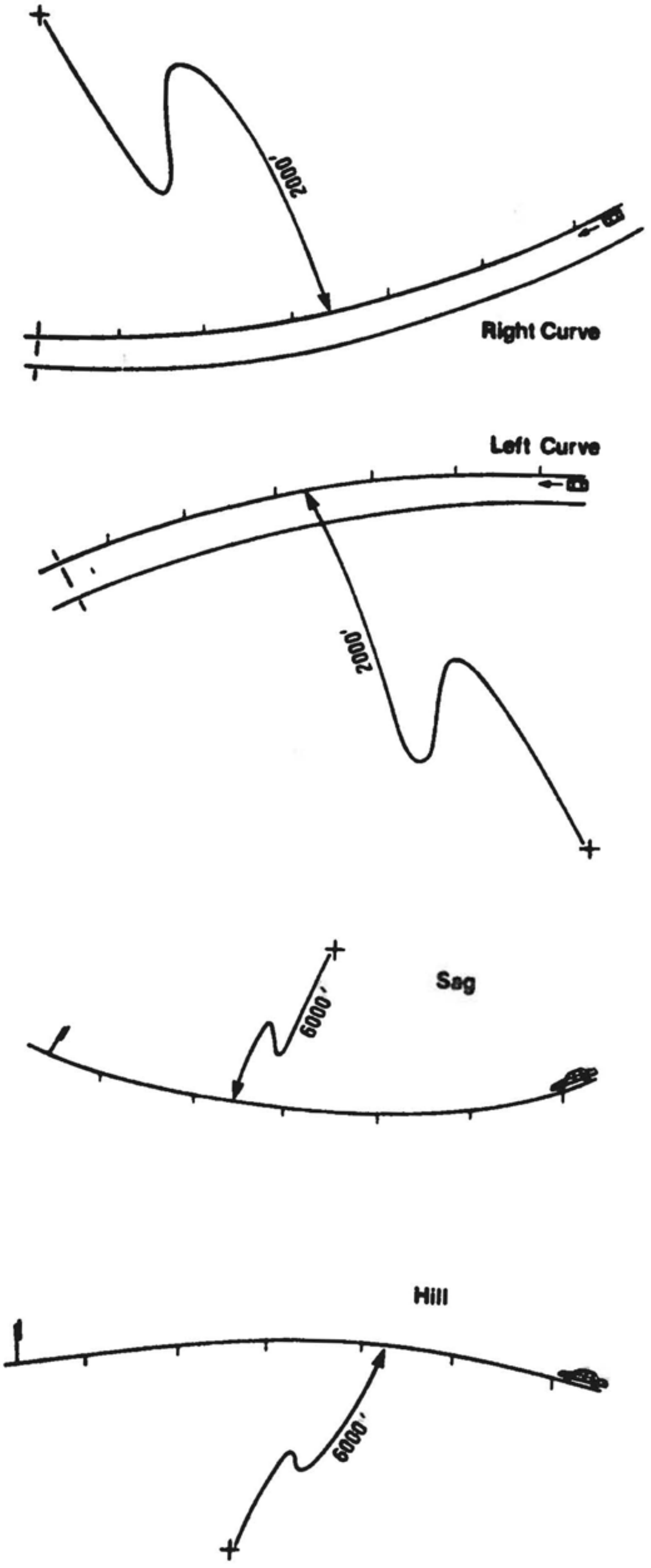


Figure 2.
(Cont.)



SIGN LUMINANCE ESTIMATES

Estimates of sign luminances for the three retroreflective materials at six approach distances, and three sign positions are provided in Table 3,4,5,6, and 7. The tables are arranged for separate roadway geometrics, tangent, right and left curve, sag and hill; to permit separate assessment for these approach conditions.

Two estimates are provided for signs seen at angles. Approximately six percent of signs of a recent inventory of entrance angles of over 1300 signs by Austin, are seen by the motorist at an angle of 30 degrees or greater. Estimates are provided for shoulder signs angled 30 degrees away and toward the road in Tables 8 and 9 respectively. It must be noted that the photometrics of the retroreflective materials are represented with reasonable accuracy for the entrance and observation angles actually encountered on all of the above approaches.

TABLE 3 STRAIGHT TANGENT ROAD

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

	Distance -- meters/feet					
Sign Mounting	366/1200	305/1000	244/800	183/600	122/400	61/200
	1.75	2.56	4.06	6.90	9.11	3.19
Right	4.34	6.45	10.47	18.07	21.75	4.30
	12.38	18.86	31.28	57.59	61.74	8.62
	0.79	1.04	1.42	2.07	3.24	3.74
Left	1.97	2.65	3.69	5.20	7.04	6.42
	5.66	7.59	10.22	15.15	23.99	14.03
	0.62	0.77	0.96	1.23	1.60	1.66
Overhead	1.55	1.96	2.49	3.19	3.54	1.36
	4.52	5.84	7.38	9.83	10.18	2.48

TABLE 4 RIGHT CURVE

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

	Distance -- meters/feet					
Sign Mounting	366/1200	305/1000	244/800	183/600	122/400	61/200
Right	0.18	0.26	0.62	1.35	3.65	5.19
	0.46	0.67	1.63	3.76	9.16	4.30
	1.16	2.31	6.08	12.24	21.03	8.62
Left	0.17	0.49	1.38	5.97	4.37	4.06
	0.44	1.31	3.79	16.81	10.80	7.14
	0.87	2.96	10.65	51.75	34.73	14.25
Overhead	0.17	0.31	0.59	0.96	1.33	1.64
	0.45	0.82	1.61	2.71	3.17	1.26
	0.91	2.10	5.59	9.20	7.88	2.19

TABLE 5 LEFT CURVE

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

	Distance -- meters/feet					
Sign Mounting	366/1200	305/1000	244/800	183/600	122/400	61/200
Right	0.15	0.21	0.59	1.92	3.21	5.19
	0.40	0.57	1.52	4.71	7.09	4.30
	0.81	1.44	4.40	14.55	23.49	8.62
Left	0.15	0.21	0.48	0.86	1.75	2.22
	0.41	0.57	1.24	2.03	3.86	3.25
	0.80	1.41	3.69	6.46	13.03	6.43
Overhead	0.15	0.24	0.45	1.03	1.39	1.65
	0.40	0.64	1.18	2.52	3.02	1.43
	0.75	1.37	3.22	8.49	9.93	2.90

TABLE 6 SAG

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

Sign Mounting	366/1200	305/1000	244/800	183/600	122/400	61/200
	*	0.25	0.42	0.74	1.61	5.19
Right	*	0.66	1.10	1.96	3.71	4.30
	*	2.11	3.52	6.08	9.82	8.62
	*	0.25	0.45	0.75	1.39	2.66
Left	*	0.65	1.19	1.91	3.01	3.88
	*	2.01	3.58	5.99	9.53	8.59
	*	0.25	0.35	0.61	1.00	1.62
Overhead	*	0.66	0.93	1.60	2.16	1.20
	*	2.07	2.94	4.95	5.76	2.08

Distance -- meters/feet

*Range exceeds headlamp field

TABLE 7 HILL

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

	Distance -- meters/feet					
Sign Mounting	366/1200	305/1000	244/800	183/600	122/400	61/200
Right	*	*	*	66.12	192.89	5.19
	*	*	*	172.25	450.86	4.30
	*	*	*	504.65	1326.89	8.62
Left	*	*	*	53.65	70.69	4.57
	*	*	*	134.13	153.63	8.91
	*	*	*	353.83	547.59	19.57
Overhead	*	*	*	55.45	3.40	2.14
	*	*	*	143.12	7.79	1.97
	*	*	*	421.46	24.51	3.76

*Sign obscured by hill

TABLE 8 STRAIGHT TANGENT ROAD

SIGNS ROTATED +30°

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

	Distance -- meters/feet					
Sign Mounting	366/1200	305/1000	244/800	183/600	122/400	61/200
	1.05	1.52	2.38	3.98	5.07	2.59
Right	3.11	4.59	7.40	12.59	14.71	2.68
(Away)	7.10	10.49	17.04	30.42	34.84	6.10
	0.52	0.69	0.96	1.43	2.36	3.14
Left	1.50	2.03	2.86	4.10	5.71	5.71
(Toward)	3.38	4.48	6.28	9.75	12.53	9.88

TABLE 9 STRAIGHT TANGENT ROAD

SIGNS ROTATED -30°

Retroreflective Sign Luminance ECE Low Beam Lamps

Materials: Enclosed Lens
Encapsulated Lens
Micro-prism

Luminance in Cd/m2

	Distance -- meters/feet					
Sign Mounting	**366/1200**	**305/1000**	**244/800**	**183/600**	**122/400**	**61/200**
	1.13	1.66	2.66	4.61	6.32	4.01
Right	3.26	4.87	7.97	13.93	17.19	3.62
(Toward)	7.42	11.03	18.15	33.29	36.62	4.91
	0.46	0.60	0.18	1.15	1.70	1.63
Left	1.39	1.86	2.55	3.53	4.59	3.67
(Away)	3.16	4.18	5.78	8.60	11.28	10.36

Tables 3 through 9 provide estimates of sign luminance for distances from 366 meters (1200 feet) to 61 meters (200 feet) in increments of 61 meters (200 feet). The luminance estimates are for a motorist in a standard size passenger vehicle and are intended to provide a number of scenarios in which various approaches can be studied. Thus various retroreflective material types can be reviewed for a specific approach condition or distance and the appropriate type chosen which satisfies a given luminance criteria.

The tables also permit comparisons for the three differing sign positions. It is altogether appropriate to consider the left shoulder position as an alternate for certain types of informational signs which might otherwise be positioned in the overhead position. As may be determined from the tables, the left side mounting frequently provides approximately twice the luminance as compared to the overhead position. The possibilities of left side positioning should not be overlooked.

Tables 8 and 9 deal with shoulder mounted signs facing 30 degrees away from and toward the roadway. A substantial number of these signs, some of which are highly critical, have been found at this entrance angle or greater. The luminance estimates of Tables 8 and 9 are derived from photometrically determined retroreflectance values appropriate for these entrance and observation angles.

CONCLUSIONS

The use of luminance as a criteria for evaluating performance of signs instead of retroreflectance, provides a means of matching materials to roadway situations and driver needs. The driver's luminance needs must account for the variability of the driver population with a suitable factor of safety to accommodate driver age, expectancy, the complexity of the surround and the criticality of the sign. These important factors are beyond the scope of this investigation. The lumninance supply may be estimated from the tables which offer the closest match to the roadway situation. Practioneers may then select from the materials those which may be expected to offer the luminance level desired through the range of distances which may be of interest. In this manner it is hoped that material selection can become a part of the design process.

REFERENCES

1. Standard Specifications for Construction of Roads and Bridges on Federal Highway Projects, FP-85, 1985, U.S. Department of Transportation, Federal Highway Administration, Washington, D. C. Pp 672-679.

2. Recommendations for Surface Colours for Visual Signalling, Publication CIE No. 39-2, International Commission on Illumination, Kegelgasse 27, A-1030, Wien, Austria. 1983.

3. Safety Colours and Safety Signs, International Standard ISO 3864-1984. International Orginization for Standardization, 1 Rue De Varembe, CASE Postale 56, CH-1211, Geneva 20 Switzerland. 1984.

4. Deutsche Norm, Retroreflektierende Materialien zur Verkehrssicherung, DIN 67 520, Deutsches Institut fur Normung e.v., Berlin, 1982.

5. Road Traffic Signs and Internally Illuminated Bollards, BS 873:Part 6:1983. British Standards Institution, 2 Park Street, London W1A 2BS.

6. Sivak, Michael, and Olson, Paul L., "Optimal and Replacement Luminances of Traffic Signs: A Review of Applied Legibility Research". Transportation Research Institute UMTRI-83-43, University of Michigan, Ann Arbor, MI 48109. Dec. 1983.

7. Morales, Juan M., "Retroreflective Requirements for Traffic Signs - A STOP Sign Case Study". FHWA/RD-87/017, Traffic Safety Research Division, HSR-30, Office of Safety and Traffic Operations R&D, Federal Highway Administration, 6300 Georgetwon Pike, McLean, VA 22101. Jan. 1987.

8. Olson, Paul L., "Minimum Requirements for Adequate Nighttime Conspicuity of Highway Signs", Transportation Research Institute, UMTRI-88-8, University of Michigan, Ann Arbor, MI 48109. Feb. 1988.

9. Perchonok, K. and Pollack, L. "Luminous Requirements for Traffic Signs". Contract No. FHWA-RD-81-158. Federal Highway Administration, Washington, D.C., Dec. 1981.

10. SAE J 599, Society of Automotive Engineers, 400 Commonwealth Drive, Warrendale, PA 15096. Vol 2, 1983. Pp 21.71 - 21.73.

11. Youngblood, W. P., and Woltman, H. L., "A Brightness Inventory of Contemporary Signing Materials for Guide Signs". Highway Research Record No. 377. Transportation Research Board, National Academy of Sciences, Washington, D. C., 1971. Pp 69 - 90.

12. Woltman, H. L., and Youngblood, W. P., "Indirect Factors Affecting Reflective Sign Brightness". Transportation Research Record No. 611. Transportation Research Board, National Academy of Sciences, Washington, D.C. 1976. Pp 35-38

13. Youngblood, W. P., and Woltman, H. L., "Relation Between Sign Luminance and Specific Intensity of Reflective Materials". Transportation Research Record 681. Transportation Research Board, National Academy of Sciences, 1978. Pp 20 - 24.

14. Elstad, J. O., Fitzpatrick, J. T., and Woltman, H. L., "Requisite Luminance Characteristics for Reflective Signs". Highway Research Board Bulletin 336. Highway Research Board, National Academy of Sciences, Washington D. C., 1962 Pp 51-59

15. Szczech, T.J., Technical Report No. 004, TCM/3M, 3M Co., St. Paul, MN 55144. Dec. 1987.

16. Standard Test Method for Coefficient of Retroreflection of Retroreflective Sheeting. ASTM E 810-81, American Society for Testing and Materials, 1916 Race St., Philadelphia, PA 19103.

17. Manual On Uniform Traffic Control Devices for Streets and Highways, U.S. Department of Transportation, Federal Highway Administration, Mar. 1986. Pp 2A-9 - 2A-10.

"METHOD FLOATING CAR"; A RESEARCH METHOD TO STUDY THE SPEED-BEHAVIOUR AND ROUTES OF CARDRIVERS IN RESIDENTIAL AREAS.

Marien G. Bakker
Ministry of Transport and Public Works,
Transportation and Traffic Engineering Division,
The Netherlands.

INTRODUCTION

Since January 1984 it has been possible for the local authorities in the Netherlands to assign an area as a 30 km/h zone. The safety and the freedom of movement of the other road users in residential areas require a speed lower than the maximum speed of 50 km/h within built-up areas. The 30 km/h sign will in itself not guarantee that the compulsery maximum speed is complied with. Only if the character of the road is such that the compulsery speed is a logical consequence of the configuration, one can expect compliance and thus the desired speed.

The effects of the 30 km/h zone regulation are tested within 15 experimental zones. This research is carried out in terms of reference of the Ministry of Traffic and Transportation. One of the main questions which should be answered is:

- what is the effect of the 30 km/h regulation on the speed behaviour of motorized traffic?

To answer this question it is preferable to study the speed behaviour in the total zone. The methods now in use are however, focussed on spot measurements by means of radar and video-observations. For getting a view of the speed behaviour over the total zone a new method is desirable.

Within the Division of Transportation and Traffic Engineering of the Ministry of Transport and Public Works such method has been developed. In this paper this "Method Floating Car" will be discussed.

METHOD

Since 1981 the Division of Transportation and Traffic Engineering has had the Floating Car at its disposal (Kleinloog 1981). The Floating Car is an instrumented car for the recording of speed, distance and fuel consumptions data. Its a normal car like all others. The covered distance of the car is recorded as a function of time. So every metre of distance is registered with an accuracy of 4 microsec. on a tape recorder. The same principle is in use by the registration of fuel consumption data (registration of CC) and non-automatical events like the start of a measurement, the passage of traffic lights or a crossing and the cause of desceleration etc. These events we call "the Pocket events". The configuration of the system is given in scheme 1.

Scheme 1. The configuration of the Floating Car system.

```
- Position ------------> I            I ------> Tape Recorder
  Puls Generator         I  Central   I
  (1/M.)                 I  Micro     I
- Time (micro sec) ----> I  Computer  I
- Fuel-    ------------> I            I <----- Pocket Events
  Consumption (1/CC).                          (Non-Automatic)
```

This system is the base of the "Method Floating Car". The principle of this method has been based on "Car-Following". Any car entering a 30 km/h zone is followed by the Floating Car. Inside the zone every crossing has got a number of two figures. The speed behaviour and route of a car can be recorded by following such a car which a certain strategy, and registering the numbers of the crossings while passing them. During the measurement, disturbings like descelerations caused by other traffic and the moment of parking of the followed car, can also be registered. After the registration an analysis will then be made on base of the data found.

Two "following strategies" can be done. At first, instantaneous following at a fixed distance and second following with a time gap of a few seconds. Which strategy yields superior performance? Both strategies have there pros and cons. The estimation of a distance between two cars is easier than the estimation of the time gap between them. To keep an instantaneous distance at curves and crossings is very difficult. Following with a time gap is easier because the action which has to be taken will come a few seconds later.
It's also a question at what distance or time gap the following strategy should be undertaken. If the distance or time gap between the two cars is too short the Floating Car might influence the speed behaviour of the other car. If the distance or time gap is too big the estimation of this distance or time gap will be more difficult and there is a possibility that other cars might come in between.
In the next paragraph the results of a study into these aspects of the method will be shown. Also the aspect of influencing the other car by following with the Floating Car will be explained.

The analysis of the data has been pointed at the effects of the 30 km/h zone on speed behaviour, route choice, travel distance/traveltime and "character" of the trip (through- or bound traffic) and fuelconsumption. The data has been collected in a evaluation study. In addition in the period before the reconstruction as well as in the period after, a research has been done. The "following measurements" have been undertaken from the main entrances of the zone. A clear view can be obtained if "following measurements" will be done from all of the main entrances during comparable timeperiods . The ratio between the number of through- and bound traffic, and the route choice of it will be clear. The changes in both items after the reconstruction will give a close view of the effects on it.

The analysis of the speed behaviour are been pointed at the links between two crossings and at the total routes. The maximum speed, the 85percent- the average- and the minimum speed are analysed on link level. On route level items like maximum- and average speed at the route, the travel- and time distance and fuelconsumption seem interesting. In one of the last paragraphs some results will be shown of the study in a part of the 15 experiments in 30 km/h zones.

METHOD IN PRACTICE

In 1984 the "Method Floating Car" was tested extensively in two experimental residential areas in Den Bosch and Delft. The test items were:

- the handling of the system;
- the accuracy of the system;
- the capacity of the method;
- the "following strategy";
- the "influencing" of the followed car.

The first three items will be described shortly (Bakker, 1985). The other two itms extensively because a seperate study was necessary to get know something more about these items.

The handling of the system.

The handling of the Floating Car system is easy. At first one has to put some data in the computer like date, time of day, a definition for codes in use, and a verification mark. After that the system can be loaded and the measurement can be started. During the measurement the task of the observer is to register the start, the type of car they follow, the numbers of the crossings they pass, some influencing factors, like other traffic and parking cars, and the end of the measurement. The registration can be done by means of a box with eight push-buttons. With this box figures from 1 till 88 can be made. The most difficult task for the observer is to fix continuously his position in the area. The driver's task is: "follow any car which enters the area with a certain strategy, take care of the safety of one'self as well as the safety of the other road users".
As concluded one need some experience with the system and the experimental area in which the research will be done; one must have feelings for direction.

The accuracy of the system.

The system is very accurate, more than 99%. This was one of the conclusions of the test procedures in 1981, the year of its creation. With the verification mark the accuracy can be adjusted. During the tests in Den Bosch and Delft it seems that the distances given by the authorities in these cities agree with the results of the Floating Car.
After the measurements corrections can be made. So mistakes made during the registration are not hopeless. Mistakes like wrong crossing numbers or pussing the number while the crossing is already passed can be changed.

The capacity of the method.

The "Method Floating Car" is very intensive. Within one hour no more than 10 till 20 car-following measurements can be done. This number depends on the frequention of cars entering the experimental area, the time of day, the kind of research etc. The results of one day might be 120 measurements. The time necessary for the futher analysis of the data of one day measuring will cost two days for one person.
For the total evaluation of a zone 6 days will be necessary. But then a lot of data about the speed behaviour, distance and traveltime, route choice, character of the trip, and fuelconsumption has come to one's disposal.

The "following strategy".

As we saw two stategies are in use, "instantaneous following at a fixed distance" or "following with a fixed time gap". Both distance and time gap can be varied in length. Which strategy yields superior performance and what is the accuracy of speed measurement with the car-following-method?
To answer these questions a study has been started in cooperation with TNO Institute for Perception. The study was based on runs with two instrumented cars, the "Icarus" of TNO and the "Floating Car. The speed of both cars was simultaneously registered. The testruns were driven on a traffic-training-centre near Gorinchem. In total 72 runs in which the Icarus was followed by the Floating Car were carried out. Aspects as Following-Strategy, length in distance and time and maximal speed of the first car were systematically varied.
The results of this study were (Tenkink, 1986):

- Following with a time delay is more accurate than instantaneous following at a fixed distance;
- Following at a shorter length, in distance or time, gives better results;
- The method remains valid for speeds up to at least 70 km/h (speeds above this level couldn't be reached on the traffic-training-centre);

Under the most favourable condition,

- It is possible to measure 90% of the variance of speed changes of the leading car ;
- The systematical deviation in the estimation of maximal speeds on each straight section may be less than ± 1 km/h or less than 1% of the driven speed;
- The standard deviation in the measurement of the maximal speed limits accuracy to 6% of this speed.

The "influencing" of the followed car.

The experiences in practice were positive. Most of the drivers (between 90 - 95%) payed hardly any attention at all to the Floating Car which was following. The reason why is that people driving through a zone are mostly in a hurry, or they are looking for a parking place. It also depends on the spot where the car-following measurements are started. The possibility of influencing might be less if the spot the Floating Car is positioned before starting the measurement is not conspicuous. It is prefarable to start the following measurement coming from an side street.

To survey the "influencing" and to which extent, a study was started in cooperation with TNO Institute for Perception and De Meent Office for Research and Advice. The main goal of this study was to find the differences in speed behaviour between drivers followed by a Floating Car and drivers who are not. The variables were: the length in following-time between the two cars (3 and 4.5 sec) and the length of the route (250 - 450 m). The study was carried out in two days. On the first day the speed, the number-plate and the circumstances were registered on several spots in a residential area. On the second day the drivers registered on the first day were followed by a Floating Car (total number: 30). The same items were registered. The hypothesis was: "drivers show the same speed behaviour under the same circumstances". Change one of these circumstances ("following") and you know the influence of following on the speed behaviour.

The concept result of this study (Meent, 1988) is;

- There is no significant difference in speed behaviour between drivers followed by a Floating Car and drivers who are not.

This result agrees with one of the conclusions of a evaluation study in Lelystad (AGV,1987).

RESULTS

The method "Floating Car" has already been used in the experiments in 30 km/h zones and in some other experiments, like the evaluation of roundabouts, and a road through a shopping centre. The results of 4 experiments in 30 km/h zones are:
After the reconstruction the average speed in the total zone decreased from 28 km/h to 24 km/h. The speedreduction measurements had the biggest effect on the speed of through traffic. The reduction of the speed of bound traffic was less. The average speed of through traffic is 3 km/h to 4 km/h more higher than the speed of bound traffic. Before the reconstruction maximum speeds of more than 60 km/h were usual. After the reconstruction an incidental speed of more than 50 km/h still exists. In two of the zones however the maximum speed never exceeds 45 km/h.
The traveltime in the zones increased slidely after the reconstruction. The ratio between through- and bound traffic agree with the results of the other studies in the same experiments.

DISCUSSION

After the studies carried out by TNO and de Meent and after the several experiments in 30 km/zones we conclude the "Method Floating Car" can be used succesfully in evaluation studies on speed behaviour in areas, for measuring effects of reconstructions in routes and parts of routes. Conclusions about speed behaviour, route choice, fuel consumption , the "character" of the trip, travel distance and time delay can be given.
The method isn't accurate enough for the evaluation of for example the speed behaviour nearby a traffic humb. For these kind of measurements it is necessary to develop an instrument for instantaneous distance measuring. With this instrument it will be possible to convert differences in speed of two cars.

REFERENCES

AGV (1987). Reconstructie kruispunt Polderdreef/Atol in Lelystad. Adviesgroep voor Verkeer en Vervoer, Nieuwegein, 1987.

Bakker, M.G. (1985). De ontwikkeling van een meetmethodiek, waarmee het snelheidsgedrag van automobilisten in woongebieden kan worden vastgesteld. In: Verkeerskundige Werkdagen 1985. Studie Centrum Verkeerstechniek, Driebergen-Rijsenberg, 1985.

DHV (1987). Verwerking meetgegevens Floating Car; onderzoeken Den Bosch, Heerde, Zuidwolde, en Deventer. DHV Raadgevend Ingenieursbureau BV, Amersfoort, 1987.

Kleinloog, J.W. en Velthoen, R.A.L. (1981). Floating Car Registratie Systeem. Technisch Physische Dienst TNO, Delft, 1981.

Meent, de (1988). Onderzoek naar beinvloeding van bestuurders van personenauto's door achtervolging met de Floating Car. Advies en onderzoeksburo de Meent, Boxtel, 1988.

Tenkink, E. en Hoekstra, W. (1986). Een verkennend onderzoek naar de nauwkeurigheid van snelheidsmetingen met de Floating Car. Instituut voor Zintuigfysiologie TNO, Soesterberg, 1986.

A NEW METHOD FOR TRAFFIC SAFETY RESEARCH ON DRIVER DISTRACTION

G. Blikman
Delft University of Technology
Department of Transportation Planning and Highway Engineering
Traffic Safety Division

ABSTRACT

A traffic safety research method is introduced based on the the three driving tasks: tracking (keeping course), accounting for the other traffic and the traffic environment, and following a planned route. In this method accountance is made for the characteristics of unsafety and the way in which drivers execute these tasks. As unsafety is the result of a combination of the most critical circumstances, safety research should not be conducted under average conditions or with test subjects representative for the entire population. Instead, the most critical situations should be tested.
A laboratory research on the possible distraction caused by in-car visual route guiding advices serves as an illustration.

1 INTRODUCTION

This paper describes a newly developed method for determining the effects on traffic safety of man-machine interfaces, based on the driving tasks. This method, that also could be applied for researching the effects on traffic safety of out-car elements of the traffic system that influence driver behaviour, is illustrated with a research on the possible distraction of a driver caused by the presentation of in-car visual (route guiding) information.

2 RESEARCH METHODS

2.1 Introduction

Researches on the distraction of a driver by e.g. man-machine interfaces often are carried out by either measuring the mental load of a driver or by determining the deviation of the lateral position of a test car. The most optimal interface then is considered to be the one that requires the least mental load or causes the smallest deviation of the lateral position on a sample of representative test subjects.
However, from a traffic safety point of view both methods do not entirely justice to the tasks a driver has to accomplish during a journey. Furthermore it is questionable whether test subjects should represent the entire population.
This motivated the development of a new research method, which will be presented in paragraph 2.3. It is based on an analysis of the driving tasks that is treated in paragraph 2.2.

2.2 The driving tasks

Car driving is a triple task (1,9,10); under all circumstances a driver has:

- to keep his vehicle on the road (tracking)
- to account for the other traffic and the traffic environment
- to follow his planned route

These tasks are described briefly in the following.

The tracking task

Course holding is not and can not be carried out by keeping a straight line. Due to the vehicle characteristics, vehicles slowly deteriorate from their straight line and generally drivers accept such a deterioration until they come too close to the borders of their lane. Then, drivers correct their course towards the centre of the lane. Thus they drive along zigzagging.

This strategy of drivers implicates that drivers have at their disposal a certain time period between two successive course corrections, depending on the speed of the car, the heading angle, the lane width, and the car width (2).

The tracking task is carried out almost automatically and only requires a very low attention level.

Thus, to test whether a driver is distracted from his tracking task by e.g. a man-machine interface, it is sufficient to check whether the time during which drivers are visually occupied by the man-machine interface, is larger than the time they have, under certain circumstances, between two successive course corrections, which times can be derived from the existing literature.

The task to account for the other traffic and the traffic environment

Within the driver task to account for the other traffic and the traffic environment, distinction should be made between two entirely different situations of the traffic process: the encounter situation and the incident situation.

In the encounter situation a driver meets an approaching vehicle, a pedestrian, a bend in the road, and the like. These are discontinuities that normally approach the driver gradually and leave him ample time for a decision and an adequate reaction.

An incident situation does not leave the driver much time for an adequate reaction. After the sudden detection of a danger, followed by the recognition of the danger, e.g. an other vehicle on a collision course, the driver has left very little time to successfully carry out an emergency manoeuvre. Mostly, the driver reacts instinctively, usually through a combination of braking and

swerving.
The task to account for the other traffic is not carried out automatically. It requires a cognitive processing: a driver has to take decisions as whether to stop or to accalerate, whether to give or to take right of way, and the like.

In the driver task to account for the other traffic and the traffic environment peripheral vision is very important.
There is a considerable difference between the peripheral task in the encounter and in the incident situation. In the encounter situation the peripheral field serves to detect an object that approaches during a long time and that only slowly reaches it's full distinction. In the incident situation the object is mostly seen very shortly, often almost as in a flash, but with full distinction.

The task to follow a planned route
The driver task to follow a planned route requires relative little time, but may under circumstances consume the larger part of a driver's attention (e.g. when searching for street name signs during darkness and rain). In it self this driver task is not very critical to traffic safety: only if excessive detours cause a larger contribution and a longer exposure to the traffic risks, the traffic safety will be affected negatively.
More important, however, is that the execution of this task may disturb the safe execution of the other two driver tasks: a driver in search of his route may be inclined to neglect his tracking task and his task to pay attention to the other traffic. New electronic route guiding systems that assist the driver in the following of his planned route therefore have a large potential in reducing this part of traffic unsafety.

2.3 Research method based on the driving tasks

From the above paragraph, it follows that mental load is not the most optimal criterion for the affection of the driving tasks. For, it is not desirable that the mental load is always very small. Indeed, during the tracking task the mental load or the attention level required to carry out the tracking task should be small in order not to exhaust the driver, but in the task to account for the other traffic and the traffic environment (especially in incident situations) a large mental load is inevitable to be able to successfully carry out the required manoeuvre. Accidents often just seem to happen during moments of a low attention level of the driver. It also easily can be seen why the deviation of the lateral position is not optimal either: it only accounts for the tracking task, but not for the task to account for the traffic and the traffic environment, which probably is far more critical from a traffic safety point of view.

In the light of the above, a different approach is proposed to test whether a driver is distracted from his driving tasks. Instead of indirectly measuring the distraction through variables of which the relations to the driving task are not exactly known, the distraction is to be measured directly through the performance of the driver on his driving tasks.
As any distraction from the tracking task can be determined via the times during which drivers are visually occupied by e.g. a man-machine interface, the attention can be focussed on a distraction from the task to account for the traffic and the traffic environment.
A verdict on the tested man-machine interface then can be given on the basis of a comparison of the driver performance during the times that the driver was occupied by the man-machine interface and the driver performance when no attention was paid to the man-machine interface.
In chapter 3, the above method is illustrated with a research on the possible distraction of a driver caused by the presentation of in-car visual route guiding information.

2.4 Characteristics of unsafety its consequences for research

Unsafety always occurs in exceptional situations and is the result of a combination of critical circumstances (1). Under average conditions accidents hardly ever occur. Average drivers in well maintained cars on roads with sufficient skidding resistance, in dry weather, etc.

only have small chances on an accident. But these chances increase considerably if e.g. the attention level of a driver is low as a result of use of alcohol, drugs or medicine and / or the tyres of the car have insufficient profile, the roads are slippery, and it is raining heavily.
As a consequence of the above, traffic safety research should not test average situations or test subjects representative for the entire population, but most critical situations and representatives of that part of the population that is most critical to the tasks to be executed in the tests.
If the most critical circumstances can be dealt with safely, the less critical situations surely will not encounter any difficulties.

The above requirements also have been incorporated in the illustrative research that is described in chapter 3.

3 DISTRACTION BY IN-CAR VISUAL ROUTE GUIDING PICTOGRAMMES

3.1 Introduction

At the moment many new electronic in-car systems are envisaged that should provide the future driver with additional information. One of the first of these systems to come onto the market is an electronic route guidance system like Philips' Carin system.
In this Carin system the route guidance information is presented to the driver primarily aurally for reasons of traffic safety, as the driver task is already heavily loaded with visual information. For cases in which the advice has not been heard or understood, the aural advice is supported by a simple visual route guiding pictogramme on a dashboard-mounted flat panel display. These pictogrammes are schematic representations of junctions and routes to be followed, comparable to the pictogrammes on sign posts.

A route guiding system like Carin could, after implementation into a car, cause a number of effects on traffic safety, both positive and negative, in various situations of the traffic process. It lies outside the scope of this paper to discuss all possible effects that have been studied extensively (3,4,5). This contribution will be limited to the research on a distraction from the driving tasks that may be caused by the visual route guiding pictogrammes.

Of importance for traffic safety is the question whether the presentation of visual route guiding messages which assist a driver in his task to follow a planned route, prevents him from performing his other two tasks.
This has been studied in a series of preliminary laboratory experiments that are to form the basis of real world studies. In this study the attention was focussed on a distraction from the task to account for the traffic and the traffic environment (see also 2.3).
It was taken into account that a driver with a choice between concentrating on a route guiding advice and avoiding an incident or accident, will choose the latter alternative. Of importance therefore was whether an encountering vehicle or an incident is detected less well when a driver is paying attention to a route guiding message.

3.2 Experimental setting

In the laboratory experiments no aural route guiding advices were given as the situations in which drivers do not hear or understand anything of the aural route guiding advice can be marked as most critical and should therefore be tested (see 2.4).

During the experiments the test subjects were seated behind the steering wheel of a car from which the upper half had been removed to eliminate any dead angles and reflections in the windows.
The whole experiment was controlled by a personal computer (IBM-AT).

In the laboratory the driving tasks were simulated as accurately as possible.
The visual route guiding information was projected on a small display that was mounted into the dashboard.
To test whether the presentation of visual information distracted a driver from his task to account for the other traffic, the peripheral reaction capability during the presentation of pictogrammes was

compared with the achievements when no pictogrammes were shown.

To check whether subjects were actually watching a presented pictogramme, the head and eye movements of subjects were registered through a video camera.

In the set-up of the tests and the evaluation of the results, differences in the individual subjects (age, sex, level of education and profession) and the subjects' experiences with the pictogrammes have been accounted for.

The testing conditions will be discussed in further detail in the following paragraphs.

3.3 Requirements

3.3.1 Task to account for the other traffic and the traffic environment

In the task to account for the other traffic and the traffic environment, a driver has to watch and react upon what he sees, in which task peripheral vision is most important. In the laboratory this was simulated by requiring test subjects to registrate the random lumination of one of a set of peripheral lights that were situated in a semi-circle round test subjects at a distance of approximately 5 m and a height of ± 1.1 m. Behind the peripheral lights wooden fences were placed to create a neutral, non-reflecting background.
The angles of the peripheral lights with the line of sight remained constant at 80°, 50°, and 20° with the line of sight, which corresponds with the most critical situations in real traffic (when two vehicles are in a collision course the angles remain constant also).

Distinction between the encounter situation and the incident situation was created through different lumination patterns of the peripheral lights and different required reactions on those luminations.

In the **encounter situation** a driver has to watch the traffic that approaches him gradually with often low relative speeds. These conditions could be simulated by having yellow peripheral lights gradually reach their full brightness, followed by a lumination of some seconds. Thus, just as in real traffic, the illumination level of the encountering object on the driver's eye is increased slowly to it's maximum level.
To be able to register even small differences, the maximum lumination level was adjusted in such a way that under normal circumstances the luminations just could be detected and that in more extreme situations a part could be expected not to be noticed.

Subjects were required to react on the lumination of a yellow light by pushing the horn lever situated in the centre of the steering wheel.

In an **incident situation** drivers do not immediately recognize the danger of an accident. At first there is only the sudden detection of e.g. a vehicle in the corner of an eye, immediately followed by a head turning and a foveal recognition of the vehicle being on a collision course. In the laboratory this was simulated by a set of two lights which were luminated shortly after each other: the flashing of the yellow light had to draw the attention of a driver, while the lumination of a small red light next to it required a foveal recognition task.
The brightness of the red lights had to be so small that a lumination could not be detected peripherally, but had to be recognized foveally.

Not always the flashing of the yellow lights was accompanied by the lumination of the small red light very near the yellow one. For this could have induced test subjects to react only on the flashing of the yellow lights without taking notice of the red light.

Subjects were to react on the flashing of a yellow light and the subsequent lumination of a red one by pressing the brake pedal. When the flashing of a yellow light was not succeeded by the lumination of a red light, subjects were required not to take any action.

3.3.2 Tracking task

Although any distraction from the tracking task under varying conditions was no part of the tests, the subjects were to perform a tracking task during the tests in order to create realistic testing circumstances.

The tracking task was simulated through a traffic simulator. It was decided to execute the simulation through a great schematisation.
Generally speaking, the simulation of the tracking task should either be very simple or an exact copy of the reality. A compromise between those two is not to be preferred as this might induce test subjects to consider the simulation an exact copy of reality and behave as such. The slight deviations from reality then may greatly influence the test results.
The schematized tracking task should, just like in real traffic, be able to be conducted automatically and learned fastly, and deteriorations from a straight line should be allowed. Test subjects only were to keep their vehicle within a lane.
A schematisation that met these requirements was the task to keep a figure on a monitor between two paralel lines (figure 1). The dimen

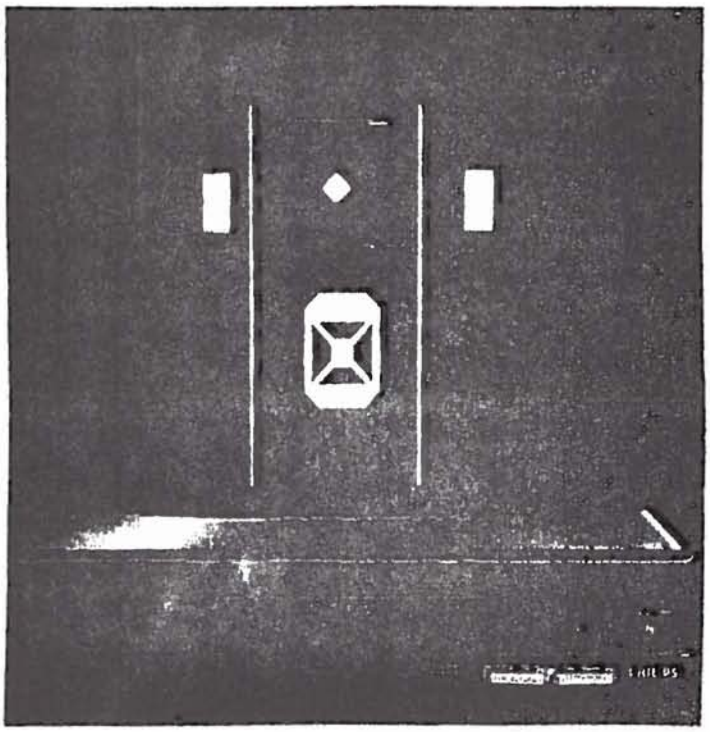


Figure 1:The simulation of the traffic task

sions were calculated thus that the ratio between the width of the figure to be controlled and the distance between the two paralel lines was the same as in real traffic.

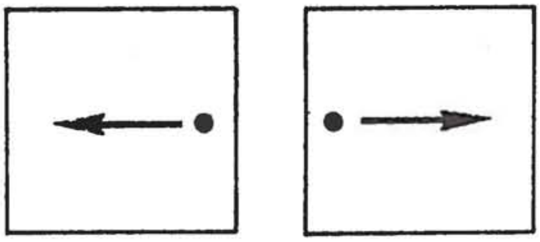
Of course the execution of the simulated tracking task showed differences with the real world tracking task (in the simulation e.g. hardly any use was made of peripheral sight). This was justified as the tracking task was no subject of the test. It only served to occupy the test subjects with a perceptive-motorial task.
The monitor was situated at approximately 5 m in front of the test subjects in order to enable the tracking task to be executed with unaccomodated or hardly accomodated eyes, in accordance with real traffic situations (8).
The tracking task was executed through a steering wheel. The steering apparatus was such that there was no exact neutral position at which the lateral speed was zero. For, in real traffic vehicles also have a natural drift that causes a small lateral speed to the left or to the right. Therefore the figure on the screen was given a minimal lateral speed.

The effects of speed were accounted for by relating the lateral speed to the longitudinal speed, which could be adjusted by the test subjects themselves through means of an accelerator pedal in order to enable them to choose their own speed and stress level, just as in real traffic situations.
To be able to check whether the tracking task was executed correctly during the experiments, all lane excessions were registered.

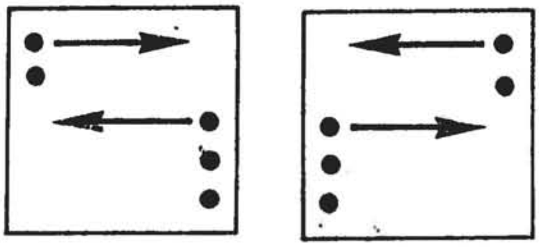
3.3.3 Route guiding messages

It turned out to be difficult to use real route guiding advices in the simulated tracking task (messages like "Take the second exit at roundabout" had no meaning in the schematized tracking task).
Therefore, the route guiding pictogrammes were schematized. In order to provoke a simular reaction as in real traffic, these schematized route guiding messages should contain information on the action that

has to be undertaken and on the place where that action has to be undertaken.
To indicate the place where a route advice has to be followed, dots were generated that moved from top to bottom over the monitor. The velocity of these dots simultaneously were an indication of the speed at which was driven.
Examples of a simple and a complex schematized pictogramme and their meaning are presented below.



Drive left/right at the first dot[1]



Drive left/right at the third dot and then right/left at the second dot

Subjects were free in their choice to watch a pictogramme. Therefore video-recordings were made of head and eye movements of the test subjects in order to check whether a presented pictogramme was actually watched.

3.4 Variables

3.4.1 Introduction

During the testing of the peripheral reaction, allowance was made for

---

[1] This message requests subjects to abandon the tracking task and bring the figure on the monitor that is controlled through the steering wheel to the left/right paralel line when the first of the dots that move from top to bottom has reached the controlled figure.

the variables that could influence the results of the experiments. These variables, that were either neutralized or were tested in their most extreme manifestations, are mentioned below and will be discussed in the following paragraphs.

- experience with the peripheral reaction task
- complexity of the peripheral reaction task
- experience with the tracking task
- complexity of the tracking task
- experience with the route guiding pictogrammes
- complexity of the route guiding pictogrammes
- characteristics of the test subjects

3.4.2 Experience with the peripheral reaction task

The peripheral reaction task was to simulate encounter and incident situations in real traffic. In both these situations a driver does not have to think long over the action to be undertaken in order to avoid an incident, respectively an accident. This also had to be the case in the laboratory experiments. Therefore the test subjects were given ample opportunity to practise the required reaction on the lumination of the yellow and red peripheral lights.

3.4.3 Complexity of the peripheral reaction task

The peripheral reaction capability was tested for three lumination patterns (one to simulate encounter situations, one to simulate incident situations, and a dummy pattern) in the ratio 1 : 1 : ½. The reaction was tested with six peripheral lights, set at 20°, 50°, and 80° on either side of the line of sight. As vision is symmetrical (8), it was justifiable to examine the results of the right and left light at each angle with the line of sight together, leaving three levels of the variable to be researched.

3.4.4 Experience with the tracking task

Test subjects had to be able to perform the tracking task automatically. To meet this condition subjects were to practice the tracking task before the tests until their skill was sufficient.

3.4.5 Complexity of the tracking task

The complexity of the tracking task, which increases as the speed becomes higher, certainly influences the peripheral detection capability: when the attention largely has to be focused on tracking, the peripheral field may narrow. However, not the peripheral detection capability itself was of importance, but the influence of the presentation of pictogrammes on the peripheral detection capability.

Therefore it was of interest whether any deterioration of the peripheral detection capability was larger at high speeds than at low speeds.

To test the effects of speed, one group of subjects was allowed a higher maximum speed than the other ones (see 3.4.8).

3.4.6 Experience with the route guiding pictogrammes

Experience with the route guiding pictogrammes will result in shorter interpretation times and may also result in a smaller amount of cognitive processing required. Subjects who are experienced with a pictogramme will tend to gloss over familiar elements and concentrate on the changeable parts.

To prevent subjects from becoming experienced already in the first series of tests, the duration of the tests was limited for each subject to approximately three quarters of an hour.

3.4.7 Complexity of the route guiding pictogrammes

The complexity of the route guiding pictogrammes is the most important factor for the duration of the interpretation time (6,7,10) and could also have effect on the concentration with which it is watched, which in it's turn may affect the peripheral detection capability through a narrowing of the peripheral field.

The basic question that had to be answered was whether complex route guiding pictogrammes distract drivers more than simple pictogrammes.

Therefore it was sufficient to test a (series of) simple route guiding pictogrammes and a (series of) complex pictogrammes, thus leaving only two levels to be tested.

It should be remarked here that the complex pictogrammes to be tested still were relative simple with only two units of information, as in the Carin system pictogrammes were envisaged with at the most two units of information.

3.4.8 Characteristics of the test subjects

The characteristics of test subjects are of great importance for the way in which tasks are performed. For the experiments those test subjects were selected that were expected to be most critical to the execution of the tests.

The tasks that had to be executed in the experiments were tracking, watching the peripheral lights and reacting upon their lumination, and interpreting the route guiding pictogrammes and following a route. These tasks were to be executed mainly simultaneously and only partly sequentially.

Tracking task

Old people have more difficulty in tracking than young ones, but they compensate their infirmities by driving slowly. Young, especially male drivers show a tendency to speeding and taking high risks. This is confirmed by the accident rates: elderly people seldom or never are involved in accidents with only one vehicle. The major part of these accidents are caused by young, mainly male, inexperienced drivers.

Task to watch the peripheral lights and react upon their lumination

The peripheral detection capability deteriorates with age. This is a handicap even more as old people also have difficulty in turning their head, which otherwise could have compensated the deteriorated peripheral vision.

The subjects have to react by pushing a horn lever or by activating the brake pedal. Although the reaction velocity differs from one person to another and is dependent on temporary circumstances, it can be stated that it deteriorates with age (11).

Task to interprete the route guiding pictogrammes

Elderly people have more difficulty in learning something new and in remembering something that has been learned recently. As a result of this their level of experience with the route guiding pictogrammes will rise only slowly, if it rises at all (the pictogrammes seldom will be necessary to inform drivers).

The accommodation of the eyes, which is necessary to focus the eyes from a point in the distance onto the display, becomes more difficult as people grow older.

The level of education and the profession of subjects may play a role in the speed with which the pictogrammes are learned. It is not inconceivable that people with a low level of education and a profession in which little thinking is necessary experience greater difficulties in interpreting the schematized route guiding pictogrammes.

Multiple tasks

The ability to perform a multiple task deteriorates with a climbing of the years. Old people prefer to execute actions sequentially rather than simultaneously. This may lead to a concentration on the pictogrammes and a neglection of the peripheral reaction task.

Conclusions

Overlooking the above, three groups of people were, for different reasons, marked as possibly critical. These are:

I : young male drivers (aged 18 - 24)
II : elderly people (60+)
III: people with a low level of education and a profession that requires little cognitive processing

To enable a comment on these groups, they were compared with a reference group of non-critical drivers, consisting of:

IV : people of moderate age (35-45), of both sexes, and with a moderate to high level of education and a profession in accordance with the education

The number of test subjects was 24.

3.5 Procedure

The route guiding pictogrammes to be presented during each test were ordered randomly and consisted of 50% simple and 50% complex pictogrammes.
The time intervals between the presentations of the route guiding pictogrammes were varied randomly and given a uniform distribution with a minimum and maximum of respectively 15 and 45 s. However, when test subjects were not able to follow-up a route guiding advice within the planned time interval, this was registered by the computer and the presentation of the next route guiding advice delayed until the previous advice had been executed.

During each test the peripheral lights were luminated both when route guiding pictogrammes were presented and when no such pictogrammes were shown.
The events at which a lumination was accompanied by the presentation of a pictogramme consisted of 15 different combinations of the variables "situation of the traffic process", "angle with the line of sight", and "complexity of the route guiding pictogramme". Each of these combinations was tested twice during one experiment. In order to avoid a high level of expectancy, these 30 events were distributed randomly over the 80 presentations of a route guiding pictogramme.
The luminations at times when no route guiding pictogrammes were presented, comprised of 7½ combinations. Each of these combinations also was tested twice per test subject. These 15 tests were situated in the largest intervals between two successive pictogrammes.

Each of the 24 selected test subjects was send a description of the test in advance.
Before the start of the tests subjects were shown the set-up of the laboratory and the whole procedure was explained to them aurally. Subsequently they were given opportunity to study a sheet with the route guiding pictogrammes to be expected. They were instructed to accurately follow-up the route guiding advices (as they would have done in real traffic).
Before actually beginning the tests, subjects could practise the tasks to be performed. Only when they turned out to be able to control all tasks simultaneously, the real tests were started.

4 RESULTS

For all test subjects the reaction times in the incident situations were considerably shorter than in the encounter situation. The young males were the fastest to react; the elderly drivers the slowest.
Lane excessions hardly occurred in the tests, except on purpose when test subjects were following-up a route advice.

The results of the tests are displayed in the tables 1, 2, 3, and 4. Table 1 presents major increases of the reaction time averages (per group of test subjects and for all subjects). Incorporated in this table are reaction failures that have been attributed the maximum reaction time value of 5 s. Table 2 shows the statistical significant reaction time changes. In table 3 the reaction failures have been omitted from the results. These have been placed in table 4.

Overlooking the outcome of the laboratory experiments, the conclusion seems justified that the in-car presentation of pictogrammes under circumstances does lead to a deterioration of the reaction capability.
Most important for the extent of the deteriorating effects was the situation of the traffic process. In the encounter situation hardly any deteriorations occurred in contrast with the incident situation. It should be considered, however, that these statements are based on

Table 1: **Changes in reaction time averages larger than 15%**

Group	angle peripheral light	Encounter situations (Lumination pattern 1)		Incident situations (Lumination pattern 2)	
		Simple pictogrammes	Complex pictogrammes	Simple pictogrammes	Complex pictogrammes
I	80°			+ 32%	+ 63%
	50°			- 33%	+ 33%
	20°			+ 27%	+ 47%
II	80°			+ 30%	
	50°			- 30%	
	20°	+ 24%	+ 17%	+ 41%	+ 24%
III	80°				
	50°	+ 19%			+ 19%
	20°				+ 44%
IV	80°			+ 23%	+ 18%
	50°			+ 19%	+ 38%
	20°			+ 30%	
all	80°			+ 17%	+ 25%
	50°				+ 27%
	20°			+ 22%	+ 33%

Table 2: **Statistical significant ($\alpha = 0.05$) reaction time increases**

Group	angle peripheral light	Encounter situations (Lumination pattern 1)			Incident situations (Lumination pattern 2)		
		H_1 1)	H_2 2)	H_3 3)	H_1 1)	H_2 2)	H_3 3)
I	80°					+ 63%	
	50°						+ 100%
	20°				+ 27%	+ 47%	
II	80°						
	50°						+ 53%
	20°	+ 24%					
III	80°						
	50°						
	20°						+ 44%
IV	80°				+ 23%		
	50°						
	20°						

1) H_1: Reaction times at simple pictogrammes larger than without any pictogrammes shown

2) H_2: Reaction times at complex pictogrammes larger than without any pictogrammes shown

3) H_3: Reaction times at complex pictogrammes larger than at simple pictogrammes

Table 3: Changes in reaction time averages larger than 15% when reaction failures are omitted from the results

Group	angle peripheral light	Encounter situations (Lumination pattern 1)		Incident situations (Lumination pattern 2)	
		Simple pictogrammes	Complex pictogrammes	Simple pictogrammes	Complex pictogrammes
I	80°				
	50°			- 26%	
	20°			+ 27%	
II	80°			+ 50%	+ 22%
	50°				+ 37%
	20°				+ 24%
III	80°				+ 25%
	50°				+ 25%
	20°				+ 28%
IV	80°			+ 20%	
	50°				
	20°			+ 30%	
all	80°			+ 16%	+ 26%
	50°				+ 21%
	20°			+ 17%	+ 17%

Table 4: Changes in percentage reaction failures larger than 15%

Group	angle peripheral light	Encounter situations (Lumination pattern 1)		Incident situations (Lumination pattern 2)	
		Simple pictogrammes	Complex pictogrammes	Simple pictogrammes	Complex pictogrammes
I	80°			+ 22%	+ 33%
	50°	- 16%			+ 20%
	20°		+ 20%		+ 17%
II	80°	- 23%		+ 21%	
	50°	- 17%	+ 33%	- 24%	
	20°	+ 22%			
III	80°	- 19%			
	50°	+ 44%	+ 17%		
	20°				
IV	80°				
	50°		+ 16%		+ 20%
	20°				
all	80°				
	50°				
	20°				

the presentation of simple and complex pictogrammes that still were relative simple. A less adequate choice for the complex pictogrammes might well have resulted in deteriorated reactions in the encounter situation also.

The deteriorated reactions in the incident situation are very serious, as they can not be eliminated by changing some variables. At the most the deteriorations can be decreased slightly by an adequate selection of pictogrammes. For, the test subjects of the groups II and III (and only of those two groups) experienced a smaller deterioration at the presentation of simple pictogrammes.

The consequences of most deteriorations occurring in incident situations are severe, as in the incident situation life-saving emergency manoeuvres have to be executed timely and adequatly, whereas in encounter situations drivers still have left some seconds to react. Fortunately incident situations happen far less frequent than encounter situations.

Most deteriorations occurred when the stimulus was given from a peripheral light at 20^o with the sight line. Presumably this is caused by the good sight drivers normally have over that location, an advantage that is lost when drivers watch the dashboard. Especially elderly drivers suffer under a large deterioration of the reaction capability for stimuli from lights at 20^o. This is expected to be related to the deteriorated peripheral vision of elderly people.

The larger part of the traffic a driver has to account for, approaches from an angle of some 20^o-50^o with the sight line. A larger sensitivity to deteriorations for stimuli from $\pm$ 20^o would therefore be most unhappy.

A possible way to improve the performance for stimuli from 20^o, could be a higher positioning of the pictogramme display in the car.

The relation between the complexity of the pictogrammes and the extent of the deteriorations of the reaction capability was not as manifest as expected. This may have been due to the high reality level of the complex pictogrammes that still were relative simple. More complex pictogrammes might well have caused larger deteriorations.

In normal situations the tested complex pictogrammes seem to cause no

extra deterioration of the reaction capability, but under difficult circumstances they do. Thus, the complexity of pictogrammes turned out to have no effect for the test subjects of the reference group, but a moderate effect for subjects of groups I, II, and III in incident situations.
The results lead to the conclusion that the pictogrammes should be very simple and quickly interpretable, as even the tested complex pictogrammes caused, under circumstances, larger deteriorations of the reaction performance than simple pictogrammes. It should be kept in mind though that even the simple pictogrammes caused deteriorations.

The experiments show that not all drivers are equally distracted by the presentation of pictogrammes.
Young males most often demonstrated a deteriorated reaction performance, resulting not so much in longer reaction times as in more reaction failures. Cause for the deteriorated reaction capability presumably is the higher task load as a result of the self-selected higher speed, in combination with the inexperience of the subjects.
Elderly drivers showed a deteriorated reaction performance (that was already low without pictogrammes being shown) at complex pictogrammes in the incident situation and for stimuli coming from lights at 20°. The latter may be related to a typical characteristic of elderly people: a deteriorated peripheral vision.
Drivers with a low educational level and a profession that requires little cognitive processing experienced few deteriorations of the reaction capability. They seemed to be affected especially by the presentation of complex pictogrammes in the incident situation.
Subjects of the reference group did not prominently distinct themselves on one or more of the aspects that may influence the deterioration of the reaction performance. In the incident situation they experienced most major reaction time increases.

Although the laboratory experiments demonstrated that, under circumstances, the presentation of pictogrammes does lead to a deteriorated reaction capability and therewith to a decrease of traffic safety, this should not immediately lead to the conclusion that visual route guiding displays therefore must not be build into cars. The reasons herefore are threefold.

First, when aural informations serves as a prime information source, a driver seldom will have to use the visual route guiding information.
Second, the pictogrammes cause most deteriorations at relative seldom occurring situations, when an incident arises just when a driver is attending to his display.
Third, the distraction by visual information displays is only one of a whole series of possible positive and negative effects on traffic safety of route guiding systems. An enormous advantage e.g. is that search behaviour with the absence of a navigation system (searching for street name signs, and the like), that probably is more distracting than the presentation of pictogrammes, can be avoided. The distraction caused by visual information displays will have to be balanced with all other effects of the system.

Furthermore, the laboratory experiments were indicative. It is recommended to verify the results in any form of real world tests.

REFERENCES

1. Asmussen, E.
   Inleiding Verkeersveiligheid
   Delft University of Technology
   Delft, the Netherlands, 1987

2. Blaauw, G.J.
   Car driving as a supervisory control task, pp 139
   Graduating paper for a doctor's degree
   Instituut voor Zintuigfysiologie TNO
   Soesterberg, the Netherlands, 1984

3. Blikman, G.
   The traffic safety of the Carin car information and navigation system IA - Literature study: Main report, pp 80
   Delft University of Technology
   Delft, the Netherlands, 1988

4. Blikman, G.
The traffic safety of the Carin car information and navigation system IB - Literature study: Description of reports and articles, pp 126
Delft University of Technology
Delft, the Netherlands, 1988

5. Blikman, G.
The traffic safety of the Carin car information and navigation system II - Distraction by in-car visual route guiding pictogrammes, pp 96
Delft University of Technology
Delft, the Netherlands, 1988

6. Easterby, R.S. & Cox, D.E.H. & Hughes, A.W.
The perception of variable message symbolic road signs: reaction time studies, pp 49
University of Aston in Birmingham, Applied Psychology Department
Birmingham, England, 1977

7. Erke, H. & Richter, B. & Richter, S.
Psychologische Untersuchungen zu Unfällen auf innerörtlichen Strassen im Zusammenhang mit Merkmalen der Vorwegweisung.
Köln, Federal Republic of Germany, 1974
in: Forschungsheft zur Verkehrssicherheit 20 (1974) nr.2.

8. Grambergen-Danielsen, B.
Sehen und Verkehr
Heidelberg, Federal Republic of Germany, 1967

9 Holloway, C.H. & Wright, P.
Assesment of the attentional demands of road information transmitted aurally, pp 28
Medical Research Counsel, Applied Psychology Unit, Cambridge
Cambridge, England, 1980

10. Jansen, W.H.
Routeplanning en -geleiding - een literatuurstudie, pp 51
Instituut voor zintuigfysiologie TNO
Soesterberg, the Netherlands, 1979

11. Triggs, T.J. & Harris, W.G.
Reaction time of drivers to road stimuli, pp. 97
Monash University, Department of Psychology
Melbourne, Australia, 1982.

SIMULATION OF CASUALTIES IN PERSON TRANSPORT SYSTEMS

Åke Forsström
Department of Injury Prevention
Chalmers University of Technology
Göteborg
Sweden

1 RESEARCH PROGRAM

1.1 Problem area and research findings

Simulation of casualties in person transport systems may be regarded as belonging to the research territory of transport safety. It is, however, a part of both a social and a technical problem area. The first area concerns questions of accidents and injuries, which is a part of the social pathology component among criteria of social well-being. Transport facilities are on the other side a general socio-economic well-being component (Smith 1973). The second area deals with various technical means of travelling, the generated traffic, the network and the organization that is needed to support a social and spatial structure. This problem area involves several research territories such as transport technology, traffic and urban planning and human geography (Dicken and Lloyd 1981).

Safety against injuries is a highly relevant issue in the industrial and transport society of the western world. Injuries cause around 40 per cent of all lost years of future worklife in the United States. The rest is due to the combined effects of all diseases together (Morb. Mort. Weekly Rep. 31:599 1982). In Sweden the corresponding figure is around one third (Causes of Death 1984. SOS).

> "Injuries are the leading cause of death and disability in children and young adults." *Injury in America A Continuing Public Health Problem (1985). p. 1.*

In the age interval up to 25 years more than half of the lethal injuries are related to transports. Transport accidents accounts for more than a third of total injury losses of future worklife in the whole population or one tenth of all death-caused losses (Causes of Death 1984 SOS). A historical study of death-causes from 1860 to 1960 shows that various transport systems take their contribution of accidents and deaths at different times. To-days dominant cause of death, motor vehicle accidents, has replaced drowning accidents from its first position during a period from 1920 to 1960. Neither horse nor railway accidents have been a major cause of violent death in Sweden (Odeén 1977).

The problem area is not limited to violent deaths. The lack of statistical sources prevent however the dealing with permanent and non-permanent disabilities. A recent national sampling survey of accident patterns in Sweden estimates the total number of casualties per year to 700,000 cases. They are distributed by five categories. An interval of 75-80 per cent occur in residential areas , in school and during leisure time, while 15-20 are occupational accidents and 4-12 per cent happen in the traffic (road traffic accidents). Distributed by activity, a non-overlapping category as "walking and running", is the number one cause of injuries (15 per cent). This category excludes all sports and playing activities (Danielsson 1986.)

It is possible to calculate the number of transport casualties by cross tabulating "walking and running" with transport elements such as road, street, pavement etc, and the amount of road traffic accidents with personal injuries. This system-oriented and wider concept of transport accidents thus includes nearly 60,000 road traffic cases in the total sum of more than 130,000 casualties (see app. 1). This estimate can be compared with another calculation based on a study of commuting accidents, with a figure of 150,000 transport casualties. The road traffic accident share in this study is 56,000 casualties (Forsström 1982).

1.2 Terms of concepts

Obviously, the road traffic accident concept is too narrow to cover all accidents that occur during travels and other personal movements. This term was once created to evaluate the malfunctions of the motor vehicle system. It is thus not possible to judge the efficiency of different person transport systems within this frame . Existing accident report systems of public means of transportation take only causalties during travel into account . In order to cover a chain of transport movement it is advisable to include walking accidents and casualties. The importance of walking accidents as a part of the total number of casualties during a bus ride was first noticed by Jørgensen, Led och Lewy (1974).

Three new terms are suggested. A person who is slipping, stumbling and treading when an accident occurs without interference from any traffic element is engaged in a **walking single accident.** An accident emanating from the interference of two pedestrians is a **walking accident.** These accident types together with all other vehicle accident types comprise the total set of **person transport accidents.**

1.3 The problem

It has been proclaimed with certainty that society would benefit if people would travel more extensively with public means of transport. Instead the car transport system with vehicles, streets, roads and traffic signs and signal systems is the dominating mode. The consequences of this dominance are numerous deaths and injuries in traffic. There have been arguments for a transfer of commuters from car to public transport systems in order to save lives and health. This proponents consider the society to be suboptimally organized. (Näätanen and Summala 1976, Whitelegg 1982, the Swedish government proposition 1982/85:50).

Is this view really irrefutable? No it is not. First, the numerous losses in motor vehicle traffic is a consequence of the massive exposure of the vehicle in the person transport system; second, the experience of losses are based on an inadequate accident concept; third, there are studies with contradictory conclusions. One cannot exclude the possibility that, at the actual distribution of passengers on cars and on public means of transport, the relative societal health losses are fairly equal (Forsström 1982).

The problem of this paper can be formulated in the following manner: What is the likely development of losses during person transports when the distribution of passengers in an urban region changes between cars and public means of transport?

The purpose of this paper is to suggest steps in the development of a simulation model which describes these changes.

2 AN ALTERNATIVE THEORICAL APPROACH

An alternative theoretical approach may be classified as a transport safety theory as opposed to traffic safety theory. A dominant feature of the latter theory is the opinion that interference between different vehicles is a major complex of accident causes. Transport safety theory makes the opposite emphasis. Single accidents of all kinds are considered relevant to safety, no matter how unimportant they seem to be. The philosophy of this approach is that frequency counts.

A transport safety conception have its roots in system theory and in economics. In order to study the malfunctions of the motor vehicle system it is not enough to delimit that system and observe it. It is also necessary to treat involved lateral person transport systems, such as walking and going by bus, tram or underground.

From an economic point of view costs are related either to firm or to society. These cost concepts must be completely kept apart. One cannot evaluate societal costs with a firm-related cost concept. The road traffic accident concept has a limited substance in this respect. Misapprehensions regarding this concept have probably led economists to use it in many reports considering the societal cost of (road) traffic accidents. In the discussions following those reports the participants almost always place road traffic on an equality with traffic or transport. The conclusions of the participants invariably involve misunderstandings of the real losses from a change of systems regarding travelling.

A societal cost approach would imply that the transport accident concept is supplemented by the occupational accident concept. Thus, the safety of each transport subsystem is evaluated from three angles:

(1) The risk to injure the passengers of its own sub system
(2) The risk to be injured by the vehicles and passengers of other subsystems.
(3) The risk that the staff of its own subsystem is injured.

The first risk consists of two components. The safety of a passenger category depends on the risk of being injured in a single accident and also of coming into interference with a vehicle from its own subsystem. The second risk have as many compnents as there are further transport subsystem or modes. The injury risk of the staff is an occupational risk and is related to the time of exposure.

The risks are later transformed to number of casualties and evaluated by the severeness of injury.

3 FLOW MODELS OF TRANSPORT SYSTEM CASUALTIES

3.1 Smeed's traffic safety world.

In Smeed's famous formula from 1949 the population at risk, the exposure to motor vehicles and the expected number of deaths in traffic are related to one another in a certain relationship. Already the participants of the discussion, which is reported in connection with his paper, express a critical opinion of the formula validity. This criticism is resumed many times. Smeed responds as late as 1972, and concludes, based on new data, that the formula "still gives a fairly good estimate of the number of deaths in most countries". (Smeed 1982, p 311). Haight (1987, p 13) maintains that there are "numerous data sets which do not come near to satisfying the formula". Although, he admits it to be a really important early achievement, he insists that it lacks a theoretical base.

This paper aims at an organization of some theoretical elements of transport safety. This attempt starts partly with Smeed's formula. Attention is , however, not primarily directed to time series analysis or the exponent values. It is rather the nature of the relationships between population at risk and accident risk factors which are at the focal point. Forrester's (1969) early study of the urban crises in the United States serves as one example of systems theory application to complex dynamic social and technical processes. The processes of transports and the consequent accident processes are treated in this paper also with a systems approach. The elements of Smeed's formula are transformed to properties in a system model, such as stocks, flows, feedback loops inclusive goods and programmed activities.

Smeed's formula in his paper from 1949 is expressed as

$$D = 0.0003 (N \cdot P^2)^{1/3} \qquad (1)$$

where D = the number of traffic deaths in a country, P = the number of inhabitants and N = the number of cars.

The formula contains the constant 0.0003 which is replaced by k. It is rewritten,

$$D = k \cdot N^{1/3} \cdot P^{2/3}. \qquad (2)$$

The simplest way to develop the formula is as a flow diagram. The population P and the number of deaths D are represented by stocks which are joined by a flow of deaths per period (year) controlled by a flow regulator. It expresses the risk level by knowledge of the number of cars through an information line. Model 1.1 represents the formula.

MODEL 1.1

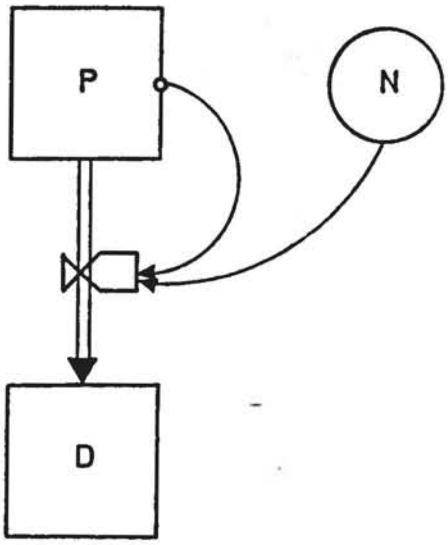


Fig. 1. A model of the traffic safety world according to Smeed's formula.

The flow model initiates some changes of Equ. (2). The amount of exposure to danger is approximated by the amount of person transports per year. The risk to be killed is dependent on the extent of vehicle traffic . That is,

$$D = f[TT(P), TC(N)] \qquad (3)$$

where TT(P) = the number of person transport kilometres and TC(N) = the number of vehicle traffic kilometres. Let $k = k_1 \cdot k_2 \cdot k_3$. Equ. (3) may be expressed and rewritten in Smeed's terms as

$$D = k_1 \cdot k_2 \, P^{2/3} \cdot k_3 \cdot N^{1/3}, \qquad (4)$$

with the following suggestions, k_1 = the risk of being killed in a traffic accident, k_2 = the number of person transport kilometres per year and k_3 = the number of vehicle traffic kilometres per year.

However, the model needs refinement. The risk of being killed in a traffic accident k_1 is not a single entity. It is preferable to distinguish between motor vehicle passengers and other passengers. The degree of motorization of the society is an expression of an actual transport policy. Model 1.1 is altered.

3.2 The dichotomy of traffic safety regarding motorists and non-motorists.

There are two stocks of population at risk, one of car drivers and car passengers P(N), and another of unprotected road users P(non-N). The rate of motorization N/P is continually increasing, perhaps it slows down when the number of deaths D is increasing. The two death rates are functions of two different risk levels, the amount of vehicle kilometres and of the number of person kilometres for car users and unprotected road users respectively.

MODEL 1.2

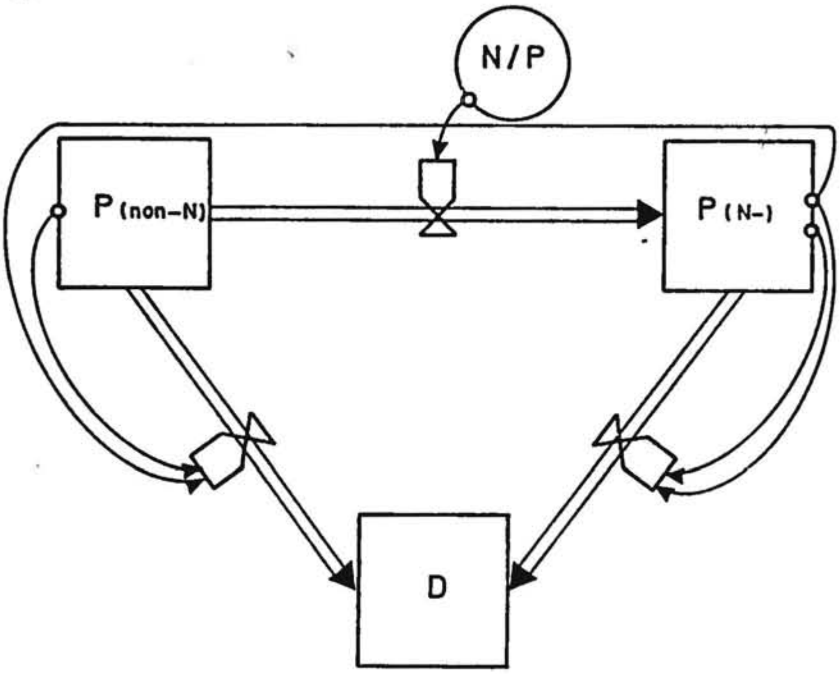


Fig 2. The refined model expressing Smeed's formula

The number of deaths in traffic is now a sum of two different types of accident processes. Equ. (3) is rewritten to show these two processes. The risk of accident is split into two different risk levels k_{11} and k_{12}.

$$D = f_1\{P(\text{non-}N)TT, P(N)TC\} + f_2\{P(N)TT, P(N)TC\} \qquad (5)$$

Smeed used the number of deaths as the only reliable measure of the malfunctions in the motor traffic system. A societal view of losses implies that all losses are included. The casualties, irrespective of severeness, are denoted by C.

A city region of considerable size, must be equipped with a public transport system in order to function (Thomson 1977). The transport policy of such a region is partly based on varying shares of passengers using car and public means of transportation (PMT). An other important element in this policy is the number of casualties produced by each system. Model 1.2 is thus transformed into a different shape, model 2.1.

3.3 The safety of a person transport system including car users and users of public means of transport.

This model allows for an evaluation of the casualties caused by the car system and the PMT-system, respectively. Each stock of casualties is a result of two different risks, which are dependent on the traffic produced by each mode in the transport system $(\rho \mid TC)$. The risks are exposed to the amount of transport in each system (TT).

MODEL 2.1

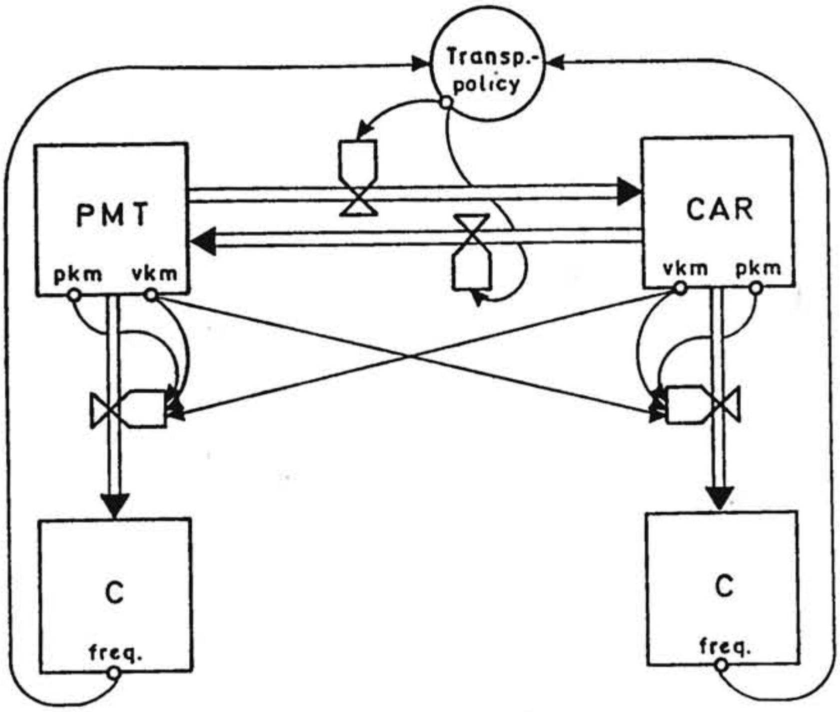


Fig. 3 A person transport system with two modes producing casualties.

Four equations show these relationships. The first two equations describe the losses among PMT-users, C(PMT).

$$C(PMT;Pmt) = \{\rho(pmt) \mid TC(Pmt)\} \cdot TT(Pmt) \qquad (7)$$

$$C(PMT;Car) = \{\rho(pmt) \mid TC(Car)\} \cdot TT(Pmt) \qquad (8)$$

Equ. (7) show the losses produced by the pmt-system itself, while the losses in Equ. (8) are caused by the car system.

$$C(CAR;Pmt) = \{\rho(car) \mid TC(Pmt)\} \cdot TT(Car) \qquad (9)$$

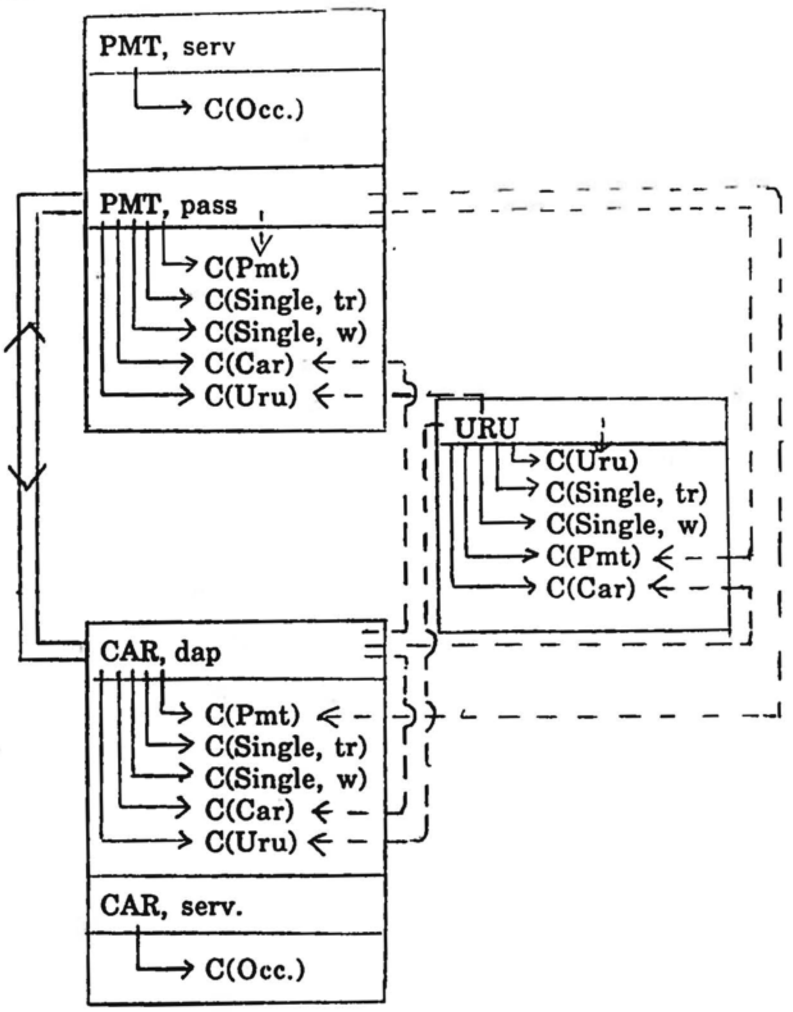
$$C(CAR,car) = \{\rho(car) \mid TC(Car)\} \cdot TT(Car) \qquad (10)$$

The losses of the car users are partly caused by the pmt-system (Equ. 9) and partly by the car-system itself (Equ. 10). These four equations describe only interference accidents. Single accidents remains to be described.

$$C(PMT;single) = \rho(pmt;\ single) \cdot TT(Pmt) \qquad (11)$$

$$C(CAR;single) = \rho(car;single) \cdot TT(Car) \qquad (12)$$

MODEL 2.2



Abbreviations:

C = Casualties
Pmt = Public means of transport
Pass = Passengers
Dap = Driver and passenger
Uru = Unprotected road user
Occ = Occupational accident
Serv = Service personnel
tr = Travelling
w = Walking

Legend-examples:

PMT, serv - Population at risk, service personnel maintaining pmt vehicles
C(PMT,Pmt) - Casualties among pmt-users in accidents with pmt-vehicles
——→ - Relation between population at risk and casualty category
⇄ - Flow of travellers in either direction due to changes in transport policy and behavior
- - → - Risk change due to shifting volumes of interfering traffic

Fig. 4. A diagram of the most relevant relations (risks) between the exposed category of population in terms of transport volume and the traffic volume of the interfering category of vehicle on one side and the amount of casualties on the other.

As indicated in section 1 and in appendix 1 single accidents with casualties are frequent during the walking moments of person transports. There are thus two types of single accidents, walking and travelling single accidents. This feature, the unprotected road users and the occupant injuries are included in model 2.2. A simplified version of the model is reproduced in figure 4.

3.4 The concept of transport safety replaces the concept of traffic safety.

The model consists of three transport systems. The system of unprotected road users are added to the car and the PMT systems. In each system five different transport accident processes with personal injury ococur. In the first process the system users are injured in interference with the vehicles of their own system. The second process produce single accidents during travels and the third single accidents during walking transport moments. In the fourth and the fifth process system users interfere with vehicles from each of the other two systems. One occupational accident process is added to the car and the PMT system, respectively. There are fifteen transport accident processes and two occupation accident processes. in the model. Nine of these processes depend on the amount of traffic produced in some of the other systems.

The outcome of casualties of each system is thus related not only to person transports of this system but also to the traffic produced by other systems. The person transports and the amount of traffic are in turn dependent on the distribution of passengers on the three transport system, which is the outcome of the transport policy at a given point of time. This complex web of relations forms a simulation model by which the unpredicteable total outcome of casualties during periods of varying transport policies might be anticipated.

4 SUMMARY AND CONCLUSION

It is concluded that it is possible to evaluate the societal losses of transport policy changes by using a transport safety concept in a simulation model. The losses are associated with each one of three major person transport systems in the model. Transport policy is expressed as a specific distribution of passengers on different transport systems. The transport safety concept is based on a judgement of accident risk on all moments of person transports including walking single accidents.

REFERENCES

Causes of death 1984. SOS. National Central Bureau of Statistics (1986). Stockholm: Liber.

Danielsson, K. (1986). *Increase Safety. Accident Patterns in Sweden 1982-1984.* National Board for Consumer Policies Report P1.

Dicken, P. & Lloyd, P.E. (1981). *Modern Western Society A Geographical Perspective on Home, Work and Well-being.* London: Harper & Row.

Forrester, J. (1969). *Urban Dynamics.* Cambridge Mass.: MIT Press.

Forsström, Å. (1982). *Commuting Accidents. A Study of Commuting Accidents and Casualties in some Swedish Regions During 1971.* The Departments of Geography, University of Gothenburg Series B No 69.

Haight, F.A. (1987). International cooperation in traffic safety. *IATSS Research*. Vol. 11 1987, New series No. 1.

Injury in America. A continuing Health Problem. (1985). Committee on Trauma Research Washington, D.C.: National Academy Press.

Jørgensen, N.O., Led, J.C., and Lewy, P. (1974). *Sökerhet.* Nordiska ministerrådets sekretariat, Nordkolt delrappport 4, juni 1974, Oslo.

Morbidity Mortality Report 31:599 (1982). Centers for Disease Control.

Näätänen, R. och Summala, H. (19796). *Road-user behaviour and traffic accidents.* Amsterdam-Oxford: North-Holland Pub. Co.

Odén, B. (1977). *Historiskt perspektiv på riskpanoramat i ett föränderligt samhälle.* The Project Risk Generation and Risk Assessment in a Societal Perspective, rep 13-77. Com. of Future Research, Stockholm.

Smeed, R.J. (1949). *Some Statistical Aspects of Road Safety Research. Journal of the Royal State Soc. Ser. A,* Vol CXII, Part I, pp 1-34.

Smeed, R.J. (1972). *The usefulness of formulae in traffic engineering and road safety. Accid. Anal. & Prev.* Vol. 4, pp 303-312.

Smith, D.M. (1973). *The Geography of Social Well-being in the United States.* New York: McGraw Hill.

Swedish Government. Proposition 1982/83:50.

Thomson, J.M. (1977). *Great Cities and Their Traffic.* London: Victor Gollan cz Ltd.

Whitelegg, J. (1983). *Road Safety: Defeat, Complicity and the Bankruptcy of Science.* Accid. Anal. & Prev. Vol. 15, No 2, pp. 153-160.

Appendix 1

INJURY PATTERN IN SWEDEN

The number of casualties by accidents in Sweden per year during the period 1982-1984 are estimated in a sample survey performed by the National Board for Consumer Policies. The sample contains 6,755 persons in five of Sweden's 24 counties. They were interviewed during two periods of two weeks.

There were 700,000 casualties among the Swedish population of 8.3 million inhabitants in this period. They are distributed by place of accident and by activity. Places of accidents are work, school, residential area, leisure facilities, and the traffic environment. Work includes all working areas and buildings. School is defined as yards, pathways and buildings within the school area. The residential area consists of home, yard, and estate area. Leisure facilities are buildings and all sorts of establishments of art, entertainment, sports etc. The activities includes first of all "going, walking, and running", which represent 15.3 per cent of all casualties. Next in rank are out-door playing and in-door playing, with 8 per cent each. They are followed by various sports activities, out-door and in-door homework, cardriving, and hobbies.

The sample of 6,755 cases are distribuated by place as follows,

Work	School	Residential area	Traffic environment	Leisure
1,185	1,351	2,026	540	1,653
Walking:	1,034	227 in-door	Jogging etc.	225
		-196 estate - - - -	+196	
		- 14 yard- - - - - -	+ 14	
		———	+ 16 park - - - - - - - - - - -	-16
		1,816 inside resid	+ 9 pavement - - - - - - - -	- 9
			+139 pathway - - - - - - - - -	-139
			+208 street, road - - - - - - -	-208
			———	
	1,034 =	227	+582 walking casualties +540 vehicle casualties	+225
			———	
			1,122 transport casualties	

The headline "Leisure" serves as a main title for miscellaneous items. By separating walking-activities with a transport function from "Leisure" and from "Residential area" it is possible to estimate the total number of transport casualties.

This total is thus composed by 540 casualties in road traffic accidents when the passenger is using a vehicle and by 582 casualties when the injured person is walking whether the injury is caused by a road traffic-accident or by another transport related accident. The total number of transport casualties is thus 1,122 cases in the sample.

Further more, threre are a number of vehicle and walking casualties which are not outdoor transport-related, but rather grouped as due to leisure or residence accidents. From above, threre are obviously 227 cases of indoor-walking accidents and 225 cases of jogging and running accidents mainly appearing in orienteering competions and trainings. It is also to be added 350 cases of accidents when using vehicles of different kinds in leisure activities. In total. there are 802 walking and vehicle casualties in leisure and home activities.

The conclusion is that a wider accident and injury concept has to be defined. It is not enough to use the road traffic accident concept or the motor vehicle accident concept, nor is it satisfactory to talk about transport accidents and casualties. Beside the transport function there is also the competion function, which are joined by the movement activity. The various kinds of accidents and casualties mentioned above, are suggested to be grouped under the heading of movement accidents.

Apppendix 2

CALCULATED RISKS OF INJURY PER YEAR IN SWEDEN 1982-1984

In order to be able to evaluate the safety and danger of different aspects of human life two things are necessary to know. The first is some sort of relation between the number of casualties in certain situations in life and the exposure of the population in the respective situations. This relation is customarily expressed as a risk per performed process of activities. Sometimes, it is convenient to use time-units as a common denominator.

The second aspect is an evaluation of the losses due to the casualties and the accidents. In this context, the various restrictions on performance of human activities are essential. They are described as the permanent or the non-permanent disabilities and as deaths. These restrictions are measured as the time of the disability and as the remaining life-time. Losses are not regarded in this calculation.

Table 1 Time of exposure per year by daily human activities in Sweden.

	ACTIVITIES						
	Work	School	Travel	Leisure		Residence	Total
Population at risk in thousands	4,224	1,353	7,748	4,073	4,256	8,329	8,329
Absence	151	?	?	?			
Average activity time:							
Hours per day	7.5	6	1.2	1	4	9,578	16
Days per year	228	180	365	365	365	365	365
total per person	1,710	1,080	438	365	1,460	3,496	
Total activity time in million hours	6,965	1,461	3,393	1,487	6,214	29,121	48,641

Sources: Living Conditions, Report 47, SOS.
Swedish Travel Patterns, 1984/85, T11 SM 8701, Statistics Sweden.
Notes: Travelling population account the ages of 15-84 years.

Time-budget considerations are useful when the amounts of eposure are calculated. The total time-budget of the Swedish population was 73,286 million hours in 1983, of which 48,857 million was spent as active time.

The activities during these hours grouped under the headings of work, school-work, travels, leisure, and home-life. The number of casualties in Sweden per year during 1982-1984 are shown under the same headings from a nationwide survey (Danielsson 1986).

Table 2. Risk of injury distributed by human daily activities in Sweden.

	ACTIVITIES					
	Work	School	Travel	Leisure	Residence	Total
Exposure in million hours	6,965	1,461	3,393	7,701	29,121	48,641
Casualties in thousands	122.4	139.9	112.2	115.2	209.9	699.6
Risk (cas./ mill.hours)	17.5	95.8	33.1	14.9	7.2	14.4

Notes: School includes children between 7-18 years of age. Travel including casualties on yard or estate area amounts to 133,000 cases.

It is possible to distinguish the risk of commuting accident with personal injury from the risk of travelling. There were 13,991 cases in the year 1983 according to the official statistics (Occupational injuries 1983 SOS). The average journey-to-work single and return lasted 40 minutes (Living conditions, Rep. no 47). The employed labour force consisted of 4,073 thousand people (Arbetsmarknadsstatistik 1-4 kvartalet 1983). The number of working days are 212 excluding sickleave etc. Thus the yearly commuting time of the Swedish population is around 575.7 million hours. The risk of commuting is around 24 casualties per million hours, while the other purposes of travelling have a risk of 35 cases per million hours.