## ision of Research \&

 ents. HRB Proc. 32, 2. Heft 201, 1976. ge daily trafic. Ace. Submitted to Accid. ssure to risk. Ph.D. c Quart. XXI, 3 July on of the relationship Verkehr, March 1966. = volume on a section
# RELATIONSHIPS BETWEEN ROAD ACCIDENTS AND HOURLY TRAFFIC FLOW-II 

PROBABILISTIC APPROACH

Avishal Ceder<br>Transportation Research Institute. Teclinion-Israel Institule of Technology, Hajfa, Israei

## '88 FEB -2 A9:28

QFFICE OF SECRETAR boCkETMN \& FRviCh. BRAND $1600 \leqslant q \leqslant 1900$ (the maximum measured hourly flow on section $13-6$ is $1900 \mathrm{veh} / \mathrm{ht}$ ).

## 3. FREEFLOW AND CONGESTED-FLOW MODELS

Under free-flow conditions, both single- and multi-vehicle accidents occur where the proportion of single-vehicle accidents is decreasing with $q$. Based on the data exhibited in Part 1 of this work, it can be surmised that on four-lane divided roadways, the majority of multivehicle accidents are rear-end collisions over all $q$ ranges. It is, therefore, reasonable to claim that a potential multi-vehicle accident is associated with those traffic situations in which $h<T$, where $h=$ headway in seconds (front bumper to front bumpet) and $T=$ a time lag of the driver-vehicle system in seconds (sufficient time in order to completely perceive, interpret, decide and act, and for the vehicle to respond).

The time lag is an essential variable both in the event of an emergency deceleration (to avoid rear-end collision), and in the event of a risky manocuvre (to avoid angle collision). On four-lane divided roadways, head-on collisions rarely occur, and can be distegarded. Hence, under free-flow conditions, the probability of a multi-vehicle accident is particularly dependent on the interaction of two events, $A$ and $B$ :

Event $A: h<T$ in a car-following mode:
Event B: a risky situation (e.g. the leading car performs a hazardous manoeuvre, or the driver of the following car drastically reduces his attention, or an eaternal factor interferes with one of the vehicles).

According to the muttiplicative law of probability:

$$
P\left\{\begin{array}{c}
\text { multi-veh }  \tag{I}\\
\text { accident }
\end{array}\right\}=P(A \cap B)=p(A) \cdot p(B \mid A) \text {. }
$$

The measure $A_{t}(q)$ which is the accident rate (acc/10 veh $\cdot \mathrm{km}$ ) is explained in Part I of this work. This measure, divided by $10^{\circ}$, can be used as a probability measure for an accident in each veh-km within the interval $q \pm \Delta q$. It is certain that the event $A$ is dependent on the hourly flow $q$ and therefore, $p(A)=p(h<T \mid q)$. The expression $p(B \mid A)$ in eqn (1) depends also on $q$ and it is represented by a power function. These interpretations lead, from eqn (1), to:

$$
\begin{equation*}
A_{i}^{\prime},(q)=p(h<T \mid q) \cdot \sigma_{1} q_{1}^{b_{1}} \tag{2}
\end{equation*}
$$

Where $A_{\text {if }}(q) \cdot 10^{*}$ is the (probability) measure for a multi-vehicle accident in each veh-km and $a_{1} q^{f_{1}} \cdot 10^{\text {a }}$ represents the probability of a risky situation within the flow $q \pm \Delta q$ such that

## ds in which 95\%

 the propertion of ds are conisidered nes the boundarybver, running-offpilisions (in comthree four-lane, this study|Ceder travel (two lanes) ved only on one ration seems also bach et al., 1970;
g night) accidents Certainly, for the hd cross-sectional d congested-flow tervals within the within the range veh/hr).
occur where the exhibited in Parl 1 najority of multiasonable to claim ns in which $h<T$. time lag of the erceive, interpret,
eleration (to avoid gle collision). On sregarded. Hence, cularly dependent

## manoeuvre, or the

 tor interferes with dent on the hourly depends also on $q$ qn (1), to: neach veh-km and $q \pm \Delta q$ such that$h<T ; a_{1}, \beta_{1}$ are constants and certainly all the probability expressions are constrained so as to be less than 1 .

Another approach in the interpretation of $p(B \mid A)$ is through the definition of $A_{d v}(q)$, from Part 1 of this study, which is the accident density (acc/ $16^{3} \mathrm{~km}$ ) per one hour exposure of traffic flows within the interval $q \pm \Delta q$. Under free flow of $q$ veh/hr, the number of headways which are potential for accidents is $q \cdot p(h<T \mid q)$. Assuming a steady flow for one hour:

$$
A_{d *}(q)=q \cdot p(h<T \mid q) \cdot \gamma(q)
$$

where it is clear that $\gamma(q) \cdot 10^{-3}$ is the probauility for being in a risky situation when $h<T$. Since $A_{\text {ds }}(q) \cdot 10^{-1} / q$ is the number of accidents per veh- km , it turns out that $\gamma(q)=p(B \mid A)$.

For single-vehicle accidents under free-flow conditions, one cannot assume a clear cut intersection between events. Equation (3) represents the power function for these accidents:

$$
\begin{equation*}
A_{F}^{*}=a_{2} q^{B_{2}} \tag{3}
\end{equation*}
$$

where $A ; \cdot 10^{-6}$ is the (probability) measure for single-vehicle accidents in each veh-km and $\alpha_{2}, \beta_{2}$ are constants. The total measure for free-tlov "-cidents is the summation of eqns (2) and (3):

$$
\begin{equation*}
A_{, f}=p(h<T \mid q) a_{14}{ }^{1}+a_{2} q^{h_{1}} . \tag{4}
\end{equation*}
$$

For the complimentary congested-flow conditions, a simple power function is selected as a model (due to already determined criterion, mentioned in the previous section):

$$
\begin{equation*}
A_{r c}=a_{3} q^{\beta_{3}} \tag{5}
\end{equation*}
$$

where $A_{1 s} \cdot 10^{-6}$ is the measure for mulli-vehicle accidents in each veh-km and $\alpha_{3} . \beta_{3}$ are constants.

## 4. Generalized headway models

The first expression in eqn (4) includes the headway probability distribution which basically is determined by measurements of headways between successive vehicles in a single-lane stream. While substantial literature has developed regarding the mathematical description of this distribution(Edic, 1974), only a few studies are concerned with model parametrization in respect to different a vaiues.

The search for a generalizei headway model in terms of $q$ dependency has led to three different models-each constructed with two components associated primarily with free and constrained vehicles. The first, reported by Grecco and Sword [i968], is an empirically-based model which considers Schuhl's distribution (best fitted distibution anovg: Schuhl, Gamma, Erlang and Pearson type III). Their results, from measurements on four-lane divided roadways. include an hourly Now variable on a per lane basis, $q_{1}$, and takes the foral:

$$
\begin{equation*}
p(h \geqslant t)=155 \cdot 10^{-5} \cdot q_{1} \cdot e^{(1-1225}+\left(1-115 \cdot 10^{-5} q_{1}\right) e^{\left.e t i 012_{Q_{1}}-20\right)} \tag{6}
\end{equation*}
$$

where $t \geqslant 1$ second and the parameters were derived according to the range $0<q_{1} \leqslant 700$ veh $/ \mathrm{ht}$.
The second study, reported by Dawson and Chimini(1968), describes what thay call the hypertang headway model. Their model is a lineat combination of a translated exponential function and a translated Erlang function:
where $\delta_{1}, \delta_{2}$ are the minimum headways and $\gamma_{1}, n_{2}$ are average headways for free and constrained vehicles, tespectively: $k$ is an index that indicates the degree of nonrandomness in
the constrained headway distribution; and $\alpha_{1}, \alpha_{2}$ denote the proportion of free and constrained vehicles, respectively $\left(\alpha_{1}+\alpha_{2}=1\right)$. The parameters of eqn (7) were evaluated for one-lane fiows (on a four-lane divided roadway) ranging from $158(k=11$ to $957(k=6)$ veh/hr. Though eqn (7) does not include the hourly flow variable, the adjustable parameters for nine different flow levels provide sufficiently adequate data for this work.

The third work, recently reported by Wasielewski[1979], is based on the semi-Poisson headway distribution model. The estimate of the total headway probability density function, $f(t)$, is given by:

$$
\begin{equation*}
f(t)=\Phi \dot{g}(t)+A \lambda \mathrm{e}^{-\lambda} \int_{0}^{1} \tilde{g}(u) \mathrm{d} u, \tag{8}
\end{equation*}
$$

with

$$
\Phi=1-A \lambda \int_{0}^{\infty} e^{-\lambda} \int_{0}^{t} \hat{g}(u) d u d t
$$

where $\Phi$ is the proportion of following (constrained) vehicles; $\hat{g}(t)$ is an estimate for $g(t)$ and the latter is the probability density function of the constrained vehicles; $A$ and $\lambda$ are parameters which are evaluated from the observed data (in those situations in which the vehicles are not under the car-following mode). The findings of Wavielewski introduce the function $\tilde{g}(t)$ and indicate that no significant disagreemert is foun: beiween $f(t)$ and the observed total headway probabulity density function; also, it is interesting to note that the flow dependence is considered only through the parameters $\Lambda$ and $\lambda$. These findings are based on 42,000 observed headways regarding, 12 groups of hourly flows ranging from 922 to 1985 veh/hr per lane on a six-lane divided roadway.

The above three reviewed models are used bere to evaluate the expression $p(h<T \mid q)$ in eqn (4). The value of $T$ defined in the previous section is censidered as 2 sec . Generally, $T$ is a distributed variable and ranges from 0.5 to even 4.0 sec , depending on the complexity of the driving situation|Greenshields, 1965]. Let us recall that $h<T$ is considered as a potential situation for a isulf-vehicle accident where for $h \geqslant T$, it is improbable that such an accident will occur. The value of $7=2$ is substituted in equs (6) and (7), and in the numerical integration of eqn (8), in order to obtain the function $p\left(h<2 \mid q_{1}\right)=1-p\left(h \geqslant 2 \mid q_{1}\right)$, where $q_{1}=$ the flow on a per lane basis.

The results are demonstrated in Fig. 1. Iu the upper part of this figure, the results of each model are exhibited separately where the flow levels correspond to the based-data of each model. In the lower part of Fig. I, a regression line based on a power function is indicated for all the models' results, namely:

$$
\begin{equation*}
p\left(h<2 \mid q_{1}\right)=0.011 q_{i}^{0,472} \tag{9}
\end{equation*}
$$

with standard error (SE) of 0.014 probability units. It is rather interesting to note that the regression line for the hyperlang and empirical-based models only, in which $q \leq 957 \mathrm{veh} / \mathrm{hr}$, is almost like eqn (9). That is, $p\left(h<2 \mid q_{1}\right)=0.011 q_{1}^{0,48}$ winh $S E=0.024$. The interpretation of the latter result is that extrapolation of the first two rodels fits very well an independent model which is calibrated with data characterized by the rage $922 \leqslant q_{1} \leqslant 1985$ veh/hr. This finding supports and strengthens the generality of eqn (9).

## S. REGRTSSION RESULTS AND DISCUSSION

the selected data are described in Section 2 with the separation criterion between hourly free-flow and enngested-flow periods. The fitted power function to mult-vehicle accident data under free flow conditions is:

$$
\begin{equation*}
A_{i F}^{\prime}(q)=4.3 \cdot 10^{-5} \cdot q^{1.32} \tag{10}
\end{equation*}
$$

with $\mathrm{SE}=0.19$ (acchit veh-km). The breakdown of eqn ( 10 ) in accordance with eqns (2) and

## ee and constrained

 for one-lane flows hr. Though eqn (7) nine different flowthe semi-Poisson y density function,
timate for $g(t)$ and nd $\lambda$ are parameters the vehicles are not e function $\dot{g}(t)$ and rved total headway low dependence is on 42,000 observed eh/hr per lane on a
$\operatorname{ssinn} p(i)<T \mid q)$ in c. Generally, $T$ is a c complexity of the ered as a potentiat uch an accident vill erical integratio. $q_{1}=$ the flow on a the results of exch based-data of each tion is indicated for

## ng to note that the

 :h $q \leqslant 957 \mathrm{veh} / \mathrm{hr}$, is interpretation of the independent model veh/hr. This findinginn between hoarly en ich accident data
e with eqns (2) and


Fig. 1. The probability of headway less than 2 sec according to the empritical-based results of Grecco and Sword [1968), the Hyperlang Model of Dawson and Chimini( 1968 ), and the semi-Poisson Model of Wssielewrkif1979): in the lower figure, a regression model is shown for the three sets of results.
(9), reveals that $\alpha_{1}=5.42 \cdot 10^{-3}$ and $\beta_{1}=0.848$ : This breakdown presumes that eqn (9) cin be applied to two-lane flows by considering separately each lane behaviour, i.e. $p(h<2 \mid q)$ is hased on eqn (9) with $q_{1}=\frac{1}{q}$. Consequently, $\alpha_{1}$ and $\beta_{1}$ of eqn (2) determine the expression $p(B \mid A)$ belong to eqn (1).

For single-vehicle accidents, the following formula is obtained through regression:

$$
\begin{equation*}
A_{F F}^{*}(q)=232.27 \cdot q^{-1.15} \tag{!?}
\end{equation*}
$$

with $\mathrm{SE}=0.34\left(\mathrm{acc} / \mathrm{i} 0^{6} \mathrm{veb}-\mathrm{km}\right)$. Equation (II) is associated with eqn (3) and eqn (4) is futfiled through the summation of eqns $(10)$ and ( 11 ), i.e. $A_{t, F}=A_{i p}^{\prime}+A_{\text {,F }}^{*}$ which is the total measure for free-flow accidents. The left side of Fig. 2 illustrates both the data points and the free-flow model.


Fig. 1. The data and regression models ior free-flow and congested flow conditions.

The optimum conditions, by differentiating, yield $q$ rof $=503$ veh/hr, which is similar to that found in Part I of the work (for dala composed also of night accidents and without separating free and congested conditions).

For congested flow conditions, eqn ( $\$$ ) through regression, takes the form:

$$
\begin{equation*}
A_{r c}=7.21 \cdot 10^{-68} \cdot q^{14.45} \tag{12}
\end{equation*}
$$

with $\mathrm{SE}=0.06\left(\mathrm{acc} / 10^{6}\right.$ veh -km$)$, and is demonstrated on the right side of Fig. 2. Despite the small number of data points, it is possible to observe a sharp increase in $A_{r} C$ as $q$ increases.

In traffic flow theories (Edie, 1974), a congested-flow behaviour refers particularly to a low, slow and congested stream of vehicles. Under these conditions, the time headway (not the spacing between vehicles) is usually higher than that observed under high flow levels and therefore, the probability of collisions is reduced. Perhaps this explanation can cast light on the results in Fig. 2 which show a diminishing lendency of $A_{r} c$ as $q$ decreases. A study on the attentional demands of drivers [Ceder, 1977], also indicates the increase in collision risk under peak flow conditions. This study, based on a driver's uncertainty model, shows that under peak flow conditions (small spacing with relatively high speed), drivers tend to absorb information incompletely. This mode is characterized as overload attention. The latter might explain the relatively high probability of being involved in a collision at such flow conditions.

Generally, traffic engineers attempt to manage traffic at high flows in order to enable movement of as many units of car as possible in a unit of time. Their belief in a productivity measure such as this results in neglect of the safety component, which is clearly indicated in Fig. 2. It is desirable to approach a weighting objective function which will balance increased savings in travel time with an increased accident rate as the flow level increases.

## 6. PROBABILISTIC ASPECTS

This section further examines probabilistic interpretations of the accident measures. These aspects are an essential input for both simulation studies and theoretical models of traffic laccidents.

Equation (1) considers the intersection between the two events $A$ and $B$. While event $A$ has been widely investigated, event $B$ is a complex one and depends on the driver population, human factors and other elements which can hardiy (if at all) be predicted. An attempt is made here to examine event $B$ given that event $A$ occurs, based on the investigated data. The component $p(B \mid A)$ in eqn (1) takes the form:

$$
\begin{equation*}
p(B \mid A)=5.42 \cdot 10^{-9} \cdot q^{0.848} \tag{13}
\end{equation*}
$$

which is the probability of being in a risky situation given that the headway (between two vehicles in a single lane) is less than two seconds. For example, if one counts 491 vehicles in a single lane during one hour, one can expect to observe (or measure) 100 out of 490 headways to be characterized bv $h<2 \sec$ (using eqn 9). The probability, for those vehicles involved in these 100 headways, of being in a risky situation for one kilometer of driving is $1.87 \cdot 10^{-6}$ (substituting $q=982$ in eqn 13). In other words, this is the probability that the situation becomes an actual accident from a potential accident.

Two additional probabilistic aspects which can be derived from the results of this study are:
(1) determination of the number of kilometers with an hourly flow $q$, for a given probability such that (at least) one accident will occur.
(2) determination of the number of hours with an hourly flow $q$, for a given probability such that (at least) one accident will occur.

For both aspects, the determined quantity (kms or hrs) does not necessarily maintain the continuity properly (e.g. for the first aspect it gives the number of kms exposed to $q$ in one year for a given probability).

In fact, for both aspects repeated independent trials (Bernoulli trials) are performed. Considering the first aspect, one inspects whether or not (at least) one accident wilk occur at each veh-km under the flow $q$, presuming independence between each two inspections. For large numbers of veh-kms the description of the first aspect approaches the normal distribution
h is similar to that 1 without separating m:
(12)
2. Despite the smali increases.
articularly to a low, headway (not the gh flow levels and an cast lighi on the es. A study on the collision risk under pws that under peak absorb information might explain the litions.
in order to enable ef in a productivity y indicated in Fig. 2. increased savings in


Fig 3. The tesultant telationship between the probability for at least one accident an $n$ for various $q($ veh/ht) values.
(approximation to the binomial distribution). Also, since the analysis is not succersive with respect to $q$, a vehicle involved in an accident is, theoretically, not excluied from further examinations fotherwise, the appropriate distribution is geometric rather than normal). Thus, the analysis of previous sections enables a definition:
$A,(q) \cdot 10^{-6}=$ the probability of being involved in an accident in each veh-km within the flow rang" $q \pm \Delta q(\Delta q=100 \mathrm{veh} / \mathrm{hr})$
$X=$ number of accidents-normally distributed $\left(\mu_{1}, \sigma^{2}\right)$
where

$$
\mu_{1}=n \cdot A_{r}(q) \cdot 10^{-6} ; \quad \sigma_{1}^{2}=n \cdot A_{r}(q) \cdot 10^{-6} 11-A_{r}(q) \cdot 10^{-6} 1
$$

an $1 n$ is the number of kilometers travelled by a vehicle at a given flow (within the range $q \geq \Delta q)$. The investigated probability is $p(X \geqslant| | q)$. Figure 3 illustrates this probability as a function on $n$ for different flow levels, based on eqns (10)-(12). For example, at the probability of $90 \%$, a flow of $1000 \mathrm{veh} / \mathrm{hr}$ needs to cover 40 million km so that (at least) one single-vehicle accident will occur, in comparison with 8.5 million km for a multi-vehicle accident. Figure 4 illustrates the functional dependency between $n$ and $q$ for three probability levels: 30,70 and $95 \%$. That is, both Figs. 3 and 4 demonstrate the resultant relationship between $n, q$ and $p(X \geqslant \mid q)$-each in a different manner.

Similarly, for the second aspect:

$$
\begin{aligned}
A_{d r}(q)=A_{t}(q) \cdot q \cdot 10^{-6} & =\text { the probability of being involved in an accident on any kilometer } \\
& \text { exposed to one hour of flow within the range } q+\Delta q \\
y & =\text { number of accidents-normally distributed }\left(\mu_{2}, \sigma \dot{)}\right)
\end{aligned}
$$

where

$$
\mu_{2}=m \cdot A_{d *}(q) ; \quad a^{2}=m_{2} \cdot A_{d \times}(q)\left[1-A_{d *}(q)\right]
$$



Fig. 4. The resultant relationship between $n$ and $q$ for three levels of probabilities (for at least one accident).
and $m$ is the number of hours which experience a flow (within the range $q \pm \Delta q$ ) at a given kilometer. The investigated probability $p(y \geqslant 1 \mid q)$ is shown continuously in Fig. S, and for only three levels in Fig. 6.

## 2. SUMMARY

This study which is the last, and phase IV of the entire research (shown schematically in the first figure in Ceder and Livneh, 1981), attempts primarily to consider accident data regarding


Fig. 5. The resulant relationship between the probability for at least one accident and an for various a (veh/hr) values.
the se proba relati
ge $q \pm \Delta q$ ) at a given in Fig. 5, and for only
n schematically in the ccident data regarding


Fie. 6. The resultant relationship between $m$ and $q$ for three leve's of prohabilities (for at least one accident).
the separation between free-flow and congested-flow conditions. The adopted and developed probabilistic approach outlined in this paper is another way of tackling the determined relationships between the accident rate and the hourly traffic flow.

Based on a predetermined criterion, the congested-flow conditions are separated from the free-flow conditions, and subsequently applied an d correlated to accident data in Fig. 2. For the free-flow data, the total accident rate curve follows the known U -shaped configuration with respect to the hourly flow, which is the result of combining a convex downward and a convex upward curve for single and multi-vehicle accidents, respectively. For congested-flow data, the (multi-vehicle) accident rate is sharply increased with hourly flow. This outcome suggests, from a safety viewpoint, avoidance of high flow levels in contrast to the general traffic engineers' desire to move as many cars as possible in a unit of time (i.e. to approach a capacity level). A balanced traffic productivity measure might then attempt to maintain the stream of vehicles (assuming it is under control) at a buffer point below its m2ximum range.

The probabilistic aspects utilize a generalized headway model fitted to three different models from the literature. This headway model is hourly flow dependent and it represents the probability that two velicies which are, even instantaneously, under the car-following mode, are in a potentially hazerdous situation. The remaining underlying probability aspects are based on the delermined accident modeis for the free-flow and congested-flow conditions. It is believed that such models are essential input both for simulation studies and for theoretical models of road traffic accidents.

## REFERENCES

Brilon W. . Unfallgeschehen und Verkehrsablauf. Forschung Strassenbau und Strassenverkehrsiechnik, Heft 201, 1976. Ceder A. Drivers' eye movements as related to attention in simulated traffic flow conditions. Human Factors 19(6), 571-581, 1977.
Ceder A. Stable phase piane and car-following behaviour as applied to a macroscopic phenomenon. Transpn Sci. 13(1). 64-79, 1979
Ceder A. and Livneh M., Relationships between raad accidents and hourly trallic flow-1. Analysis and interpretation. Accid. Anal \& Pres. 14, 19-14, 1982
Dawson R.F. and Chimini L. A. The hyperlang probability distribution-a generalized traflic headway model. Highway Res Rec 260, 1-14, 1968
Edie L. C. Flow theories In Traffer Science (Edited by Gazis D. C). Wiley, New York, 1974
Grecce W. L. and Sword E. C., Prediction of paramelers for Schuhl's headway distribution. Traffic Engng 38 5), 36 -38. 1968.

Greenshields B. D. The driver. In Traffic Engincering Handbook (Edited by Bacrwald J. W ). Institute of Traffic Engineers, Washington, D.C. 1965
Leuizhach W., Siegener W. and Wiedemann R. On the connection between traffic accidents and traffe volume on a section of \& German Aulntalin. Accid. Anal. \& Prev. 2, 143-146 (in German), 1970.
Wasielewski P. Car-following headways on freeways interpreted by the semi-Poisson headway distribution model. Transpn Sci. IM (1), 36-55, 1979.

Emile L. Julian, Wcting Chief Docketing and Service Branch

SUBJECT:
SEABROOK EXHIBITS

Any documents filed on the open record in the SEABRook proceeding and made a part of the official hearing record as an exhibit is considered exempt from the provisions of the United States Copyright Act, unless it was originally filed under seal with the court expressiy because of copyright concerns.

All of the documents sent to $T I$ for processing fall within the exempt classification.

