TECHNION-ISRAEL INSTITUTE OF TECHNOLOGY

TRANSPORTATION RESEARCH INSTITUTE

# RELATIONSHIPS BETWEEN ROAD ACCIDENTS AND HOURLY TRAFFIC FLOW:

## I. ANALYSES AND INTERPRETATION

by

AVISHAI CEDER & MOSHE LIVNEH

## II. PROBABILISTIC APPROACH

by

AVISHAI CEDER



PUBLICATION NO. 80 - 29

Haifa March 1980

# RELATIONSHIPS BETWEEN ROAD ACCIDENTS AND HOURLY TRAFFIC FLOW: I. ANALYSES AND INTERPRETATION

AVISHAI CEDER & MOSHE LIVNEH

Transportation Research Institute, Technion-Israel Institute of Technology, Haifa, Israel.

ABSTRACT

This research extends the investigation of the relationships between measures of accidents and traffic flows, and considers the hourly flow instead of the average daily traffic (ADT). The latter was already reported, and serves as a basis for further clarification of the interactions between various levels of traffic flows and road accidents. Eight four-lane sections during an 8-year period provide the adequate data based on carefully predefined criteria. Power functions are fitted and classified according to : (i) time-sequence analysis for each roadway section; and (ii) cross-sectional analysis on a one year basis. The results are presented, separately for multi and single vehicle accidents, in a matrix-format, and a linear dependency is observed between the power and the logarithm of multiple constant. This is done in a similar fashion to the previously reported study of the relationship between road accidents and ADT. The results for each type of analysis and type of accident are discussed, and three examples of a practical application are given.

## 1. INTRODUCTION

This research was developed as a part of the establishment of safety evaluation procedures for the analysis and interpretation of road accidents in Israel. The format of the entire research, based on gathering nationwide data, is shown in Fig. 1. The basis for such research is the availability of an adequate data bank consisting of effective reporting, storage, retrieval and compilation systems. Four major phases are indicated in Fig. 1 : (1) phase I investigates the relationships (power functions) between two measures of total accidents (density and rate), and average daily traffic (ADT) on four-lane and two-lane interurban road sections; (2) phase II extends the investigation of phase I for the four-lane sections through separate consideration of single and multi-vehicle accidents; (3) phase III examines deterministic relationships between two weighted accident measures and the hourly traffic flow for each type of accident; and (4) phase IV attempts to go thoroughly into the relationships between measures of accidents and hourly traffic flow by separating the traffic stream to free-flow and congested-flow modes, and by interpreting the results in a probabilistic manner.

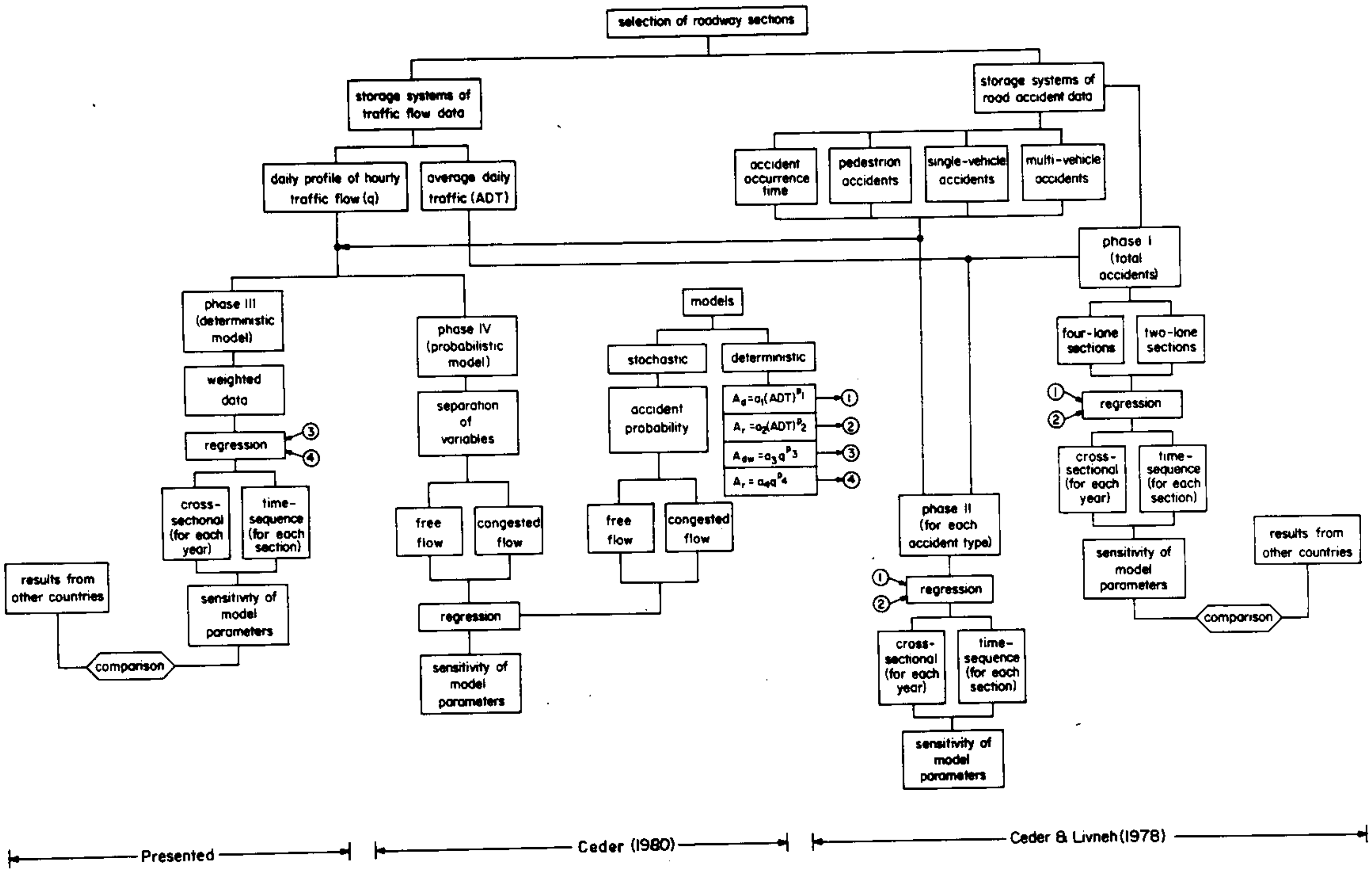



Fig. 1: The entire overall research plan.

Phases I and II are reported in Ceder and Livneh [1978]. Phases III and IV are described sequentially in this (part I) and in the following (part II) papers.

In the past, some aspect of road geometry has been identified as a dominant factor in accident causation at a given location. Thereafter, an attempt is made to interpret the frequency and number of accidents using the ADT value as additional information to the road geometry. Nevertheless, when eliminating considerations of geometry, the ADT by itself cannot be used to explain the overall interaction between traffic flow characteristics and accidents. For that purpose, one should approach the actual traffic flow observed at the time of the accident. In addition, the level of risk associated with traffic flow can be determined only on the basis of smaller time intervals than daily periods. Consequently, this work attempts to clarify and improve the understanding of the relationships between measures of accidents and hourly traffic flow, which is more fundamental than the accident/ADT relationship. The analysis and interpretation are performed in a similar fashion to that outlined in phases I and II of the study.

## 2. SOME PREVIOUS STUDIES

The common measure of accidents considered in relation to hourly traffic flow, $q$, is the 'accident rate' $A_r$. This rate is usually the number of accidents per year per million (or $10^8$) motor vehicle-kilometers (or miles). That is, the annual number of accidents per the annual amount of exposure. Several forms of relationship between $A_r$ and $q$ have been found in the literature. The variety of $A_r - q$ dependency is probably due to different (i) types of accidents; (ii) ranges of flows in the analysis; and (iii) road designs.

Belmont [1953], found for two-lane sections that $A_r$ (during daylight only) increases almost linearly with q, whereas Smeed [1955], has shown that $A_r$ for total accidents has a small variation for different annual q values. Nonetheless, Smeed pointed out that $A_r$ values for single-vehicle accidents have the tendency to decrease with the increase of q, and the opposite tendency for multi-vehicle accidents.

Leutzbach [1966] and Gwynn [1967], have concluded for four-lane divided sections that a U-shaped dependency exists between $A_r$ (for total accidents) and q, where the minimum $A_r$ values are obtained for q values between approximately 600 to 1300 veh/hour per two lanes. Thereafter, Baker & Gwynn [1968], noted that $A_r$ (total accidents) increases rapidly below q = 550 veh/hour per two lanes, but has little variation beyond this flow value.

Pfundt [1969], has compared three types of day and night accidents : rear-end collisions due to blocked lane(s); rear-end collisions due to slow and disabled vehicles; and single-vehicle accidents due to loss of control. For the first type of accident, $A_r$ tends to have a convex upward curve with q (particularly at night); for the remaining two types, the curves are convex downward. In a different study, Leutzbach et al. [1970], have shown that on a four-lane Autobahn section in Germany, the resultant U-shaped curve in the $A_r$ - q plane is mainly attributed to rear-end accidents, whereas q values have little effect upon $A_r$ values of single-vehicle accidents. Chapman [1971], analyzed accident and flow data from England, and generally agrees with Pfundt's findings.

Recently, Brilon [1976], in a study of 8 four-lane sections on German Autobahns, found similar results (U-shaped curves) to those reported by Leutzbach et al. [1970]. In addition, Brilon hypothesizes that the minimum $A_r$ value is obtained for the most frequent range of $q$. This hypothesis is examined, among other analyses, in a following section. It should be emphasized that all the above mentioned studies consider the relationship between $A_r$ and $q$ only on the basis of roadway classification (cross sectional analysis).

## 3. DATA CLASSIFICATION & ANALYSIS ORIENTATION

The data, based on carefully defined criteria, were selected with the aid of the data bank of the Israel Central Bureau of Statistics. As shown in Fig. 1, the data includes (i) fatal and injury accidents : on 4-lane interurban road sections, including the type and hour of day of the accidents; and (ii) hourly traffic flow : a daily profile (accomplished by fixed counters), by hour of day, for each road section. The overall data were gathered for an 8-year period (1967 - 1975), excluding the war year 1973, and in order to eliminate any undesirable and/or unknown influence of external parameters, four major criteria have been imposed :

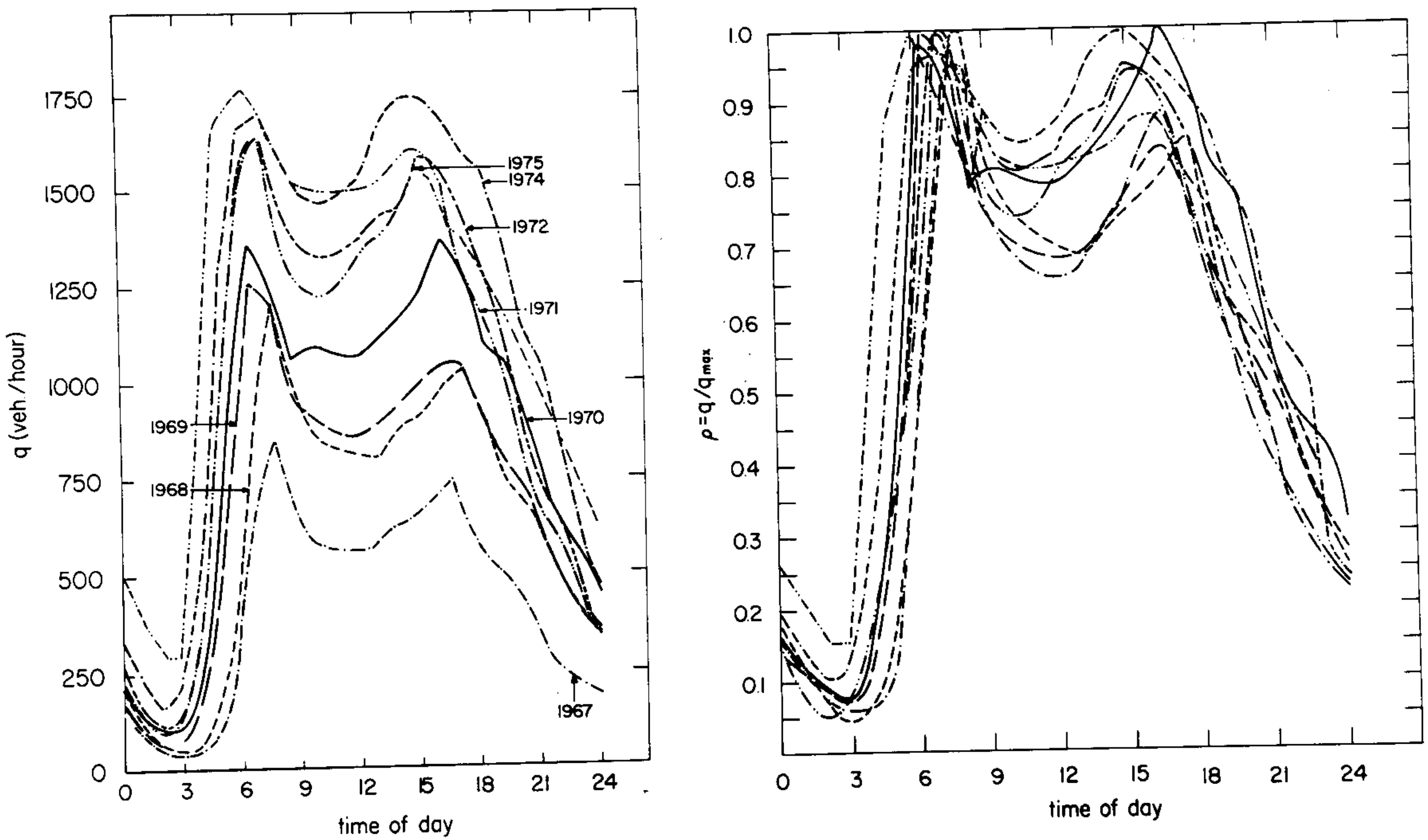
(1) that the roadway sections exclude geometric design elements which disturb the traffic flow (steep grades, curves, roadside obstacles, etc.), and be isolated from intersections, entries and exits (to eliminate the influence of cross-traffic);

(2) that no changes were made in the roadway section characteristics : section length, pavement width, and shoulder width, during the period 1967 - 1975 (to ensure a comparable basis of the data);

(3) that there be different daily profiles for $q$ : two daily peaks, one peak or no peak flow (to establish generality); and

(4) that there be similar daily profiles for $q$, excluding weekends, for each roadway section during the period of 1967 - 1975 (to ensure steady travel characteristics).

In order to clarify criterion (4), an example of average daily $q$ profiles of a roadway section is exhibited in Fig. 2. The left illustration of Fig. 2 shows that the general tendency of each year-profile is eminently preserved. However, there are changes in the levels of each year-profile as the result of increasing ADT values with time. The latter effect on daily $q$ profiles can be eliminated, to some extent, by the use of normalized $q$ values. That is, taking $\rho = q/q_m$ instead of $q$, where $0 < \rho \leqslant 1.0$ and $q_m$ is the maximum $q$ value, demonstrates the similarity between the average daily $q$ profiles, as shown on the right illustration of Fig. 2. It is worth noting that criterion (4) has also been applied to an examination of four seasonal average daily $q$ profiles for each year, to ensure the steady characteristics of the travel pattern on each selected roadway section. By the above described selection




<u>Fig. 2:</u> An example (Roadway 13, section 6), of daily profiles of $q$ and $\rho$ (normalized values) for an 8-year period.

procedure, a total of 8 four-lane divided roadway sections were selected.

The accident and q data collected from the 8 roadway sections during the 8-year period are analyzed on a separate basis (described in detail by Ceder & Livneh, 1978), as follows :

(i) time-sequence analysis of data for specific roadway section over an extended period of time (8 years);

(ii) cross sectional analysis of data for a given period of time (one year) for a group of roadway sections (8 sections) having the same roadway classification.

This approach enables, in essence, the consideration of dynamic and environmental effects inherent in the relationship between measures of accidents and measures of traffic flow. Let us explain in the next two paragraphs the differences between a consideration of ADT and q values with respect to the time-sequence and cross sectional analyses.

In the time-sequence analysis (for a specific roadway section), each ADT value refers to a different measure of accidents, due to their possible mutual changes over the considered period of time; hence, they constitute, say, eight data points for the 8-year period. The time-sequence analysis with q values is based for each year on a number of data points (dependant on the considered q range and intervals), whereas only one data point is presented when considering the accident/ADT relationship. The ADT value increases, usually, with time, and this increase is reflected in the time-sequence analysis. The q values, however, could have only an upper bound, and therefore, it is perhaps only possible to notice the changes with time for high q values. Nevertheless, the time dependency, in the time-sequence analysis, for the relationship between measures of accidents and q, is relatively negligible, due to overlap among the q ranges of the considered years.

In the cross sectional analysis (for a given year), each ADT value refers to a different measure of accidents, due to the various roadway sections; hence, they constitute, say, eight data points for 8 roadway sections. Similar to explanations for the time-sequence analysis, when considering q values, one obtains several data points for each roadway section rather than single data point (for the accident/ADT relationship). To sum up, for the time-sequence analysis, the relationship between measures of accidents and q emphasizes the uniqueness of each roadway section, while for the cross sectional analysis, the uniqueness of each year (or other given period) is emphasized.

## 4. MEASURES & MODELS

The data were gathered, basically, by matching each type of accident with the average q value at the time of the accident. The q value is considered with respect to the interval $q \pm \Delta q$, where $\Delta q = 50$ vehicle/hour per direction of travel (two lanes). That is, each q range was divided into intervals of 100 veh/hour on a two-lane basis.

The measures of accidents were carefully defined and selected as follows :

i => denotes either a particular year (for the time-sequence analysis) or a particular roadway section (for the cross sectional analysis).

$$A_d(q) = \sum_i \frac{N_i(q)}{L_i} = \begin{Bmatrix}\text{accident density (acc/km)} \\ \text{for the interval } q \pm \Delta q\end{Bmatrix}$$

where :

$L_i$ = the length, in kilometres, of roadway section i ; in the cross sectional analysis i denotes a section and $L_i$ is a variable, while in the time-sequence analysis i denotes a year and then $L_i = L$, i.e. has a constant value.

$N_i(q)$ = total annual number of accidents (of a given type) which occurred during the five work days of each week for the interval $q \pm \Delta q$.

Another component which should be taken into account is the exposure time of each $q \pm \Delta q$ interval. Hence :

$$T(q) = \sum_j \sum_i t_{ij}(q) = \begin{Bmatrix}\text{the annual exposure time of} \\ \text{traffic flows within the interval } q \pm \Delta q\end{Bmatrix}$$

where :

$t_{ij}(q)$ = the daily exposure time for the interval $q \pm \Delta q$ at the jth day, $j = 1,2,\ldots,261$ (excluding weekends).

An example of $T(q)$ distribution is shown in Fig. 3, where the upper illustration is for the time-sequence analysis (one section, over an 8 year period), and the lower illustration is for the cross sectional analysis (one year for 8 sections).

As the consequence of the above definitions, two accident measures were selected for this research :

$$A_{dw}(q) = \frac{A_d(q) \cdot 10^3}{T(q)} \tag{1}$$

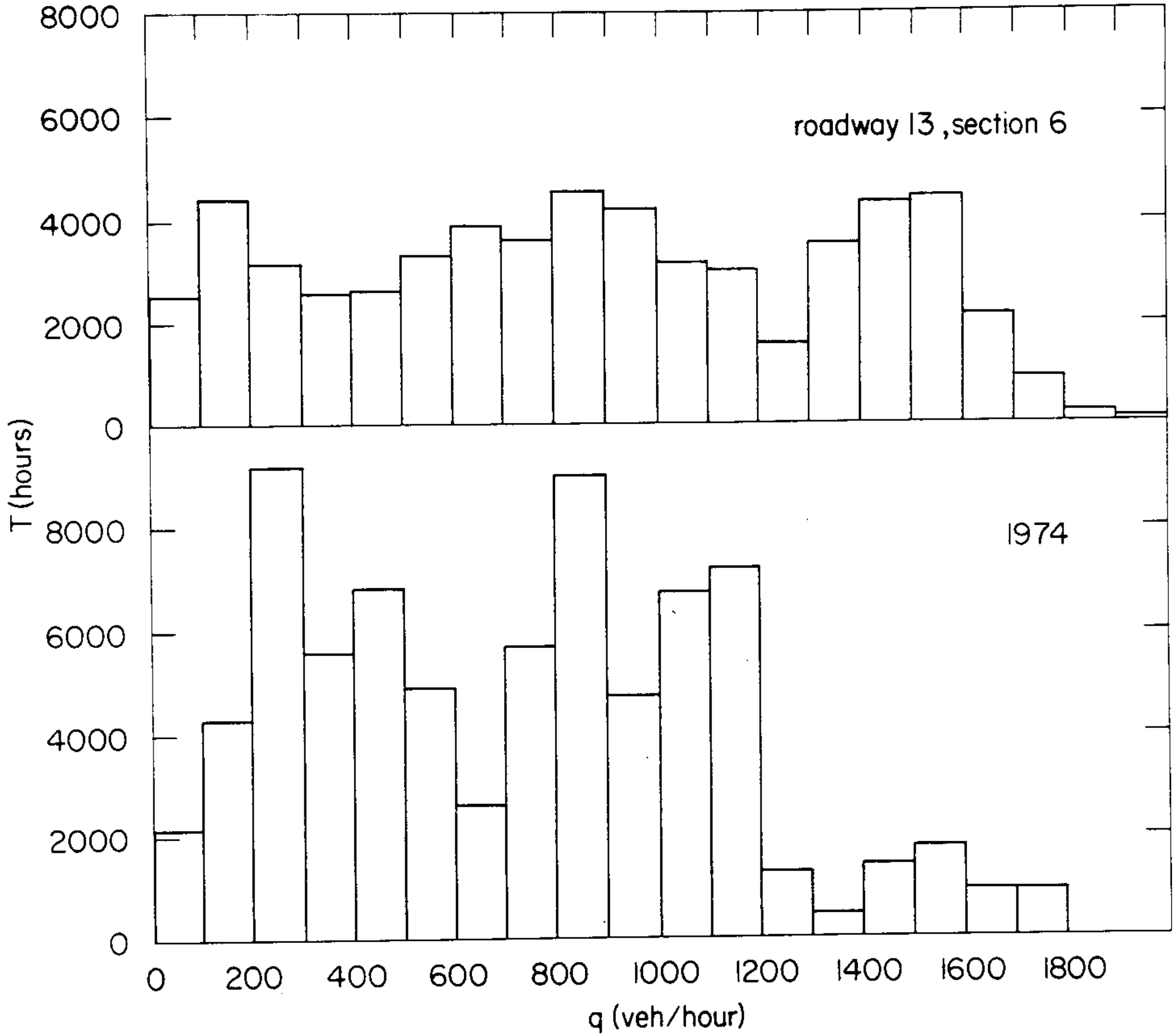



Fig. 3: An example of $T(q)$ distributions for the time-sequence (upper part) and cross sectional (lower part) analyses.

$$A_r\,(q) = \frac{A_{dw}\,(q)\cdot 10^6}{q} \tag{2}$$

where :

$A_{dw}(q)$ = weighted accident density (acc/$10^3$ km-hour), which means the accident density (acc/$10^3$ km) per one hour exposure of traffic flows within the interval $q \pm \Delta q$; and

$A_r(q)$ = accident rate (acc/$10^6$ veh-km).

The models which were found appropriate for the analyses made are power functions (same as the models in phases I and II, see Fig. 1) :

$$A_{dw}(q) = a_3 q^{p_3} \tag{3}$$

$$A_r(q) = a_4 q^{p_4} \tag{4}$$

where $a_3$, $a_4$ and $p_3$, $p_4$ are constant parameters which are determined by regression technique. Note that the powers $p_3$ and $p_4$ determined the functional tendency (certainly for positive $a_3$, $a_4$ values) :

$p_3$, $p_4 > 1$ => convex upward;
$p_3$, $p_4 = 1$ => linear upward;
$1 > p_3$, $p_4 > 0$ => concave upward;
$p_3$, $p_4 = 0$ => constant linear; and
$0 > p_3$, $p_4$ => convex downward

The models represented by Eqs. (3) and (4) are fitted separately for single- and multi-vehicle accidents. The sum of these models for each category of analysis (time-sequence and cross sectional), might reveal the possible dependency between the total accidents and $q$, and might also yield the earlier mentioned U-shaped function. If the latter is the result, then one can arrive at equations satisfying optimum conditions, i.e. $dA_{dw}/dq = 0$

or $dA_V/dq = 0$, which yield the following optimum parameters :

$$q_{do} = \left(\frac{-p_3'' \ a_3''}{p_3' \ a_3'}\right)^{1/(p_3'-p_3'')}$$

$$q_{ro} = \left(\frac{-p_4'' \ a_4''}{p_4' \ a_4'}\right)^{1/(p_4'-p_4'')} \qquad (5)$$

where the prime represents the parameters of Eqs. (3) and (4) for single-vehicle accidents and the double prime - for multi-vehicle accidents. The substitution of $q_{do}$ and $q_{ro}$ (provided that each is a positive value) in Eqs. (3) and (4), respectively, gives accordingly the optimum weighted accident density measure, $A_{dwo}$, and the optimum accident rate measure $A_{ro}$.

5. RESULTS AND FINDINGS

The results of the time-sequence and cross sectional analyses are summarized in Tables 1 and 2, respectively. These regression results, by accident type, refer to Eqs. (3), (4), and (5), and are accompanied by the standard error values ($SE_{dw}$ and $SE_r$ in units of $A_{dw}$ and $A_r$, respectively).

The results given in Tables 1 and 2 are shown in Figs. 4 and 5 for multi and single-vehicle accidents, respectively, and also an attempt is made in Fig. 6 to show the summation of the curves. Note that in Figs. 4 - 6 all the curves are within the actual q range. From these results and analyses, four major findings could be emphasized :

Table 1. Regression results of the time-sequence analysis.

| Roadway number | Section number | Accident type | $A_{dw} = a_3 q^{p_3}$: $a_3$ | $p_3$ | $SE_{dw}$ | $q_{do}$ | $A_{dwo}$ | $A_r = a_4 q^{p_4}$: $a_4$ | $p_4$ | $SE_r$ | $q_{ro}$ | $A_{ro}$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 | 9 | s* | 20.0 | -0.89 | 0.235 | 263.24 | $2.8 \cdot 10^{-4}$ | 20.31 | -0.56 | 0.43 | 435.54 | 0.81 |
| | | m** | $1.29 \cdot 10^{-5}$ | 1.59 | 0.366 | | | $2.9 \cdot 10^{-9}$ | 2.90 | 1.05 | | |
| 1 | 10 | s | 550.0 | -1.41 | 0.356 | 413.16 | $1.9 \cdot 10^{-4}$ | $35.07 \cdot 10^{6}$ | -3.04 | 1.92 | 539.41 | 0.31 |
| | | m | $2.04 \cdot 10^{-6}$ | 1.88 | 0.433 | | | $5.07 \cdot 10^{-12}$ | 3.82 | 1.85 | | |
| 11 | 8 | s | 0.03 | 0.34 | 0.362 | - | - | 37.02 | -0.67 | 0.54 | - | - |
| | | m | $5 \cdot 10^{-4}$ | 1.16 | 1.130 | | | 2.72 | -0.09 | 1.79 | | |
| 11 | 9 | s | 0.27 | -0.17 | 0.103 | - | - | 24.75 | -0.73 | 0.18 | - | - |
| | | m | $2.2 \cdot 10^{-4}$ | 1.19 | 1.930 | | | 4.15 | -0.37 | 0.23 | | |
| 13 | 3 | s | 0.17 | -0.12 | 0.802 | - | - | 9.81 | -0.57 | 0.61 | - | - |
| | | m | $4.86 \cdot 10^{-3}$ | 0.72 | 0.401 | | | 4.80 | -0.28 | 0.80 | | |
| 13 | 5 | s | 120.0 | -1.00 | 0.667 | 252.31 | $7.7 \cdot 10^{-4}$ | $3.86 \cdot 10^{9}$ | -3.77 | 1.48 | 484.20 | 0.70 |
| | | m | $3.33 \cdot 10^{-5}$ | 1.64 | 1.460 | | | $2.14 \cdot 10^{-8}$ | 2.71 | 1.40 | | |
| 13 | 6 | s | 1.0 | -0.36 | 0.160 | 114.00 | $1.8 \cdot 10^{-4}$ | $5.14 \cdot 10^{4}$ | -2.09 | 1.15 | 424.23 | 0.33 |
| | | m | $1.37 \cdot 10^{-4}$ | 1.26 | 0.733 | | | $5 \cdot 10^{-7}$ | 2.10 | 1.21 | | |
| 21 | 4 | s | 10.0 | -0.66 | 0.567 | 112.91 | $7.3 \cdot 10^{-4}$ | $6.07 \cdot 10^{6}$ | 2.99 | 1.78 | - | - |
| | | m | $2 \cdot 10^{-3}$ | 1.05 | 0.201 | | | 2.18 | 0.03 | 0.36 | | |

* single-vehicle accidents

** multi-vehicle accidents

Table 2. Regression results of the cross sectional analysis.

| Year | Accident type | $A_{dw} = a_3 q^{p_3}$ | | | | | $A_r = a_4 q^{p_4}$ | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | $a_3$ | $p_3$ | $SE_{dw}$ | $q_{do}$ | $A_{dwo}$ | $a_4$ | $p_4$ | $SE_r$ | $q_{ro}$ | $A_{ro}$ |
| 1967 | s* | $6.33 \cdot 10^{3}$ | 0.56 | 0.146 | - | - | 7.68 | -0.46 | 0.36 | 258.98 | 0.67 |
| | m** | $6.5 \cdot 10^{-6}$ | 1.75 | 0.167 | | | $5.8 \cdot 10^{-11}$ | 3.77 | 2.54 | | |
| 1968 | s | $1.62 \cdot 10^{-2}$ | 0.31 | 0.023 | - | - | 25.30 | -0.76 | 0.26 | - | - |
| | m | $3.7 \cdot 10^{-3}$ | 0.62 | 0.267 | | | 0.83 | -0.08 | 0.44 | | |
| 1969 | s | $3.13 \cdot 10^{-4}$ | 0.94 | 0.133 | - | - | 0.37 | -0.06 | 0.22 | - | - |
| | m | $4.66 \cdot 10^{-5}$ | 1.41 | 0.167 | | | 0.07 | 0.35 | 0.37 | | |
| 1970 | s | $3.33 \cdot 10^{-2}$ | 0.16 | 0.102 | - | - | 1151.52 | -1.54 | 0.80 | 338.82 | 1.17 |
| | m | $2.98 \cdot 10^{-4}$ | 1.21 | 0.567 | | | 0.28 | 0.22 | 0.80 | | |
| 1971 | s | 10.0 | -0.74 | 0.403 | 271.02 | $2.7 \cdot 10^{-4}$ | $5.33 \cdot 10^{7}$ | -3.29 | 0.26 | 449.39 | 0.25 |
| | m | $4.16 \cdot 10^{-5}$ | 1.41 | 0.562 | | | $2.33 \cdot 10^{-7}$ | 2.19 | 0.89 | | |
| 1972 | s | $9.66 \cdot 10^{-3}$ | 0.44 | 0.213 | - | - | 15.12 | -0.63 | 0.38 | - | - |
| | m | $3.36 \cdot 10^{-3}$ | 0.83 | 0.833 | | | 3.81 | -0.18 | 0.93 | | |
| 1974 | s | 0.10 | 0.01 | 0.333 | - | - | $2.25 \cdot 10^{5}$ | -2.22 | 0.36 | - | - |
| | m | $6.1 \cdot 10^{-3}$ | 0.69 | 0.623 | | | 5.75 | -0.30 | 0.93 | | |
| 1975 | s | 3.43 | -0.58 | 0.202 | 204.37 | $2.4 \cdot 10^{-4}$ | $1.43 \cdot 10^{6}$ | -2.69 | 0.26 | 929.07 | 0.46 |
| | m | $2.1 \cdot 10^{-4}$ | 1.12 | 0.803 | | | 0.24 | 0.09 | 0.83 | | |

* single-vehicle accidents

** multi-vehicle accidents

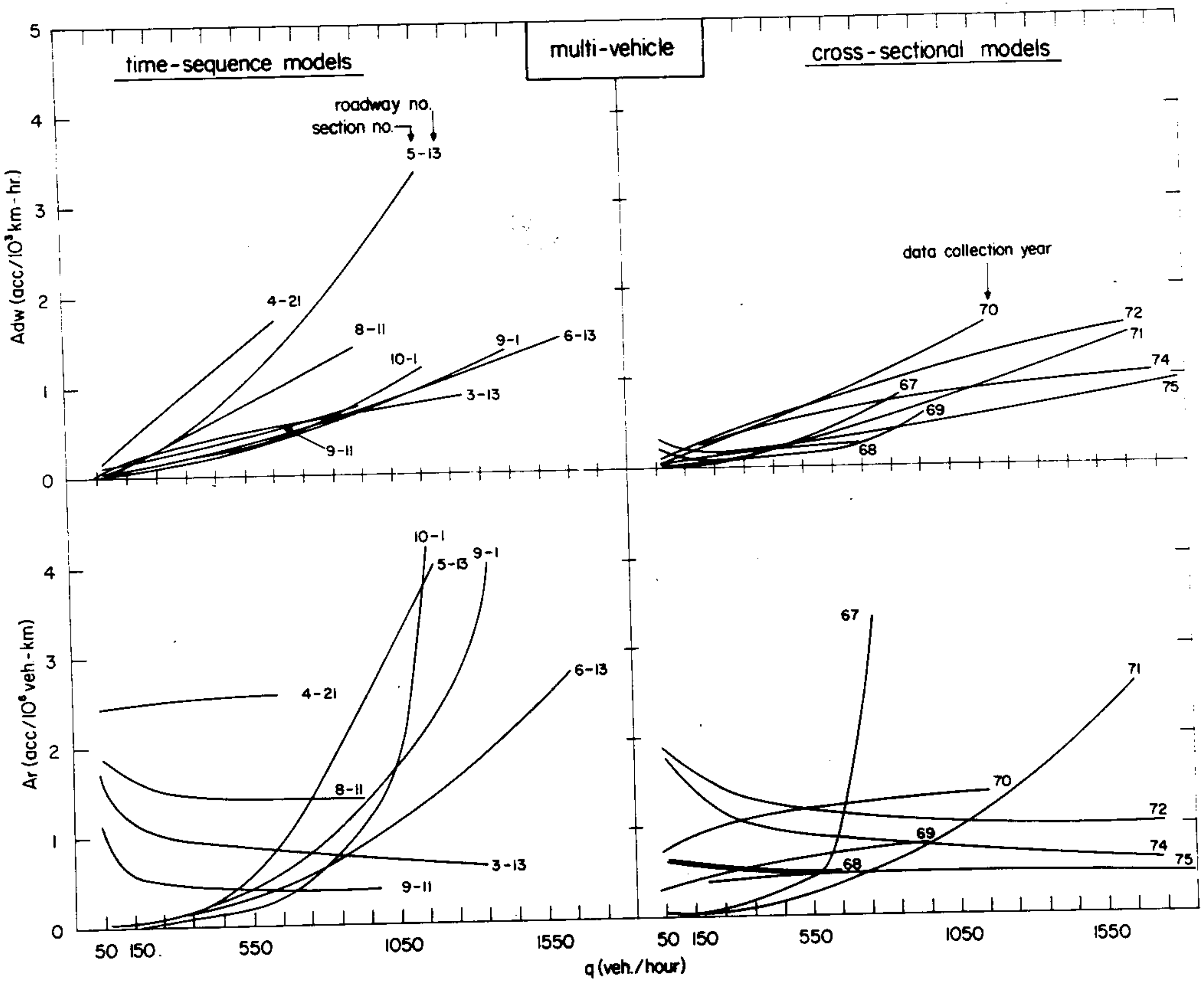



Fig. 4: The results of the time-sequence and the cross-sectional models for multi-vehicle accidents.

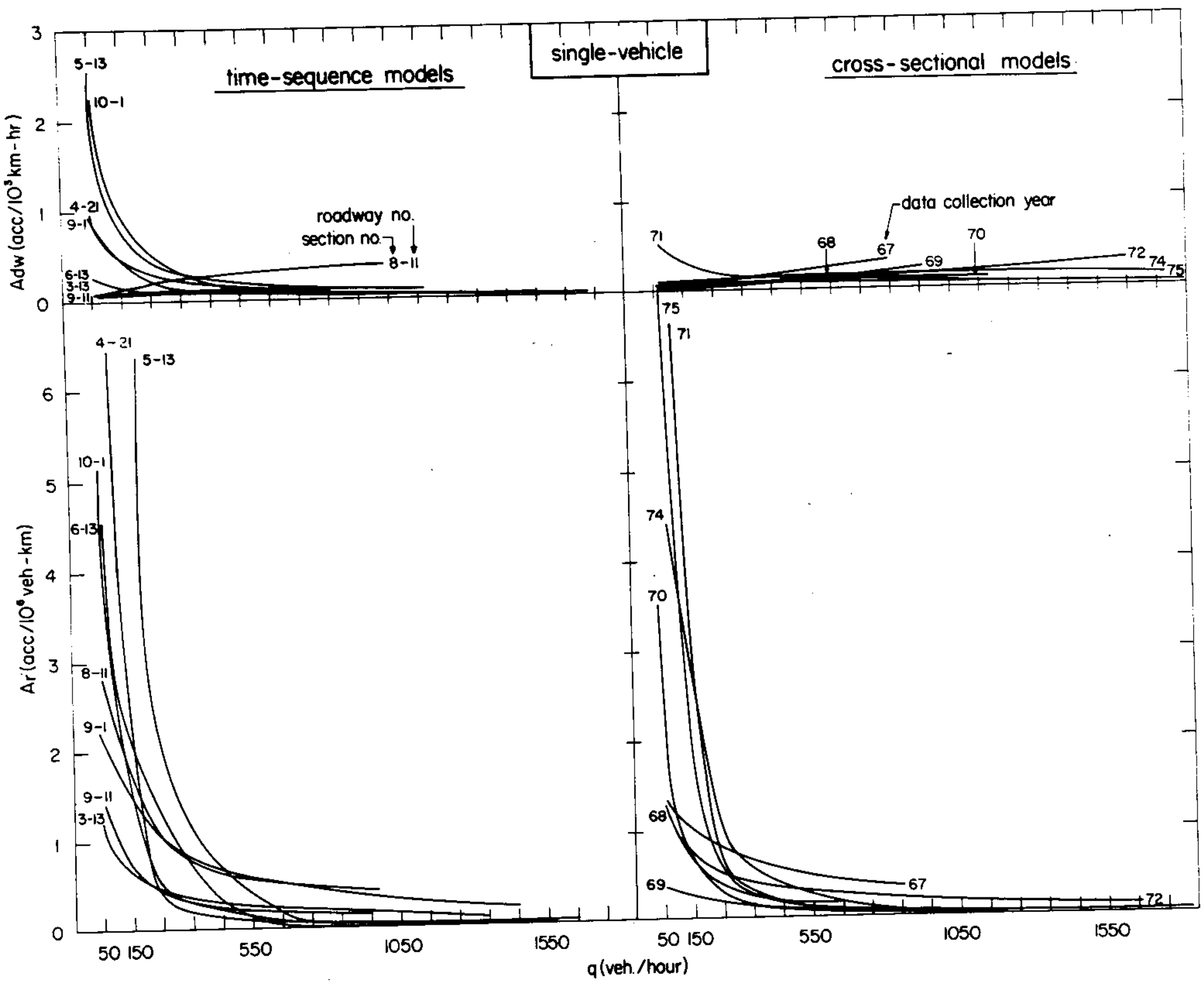



Fig. 5: The results of the time-sequence and the cross-sectional models for single-vehicle accidents.

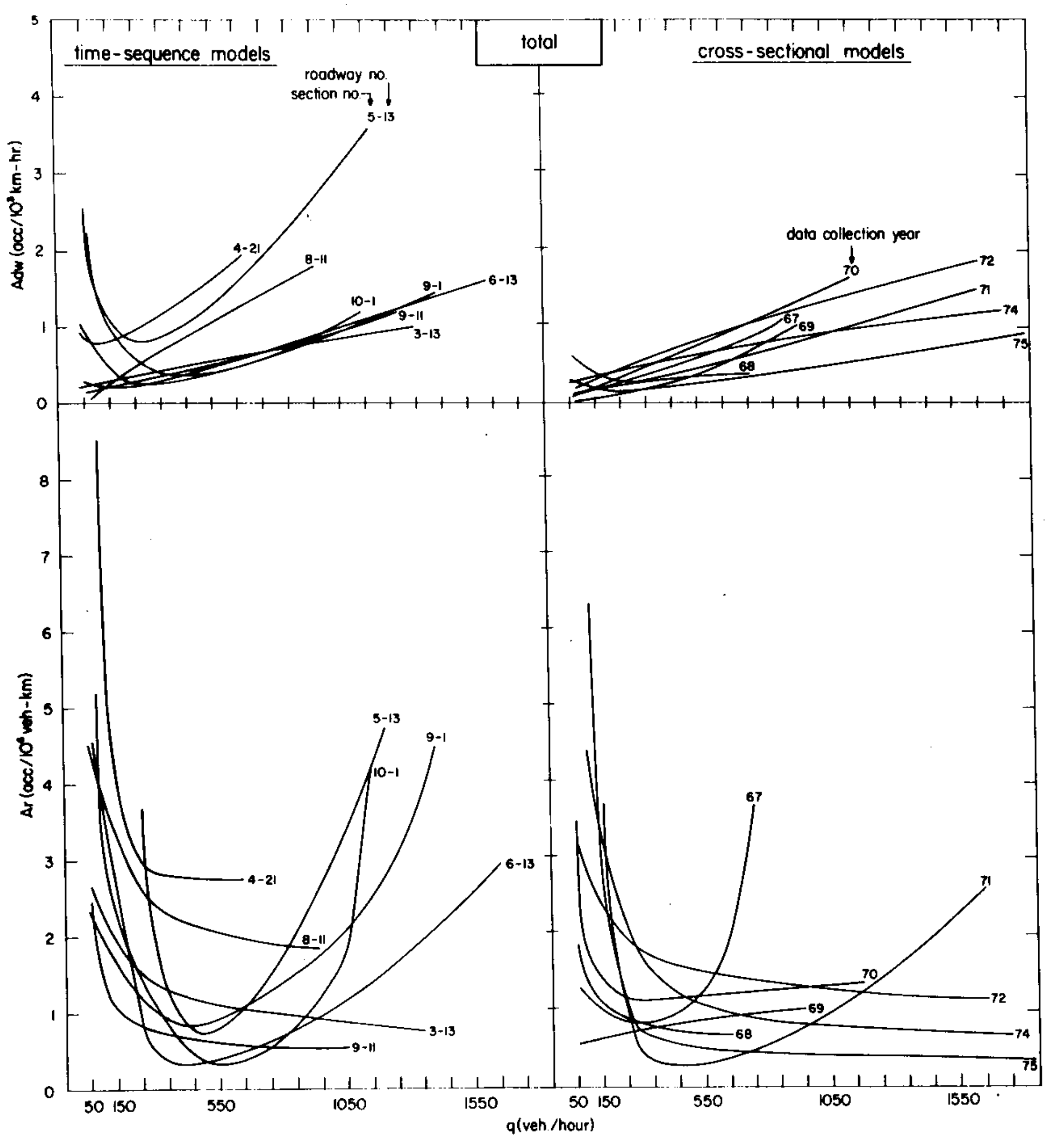



Fig. 6: The results (summation) of both multi- and single-vehicle accidents.

(a) <u>For the multi-vehicle accident models</u> all $p_3 > 0$ with a convexity tendency for the time-sequence models $(1 < p_3 < 2)$, and mixed functional tendency for the cross sectional models. While $A_{dw}(q)$ is always increasing with $q$, the $A_r(q)$ is either increasing or slightly decreasing with $q$ $(-0.37 < p_4 < 3.82)$. There are two interesting observations : (i) three out of four time-sequence models in which $A_r(q)$ does not increase sharply with $q$ are characterized by low upper bound value of the $q$ range; and (ii) the two cross sectional models obtained for two years before the energy crisis (1971, 1972), indicate lower safety level than the two models after the energy crisis (1974, 1975), where all four models have similar $q$ ranges.

(b) <u>For the single-vehicle accident models</u> all $p_4 < 0$ and, hence, indicating convex downward curves in the $A_r - q$ plane. On the other hand, there is a mixed functional tendency for the $A_{dw} - q$ relationship. The comparison between single - and multi-accident models demonstrates less curve-dispersion of the former for both the time-sequence and cross sectional analyses.

(c) <u>For the summation of the single - and multi-vehicle accident models</u> half of the $A_r(q)$ models are characterized by U-shaped curves, and the remaining half by convex downward curves. Three out of four of the $A_r(q)$ models which do not have U-shaped curves are indicated by upper bound on the $q$ range below 1,000 veh/hour per two lanes, and hence, might be (one of) the reasons for this observation. Nevertheless, the optimum average $q_{ro}$ value for the minimum $A_r(q)$ value is 500 veh/hour for those U-shaped models. This optimum value results from opposing tendencies of multi-vehicle accident models ($A_r$ increasing with q) and single-vehicle accident models ($A_r$ decreasing with q). The average $q_{ro}$ value is below these (600 to 1300 veh/hour) which were determined by Leutzbach [1966] and Gwynn [1967] - probably due to a higher upper bound on the q range (about 3000 veh/hour) than that indicated in Fig. 8 .

(d) The hypothesis made by Brilon [1976], that the minimum $A_r$ value (for total accidents) is obtained for the most frequent range of q, is not strengthened by our data. In fact, the opposite is the case, since out of sixteen data sets of T(q) distributions, none agrees with Brilon's hypothesis.

## 6. MATRIX REPRESENTATION OF THE RESULTS

Following a similar method to that outlined by Ceder & Livneh [1978], (phases I and II indicated in Fig. 1), the results are plotted on the matrix : the y-axis is the logarithm scale of $a_3$ or $a_4$, and the x-axis represents the scale for $p_1$ or $p_2$, respectively.

The log $a_3$ values versus $p_3$ values are illustrated in Fig. 7, and the log $a_4$ values versus $p_4$ values - in Fig. 8 From these figures, one can eminently observe the linear dependency that exists between the variables. Consequently, a linear regression procedure was applied to each set of results. The fitted models are shown in Figs. 7 and 8 and are indicated in Table 3 with their coefficient of determination $r^2$. In addition, the F statistic is used to examine the possibility of combining the linear models of both the time-sequence and cross sectional

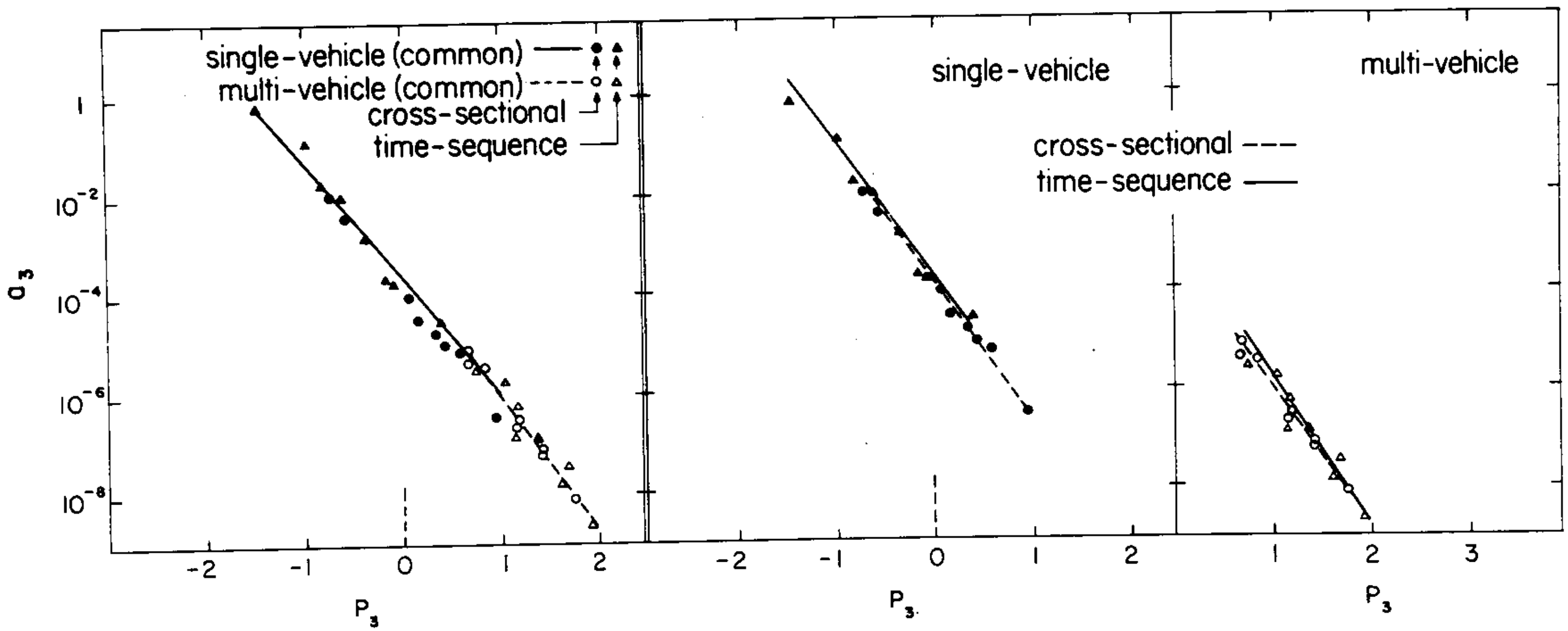



Fig. 7: Matrix representation of the model's parameters $a_3$ and $p_3$, where the left part distinguishes the results by accident type, and the remaining part - by both accident type and type of analysis.

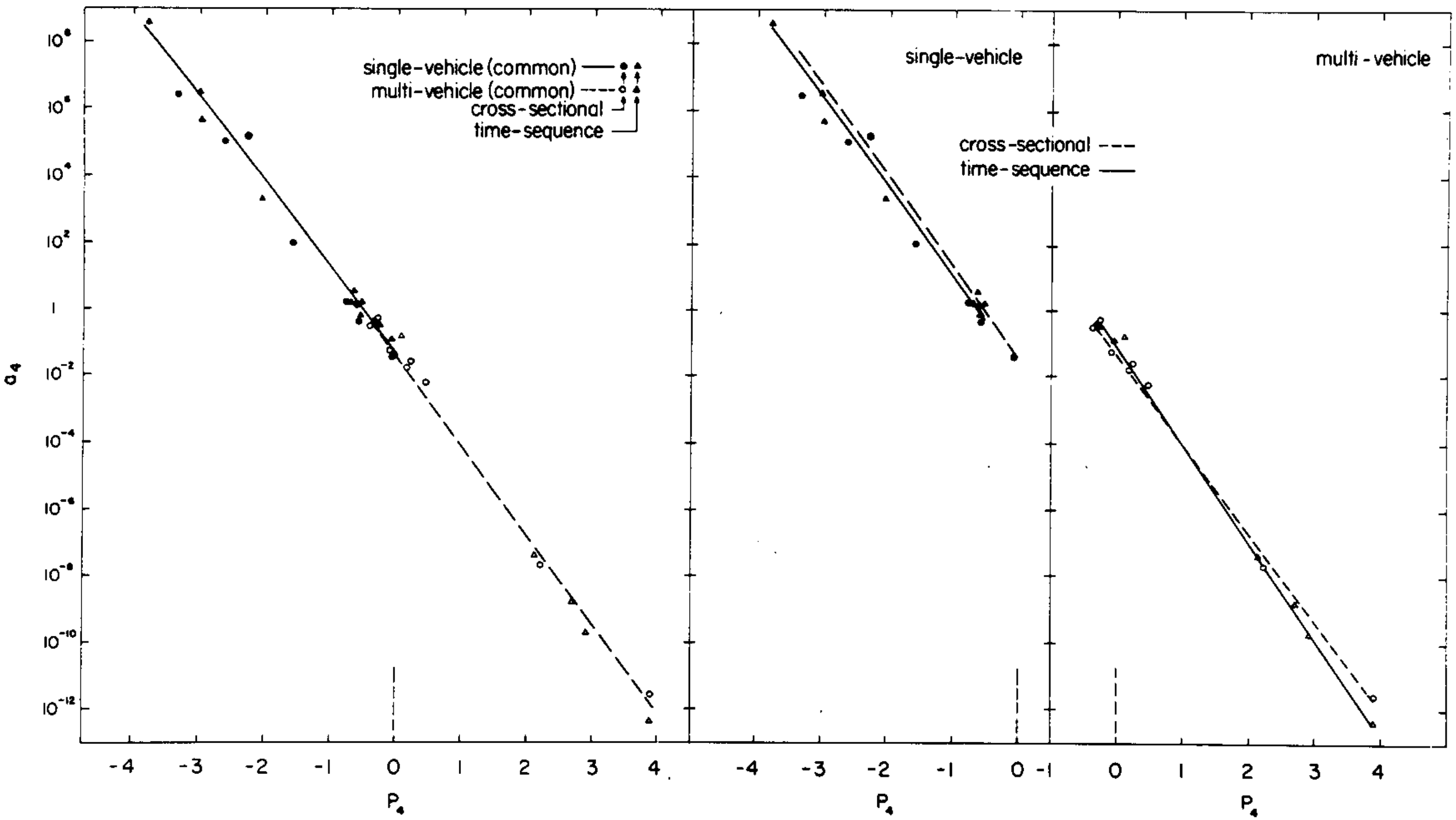



Fig. 8: Matrix representation of the model's parameters $a_4$ and $p_4$, where the left part distinguishes the results by accident type, and the remaining part - by both accident type and type of analysis.

Table 3. Results of the linear dependancy between $a_3$ and $P_3$ and between $a_4$ and $P_4$.

| model | parameters | multi-vehicle time-sequence | multi-vehicle cross-sectional | multi-vehicle common | single-vehicle time-sequence | single-vehicle cross-sectional | single-vehicle common |
|---|---|---|---|---|---|---|---|
| $\log a_3 = \alpha_o + \alpha_1 P_3$ | $\alpha_o$ | 0.40 | -0.47 | -0.24 | -0.34 | -0.94 | -0.76 |
| | $\alpha_1$ | -3.06 | -2.7 | -2.73 | -2.11 | -2.57 | -2.52 |
| | $r^2$ | 0.98 | 0.99 | 0.99 | 0.98 | 0.99 | 0.99 |
| | $P_3$ lowest | 0.72 | 0.62 | 0.62 | -1.71 | -0.74 | -1.71 |
| | $P_3$ highest | 1.88 | 1.75 | 1.88 | 0.34 | 0.94 | 0.94 |
| | $A_{dw}$* | 2.51 | 0.34 | 0.57 | 0.46 | 0.11 | 0.17 |
| | $q_d$* | 1148.15 | 501.19 | 537.03 | 128.82 | 371.54 | 331.13 |
| $\log a_4 = \beta_o + \beta_1 P_4$ | $\beta_o$ | -0.01 | -0.16 | -0.10 | -0.35 | -0.43 | -0.35 |
| | $\beta_1$ | -2.92 | -2.75 | -2.85 | -2.54 | -2.81 | -2.64 |
| | $r^2$ | 0.99 | 0.99 | 0.99 | 0.94 | 0.97 | 0.98 |
| | $P_4$ lowest | -0.37 | -0.30 | -0.37 | -3.77 | -3.29 | -3.77 |
| | $P_4$ highest | 3.82 | 3.77 | 3.82 | -0.56 | -0.06 | -0.06 |
| | $A_r$* | 0.98 | 0.69 | 0.79 | 0.44 | 0.37 | 0.45 |
| | $q_r$* | 831.76 | 562.34 | 707.95 | 346.74 | 645.65 | 36.52 |

analysis. This examination leads, indeed, to a common model through a three-stage test: (1) between the variances; (2) between $\alpha_1$'s or $\beta_1$'s ; and (3) between $\alpha_o$'s or $\beta_o$'s (see Table 3). It was found that for both the $A_{dw}(q)$ and $A_r(q)$ models, there is no significant difference between the linear models at the 95% level. The common models are specified in Table 3, and are shown on the left part of Figs. 7 and 8 with respect to each type of accident. The remaining parts of these figures illustrate, for each type of accident, separate time-sequence and cross sectional models. Some of the findings mentioned in the previous section are clearly and systematically demonstrated in this matrix representation.

Each linear model shown in Table 3 represents a family of curves which intersect at a unique point. For example, this point in the $A_r - q$ plane $(A_r^*, q_r^*)$ is obtained by :

$$A_r^* = a_4 (q_r^*)^{p_4} \quad ,$$

$$\log A_r^* = \log a_4 + p_4 \log \ (q_r^*)$$

and therefore, $\log A_r^* = \beta_o$, $\log \ (q_r^*) = - \beta_1$ , and similarly, $\log A_{dw}^* = \alpha_o$, $\log \ (q_d^*) = - \alpha_1$. These intersecting points are symbolized with an asterisk in Table 3.

## 7. EXAMPLES OF A PRACTICAL APPLICATION

Knowledge of the proper relationship (and limitations involved) between $A_{dw}$ or $A_r$ and q is important from various aspects : traffic planning, design, operation and research. This section, however, introduces examples of only one practical application. That is, the evaluation of the safety level either before and after implementation of a roadway improvement project, or after short-term operation of a new roadway section. This practical application is discussed also in phases I and II (see Fig. 1), in view of the relationship between measures of accidents and ADT.

Example 1. It is assumed that in section No. 6 of roadway No. 13, a safety improvement was carried out in January, 1976. The data collected after one year are :

| | | $A_d$ (q) | | $A_{dw}$ (q) | | $A_r$ (q) | |
|---|---|---|---|---|---|---|---|
| ranges of q | T(q) | multi-veh. | single-veh. | multi-veh. | single-veh. | multi-veh. | single-veh. |
| 0 - 500 | 1044 | 0.114 | 1.699 | 0.109 | 0.163 | 0.436 | 0.652 |
| 500 -1000 | 261 | 0.100 | 0.120 | 0.383 | 0.460 | 0.511 | 0.613 |
| 1000 -1500 | 522 | 0.548 | 0.175 | 1.050 | 0.335 | 0.840 | 0.268 |
| 1500 -2000 | 4437 | 4.138 | 0.106 | 0.933 | 0.024 | 0.533 | 0.014 |

with suitable units to those indicated for the models. The question is, whether the level of safety improved, and to what magnitude. According to this research approach, the results derived by the time-sequence analysis can be applied to this example. Hence, the evaluation procedure is based on the results indicated in Table 1 for roadway No. 13, section No. 6 for each q range along with a confidence interval. Since the power functions are intrinsically linear (can be expressed by natural logarithms, in a linear form), a 95% confidence limit can be found for the new data (after the improvement), $\ln(A_r)_{new}$ or $\ln(A_{dw})_{new}$ in the transformed plane according to :

$$\pm\ t(n-2;0.975)\cdot s\cdot \{1 + \frac{1}{n} + \frac{[\ln(A_r)_{new} - (\overline{\ln A_r})]^2}{\sum_i[\ln(A_r)_i - (\overline{\ln A_r})]^2}\}^{\frac{1}{2}}$$

for a two-sided 95% level t-test using the t table with (n-2) degrees of freedom, and where $n$ is the number of data points, $s$ is the standard error for either $\ln A_r$ or $\ln A_{dw}$, and $(\overline{\ln A_r})$ or $(\overline{\ln A_{dw}})$ is the mean value of $\ln A_r$ or $\ln A_{dw}$, respectively. Note that this confidence limit is based on the assumption that the residuals in the transformed scale are normally distributed with mean zero and constant variance. If, for example, the confidence interval is $\pm\ SE_{dw}$ and $\pm\ SE_r$ for $A_{dw}(q)$ and $A_r(q)$, respectively, then by substituting the data in the models, the following results are obtained :

| | $A_{dw}(q)$ | | | | $A_r(q)$ | | | |
|---|---|---|---|---|---|---|---|---|
| q | multi-veh. | safety change | single veh. | safety change | multi-veh. | safety change | single veh. | safety change |
| 250 | 0.140 | ** | 0.137 | *** | 0.054 | ** | 0.500 | *** |
| 750 | 0.558 | ** | 0.092 | **** | 0.545 | ** | 0.050 | *** |
| 1250 | 1.062 | ** | 0.077 | **** | 1.594 | ** | 0.017 | *** |
| 1750 | 1.622 | ** | 0.068 | ** | 3.231 | * | 0.009 | *** |

* significant improvement

** an improvement, but not significant

*** a deterioration, but not significant

**** significant deterioration

The asterisks attempt to interpret the results with respect to the confidence interval. It is worth noting that the data could also be analyzed by the $A_d(ADT)$ and $A_r(ADT)$ models specified for the considered roadway section in Table 3 of Ceder & Livneh [1978]. In the latter case, the results indicate significant improvement for multi-vehicle accidents and significant deterioration for single-vehicle accidents. Certainly, the consideration of q instead of ADT determines, more specifically, the relative changes in the safety level after the improvement. Furthermore, the knowledge of the exposure time for each q range might lead to isolation of the problematic daily hours from the safety standpoint.

Example 2. If the data from example 1 are associated with a new four-lane section (also after one year of operation), then it is only possible to select an appropriate cross sectional model. That is, the time-sequence model cannot be applied due to lack of comparable basis and/or information. The selection of a cross sectional model might be based on two criteria : (i) that it includes the upper bound q range of the considered data; and (ii) that it reflects the environmental characteristics of the considered new roadway section (usually exhibited by the latest year model which satisfies criterion (i)). Consequently, the model selected is that of 1975 and, in a similar way to example 1, the results are obtained by substituting the data in the models indicated for 1975 in Table 2 (based on $SE_{dw}$ and $SE_r$):

| | $A_{dw}(q)$ | | | | $A_r(q)$ | | | |
|---|---|---|---|---|---|---|---|---|
| q | multi-veh. | Safety change | Single-veh. | Safety change | multi-veh. | Safety change | Single-veh. | Safety change |
| 250 | 0.102 | *** | 0.139 | *** | 0.394 | *** | 0.507 | *** |
| 750 | 0.349 | *** | 0.074 | **** | 0.435 | *** | 0.026 | **** |
| 1250 | 0.618 | *** | 0.055 | **** | 0.456 | *** | 0.007 | *** |
| 1750 | 0.900 | *** | 0.045 | ** | 0.467 | *** | 0.003 | *** |

* significant improvement
** an improvement, but not significant
*** a deterioration, but not significant
**** significant deterioration

} with respect to similar roadway sections which have the same characteristics

Perhaps the major finding is the significant deterioration in single-vehicle accidents at the mid - q range. When considering the $A_d(ADT)$ and $A_r(ADT)$ models in phase I, the results are a significant improvement in the safety level of the new roadway section for multi-vehicle accidents and an improvement, but not significant, for single-vehicle accidents. In essence, in this example, the ignorance of q produces a situation in which the relative safety deterioration at the mid - q range cannot be detected.

Example 3. An alternative means of estimating the $A_{dw}(q)$ and $A_r(q)$ models is use of the interrelationship between $p_3$ or $p_4$ and log $a_3$ or log $a_4$, respectively. In fact, for any given q, T(q) and the number of accidents on a four-lane section, one can obtain a crude estimation of such models. For example, on a four-lane section, the hourly flow, during specific daily hours, increased from 500 to 800 veh/hour for two lanes (due to either a closure of a parallel road or by means of traffic direction). The question is whether the level of safety has been changed and to what magnitude. The $A_{dw}$ and $A_r$ values for q = 500 veh/hr and this q value are then substituted in the common linear models indicated in Table 3. This substitution determines the appropriate models to be used for the "after change" (q = 800 veh/hr ) data. That is, $A_r(q)$ for multi-vehicle accidents is obtained through determination of $a_4$ and $p_4$, based on $A_r = 1.10$ acc/$10^6$ veh-km :

$$\left.\begin{array}{l} 1.1 = a_4 \cdot 500^{p_4} \\ \log a_4 = -0.1 - 2.85 p_4 \end{array}\right\} \begin{array}{l} a_4 = 370 \\ p_4 = 0.936 \end{array}$$

and similarly, $A_r(q)$ for single-vehicle accidents and $A_{dw}(q)$ for multi and single vehicle accidents can be calculated. The complete 'before and after' data and results are :

| Accident type | Situation / average q | Before 500 | After 800 |
|---|---|---|---|
| multi-vehicle | $A_{dw}$ | 0.55 | 0.90 |
| | $a_3$ | 0.011 | - |
| | $p_3$ | 0.633 | - |
| | $\overline{A_{dw}}$ | - | 0.74 |
| single-vehicle | $A_{dw}$ | 0.20 | 0.20 |
| | $a_3$ | 0.024 | - |
| | $p_3$ | 0.341 | - |
| | $\overline{A_{dw}}$ | - | 0.23 |
| multi-vehicle | $A_r$ | 1.10 | 1.12 |
| | $a_4$ | 370 | - |
| | $p_4$ | -0.936 | - |
| | $\overline{A_r}$ | - | 0.71 |
| single-vehicle | $A_r$ | 0.40 | 0.25 |
| | $a_4$ | 62.55 | - |
| | $p_4$ | -0.813 | - |
| | $\overline{A_r}$ | - | 0.27 |

where $\overline{A_r}$ and $\overline{A_{dw}}$ are the results of substituting $q = 800$ in the models. The comparison between $A_{dw}$ and $\overline{A_{dw}}$ and between $A_r$ and $\overline{A_r}$ for $q = 800$ reveals that : (i) single-vehicle accidents have almost not changed, though the absolute $A_r$ value after the change decreased by 40%!; and, (ii) relative improvement is observed for the multi-vehicle accidents though the absolute $A_{dw}$ value after the change increased by 60%! For such applications, it is also advisable to check that the resultant $p_3$ and $p_4$ values are within or close to the indicated range in Table 3.

## CONCLUDING REMARKS

This research attempts to find quantitative models (power functions) to represent the possible dependency between two measures of accidents and the hourly flow for eight interurban road sections during an 8-year period.

From these attempts to search for proper relationships between measures of accidents and the hourly flow, it is apparent that the technical procedure involves a combination of two primary types of analyses : time-sequence (for each roadway section) and cross sectional (on a one year basis). For each type of analysis, the total accidents are primarily separated into multi- and single-vehicle accidents. The latter separation enables one to : (i) distinguish accident costs for each type of accident and for each q range; (ii) find the differential effects of traffic flow on each type of accident; and (iii) perform a more reliable safety evaluation for "before and after" studies.

Phases I and II of the overall study, described in Fig. 1, are concerned with the influence of ADT on the measures of accidents. However, this consideration by itself cannot explain the interactions between road accidents and traffic flow, since it is only based on a daily average.

The consideration of hourly flows provides a better understanding of these interactions. In addition, it is possible to move one step forward in order to further understand the accidents/traffic flow dependency by separating the hourly flow into free-flow and congested-flow. This separation into components of both type of traffic flow and type of accident will ultimately lead toward more accurate accident prediction based on the traffic flow. The following paper [Ceder, 1979], describes this further analysis (phase IV in Fig. 1), and attempts also to determine and compare the probabilities for each accident/flow type component.

REFERENCES

Baker, W.T. & Gwynn, D.W., Relationships of accident rates with hourly traffic volumes. Division of Research & Evaluation, New Jersey Department of Transportation, 1968.

Belmont, D.M., Effect of average speed and volume on motor vehicle accidents on two-lane tangents. HRB Proc. 32, 383-395, 1953.

Brilon, W., Unfallgeschehen und Verkehrsablauf. Forschung Strassenbau und Strassenverkehrstechnik, Heft 201, 1976.

Ceder, A. & Livneh, M., Further evaluation of the relationships between road accidents and average daily traffic. Accident Analysis & Prevention, 10, 95-109, 1978.

Ceder, A., Relationships between road accidents and hourly traffic flow: II. Probabilistic approach. Submitted to Accident Analysis & Prevention.

Chapman, R.A., Some relationships between road accident frequencies and measures of exposure to risk. Ph.D. dissertation, University College London, January, 1971.

Gwynn, D.W., Relationship of accident rates and accident involvements with hourly volume. Traffic Quarterly XXI, 3, July, 1967.

Leutzbach, W., Traffic accidents and traffic flow on the Frankfurt-Mannheim motorway-investigation of the relationship between traffic accidents and traffic flow (in German). Hessisches Ministerium fur Wirtschaft und Verkehr, March, 1966.

Leutzbach, W., Siegener, W., & Wiedemann, R., On the connection between traffic accidents and traffic volume on a section of a German autobahn. Accident Analysis & Prevention, 2, 143-146 (in German), 1970.

Pfundt, K., Three difficulties in the comparison of accident rate. Accident Analysis & Prevention, 1, 253-259, 1969.

Smeed, R.J., Accident rates. International Road Safety & Traffic Review 3(2), 30-40, 1955.

RELATIONSHIPS BETWEEN ROAD ACCIDENTS & HOURLY TRAFFIC FLOW:
II. PROBABILISTIC APPROACH

by

AVISHAI CEDER

Transportation Research Institute,

Technion-Israel Institute of Technology, Haifa, Israel

A B S T R A C T

This is a continuation of the investigation into relationships between accident rate and hourly traffic flow that was outlined in Part I of the research. The underlying study attempts to determine appropriate models for single- and multi-vehicle accident rates in conjunction with free-flow and congested-flow conditions. For the free-flow data, the total accident rate-hourly flow curve follows the U-shaped configuration. This form is the result of a convex downward and a convex upward curve for single- and multi-vehicle accidents, respectively. For the congested-flow data (characterized by multi-vehicle accidents), the accident rate increases sharply with hourly flow. The models are applied to probabilistic aspects with consideration of a fitted, generalized (hourly flow dependent) headway model. The headway model represents the probability that two vehicles which are, even instantaneously, in a car-following mode, are in a potentially hazardous situation. The approach investigated is believed to be an essential input for both simulation studies and theoretical models of road traffic accidents.

## 1. INTRODUCTION

This is the second part of a continuation work to explore the interrelationship between road accidents and hourly traffic flow. In the first part, [Ceder & Livneh, 1980], the entire research based on nationwide data, is presented in four phases, with phase IV exhibited by this paper. The availability of data makes it possible to further separate the consideration of traffic flow components according to the following chain : ADT→Hourly Flow→Hourly Free and Congested-Flow. The latter separation, being outlined here, may be approached from both deterministic and probabilistic viewpoints. The probabilistic aspects, which are emphasized in this work, are both vindicated and based on determined deterministic relationships between an accident measure and the hourly (free/congested) flow.

## 2. DETERMINATION OF FREE-FLOW AND CONGESTED-FLOW CONDITIONS

At any instant, the driver of an automobile is confronted with a plethora of visual, aural, vestibular and pressure stimuli. Though motion of his vehicle causes continuous changes in these stimuli, the driver cannot, and need not, evaluate each change in each stimulus dimension. Of major interest are those stimuli which create a potential hazardous situation. One of the known analytical tools used to assess the safety of individual cars is the car-following theory [Edie, 1974]. This assumed that a driver would react to a stimulus generated by the nearest forward vehicle. The uniqueness of this theory is that it provides a bridge between the motion of individual cars and the entire traffic flow [Ceder, 1979]. However, the car-following rules are applicable only to those vehicles which are under the influence of other vehicles in the stream - behaviour particularly noticeable under congested-flow conditions.

Consequently, it is probable that as the traffic flow becomes more congested, the vehicles are more constrained and can hardly perform a desirable manoeuvre. Such situations have a direct bearing on rear-end and chain collisions, as they result from an inappropriate time lag of a following car response to a disturbance caused by the vehicle ahead. This argument leads to the separation of free-flow and congested-flow conditions, since they have different effects upon single- and multi-vehicle accidents. While free-flow conditions are characterized by both single-and multi-vehicle accidents, congested conditions are particularly characterized by multi-vehicle accidents.

It is difficult to distinguish between free-flow and congested-flow periods when considering only the hourly flow variable. Even in a well defined free-flow condition (say, from the traffic flow theory [Edie, 1974]), one can observe a platoon of vehicles moving under a congested mode; and vice versa in a congested-flow condition. Therefore, a safety-based criterion is established : The congested hourly flow periods are determined by those periods in which 95% or more of the overall accidents are multi-vehicle accidents, and furthermore, the proportion of rear-end collisions is 85% or more of total accidents. The remaining flow periods are considered free-flow conditions.

This criterion enables one to disregard single-vehicle accidents (skids, roll-over, running-off-the-road) in congested-flow conditions, and to consider primarily rear-end collisions (in comparison to head-on and angle collisions). The criterion has been applied to three four-lane, divided roadway sections (13-3, 13-5, and 13-6) mentioned in the first part of this work [Ceder & Livneh, 1980]. It turns out that only for $q \geqslant 1600$ veh/hour per direction of travel (two lanes) this criterion is satisfied. These congested hourly flow periods were observed only on one roadway section (13-6). It is interesting to note that the criterion under consideration seems also to be adequate for high traffic flows on four-lane autobahns in Germany, [Leutzbach et al., 1970 & Brilon, 1976].

The data considered in this paper are based on 8-year daylight (excluding night) accidents and hourly flow information for the above-mentioned three roadway sections. Certainly, for the three sections no attempt is made to differentiate between time-sequence and cross-sectional analyses, as in Part I of this work, but rather to emphasize free-flow and congested-flow conditions. The variable $q$ is analyzed for free-flow periods, with 100 veh/hour intervals within the range $0 < q < 1600$; and for congested flow periods, with 50 veh/hour intervals within the range $1600 \leq q \leq 1900$ (the maximum measured hourly flow on section 13-6 is 1900 veh/hour).

## 3. FREE-FLOW 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 I of this work, it can be surmised that on four-lane divided roadways, the majority of multi-vehicle 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 bumper) 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 manoeuvre (to avoid angle collision). On four-lane divided roadways, head-on collisions rarely occur, and can be disregarded. 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 external factor interferes with one of the vehicles).

According to the multiplicative law of probability :

$$P \left\{ \begin{matrix} \text{multi-veh.} \\ \text{accident} \end{matrix} \right\} = P\,(A \cap B) = p(A) \cdot p(B|A) \tag{1}$$

The measure $A_r(q)$ which is the accident rate (acc/$10^6$ veh-km) is explained in Part I of this work. This measure, divided by $10^6$, 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, $p(A) = p(h < T|q)$. The expression $p(B|A)$ in Eq. (1) depends also on $q$ and it is represented by a power function. These interpretations lead, from Eq. (1), to :

$$A'_{rF}(q) = p(h < T|q) \cdot \alpha_1 q^{\beta_1} \tag{2}$$

where $A'_{rF}(q) \cdot 10^{-6}$ is the (probability) measure for a multi-vehicle accident in each veh-km and $\alpha_1 q^{\beta_1} \cdot 10^{-6}$ represents the probability of a risky situation within the flow $q \pm \Delta q$ such that $h < T$; $\alpha_1, \beta_1$ are constants and certainly all the probability expressions are constrained so as to be less than 1.

Another approach in the interpretation of $p(B|A)$ is through the definition of $A_{dw}(q)$, from Part I of this work, which is the accident density (acc/$10^3$ km) per one hour exposure of traffic flows within the interval $q \pm \Delta q$. Under free flow of $q$ veh/hour, the number of headways which are potential for accidents is $q \cdot p(h < T|q)$. Assuming a steady flow for one hour :

$$A_{dw}(q) = q \cdot p(h < T|q) \cdot \gamma(q)$$

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

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

$$A''_{rF} = \alpha_2 q^{\beta_2} \tag{3}$$

where $A''_{rF} \cdot 10^{-6}$ is the (probability) measure for single-vehicle accidents in each veh-km and $\alpha_2, \beta_2$ are constants. The total measure for free-flow accidents is the summation of Eqs. (2) and (3) :

$$A_{rF} = p(h < T|q)\ \alpha_1 q^{\beta_1} + \alpha_2 q^{\beta_2} \tag{4}$$

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) :

$$A_{rC} = \alpha_3 q^{\beta_3} \tag{5}$$

where $A_{rC} \cdot 10^{-6}$ is the measure for multi-vehicle accidents in each veh-km and $\alpha_3, \beta_3$ are constants.

## 4. GENERALIZED HEADWAY MODELS

The first expression in Eq. (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 [Edie, 1974], only few works are concerned with model parameterization in respect to different $q$ values.

The search for a generalized headway model in terms of $q$ dependency has led to three different models - each is constructed with two components associated primarily with free and constrained vehicles. The first, reported by Grecco & Sword [1968], is an empirically-based model which considers Schuhl's distribution (best fitted distribution among : Schuhl , Gamma, Erlang and Pearson type III). Their results, from measurements on four-lane divided roadways, include an hourly flow variable on a per lane basis, $q_l$, and takes the form :

$$p(h\geq t) = 115\cdot 10^{-5}\cdot q_1\cdot e^{(1-t)/2.5} + (1-115\cdot 10^{-5}q_1)e^{t/(0.0122q_1-24)} \tag{6}$$

where $t\geq 1$ second and the parameters were derived according to the range $0 < q_1 \leq 700$ veh/hour.

The second work, reported by Dawson & Chimini [1968], describes what they called the hyperlang headway model. Their model is a linear combination of a translated exponential function and a translated Erlang function :

$$p(h\geq t) = \alpha_1 e^{(\delta_1-t)/(\gamma_1-\delta_1)} + \alpha_2 e^{k(\delta_2-t)/(\gamma_2-\delta_2)} \sum_{x=0}^{k-1} \frac{\left(k\frac{t-\delta_2}{\gamma_2-\delta_2}\right)^x}{x!} \tag{7}$$

where $\delta_1$, $\delta_2$ are the minimum headways and $\gamma_1$, $\gamma_2$ are average headways for free and constrained vehicles, respectively; $k$ 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 $(\alpha_1 + \alpha_2 = 1)$. The parameters of Eq. (7) were evaluated for one-lane flows (on a four-lane divided roadway) ranging from 158 (k=1) to 957 (k=6) veh/hour. Though Eq. (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, $\hat{f}(t)$, is given by :

$$\hat{f}(t) = \Phi\bar{g}(t) + A\lambda e^{-\lambda t}\int_0^t \bar{g}(u)\,du \,, \tag{8}$$

with $$\Phi = 1 - A\lambda \int_0^\infty e^{-\lambda t}\int_0^t \bar{g}(u)\,du\,dt$$

where $\Phi$ is the proportion of following (constrained) vehicles; $\bar{g}(t)$ 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 Wasielewski introduce the function $\bar{g}(t)$ and indicate that no significant disagreement is found between $\hat{f}(t)$ and the observed total headway probability density function; also, it is interesting to note that the flow dependence is considered only through the parameters $A$ and $\lambda$. These findings are based on 42,000 observed headways regarding 12 groups of hourly flows ranging from 922 to 1985 veh/hour per lane on a six-lane divided roadway.

The above three reviewed models are being used here to evaluate the expression $p(h<T|q)$ in Eq. (4). The value of $T$ defined in the previous section is considered as two seconds. Generally, $T$ is a distributed variable and ranges from 0.5 to even 4.0 seconds, depending on the complexity of the driving situation [Greenshields, 1965]. Let us recall that $h<T$ is considered as a potential situation for a multi-vehicle accident where for $h \geq T$, it is improbable that such an accident will occur. The value of $T=2$ is substituted in Eqs. (6), (7), and in the numerical integration of Eq. (8), in order to obtain the function $p(h<2|q_1) = 1 - p(h \geq 2|q_1)$, where $q_1$ = the flow on a per lane basis.

The results are demonstrated in Fig. 1. In 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. 1, a regression line based on a power function is indicated for all the models' results, namely :

$$p\ (h<2|q_1) = 0.011\ q_1^{\ 0.472} \qquad (9)$$

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 $q \leq 957$ veh/hour, is almost like Eq. (9). That is, $p\ (h<2|q_1) = 0.011\ q_1^{\ 0.478}$ with SE = 0.024. The interpretation of the

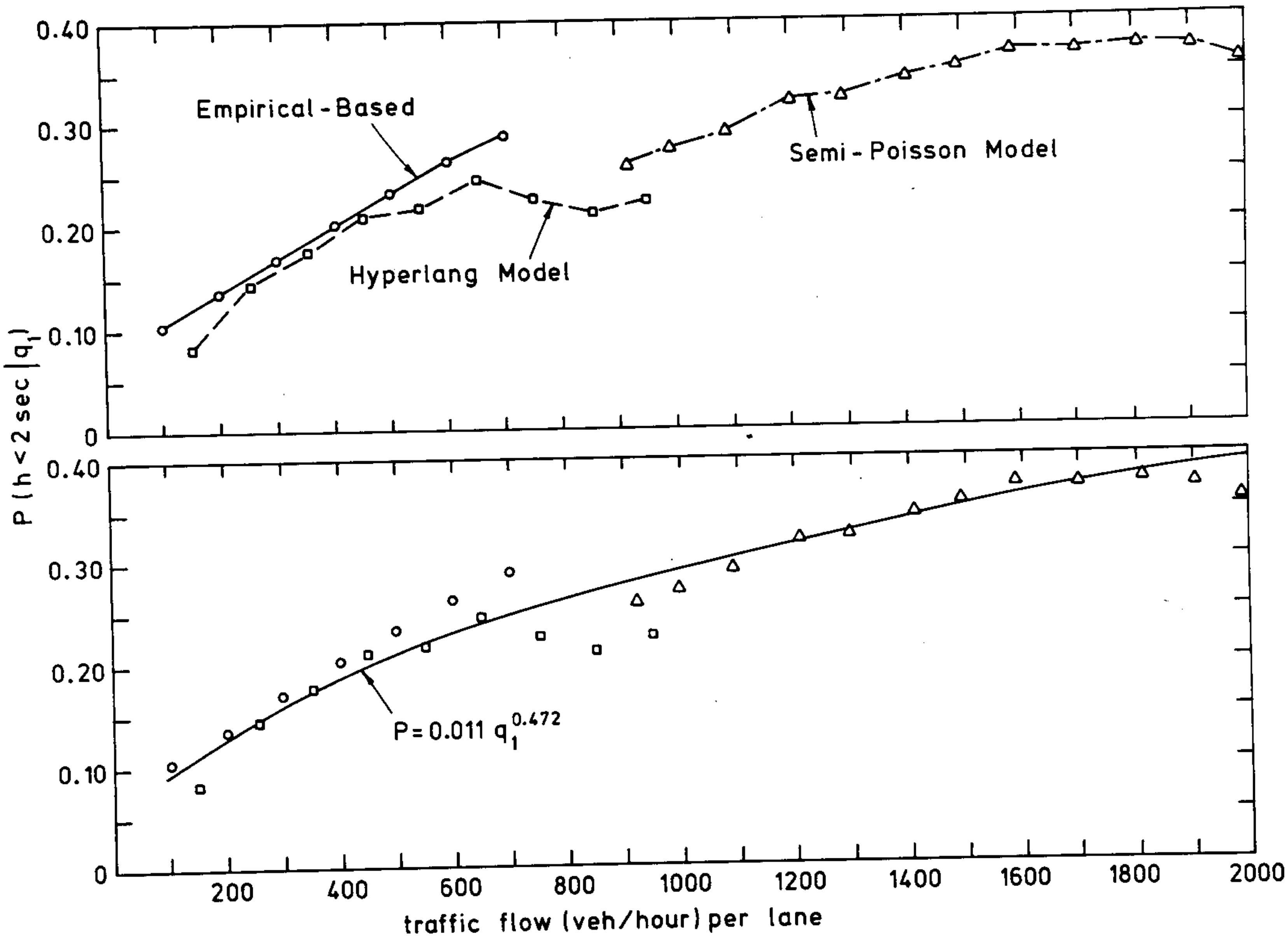



Fig. 1: The probability of headway less than 2 seconds according to the empirical-based results of Grecco & Sword [1968], the Hyperlang Model of Dawson & Chimini [1968], and the semi-Poisson Model of Wasielewski [1979]; in the lower figure, a regression model is shown for the three sets of results.

latter result is that extrapolation of the first two models fits very well an independent model which is calibrated with data characterized by the range $922 \lesssim q_1 \lesssim 1985$ veh/hour. This finding supports and strengthens the generality of Eq. (9).

## 5. REGRESSION RESULTS & DISCUSSION

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

$$A'_{rF}(q) = 4.3 \cdot 10^{-5} \cdot q^{1.32} \qquad (10)$$

with SE = 0.19 (acc/$10^6$ veh-km). The breakdown of Eq. (10) in accordance with Eqs. (2) and (9), reveals that $\alpha_1 = 5.42 \cdot 10^{-3}$ and $\beta_1 = 0.848$. This breakdown presumes that Eq. (9) can be applied to two-lane flows by considering separately each lane behaviour, i.e. $p(h<2|q)$ is based on Eq. (9) with $q_1 = \frac{1}{2}q$. Consequently, $\alpha_1$ and $\beta_1$ of Eq. (2) determine the expression $p(B|A)$ belong to Eq. (1).

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

$$A''_{rF}(q) = 232.27 \cdot q^{-1.15} \qquad (11)$$

with SE = 0.34 (acc/$10^6$ veh-km). Eq. (11) is associated with Eq. (3) and Eq. (4) is fulfilled through the summation of Eqs. (10) and (11), i.e. $A_{rF} = A'_{rF} + A''_{rF}$ which is the total measure for free-flow accidents. The left hand side of Fig. 2 illustrates both the data points and the free-flow model. The optimum conditions, by differentiating, yield $q_{roF} = 503$ veh/hr., which is similar to that found in part I of the work (for data composed also of night accidents and without separating free and congested conditions.

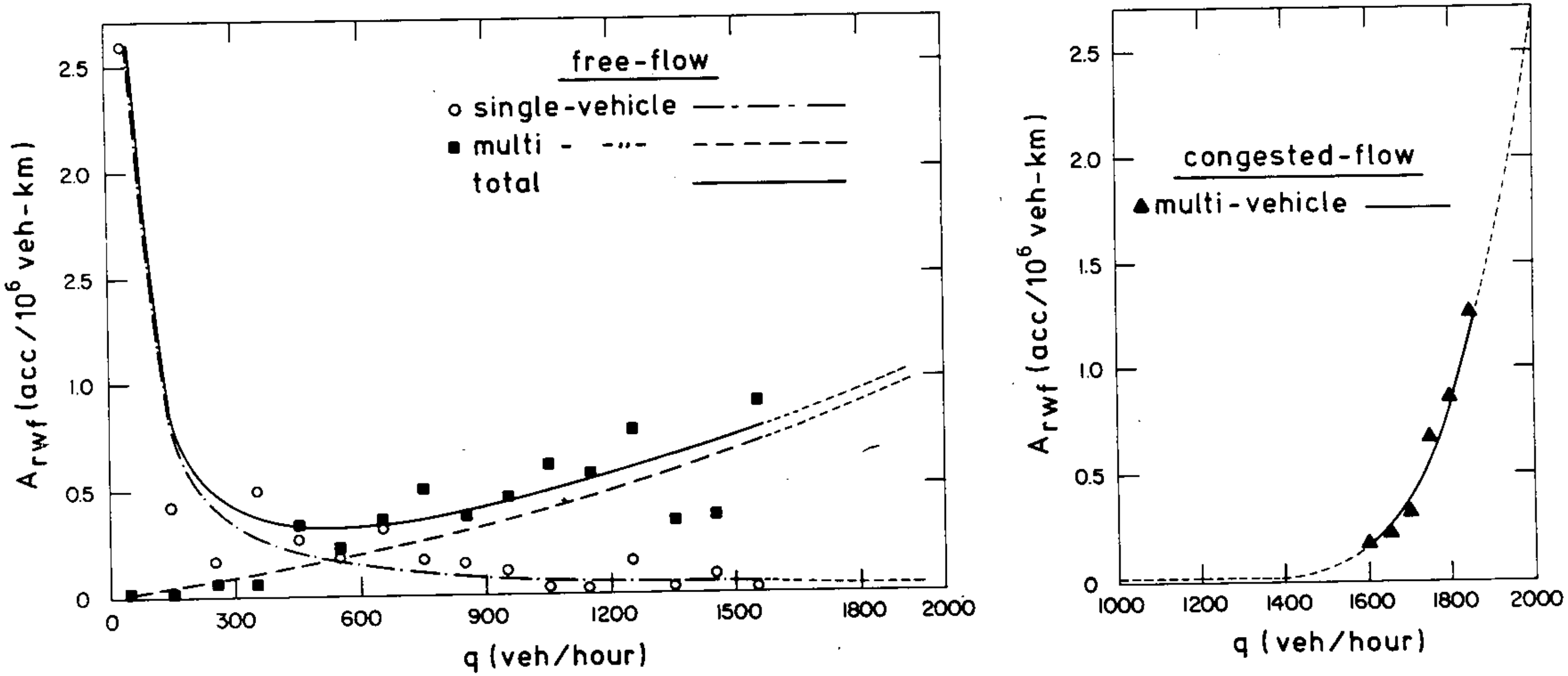



Fig. 2: The data and regression models for free-flow and congested-flow conditions.

For congested flow conditions, Eq. (5) through regression, takes the form :

$$A_{rC} = 7.21 \cdot 10^{-48} \cdot q^{14.46} \qquad (12)$$

with SE = 0.06 (acc/$10^6$ veh-km), and is demonstrated on the right hand side of Fig. 2. Despite the small number of data points, it is possible to observe a sharp increase in $A_{rC}$ 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 tendency of $A_{rC}$ 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 drivers' 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 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 accidents.

Eq. (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 hardly (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 $p(B|A)$ in Eq. (1) takes the form :

$$p(B|A) = 5.42 \cdot 10^{-9} \cdot q^{0.848} \tag{13}$$

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 by $h < 2$ sec. (using Eq. (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 Eq. (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 work are :

(i) determination of the number of kilometers with an hourly flow $q$, for a given probability such that (at least) one accident will occur.

(ii) 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 property (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 will occur at each veh-km under the flow $q$, presuming independency between each two inspections. For large numbers of veh-kms the description of the first aspect approaches the normal distribution (approximation to the binomial distribution). Also, since the analysis is not successive with respect to $q$, a vehicle involved in an accident is, theoretically, not excluded from further examinations (otherwise, the appropriate distribution is geometric rather than normal). Thus, the analysis of previous sections enables a definition :

$A_r(q)\cdot 10^{-6}$ = the probability of being involved in an accident in each veh-km within the flow range $q \pm \Delta q$ ($\Delta q$ = 100 veh/hour)

$X$ = number of accidents - normally distributed $(\mu_1, \sigma_1^2)$

$$\text{where,} \quad \mu_1 = n\cdot A_r(q)\cdot 10^{-6} \quad ; \quad \sigma_1^2 = n\cdot A_r(q)\cdot 10^{-6}\ [1-A_r(q)\cdot 10^{-6}]$$

and $n$ is the number of kilometers travelled by a vehicle at a given flow (within the range $q \pm \Delta q$). The investigated probability is $p(X \geqslant 1|q)$. Figure 3 illustrates this probability as a function on $n$ for different flow levels, based on Eqs. (10), (11), and (12). For example, at the probability of 90%, a flow of 1000 veh/hour needs to cover 40 million kms so that (at least) one single-vehicle accident will occur, in comparison with 8.5 million kms for a multi-vehicle accident. Fig. 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 $p(X \geqslant 1|q)$ - each in a different manner.

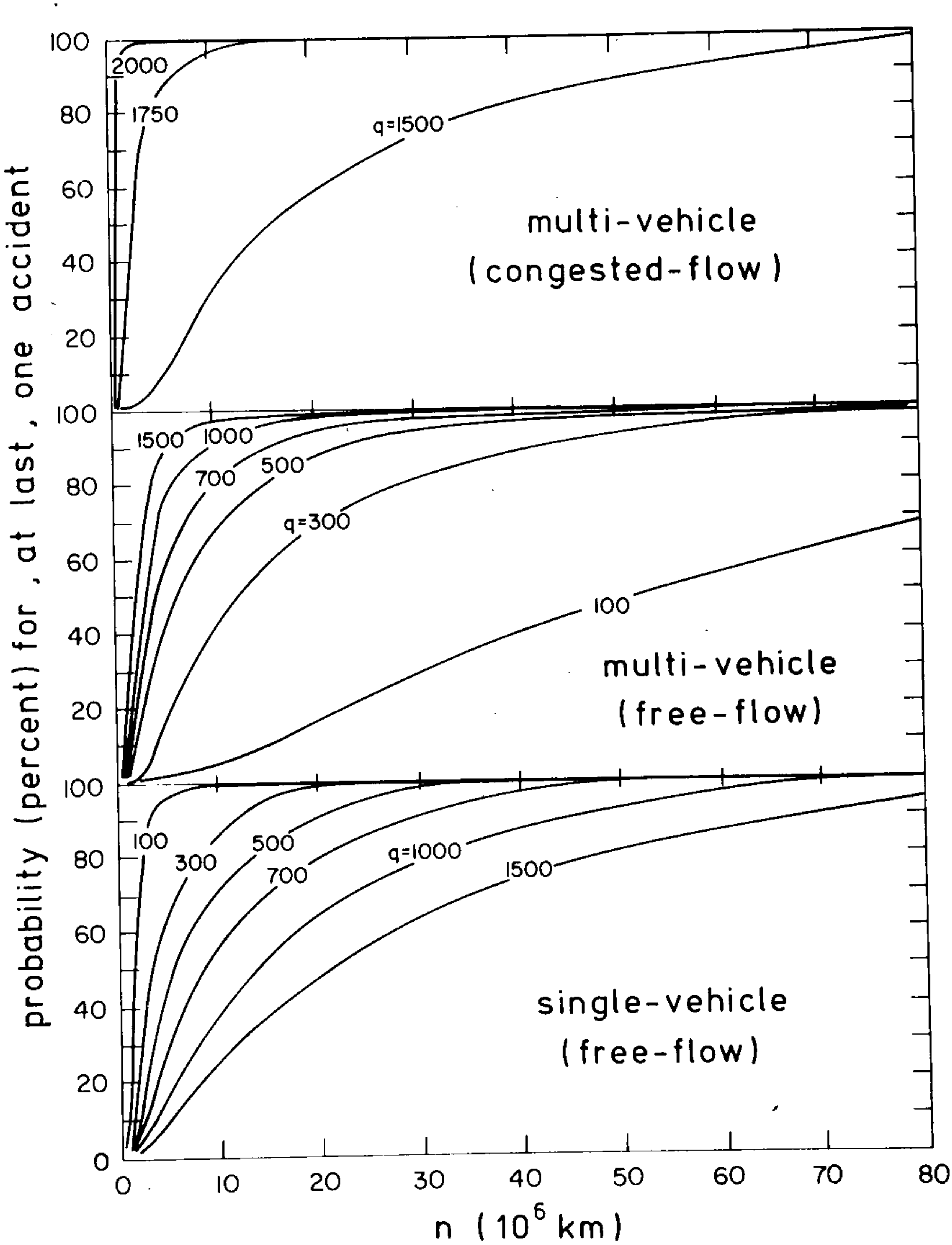



Fig. 3: The resultant relationship between the probability for at least one accident and n for various q (veh/hour) values.

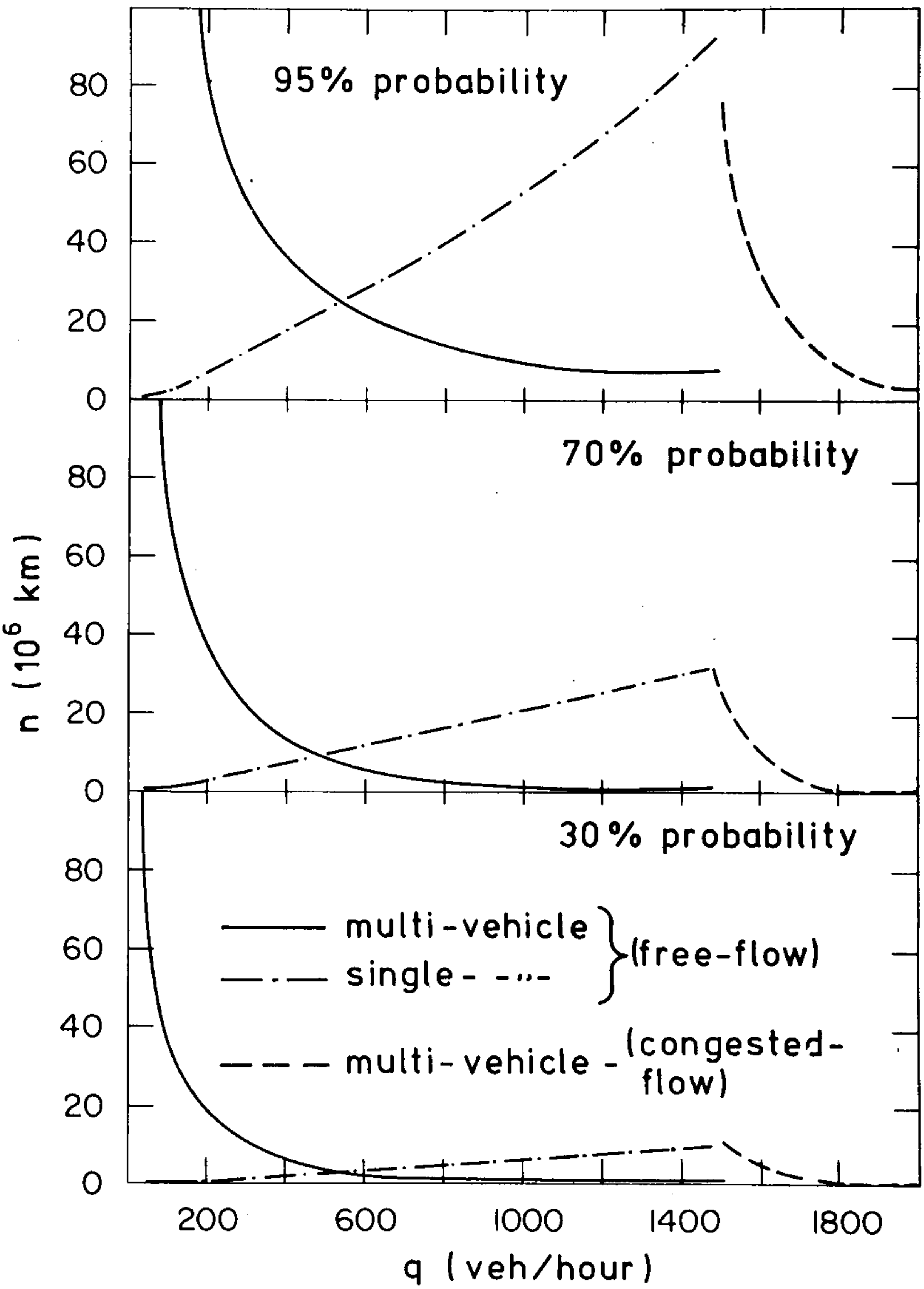



Fig. 4: The resultant relationship between n and q for three levels of probabilities (for at least one accident).

Similarly, for the second aspect :

$A_{dw}(q) = A_r(q) \cdot q \cdot 10^{-6}$ = the probability of being involved in an accident on any kilometer exposed to one hour of flow within the range $q + \Delta q$

$y$ = number of accidents - normally distributed $(\mu_2, \sigma_2{}^2)$

where :

$$\mu_2 = m \cdot A_{dw}(q) \quad ; \quad \sigma^2 = m_2 \cdot A_{dw}(q)[1-A_{dw}(q)]$$

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 $p(y \geqslant 1|q)$ is shown continuously in Fig. 5, and for only three levels-in Fig. 6.

## 7. SUMMARY

This study which is the last, and phase IV of the entire research (shown schematically in the first figure in Ceder & Livneh, 1980), attempts primarily to consider accident data regarding 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 and 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 in a unit of time (i.e. to approach a capacity level). A balanced traffic productivity measure might then attempt to maintain

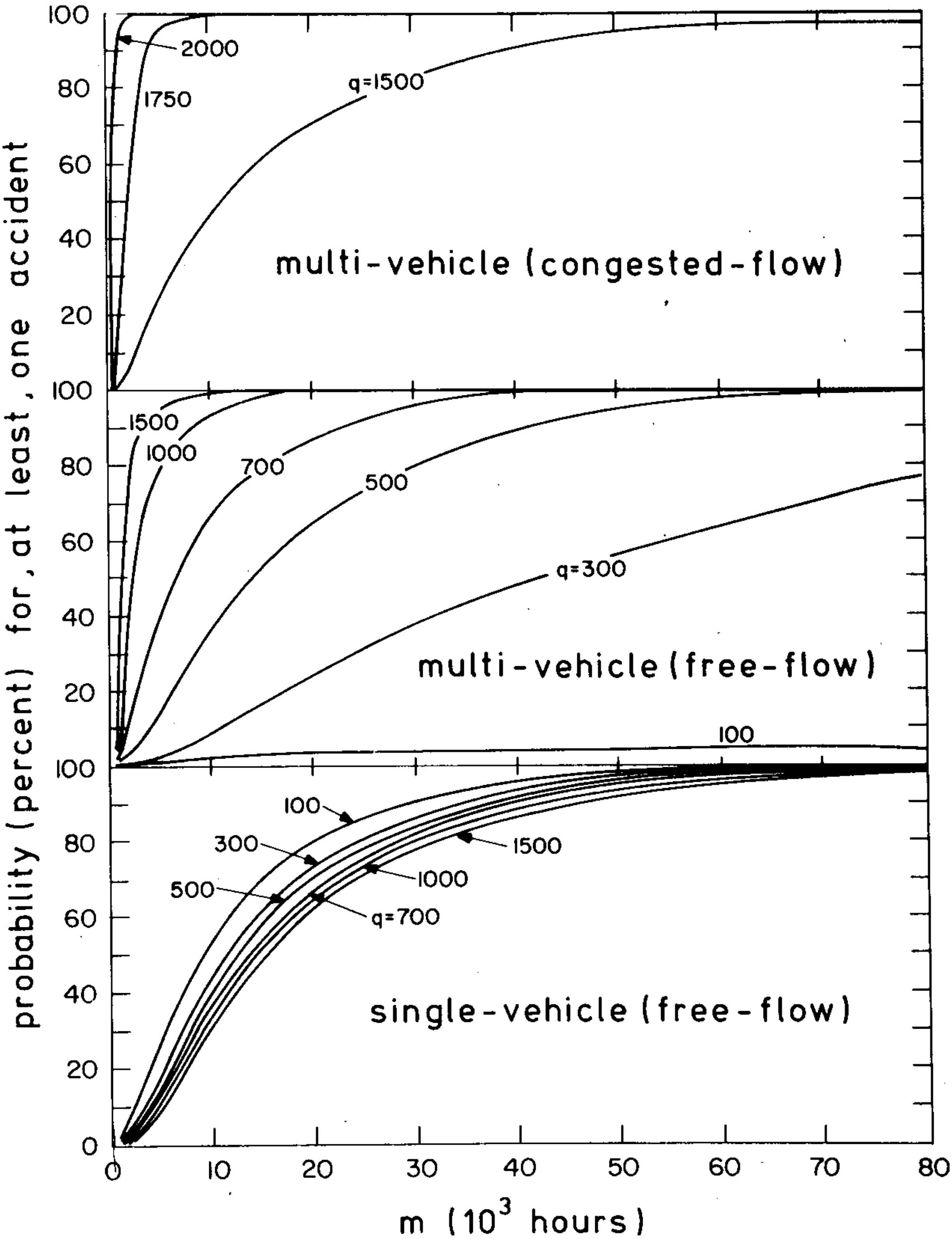



Fig. 5: The resultant relationship between the probability for at least one accident and m for various q (veh/hour) values.

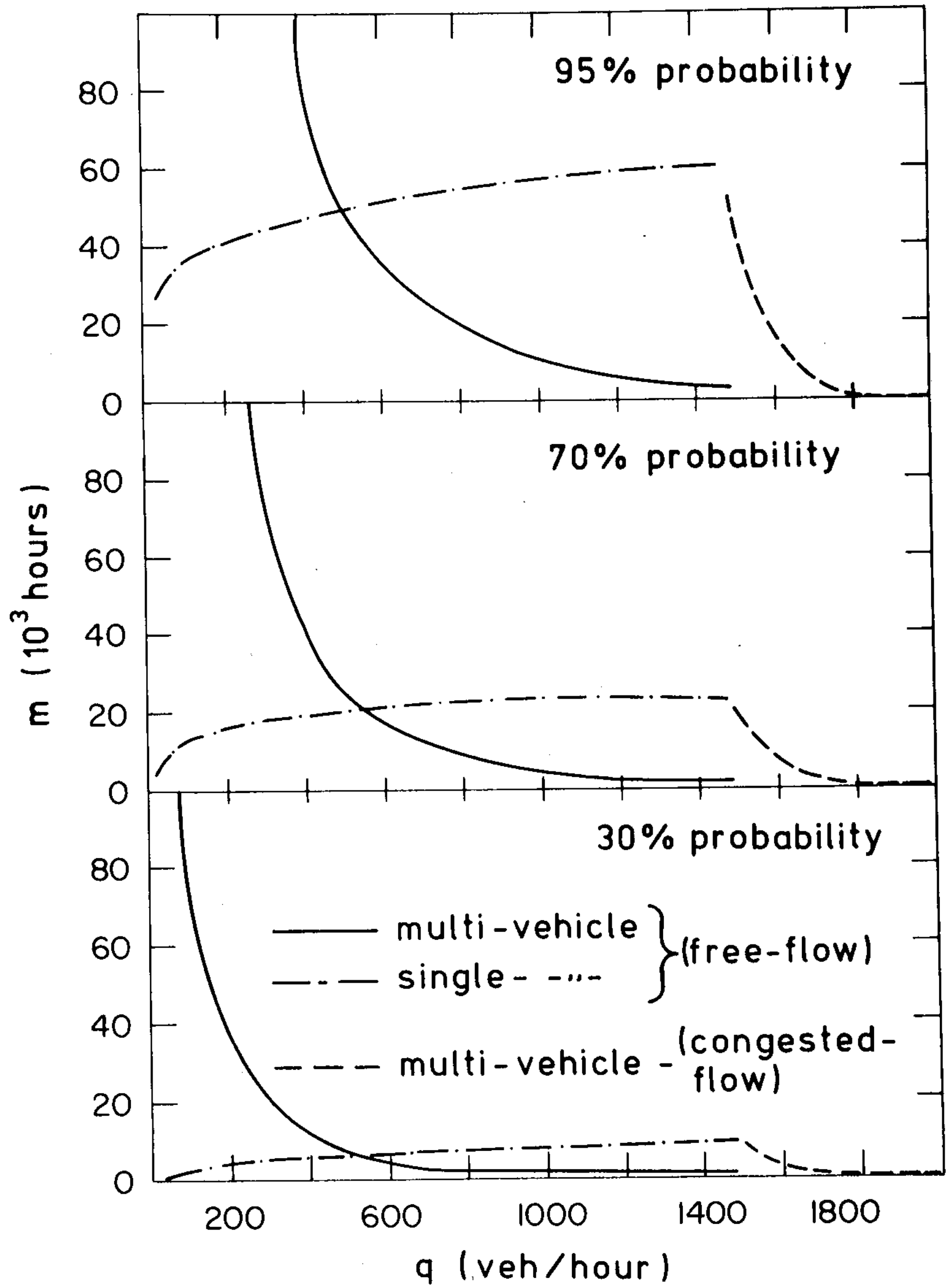



Fig. 6: The resultant relationship between m and q for three levels of probabilities (for at least one accident).

the stream of vehicles (assuming it is under control) at a buffer point below its maximum 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 vehicles which are, even instantaneously, under the car-following mode, are in a potentially hazardous situation. The remaining underlying probability aspects are based on the determined accident models 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 Strassenverkehrstechnik, Heft 201, 1976.

Ceder, A. "Drivers' Eye Movements as Related to Attention in Simulated Traffic Flow Conditions". Human Factors, Vol. 19(6), 571-581.

Ceder, A. "Stable Phase-Plane & Car-Following Behaviour as Applied to a Macroscopic Phenomenon". Transportation Science, 13(1), 64-79, 1979.

Dawson, R.F. & Chimini, L.A. "The Hyperlang Probability Distribution - A Generalized Traffic Headway Model". Highway Research Record, 230, 1-14, 1968.

Edie, L.C. "Flow Theories" in Traffic Science, Gazis, D.C., (Ed.), John Wiley & Sons, New York, 1974.

Grecco, W.L. & Sword, E.C. "Prediction of Parameters for Schuhl's Headway Distribution". Traffic Engineering, Vol. 38(5), 36-38, 1968.

Greenshields, B.D. "The Driver" in Traffic Engineering Handbook, Baerwald, J.W., (Ed.). Institute of Traffic Engineers, Washington, D.C., 1965.

Leutzbach, W., Siegener, W., & Wiedemann, R. "On the Connection Between Traffic Accidents and Traffic Volume on a Section of a German Autobahn". Accident Analysis & Prevention, Vol. 2, 143-146 (in German), 1970.

Wasielewski, P. "Car-Following Headways on Freeways Interpreted by the Semi-Poisson Headway Distribution Model". Transportation Science, 13(1), 36-55, 1979.