HIGHWAY RESEARCH RECORD

Number 208

Freeway Traffic Characteristics

3 Reports

Subject Area

22 Highway Design
52 Road User Characteristics
53 Traffic Control and Operations

HIGHWAY RESEARCH BOARD
DIVISION OF ENGINEERING NATIONAL RESEARCH COUNCIL
NATIONAL ACADEMY OF SCIENCES—NATIONAL ACADEMY OF ENGINEERING

Washington, D.C., 1967

Publication 1551

Price: $4.00

Available from

Highway Research Board
National Academy of Sciences
2101 Constitution Avenue
Washington, D. C. 20418

Department of Traffic and Operations

Harold L. Michael, Chairman
Associate Director, Joint Highway Research Project
Purdue University, Lafayette, Indiana

HIGHWAY RESEARCH BOARD STAFF

E. A. Mueller, Engineer of Traffic and Operations

COMMITTEE ON OPERATIONAL EFFECTS OF GEOMETRICS
(As of December 31, 1966)

Asriel Taragin, Chairman
Assistant Deputy Director, Office of Research and Development
U.S. Bureau of Public Roads, Washington, D.C.

Stanley R. Byington, Secretary
Traffic Systems Division, Office of Research & Development
U.S. Bureau of Public Roads, Washington, D.C.

Patrick J. Athol, Project Supervisor, Illinois Expressway Surveillance Project, Oak Park
W. R. Bellis, Director of Research and Evaluation, New Jersey State Highway Department, Trenton
Louis E. Bender, Chief, Traffic Engineering Division, The Port of New York Authority, New York, N.Y.
Ralph D. Brown, Jr., Engineer of Location and Roadway Planning, Illinois Division of Highways, Springfield
Robert R. Coleman, Assistant Director, Bureau of Traffic Engineering, Pennsylvania Department of Highways, Harrisburg
James J. Crowley, Assistant Regional Engineer, U.S. Bureau of Public Roads, Fort Worth, Texas
Harley T. Davidson, Engineer of Design Development, Connecticut State Highway Department, Wethersfield
William G. Galloway, Director, Division of Traffic, Kentucky Department of Highways, Frankfort
George F. Hagenauer, DeLeuw, Cather & Company, Chicago, Illinois
John W. Hutchinson, Department of Civil Engineering, University of Kentucky, Lexington
Harry H. Iurka, Senior Landscape Architect, New York State Department of Public Works, Babylon, Long Island
Thomas W. Kennedy, Center for Highway Research, The University of Texas, Austin
Richard A. Luettich, Planning and Traffic Engineer, Maine State Highway Commission, Augusta
Karl Moskowitz, Assistant Traffic Engineer, California Division of Highways, Sacramento
R. C. O'Connell, Highway Safety Engineer, Planning and Standards Division, Office of Highway Safety, U.S. Bureau of Public Roads, Washington, D.C.
Neilon J. Rowan, Assistant Research Engineer, Texas Transportation Institute, Texas A & M University, College Station
W. T. Spencer, Assistant Chief, Division of Materials and Tests, Indiana State Highway Commission, Indianapolis
John H. Swanberg, Chief Engineer, Minnesota Department of Highways, St. Paul

COMMITTEE ON CHARACTERISTICS OF TRAFFIC FLOW

(As of December 31, 1966)

Joseph C. Oppenlander, Chairman
Assistant Professor of Highway Engineering
School of Civil Engineering, Purdue University
Lafayette, Indiana

Patrick J. Athol, Project Supervisor, Illinois Expressway Surveillance Project, Oak Park
C. E. Billion, San Diego, California
Jack B. Blackburn, Head, Department of Civil Engineering, Kansas State University, Manhattan
Martin J. Bouman, Transportation and Traffic Engineer, Engineering Department, San Diego, California
Paul D. Cribbins, Assistant Professor of Civil Engineering, North Carolina State University, Raleigh
Kenneth W. Crowley, Senior Research Analyst, Tunnels and Bridge Department, The Port of New York Authority, New York City
J. E. P. Darrell, Minnesota State Automobile Association, Minneapolis
Olin K. Dart, Jr., Associate Professor, Civil Engineering Department, Louisiana State University, Baton Rouge
Robert F. Dawson, Civil Engineering Department, University of Connecticut, Storrs
Donald R. Drew, Head, Highway Design & Traffic Engineering Department, Texas A & M University, Texas Transportation Institute, College Station
H. M. Edwards, Chairman, Civil Engineering Department, Queen's University, Kingston, Ontario, Canada
Wayne T. Gruen, Spokane, Washington
Roy C. Haeusler, Automotive Safety Engineer, Engineering Division, Chrysler Corporation, Detroit, Michigan
John J. Haynes, Head, Department of Civil Engineering, Arlington State College, Arlington, Texas
J. E. Johnston, Deputy Director of Highways, Planning & Traffic, Utah State Department of Highways, Salt Lake City
James H. Kell, Principal Traffic Engineer, Traffic Research Corporation, San Francisco, California
Russell M. Lewis, Department of Civil Engineering, Rensselaer Polytechnic Institute, Troy, N. Y.
Jack C. Marcellis, Traffic Plans Engineer, Nashville Traffic Commission, Nashville, Tennessee
Peter A. Mayer, Automotive Safety Foundation, Washington, D. C.
William F. Petring, Principal Engineer, Vehicle Systems, Automotive Safety Research, Engineering Staff, Ford Motor Company, Dearborn, Michigan
O. J. Reichelderfer, Traffic Engineer, Pennsylvania Department of Highways, Harrisburg
August J. Saccoccio, Traffic Engineer, Anne Arundel County Department of Public Works, Glen Burnie, Maryland
Charles C. Schimpeler, School of Civil Engineering, Purdue University, Lafayette, Indiana
Curtis L. Shufflebarger, Industrial College of the Armed Forces, Washington, D. C.

Robert P. Shumate, Director, Public Services Group, Indiana University, Bloomington
David Solomon, Principal Research Engineer, Traffic Systems Division, U. S. Bureau of Public Roads, Washington, D. C.
K. A. Stonex, Automotive Safety Engineer, Technical Liaison Section, Engineering Staff, General Motors Corporation, G. M. Technical Center, Warren, Michigan
William C. Taylor, Traffic Research Engineer, Ohio Department of Highways, Columbus
Kenneth J. Tharp, Principal Systems Engineer, Cornell Aeronautical Lab. Inc., Buffalo, N. Y.
Arthur G. Wake, Director of Traffic Engineering, Bureau of Traffic Engineering, Indianapolis, Indiana
Robert J. Wheeler, Associate Professor, College of Engineering, University of Missouri, Columbia

Foreword

The very substantial capital outlays that have been spent on the design, location, and construction of freeway type facilities indicates the evident acceptance that these highways have with administrators and the driving public. Despite the challenge from other modes of transportation and the difficulties of locating freeways in urban areas, more money is being spent on these facilities than ever before. Even those cities that are going ahead with improved transit developments are continuing their freeway programs.

With the mileage in freeways have come operating problems, especially during peak traffic periods. An entirely new body of research concerned with characteristics of freeway traffic has emerged in the past few years and the three papers presented in this RECORD are but a sampling of the voluminous work recently completed or underway. This RECORD will be of chief interest to those concerned with freeway operations and design. Traffic researchers, simulation specialists, and those concerned with traffic flow characteristics will also find the material to be of value.

The first paper by a team of Texas researchers reports on the first phase of a four-year study of gap acceptance in freeway merging. This portion of the work was devoted to establishing simple statistical relationships between percent of gap acceptance and gap size. Characteristics of various types of "gap" maneuvers were compared under various traffic conditions.

Simulation is the approach used by three California researchers in the second paper, which describes a valid model for design of diamond interchanges. Chief attention was paid to the validation problems and extensive data were gathered to enable the simulation model to be more reliable. Considerable attention to the problems of data gathering by aerial camera and its reduction was given in the course of the research.

Four Texas researchers, in studying the operational effects of entrance ramp geometry on merging as described in the last paper, used data from 23 different entrance ramps. The geometric elements of acceleration lane length, angle of convergence, and ramp grade were carefully evaluated and found to have a great effect upon the merging operation. General conclusions regarding entrance ramp design were recorded.

Contents

Gap Acceptance in the Freeway Merging Process

DONALD R. DREW, LYNN R. LaMOTTE, JOSEPH A. WATTLEWORTH, and JOHANN H. BUHR, Texas Transportation Institute, Texas A & M University

This study is the first phase of a four-year program on freeway merging undertaken by the U.S. Bureau of Public Roads to furnish more detailed information on the effect that geometric variables have on the merging of ramp traffic; to develop usable distributions of traffic variables for simulation programs; and to develop an optimum ramp metering and merging control system. The report emphasizes the collection and collation of gap acceptance characteristics.

The theoretical development of models and useful parameters for describing the merging process include the derivation of the forms of the mean and variance of the delay to a ramp vehicle in position to merge and the treatment of the variability of critical gaps and gap acceptance among drivers through the identification of the representative forms for both critical gap distributions and gap acceptance functions.

By using individual record probit analyses, simple, statistically significant relations between the percent gap acceptance and gap size are established. In this approach, the characteristics of lags and gaps and single and multiple entry merges are compared, as well as the fast to slow moving merging vehicles. The probit analyses are generalized to establish a relationship between percent acceptance and gap size and vehicular speed.

•THE subject of ramp vehicles merging into the freeway stream has been treated by a number of researchers. Most of the research has been devoted to empirical studies leading to design and operational procedures. Mathematical treatment of the merging maneuver has been attempted, too, with somewhat limited success because of the complexity of the vehicle interactions. Computer technologists have contributed several digital computer simulation programs, but lack of detailed criteria on gap acceptance and merging logic has hampered progress.

In the summer of 1965, the U.S. Bureau of Public Roads undertook research to furnish detailed criteria on the merging of ramp vehicles into the freeway system. A contract was awarded to the Texas Transportation Institute. The general aim of this research is the conception of a relationship between the many variables associated with the interaction of vehicles traversing a ramp and merging onto a freeway so as to determine the effects of the following on merging operation and level of service:

1. Traffic characteristics such as gap availability, gap acceptance, speed and volume;
2. Ramp geometrics such as length, curvature, angle of convergence and grades, and acceleration lane geometrics such as length, shape, delineation and location of lateral obstructions;

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 46th Annual Meeting.

3. System considerations such as interchange type, ramp configurations, frontage roads and upstream or downstream bottlenecks, and environmental elements such as metropolitan area size, location within the city, and lighting; and
4. Control devices such as freeway lane controls, yield or merge-ahead signs, traffic signal feeding the entrance ramp, and ramp metering stations.

The underlying purpose of this research is the application of this information to: (a) in design and operation—the furnishing of more detailed information on the effect that geometric variables and traffic characteristics have on merging traffic; and (b) in simulation—the development of usable distributions of traffic variables for simulation programs. To fulfill the broad project objectives, 32 ramp-freeway connections located in 8 metropolitan areas in 6 states from coast to coast and from border to border were chosen.

There are three purposes for conducting field studies of traffic characteristics in ramp-freeway merging areas. The first is for the eventual testing and refinement of models. The second purpose for collecting data is that at the present time only very limited data of this type are available. Gap acceptance data are prime examples of this. Much of the meager gap acceptance data available is very old, based on a small sample size, or for a peculiar situation such as a left-hand ramp or stop sign control. These conditions severely limit the usefulness of these data. No data have been collected on the effect of ramp and acceleration lane geometrics on the gap acceptance characteristics of vehicles on the ramp. A third very important application of the data on merging characteristics is for simulation inputs. Specific objectives of this phase of the project research are:

1. Development of models and useful parameters for describing gap and lag acceptance and delay in the merging maneuver;
2. Presentation of gap and lag acceptance characteristics obtained from studying some 32 ramps across the United States;
3. Determination of any differences in gap acceptance characteristics for different points of entry along the acceleration lane;
4. Delineation of the roles of absolute and relative speeds in the merging process and their effect on gap acceptance;
5. Identification of gap acceptance characteristics of more than single ramp vehicles so as to determine the efficiency of platoon merging; and
6. Investigation of the effect of outside freeway lane volumes on gap acceptance.

THEORY

Definitions and Terminology

In actual practice, the entrance ramp-freeway connection may be regarded as a special case of the uncontrolled intersection. The three fundamental maneuvers performed by vehicles in the vicinity of this connection area may be identified as follows:

Lane change: The transfer of a vehicle from one traffic lane to the next adjacent traffic lane.

Merging: The process by which vehicles in two separate streams moving in the same general direction combine or unite to form a single stream.

Weaving: The oblique crossing of one stream of traffic by another accomplished by the merging of the two streams into one and then the diverging of this common stream into separate streams again.

The freeway elements associated with these maneuvers are:

Acceleration lane: An added width of pavement adjacent to the main roadway traffic lanes enabling vehicles entering the main roadway to adjust their speed to the speed of through traffic before merging.

Ramp: A connecting roadway between two intersecting or parallel roadways, one end of which joins in such a way as to produce a merging maneuver.

Frontage road: A roadway paralleling a freeway so as to intercept traffic entering or leaving the facility and to furnish access to abutting property.

Basic to the description of traffic interaction in the merging process are the following variables:

Headway: The interval of time between successive vehicles moving in the same lane measured from head to head as they pass a point on the road.

Space headway: The distance between successive vehicles moving in the same lane measured from head to head at a given instant in time.

Gap: A major stream headway that is evaluated by a minor stream vehicle desiring to either merge into or cross the major stream. The units may be those of either time or distance (space gap).

Lag: The interval of time between the arrival of a minor stream vehicle and the arrival of a major stream vehicle at a reference point or points in the vicinity of the area where the streams either cross or merge.

Space lag: The distance between a minor stream vehicle and a reference point where the minor stream crosses or joins a major stream, subtracted from the distance between a major stream vehicle and the same reference point, with both distances measured at a given instant in time.

Figure 1 is a time-space diagram illustrating the relationship between the geometric elements comprising the merging area and the movements of mainstream and ramp vehicles within the area. Distance is plotted as the ordinate, time is the abscissa, and the slope of the traces denotes speed. The traces for freeway and ramp vehicles are identified in the figure. A vehicle is regarded as a ramp vehicle as long as it remains at least partially on the acceleration lane.

It is convenient to subdivide merges into their basic types and classifications, then to determine the kind and amount of information needed for intelligent analysis of the movements. The following is a listing of merge maneuvers defined for use in this report (Fig. 2):

Optional merge: The merging vehicle voluntarily moves from the acceleration lane into the outside traffic lane.

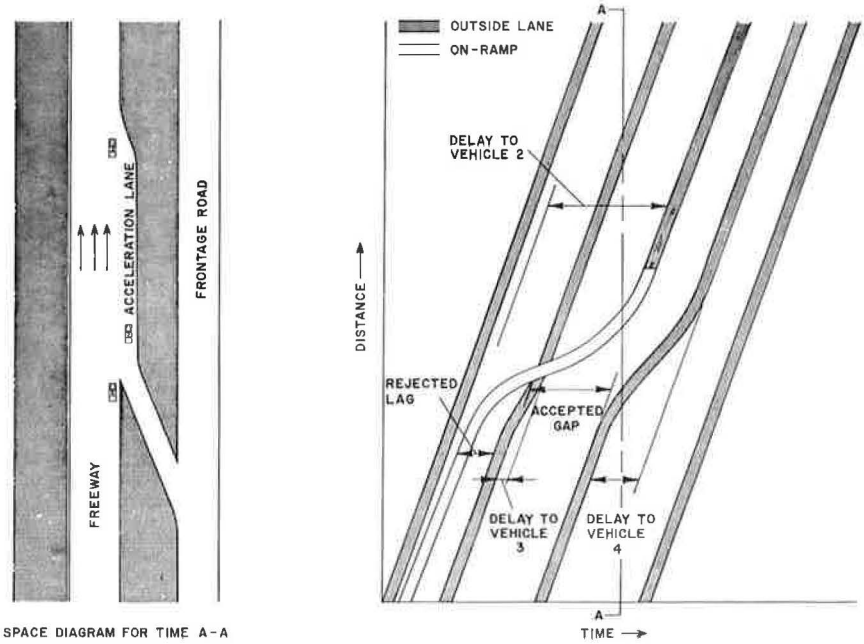


Figure 1. Time-space relationships of freeway merging maneuver.

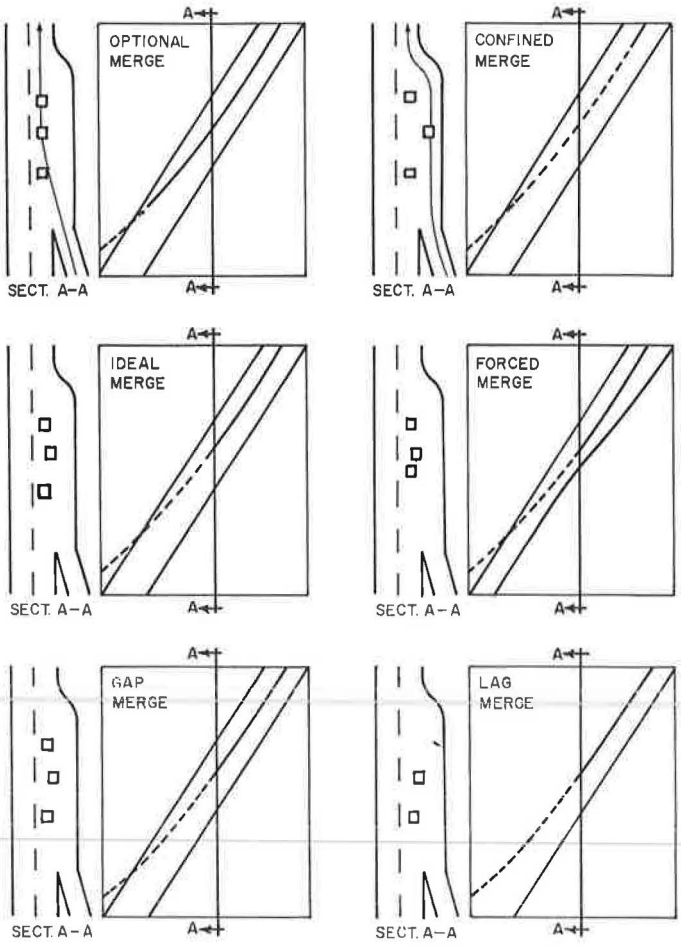


Figure 2. Types of freeway merges.

Confined merge: The merging vehicle is forced into the outside traffic lane by the presence of the end of the acceleration lane.

Ideal merge: The merging vehicle is able to enter the freeway stream without causing a freeway vehicle to reduce its speed or change lanes.

Forced merge: The ramp vehicle effects the merging maneuver into the freeway stream so that an oncoming freeway vehicle or vehicles must either slow down or change lanes.

Single entry merge: One ramp vehicle moves into a single freeway time gap.

Multiple entry merge: Two or more ramp vehicles merge into a single freeway time gap.

From Figures 1 and 2 and the just completed classification, it is apparent that a merge must be qualified as being either single or multiple entry, optional or confined, ideal or forced, and gap or lag. For example, a given merge might be described as being a "single entry, optional, ideal gap merge."

Merging Parameters

In addition to the traffic variables just described, the speed of the major stream vehicles, speed of the merging vehicles, relative speed, major stream flow, and minor

stream flow are additional variables that must be considered in any rational description of the merging process.

Whereas a variable is a quantity which can assume any value or number, a parameter is a term used in identifying a particular variable or constant other than the coordinate variables. For example, "volume" is a traffic variable and "capacity"—the maximum volume that a facility can accommodate—is the corresponding traffic parameter. Some important figures of merit for describing the gap acceptance phenomenon in freeway merging are the critical gap, percent of ramp vehicles delayed, mean duration of static delay accepting a gap, mean length of queue, and total waiting time on the ramp.

Several "critical" values have been discussed in the literature. Greenshields (1) defined the "acceptable average-minimum time gap" as a gap accepted by half the drivers. Raff (2) used a slightly different parameter, the "critical lag." The critical lag is the size lag which has the property that the number of accepted lags shorter than the critical lag is equal to the number of rejected lags longer than the critical lag. As such, the Greenshields and Raff parameters are median values.

The principal use of gap acceptance parameters is to simplify the computation of the delay duration by permitting the assumption that all intervals shorter than the critical value (lag or gap) are rejected while all intervals longer are accepted. It has been suggested (3, 4) that the mean of the critical gap distribution, not the median, should be used in delay computations. Nevertheless, the "median critical gap" remains a practical parameter because, as shall be shown, it can be readily obtained graphically.

Delay Models

There have been a number of theoretical papers (5, 6) dealing with the delay to a single waiting vehicle on the minor street of an uncontrolled intersection due to the traffic on the outside lane of an intersecting highway. It can be shown that the theory is valid for a driver on a ramp waiting to join a stream of traffic.

Most discussions of this problem have assumed that the distribution of main stream arrivals is Poisson, i.e., that the probability that a given gap is between t and t + dt seconds is given by an expression of the form qe^{-qt} where q is the flow. Raff (2) considered the delay problem as related to vehicles, whereas Tanner's analysis (5) was specifically applied to pedestrian delays. The pedestrians were assumed to arrive at random, and all waited for a critical time gap of T seconds in order to cross the highway. Although Mayne (6) showed that one could obtain results for the Laplace transform of the delay duration for other than Poisson traffic on the main stream traffic, the only case considered in detail was for this headway distribution. The derivation which follows assumes Erlang headways and as such is a generalization of the combinatorial techniques suggested by Raff (2), Tanner (5) and Mayne (6).

It may be assumed that a ramp driver waiting to merge measures each time gap t in the traffic on the outside lane of the freeway until he finds an acceptable gap T which he believes to be of sufficient length to permit his safe entry. If he accepts the first gap ($t > T$), his waiting time is zero. If he rejects the first gap ($t < T$), but accepts the second gap, his expected waiting time would be one gap interval. If it is assumed that the driver's gap acceptance does not change with time, then by induction, the individual waiting periods form a geometric distribution and the probability P(n) of any driver having a wait for n intervals each less than T seconds before merging is

$$P(n) = p^n (1 - p), \quad n = 1, 2 \tag{1}$$

where

$$p = P(t < T) = \int_0^T f(t)\, dt \tag{2}$$

and f(t) is the distribution of gaps in the major stream. It is convenient to define a normalized random variable w such that

$$w = t/T \tag{3}$$

or

$$t = Tw \tag{4}$$

and

$$dt = T\,dw \tag{5}$$

Substituting Eqs. 4 and 5 into Eq. 2 yields

$$p = P(w<1) = \int_0^1 f(Tw)\ T\ dw \tag{6}$$

The density of w may be defined as

$$g(w) = f(Tw)\ T \tag{7}$$

which when inserted in Eq. 6 gives

$$p = \int_0^1 g(w)\ dw \tag{8}$$

If the distribution of gaps on an outside freeway lane of flow q can be described by the Erlang distribution,

$$f(t) = \frac{(aq)^a}{\gamma(a)}\ t^{a-1}\ e^{-aqt} \tag{9}$$

then it follows from Eq. 3 to Eq. 8 that

$$p = \int_0^1 \frac{(aqT)^a}{\gamma(a)}\ w^{a-1}\ e^{-aqTw}\ dw \tag{10}$$

Eq. 10 is of the form of the incomplete gamma distribution, which is of course equivalent to the cumulative Poisson:

$$p = 1 - e^{-aqT} \sum_{i=0}^{a-1} (aqT)^i/i! \tag{11}$$

and

$$1 - p = e^{-aqT} \sum_{i=0}^{a-1} (aqT)^i/i! \tag{12}$$

Placing Eqs. 11 and 12 in Eq. 1 gives the probability that the first n time-gaps between vehicles in the major stream are all less than the normalized critical gap ($w = 1$), but that the $n + 1$th is greater than the normalized critical gap. This probability P(n) averaged over the distribution of gaps less than the normalized critical gap g(w) ($0 < w < 1$) gives the distribution of waiting time or delay $f(\mu)$ to a ramp vehicle in position to merge. In terms of the moment generating functions (mgf) this may be expressed (7) as

$$M_{\mu}(\theta) = \sum_{n=0}^{\infty} P(n)\, M_W(\theta)^n \tag{13}$$

where $M_{\mu}(\theta)$ is the mgf of the distribution of delay to a ramp vehicle waiting for a gap greater than the normalized critical gap. From Eqs. 1, 11 and 12, it is apparent that

$$M_{\mu}(\theta) = (1-p) \sum_{n=0}^{\infty} \left[p\, M_W(\theta)\right]^n \tag{14}$$

$$M_{\mu}(\theta) = \frac{1-p}{1-p\, M_W(\theta)} \tag{15}$$

In order to use Eq. 15, $M_W(\theta)$ must be found. It is apparent that $g(w)$ $(0<w<1)$ is a conditional probability which by definition becomes

$$g(w)\ (0<w<1) = P(w<w \mid w<1) = \frac{\int_0^w g(w)\, dw}{\int_0^1 g(w)\, dw} \tag{16}$$

Since the denominator of Eq. 16 is given in Eq. 8, one obtains the density of gaps less than the normalized critical gap as

$$g(w)\ (0<w<1) = g(w)/p \tag{17}$$

and the mgf as

$$M_W(\theta) = \frac{1}{p} \int_0^1 e^{\theta w}\, g(w)\, dw \tag{18}$$

Changing the limits of summation and substituting for $g(w)$ yields

$$M_W(\theta) = \frac{(aqT)^a}{p\gamma(a)} \left\{ \int_0^{\infty} w^{a-1}\, e^{-w(aqT-\theta)}\, dw - \int_1^{\infty} w^{a-1}\, e^{-w(aqT-\theta)}\, dw \right\}$$

If a change of variable is made such that $u = w(aqT-\theta)$ in the first integral and $u+1=w$ in the second integral, then

$$M_W(\theta) = \frac{(aqT)^a}{p\gamma(a)} \left[(aqT-\theta)^{-a} \int_0^{\infty} u^{a-1}\, e^{-u}\, du - e^{-(aqT-\theta)} \right.$$

$$\left. \int_0^{\infty} (u+1)^{a-1}\, e^{-u(aqT-\theta)}\, du \right]$$

Use of Newton's binomial identity,

$$(u+1)^{a-1} = \sum_{i=0}^{a-1} \binom{a-1}{i} u^{a-i-1}$$

in the second integral, and noting that the first integral is a gamma function gives

$$M_W(\theta) = \frac{1}{p}\left(1 - \frac{\theta}{aqT}\right)^{-a} - \frac{(aqT)^a}{p\gamma(a)} e^{-(aqT-\theta)} \sum_{i=0}^{a-1} \binom{a-1}{i} \int_0^{\infty} u^{a-i-1} e^{-u(aqT-\theta)}\, du \tag{19}$$

Since the second integral is a gamma function with parameters $(a-i)$ and $(aqT-\theta)$, the second term may be written as

$$\frac{(aqT)^a}{p\gamma(a)} e^{-(aqT-\theta)} \sum_{i=0}^{a-1} \frac{\gamma(a)}{i!\gamma(a-i)} \frac{\gamma(a-i)}{(aqT-\theta)^{a-i}}$$

or

$$\left(\frac{aqT}{aqT-\theta}\right)^a \frac{e^{-(aqT-\theta)}}{p} \sum_{i=0}^{a-1} \frac{(aqT-\theta)^i}{i!} \tag{20}$$

Substituting Eq. 20 for the second term in Eq. 19 and collecting terms, one obtains

$$M_W(\theta) = \frac{1}{p}\left(1 - \frac{\theta}{aqT}\right)^{-a} \left[1 - e^{-(aqT-\theta)} \sum_{i=0}^{a-1} \frac{(aqT-\theta)^i}{i!}\right] \tag{21}$$

Now it is possible to find the moments of the delay to a ramp vehicle in position to merge by evaluating the derivatives of Eq. 15,

$$\frac{dM_\mu(\theta)}{d\theta} = \frac{p(1-p)}{[1-pM_W(\theta)]^2} \frac{dM_W(\theta)}{d\theta} \tag{22}$$

and

$$\frac{d^2M_\mu(\theta)}{d\theta^2} = \frac{p(1-p)\,[dM_W^2(\theta)/d\theta^2]}{[1-pM_W(\theta)]^2} + \frac{2(1-p)\,[pdM_W(\theta)/d\theta]^2}{[1-pM_W(\theta)]^3} \tag{23}$$

at $\theta = 0$, where the derivatives of $M_W(\theta)$ in the expressions are given by

$$\frac{dM_W^n(\theta)}{d\theta^n} = \frac{\gamma(a+n)}{p\gamma(a)} \frac{(aqT)^a e^{-(aqT-\theta)}}{(aqT-\theta)^{(a+n)}} \left[e^{(aqT-\theta)} - \sum_{i=0}^{a+n-1} \frac{(aqT-\theta)^i}{i!}\right] \tag{24}$$

Expansion of Eqs. 22 and 23 leads to the following expressions for the mean and variance of the delay distribution:

$$\mu(w)_a = \frac{e^{aqT} - \sum_{i=0}^{a} \frac{(aqT)^i}{i!}}{qT \sum_{i=0}^{a-1} \frac{(aqT)^i}{i!}} \tag{25}$$

$$\sigma^2(w)_a = \frac{(a+1)\left[e^{aqT} - \sum_{i=0}^{a+1} \frac{(aqT)^i}{i!}\right]}{a(qT)^2 \sum_{i=0}^{a-1} \frac{(aqT)^i}{i!}} + \mu^2(w)_a \tag{26}$$

Converting from the normalized parameter w to the original variable t, the mean and variable for values a = 1, 2, 3 and 4 become:

$$\mu(t)_1 = q^{-1}(e^{qT} - 1 - qT) \tag{27}$$

$$\mu(t)_2 = \frac{e^{2qT} - 1 - 2qT - 2(qT)^2}{q(1+2qT)} \tag{28}$$

$$\mu(t)_3 = \frac{e^{3qT} - 1 - 3qT - 4.5(qT)^2 - 4.5(qT)^3}{q[1+3qT+4.5(qT)^2]} \tag{29}$$

$$\mu(t)_4 = \frac{e^{4qT} - 1 - 4qT - 8(qT)^2 - 10.7(qT)^3 - 10.7(qT)^4}{q[1+4qT+8(qT)^2+10.67(qT)^3]} \tag{30}$$

$$\sigma^2(t)_1 = \frac{e^{2qT} - 2qTe^{qT} - 1}{q^2} \tag{31}$$

$$\sigma^2(t)_2 = \frac{2e^{4qT} - e^{2qT} - 2qTe^{2qT} - 8(qT)^2\, e^{2qT} - 1 - 4qT - 2(qT)^2}{2q^2(1+2qT)^2} \tag{32}$$

$$\sigma^2(t)_3 = \frac{4e^{3qT} - 4 - 12qT - 18(qT)^2 - 18(qT)^3 - 13.5(qT)^4}{3q^2[1+3qT+4.5(qT)^2]} + \mu^2(t)_3 \tag{33}$$

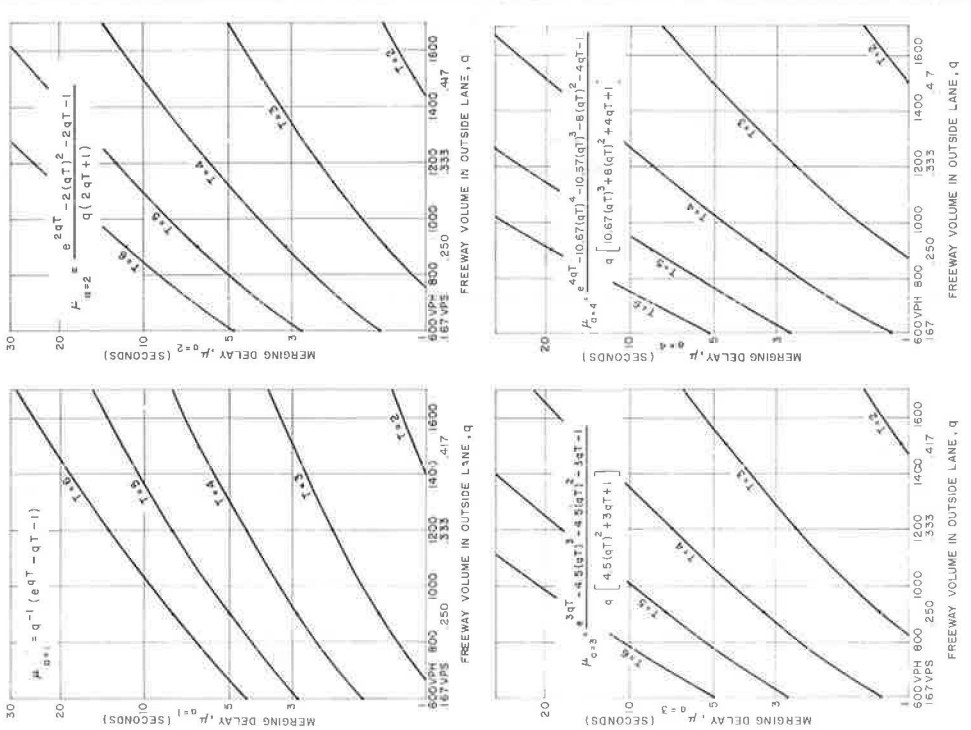


Figure 3. Merging delay in terms of the freeway flow q, critical gap T, and Erlang constant, a.

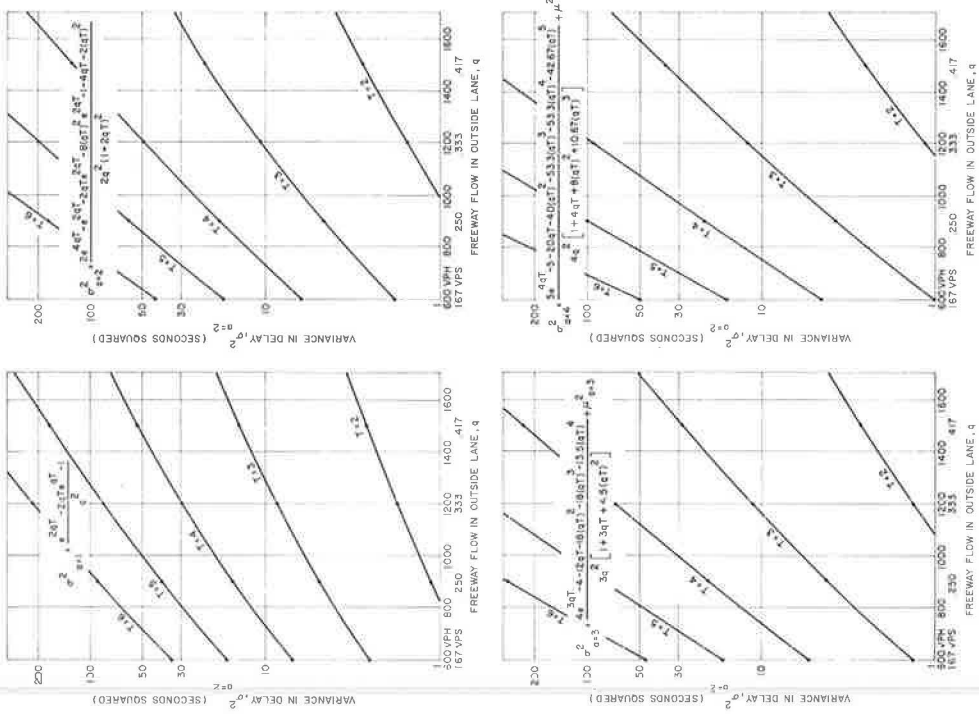


Figure 4. Variance in delay in terms of the freeway flow q, critical gap T, and Erlang constant, a.

$$\sigma^2(t)_4 = \frac{5e^{4qT} - 5 - 20qT - 40(qT)^2 - 53.3(qT)^3 - 53.3(qT)^4 - 42.67(qT)^5}{4q^2[1 + 4qT + 8(qT)^2 + 10.67(qT)^3]} + \mu^2(t)_4 \quad (34)$$

Eqs. 27 to 30 are plotted in Figure 3 and Eqs. 31 to 34 in Figure 4.

Critical Gap Distributions

The theoretical delay values of the previous section are based on a fixed critical gap for all drivers. A more realistic description of delays can be obtained by replacing the fixed critical gap with a distribution of critical gaps, f(T) (8, 9). The forms for the mean and variance of the distribution of delay become

$$M(T) = \int_0^\infty \mu(T)\ f(T)\ dT \quad (35)$$

and

$$S^2(T) = \int_0^\infty \sigma^2(T)\ f(T)\ dT \quad (36)$$

Assuming that the headway distribution on the freeway is negative exponential, for example, substitution of Eq. 27 in Eq. 35 yields

$$M(T) = q^{-1} \int_0^\infty e^{qT} f(T)dT - \int_0^\infty T\ (f)T\ dT - q^{-1} \int_0^\infty f(T)dT \quad (37)$$

Realizing that the second term defines the mean of the critical gap distribution and the integral in the last term must equal unity gives

$$M(T) = q^{-1} \int_0^\infty e^{qT} f(T)dT - \bar{T} - q^{-1} \quad (38)$$

Some representative forms for critical gap distributions are shown in Figure 5. If one assumes that the critical gaps of the drivers are distributed uniformly between c and c_1 then

$$f(T) = (c_1 - c)^{-1} \quad (39)$$

The translated negative exponential distribution

$$f(T) = (\bar{T} - c)^{-1}\ e^{-(T - c)/(\bar{T} - c)} \quad (40)$$

has also been suggested (10) as a gap distribution function, as has the Erlang distribution (11)

$$f(T) = \frac{(a/\bar{T})^a}{(a - 1)!}\ T^{a - 1}\ e^{-aT/\bar{T}} \quad (41)$$

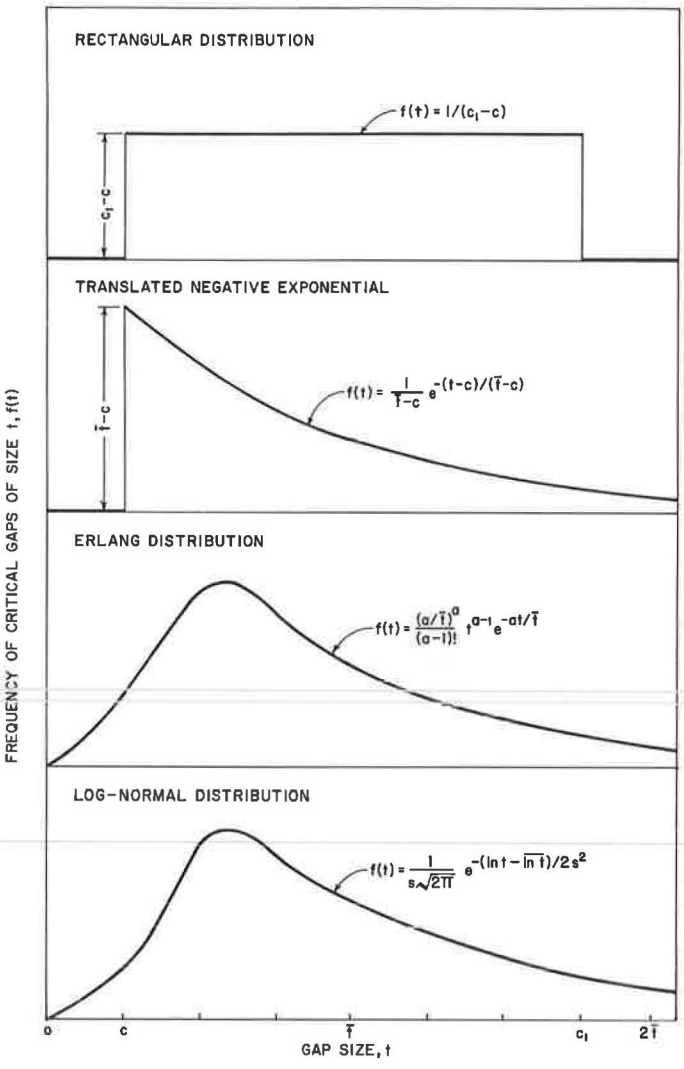


Figure 5. Representative forms for critical gap distributions.

Substitution of Eqs. 39, 40 and 41 in Eq. 38 gives respectively:

$$M(T) = (c_1 - c)^{-1} q^{-2} [e^{qc}1 - e^{qc}] - \bar{T} - q^{-1} \tag{42}$$

$$M(T) = [q(1 - q\bar{T} + qc)]^{-1} e^{qc} - T - q^{-1} \tag{43}$$

and

$$M(T) = \frac{1}{q}\left(\frac{a}{a - q\bar{T}}\right)^a - \bar{T} - q^{-1} \tag{44}$$

which correspond to the distribution of delay for the first three critical gap distributions in Figure 5.

The derivations of this section (Eqs. 35 to 44) are based on utilization of a distribution of critical gaps for all drivers, f(T). This is a concession to the obvious fact that not all drivers have the same critical gap. The difficulty lies, however, in measuring the critical gap for individual drivers in order to obtain a distribution of critical gaps. One technique to obtain such a frequency distribution (4) samples only drivers who have rejected at least one gap before merging and is based on the reasonable assumption that the driver's critical gap must lie somewhere between the largest gap he rejected and the gap he finally accepted.

Although observations of a narrow range delimit a driver's gap more closely, by the very nature of the technique observations with wider ranges have greater influence on the shape of the density function. Dawson (12) suggests two weighting procedures to overcome this shortcoming which seem promising. Still, there exist some practical as well as conceptual shortcomings in the technique which tended to rule it out as a basis of parameter determination for this project. First, the technique only considers drivers who have rejected a gap; and second, the technique is obviously not applicable to lags.

Gap Acceptance Functions

In the derivations in the preceding sections based on the critical gap concept, it is assumed that the waiting driver evaluates each inter-car time gap, choosing to merge if the gap is greater than some predetermined time gap T, or not to merge, if the gap is less than T seconds. Herman and Weiss (13) suggest that a more realistic model would be to associate with each time gap, a gap acceptance probability P(T) such that the waiting driver crosses the highway with the probability P(T) when confronted with a gap of duration T. Some representative gap acceptance functions are shown in Figure 6.

Using methods strongly dependent on renewal theory, Weiss and Maradudin (14) formulated the merging delay problem in terms of an integral equation instead of using the combinatorial reasoning approach inherent in Eqs. 1 through 44. The expression for the mean of the delay distribution is

$$M(T) = \int_0^\infty T\psi_0(T)dT + \frac{1 - P_0}{P} \int_0^\infty T\psi(T)dT \tag{45}$$

where

$$P_0 = \int_0^\infty P(T)f_0(T)dT \tag{46}$$

$$P = \int_0^\infty P(T)f(T)dT \tag{47}$$

$$\psi_0(T) = f_0(T)\ [1 - P(T)] \tag{48}$$

$$\psi(T) = f(T)\ [1 - P(T)] \tag{49}$$

and f(T) is the distribution of gaps on the outside freeway lane and P(T) is the gap acceptance function. The zero subscripts refer to the gap availability distribution and gap acceptance function for the first gap.

Weiss (13, 14) has applied his theory to some specific distributions. For example, if the distribution of gaps on the outside lane of the freeway is taken as

$$f(T) = f_0(T) = qe^{-qT} \tag{50}$$

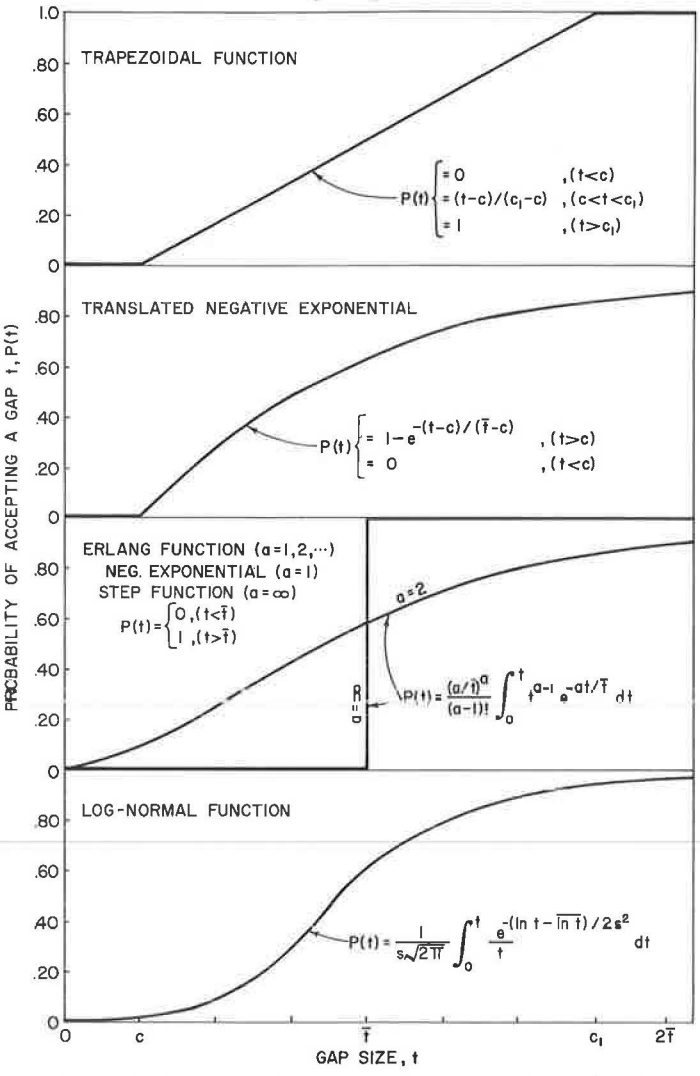


Figure 6. Representative forms for gap acceptance functions.

and the gap acceptance function is

$$P(T) = P_0(T) = 1 - e^{-\lambda(T-c)} \tag{51}$$

one obtains for the mean delay to a ramp vehicle in position to merge

$$M(T) = q^{-1}\left\langle e^{qc} - 1 - qc + q/\lambda\{e^{qc} - 1 - qc + [q/(q+\lambda)]^2\ (1+qc+c)\ (1-e^{-qc})\} + [q/(q+\lambda)+qc]\ e^{-qc}\right\rangle \tag{52}$$

It is interesting to note that Eq. 27 is, as one would expect, a special case of both Eqs. 44 and 52. Taking Eq. 44 first,

$$\lim_{a \to \infty} \left(1 - \frac{q\bar{T}}{a}\right)^{-a} = e^{q\bar{T}} \tag{53}$$

from the definition of e, the base of the natural system of logarithms. Since the variance of the Erlang distribution is zero for $a = \infty$, this may be interpreted as the case for which the critical gap in Eq. 41 is constant and has the value $\bar{T}$. Substitution of Eq. 53 in Eq. 44 gives Eq. 27. Likewise in Eq. 52, since $\lambda = (\bar{T} - c)^{-1}$ (Fig. 6), if $\bar{T} = c$ we have a step function for the gap acceptance probability and Eq. 52 also reduces to Eq. 27.

The fact that the delay obtained using the fixed value is both a special case of the delay obtained using the critical gap distribution in Eq. 41 and a special case of the gap acceptance function in Eq. 51 tends to establish, first, that the probability of a driver accepting a gap of size T, P(T) in Figure 6 is the same as the probability of that driver having a critical gap less than T (Fig. 5)

$$P(T) = \int_0^T f(t)\, dT \tag{54}$$

and second, that the mean, not the median, of both the critical gap distribution or the gap acceptance function is the correct parameter for delay computations.

Theoretically then, the critical gap may be obtained from a gap acceptance function. From a practical point of view this is desirable because it is much easier to obtain data on the gap acceptance characteristics of drivers than it is to try to measure the critical gaps of drivers directly. Moreover (4), it has been shown that the mean of the critical gap distribution and the mean of the gap acceptance function for a given ramp do in fact exhibit close agreement.

PROCEDURE

Data Collection

Since data on merging characteristics were to be collected at some 32 ramp locations from the far corners of the United States, there was a definite need for a standardized method of data collection in order to facilitate both the collection and analysis of data. The simultaneous collection of the many variables pertinent to the overall project objectives required the continuous viewing of an area of influence of about a quarter of a mile.

An aerial photographic technique was developed utilizing a 35-mm Automax data recording camera attached to a special tripod which was mounted in the rear of a Cessna P206 aircraft. The rear baggage compartment door was fitted with a large Plexiglas window. The aircraft circled above each entrance ramp in a radius of about ¼ to ½ mile during the peak traffic period and the traffic was filmed at 5 frames per second. Two 400-ft magazines were used with the camera which allowed two filming periods of 22 min each with approximately 2 min in between to switch magazines. The camera was fitted with a data chamber which included a clock, a frame counter, and a data slate where information about the study location could be recorded. The data chamber information was automatically recorded on one edge of the film.

Data Reduction

The frame number in which each vehicle in the outside lane of the freeway and in which each ramp vehicle crossed 200-ft stations (marked on the shoulder by 1- by 6-ft plastic stripes so as to be seen clearly in the films) was recorded. In this way, gaps could be determined at any point within the ¼-mile study area to an accuracy of ⅕ sec.

A battery of programs was written. These include a time-space diagram of the merging area for the entire study period using the Cal-Comp plotter of the Texas A&M

University Data Processing Center; plots of contour maps of such descriptive variables as speed, volume, density, acceleration noise, speed noise, energy, and shock wave speed; and summaries of volume-speed-density relations. These are discussed in detail in another report.

Pursuant to the objectives of this particular report, a computer program was written which identifies for each ramp vehicle the rejected (or accepted) lag, all rejected gaps, the gap finally accepted, the gap after the one accepted, the delay to the ramp vehicle, and the station of entry in the acceleration lane. All lags and gaps are referenced to the ramp nose.

The Binomial Response

If an entrance ramp driver, selected at random from a population, is given an opportunity to merge, the probability he will accept is p; the probability of rejecting the opportunity is (1 - p) or q. The opportunity to merge here may be measured by the time gap, space gap, relative speed, or some combination of these or other factors. If two drivers are given the same opportunity, and if their reactions are completely independent, the probability that both accept is p^2, and the probability that both reject is q^2; the probability that only the first accepts is pq, and the probability that only the second accepts is qp. Thus the total probabilities of 2, 1, and 0 accepting are p^2, 2pq and q^2 respectively, the successive terms in the expansion of $(p+q)^2$. In a similar manner, it may be seen that if a group of n drivers is exposed to the same merging conditions, and all react independently, the probabilities of n, (n-1), (n-2), . . ., 2, 1, 0 responding are the (n+1) terms in the binomial expansion $(p+q)^n$. The probability of exactly x acceptances is therefore

$$P(x) = \binom{n}{x} p^x q^{n-x}$$

This equation, of course, represents the binomial distribution of probabilities. The average number of gaps accepted in n opportunities is np and the average number rejected is nq. The variance of the distribution is npq.

It is apparent that the gap acceptance phenomenon involves a "stimulus" (available time gap) applied to a "subject" (the driver of a ramp vehicle). Variation of the stimulus is followed by a change in some measurable characteristic associated with the subject. This measurable characteristic—the acceptance of a certain gap in the case at hand—is referred to as the "response" of the subject.

Methods employed for the estimation of the nature of a process by means of the reaction that follows its application to living matter are called "biological assays." In its widest sense the term should be understood to mean the measurement of any stimulus (physical, chemical, biological, physiological, or psychological) by means of the reactions which it produces. One type of assay which has been found useful in many different disciplines is that dependent on the all or nothing response. The decision to accept or reject a gap is the type of response which permits no graduation and which can only be expressed as occurring or not occurring. The statistical treatment of this particular assay has been greatly facilitated by the development of probit analysis.

The Probit Method

Probit analysis is a well-established technique used widely in toxicology and bioassay work. Reference is made to two books by D. J. Finney (15, 16) in which, in the first, the technique is described and a rather lengthy computational method set forth; in the second, many examples covering a wide variety of cases are presented and the underlying statistical theory presented in some detail.

Normally, in an experiment to which a probit analysis is applied, the magnitude of the stimulus is controlled by the experimenter. Thus, a number of subjects (usually insects) are administered a stimulus (usually a poison) to which they either do or do not exhibit a certain response (usually dying). Each subject is assumed to have a

tolerance for the stimulus; if the stimulus is greater, then the subject responds (insect dies), etc. Typically, what is desired is a relationship expressing the percent killed as a function of the dose. Since this is not a linear relationship the estimation of the equation of the curve and tests of significance are severely complicated. The transformation from percentages to "probits" forces the curve into a linear relationship.

In a study of lag and gap acceptances at stop-controlled intersections, Solberg and Oppenlander (17) showed that the probit of the percent accepting a time gap is related to the logarithm of the time gap x by the equation

$$Y = a + bx$$

By means of the probit transformation the study data were used to obtain an estimate of this equation. The parameters of the tolerance distribution, mean and variance, were also determined. In particular the median gap and lag acceptance times were easily estimated from that value of x when Y = 5 (percent acceptance is 50 percent).

Whereas Solberg and Oppenlander (17) tabulated the data into groups at 1-sec intervals in order to obtain an estimate of drivers accepting this interval in the time series, in this report the data were not grouped. Such groupings are usually reserved for experiments in which the magnitude of the stimulus is controlled. In traffic studies such control is not possible, although the magnitude of the stimulus (available gap size) may be measured and recorded along with the response (accepted or rejected). Finney (16) terms such data "individual records." The probit method is just as applicable to individual records as to data from a controlled (grouped) experiment.

In the present study there were two practical considerations for not grouping the data. First, in the Solberg-Oppenlander data for intersections the effective range of accepted and rejected gaps was from about 2 to 13 sec; however, in the present investigation, for many ramps studied the range was as low as from 1 to 3 sec. Second, the wide geographic distribution of the 35 study sites and the relative travel made it difficult to film for an adequate duration at each location in order to obtain representative samples. It was quickly ascertained that, for grouping, an interval of 0.2 sec would be needed to give enough groups and several hours of data would be needed at each ramp to obtain enough gap data for each 0.2-sec interval.

There are, of course, problems associated with the individual records approach used here. As Finney notes, the chi-squared statistics computed in the statistical tests of significance are not as reliable as with the grouped data. Another problem with the individual records is that the iterative solution technique employed may fail to converge in cases where the sample size is small or where the data are particularly irregular.

RESULTS OF SINGLE VARIABLE ANALYSIS

Notation and Format

It is important to remember that what is desired is the form of the gap (or lag) acceptance function for the freeway merging maneuver. Many theoretical forms of this function were described in the section on "Theory" and are shown in Figure 6. Several researchers have shown that the log-normal function provides the best description of the gap acceptance phenomena. The purpose of the probit analysis is to transform this log-normal function into the linear form

$$Y = a + bx \tag{55}$$

where x = log t, t being the time interval (either lag or gap as the case may be) and Y is the probit of P, P being the probability of accepting a gap or lag of t (Fig. 6).

If x and Y are plotted on a grid, Eq. 55 has the form of a straight line. For convenience then, log-probability graph paper has been used throughout this section so that the relationship between t and P from Figure 6 may be viewed as a straight line. The abscissa of this graph paper (Fig. 7) is in the form of a logarithmic scale denoting the

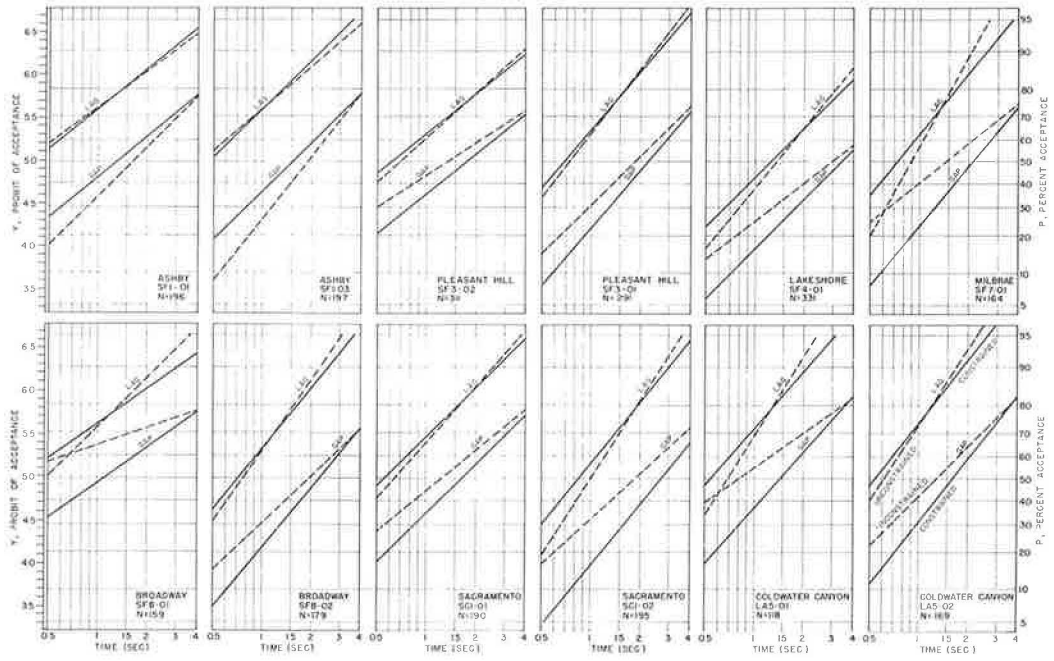


Figure 7. Lag and gap acceptance regressions for U.S. ramps.

gap (or lag) interval t in seconds; the right ordinate is a probability scale denoting percent acceptance P and the left ordinate is a linear scale denoting the probit of acceptance Y. Thus the abscissa provides the transformation between x and t and the ordinate establishes the relationship between Y and P. In the discussion which follows, a "zero" subscript will be used for lags and a "one" will be used in the subscript for gaps.

Gaps and Lags

In the past, some investigators have chosen to work with gaps (1), some with lags (2), and some with both (17). The choice of variable should not, of course, be arbitrary. Ramp drivers approaching the freeway merging area evaluate lags. In many cases, the driver evaluates the lag rather than the first gap since the lead vehicle on the freeway may have passed long before the ramp driver was in position to merge.

The reason then for studying gaps as well as lags is a practical one. It is anticipated that one important application of this research will be in the eventual development of a merging control system to help drivers execute this difficult maneuver. Presumably, such a system will be based on fitting merging ramp vehicles into openings in the freeway stream. This is most easily accomplished by a gap detector on the outside lane of the freeway located far enough upstream from the merging area to allow a metered ramp vehicle to reach the merging area at the same time as an acceptable gap.

In Figure 7, the correspondence between the effects of lag size and gap size on acceptance may be seen. Illustrated are curves for 12 films for 7 ramps taken in San Francisco (designated SF in the figure), Sacramento (SC), and Los Angeles (LA). The two nonparallel dashed lines (designated "unconstrained" in the lower right corner of the figure) define the regression of acceptance on lags and acceptance on gaps. The two solid lines (designated "constrained" in the lower right corner of the figure) illustrate the effect of forcing the lag and gap regressions to be parallel.

The results of the probit analyses for the San Francisco studies in which lags are considered the stimulus appear in Table 1. The χ^2 (chi-squared) statistic with $N_0 - 2$ degrees of freedom provides a test of the linearity of the transformed data. For all films

TABLE 1
RESULTS OF PROBIT ANALYSES ON EIGHT FILMS
Stimulus: x_0
Line Fitted: $Y = a_0 + b_0x_0$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
N_0	196	197	291	311	328	164	159	179
Sw	76.949	78.517	76.961	116.140	141.424	41.048	59.216	57.774
$\bar{x}_0$	0.112	0.049	0.220	0.268	0.274	0.105	0.067	0.056
$\bar{y}_0$	0.762	0.686	0.846	0.723	0.345	0.530	0.708	0.461
Swx_0^2	15.586	14.171	9.857	20.690	17.066	3.470	8.740	7.175
Swx_0y_0	21.298	22.265	24.058	35.178	39.035	11.871	16.247	19.104
Swy_0^2	189.883	193.564	239.783	354.345	511.978	182.492	161.794	193.001
χ^2	160.78	158.582	181.065	294.534	422.694	140.881	131.593	142.136
a_0	0.609	0.608	0.309	0.267	-0.281	0.172	0.585	0.312
b_0	1.366	1.571	2.441	1.700	2.287	3.421	1.859	2.663
m	0.358	0.410	0.747	0.697	1.327	0.891	0.485	0.764
m_l	0.150	0.218	0.519	0.451	1.085	0.670	0.261	0.570
m_u	0.578	0.602	0.962	0.941	1.573	1.100	0.698	0.959

except SF4-1 the computed χ^2 statistics are small enough to be attributed to random variation. Examination of the data for SF4-1 reveals that at several short periods during the filming, traffic was congested at the ramp. As a consequence, in several instances ramp drivers were lined up on the ramp and forced to reject large lags. Such occurrences contribute inordinately large amounts to the χ^2 statistic.

All the regression coefficients estimated are of the same order of magnitude. It is not evident, however, that all the b_0's are equal. As an approximate test of the equality of the b_0's, a χ^2 statistic with 7 degrees of freedom may be calculated as

$$\sum \frac{(S\, wx_0y_0)^2}{S\, wx_0^2} - \frac{(\Sigma\, Swx_0y_0)^2}{\Sigma\, Swx_0^2} = 27.17$$

which is clearly significant, indicating that the b_0's are not all equal. Similar tests may be made on the χ^2 for the equality of the b_0's obtained from two films on the same ramp. These tests indicate that the b_0's do not differ between films from the same ramp. Taken together, these results indicate that the regression coefficient is constant for a given ramp and some range of traffic conditions but that the b_0's may differ among ramps.

The results of probit analyses in which the first gap is considered the stimulus are given in Table 2. The only large departure from linearity occurs in SF7-1 (Fig. 7). Examination of the data shows that this heterogeneity is attributable to several short periods during which cars were lined up on the ramp. In several instances ramp drivers rejected quite large gaps. The contribution to the χ^2 from these instances is large. In each instance because the ramp driver was queued for the gap, ramp velocity was very slow compared to the speed at which freeway vehicles were traveling. As will be seen later, adjustment for ramp velocity accounts for much of this apparent heterogeneity.

In general, the b_1's are less than b_0's. This is reasonable insofar as ramp drivers are less sensitive to small differences in gap size than to small differences in lag size. However, it should be noted that inclusion of queued ramp vehicles in the analysis for gaps may cause the regression coefficient to be underestimated. A case in point is that of SF8-1 in which the relatively small b_1 may be accounted for in part by the fact that there were several occasions in which a large gap was available to several cars of which the last rejected the gap. This emphasizes the importance of separating periods of stable flow from periods of congested flow in the analysis.

TABLE 2

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Stimulus: x_i

Line Fitted: $Y = a_i + b_i x_i$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
N_i	196	197	291	311	328	164	159	179
Sw	80.146	79.427	109.890	141.276	181.660	75.080	76.572	86.609
$\bar{x}_i$	0.604	0.545	0.795	0.780	0.731	0.652	0.667	0.608
$\bar{y}_i$	0.749	0.665	0.988	0.811	0.387	0.747	0.826	0.579
Swx_i^2	6.057	5.148	7.149	14.659	16.477	6.966	7.580	7.034
Swx_iy_i	11.476	12.568	13.332	18.206	24.159	10.508	4.988	12.715
Swy_i^2	198.931	199.508	317.135	337.973	365.436	210.780	165.145	195.213
χ^2	177.188	168.826	292.273	315.362	330.014	194.930	161.863	172.229
a_i	-0.396	-0.665	-0.495	-0.158	-0.684	-0.236	0.387	-0.520
b_i	1.894	2.441	1.865	1.242	1.466	1.509	0.658	1.808
m	1.618	1.873	1.843	1.340	2.927	1.434	0.258	1.939
m_l	0.768	1.222	0.793	0.434	1.954	0.430	0.	1.049
m_u	2.306	2.398	2.742	2.234	3.781	2.308	1.279	2.665

The last three rows of Tables 1 and 2 give the 50 percent point of the tolerance (median critical lag in Table 1 and median critical gap in Table 2), the lower 95 percent confidence limit (next to last row), and the upper 95 percent confidence limit (last row) for each of the San Francisco films. Again, the results indicate differences among ramps but little difference, with the possible exception of SF8, between films of the same ramp. The 95 percent confidence limits for gap acceptance are shown in Figure 8.

Since a lag is a fraction of a gap by definition, one might expect that for any response (percent acceptance) the gap size producing that response should be a constant factor

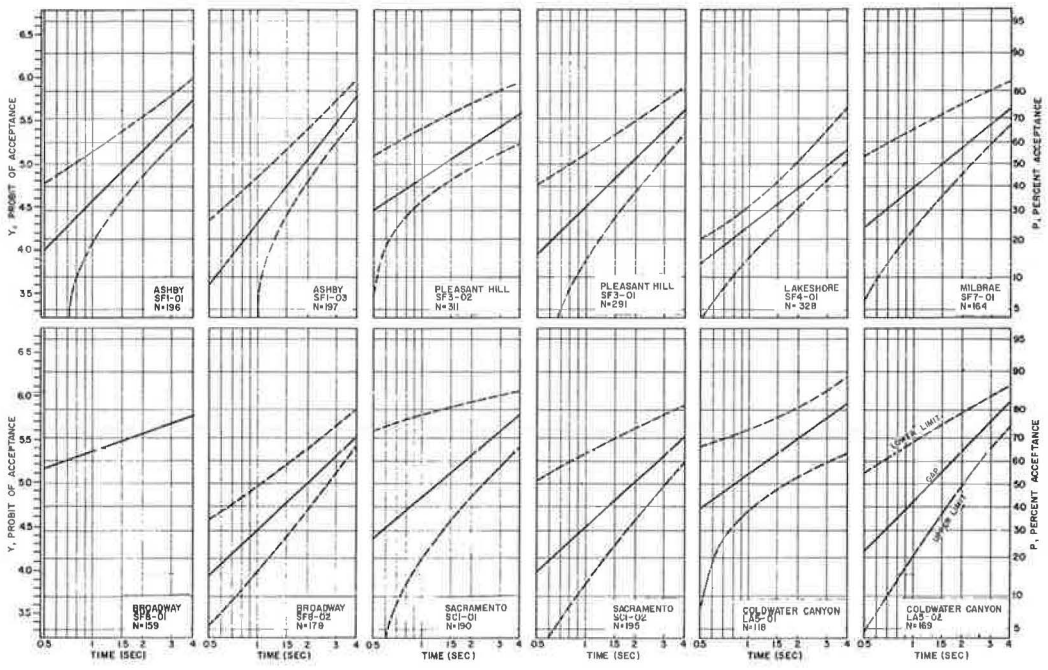


Figure 8. Gap acceptance with confidence limits.

TABLE 3

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Probit Lines for x_0 and x_1 Are Constrained To Be Parallel:

$Y_0 = a_0 + bx_0$

$Y_1 = a_1 + bx_1$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
$N_0 + N_1$	392	394	582	622	656	328	318	358
ΣSw	157.644	159.234	187.849	258.376	325.463	122.809	139.584	145.759
$\bar{x}_0$	0.093	0.025	0.236	0.285	0.283	0.149	0.117	0.073
$\bar{x}_1$	0.629	0.573	0.780	0.765	0.719	0.620	0.625	0.590
$\bar{y}_0$	0.736	0.647	0.881	0.750	0.364	0.651	0.781	0.496
$\bar{y}_1$	0.790	0.722	0.957	0.789	0.366	0.674	0.776	0.540
ΣSwx^2	21.452	18.682	18.166	36.923	35.850	13.024	18.031	15.152
$\Sigma Swxy$	32.769	34.310	40.113	56.015	68.866	29.918	23.975	32.693
ΣSwy^2	384.251	388.192	588.583	699.474	818.528	565.707	359.750	393.329
χ^2_{lin}	333.048	322.534	498.631	612.642	680.223	487.053	321.442	311.280
χ^2_{par}	1.150	2.646	1.378	1.850	6.015	9.932	6.428	2.610
a_0	0.595	0.601	0.359	0.317	-0.180	0.309	0.626	0.328
b	1.528	1.837	2.208	1.519	1.921	2.297	1.330	2.290
R	3.169	3.213	3.231	2.844	2.722	2.889	3.249	3.144
R_l	1.926	2.133	2.366	1.940	2.089	1.999	1.762	2.238
R_u	5.144	4.782	4.380	4.147	3.547	4.166	6.008	4.395

times the corresponding lag size. That is, the probit lines should be parallel. To check this and to estimate the ratio of lag size to gap size (relative potency in the probit jargon), probit analyses were performed in which the two lines were constrained to be parallel.

The results of probit analyses in which the lines for lags and gaps are constrained to be parallel are given in Table 3; χ^2_{lin} has $N_0 + N_1 - 3$ degrees of freedom and χ^2_{par} for tests of parallelism of the regression lines has 1 degree of freedom. The only significantly large χ^2_{lin} is the one for SF7-1. This is probably due to the heterogeneity of the gap data as discussed earlier. There are three significant χ^2_{par} statistics for parallelism. Each comes from one of the troublesome films discussed earlier. It is thought that inasmuch as the departure from the model in these three cases are explicable in terms of gross freeway conditions unaccounted for in the present analyses, there need be no serious doubt as to the validity of the model. The analyses of Table 3 indicate that it is not wholly unreasonable to assume that $b_0 = b_1 = b$ for each ramp, though the b's may differ among ramps. The "constrained" regression lines appear as the solid lines on the graphs in Figure 7.

In conjunction with checking the parallelism of the two probit lines, the primary purpose of the analyses in Table 3 is to estimate the "relative potencies" $R_{0,1}$ of lags to gaps: i.e., to estimate the effectiveness of lags relative to gaps in inducing drivers to merge. This is estimated as the antilogarithm of the horizontal distance between the two parallel probit lines with R being greater than unity in the case lying for lags to the left of the probit line for gaps. Thus a value of R = 2 implies that in order to induce the same percentage of response as a particular lag, the gap must be twice the size of the lag. The estimates of R in Table 3 differ little among ramps, indicating that the ratio of median critical lag to median critical gap is the same for all the San Francisco ramps.

Multiple Entries

A multiple entry occurs when two or more drivers accept the same gap. Three probit lines may be determined considering r_i the responses and x_i the stimuli. The three lines are first fitted separately as:

TABLE 4

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Multiple Entries: One or More Cars per Gap

$Y_1 = a_1 + b_1 x_1$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
N_1	120	129	131	146	176	104	106	114
Sw_1	52.986	50.771	52.742	78.079	101.772	43.244	49.389	47.045
$\bar{x}_1$	0.501	0.447	0.663	0.657	0.605	0.526	0.569	0.473
$\bar{y}_1$	0.696	0.573	0.782	0.586	0.269	0.707	0.869	0.603
Swx_1^2	2.687	2.175	2.817	7.480	8.053	3.521	3.899	2.645
Swx_1y_1	5.981	7.276	6.979	7.576	11.633	7.653	3.070	7.592
Swy_1^2	138.611	129.110	144.049	153.215	198.445	152.464	112.043	128.414
χ^2	125.298	104.770	126.759	145.542	181.641	135.830	108.626	105.623
a_1	-0.419	-0.922	-0.861	-0.080	-0.605	-0.437	0.421	-0.755
b_1	2.226	3.346	2.478	1.013	1.445	2.174	0.787	2.870
m	1.543	1.886	2.226	1.199	2.624	1.588	0.292	1.832
m_l	0.610	1.327	1.071	0.042	1.486	0.703	0.	1.161
m_u	2.177	2.312	3.056	2.371	3.594	2.289	–	2.351

$$Y_i = a_i + b_i x_i \qquad i = 1, 2, 3$$

With these lines, it is possible to estimate for any given gap size the probability that one, two or three cars will accept that available gap.

Tables 4, 5 and 6 show the results of analyses for multiple entries. The significantly large χ^2 statistic for SF7-1 and the small b_1 for SF7-1 and the small b_1 for SF8-1 are explicable as before. For each ramp there is a tendency for the b's to become successively larger: i.e., $b_1 < b_2 < b_3$ (Fig. 9). As the sensitivity of such platoons of cars to differences in gap size is affected by the decision of the last car in the platoon, this trend indicates the reasonable conclusion that the i + 1th car in line for a given gap is more sensitive to differences in gap size than the ith car. Furthermore, these analyses do not take account of ramp speed; it is expected that the i + 1th vehicle in line must slow down more than the ith vehicle in line. Thus barring any interaction of ramp speed and gap size, the probit line for three or more cars per gap

TABLE 5

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Multiple Entries: Two or More Cars per Gap

$Y_2 = a_2 + b_2 x_1$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
N_2	63	71	89	107	139	64	53	70
Sw_2	23.662	27.351	42.237	54.298	69.988	26.360	29.025	32.147
$\bar{x}_1$	0.675	0.601	0.752	0.760	0.707	0.667	0.701	0.677
$\bar{y}_2$	-0.104	-0.064	0.287	0.108	-0.253	0.118	-0.034	-0.190
Swx_1^2	0.975	1.110	2.240	3.965	4.720	1.400	2.513	2.031
Swx_2y_2	4.394	4.833	7.167	10.519	12.093	5.402	5.108	6.280
Swy_2^2	80.073	159.604	112.572	130.943	160.872	103.515	65.593	79.628
χ^2	60.271	138.561	89.641	103.037	129.889	82.671	55.211	60.210
a_2	-3.145	-2.683	-2.120	-1.909	-2.065	-2.455	-1.459	-2.285
b_2	4.507	4.356	3.199	2.653	2.562	3.857	2.033	3.093
m	4.987	4.131	4.600	5.244	6.394	4.329	5.219	5.480
m_l	4.008	3.341	3.407	4.016	5.180	3.302	3.173	4.213
m_u	6.364	5.186	5.715	6.646	8.419	5.498	8.986	7.643

TABLE 6

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Multiple Entries: Three or More Cars per Gap

$Y_3 = a_3 + b_3x_1$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
N_3	51	52	72	90	120	40	37	53
Sw_3	10.784	3.533	25.608	36.746	31.272	9.271	12.213	10.483
$\bar{x}_1$	0.883	0.868	0.833	0.876	0.908	0.920	0.869	0.871
$\bar{y}_3$	-0.210	-0.163	0.057	-0.047	-0.704	-0.325	-0.665	-0.550
Swx_1^2	0.297	0.015	0.709	1.842	1.609	0.471	0.979	0.283
Swx_1y_3	1.853	0.221	4.071	7.126	6.110	2.065	2.685	1.561
Swy_3^2	33.210	9.391	85.416	98.910	95.511	26.044	32.461	28.526
χ^2	21.650	6.135	62.041	71.343	72.309	16.991	25.098	19.916
a_3	-5.716	-13.043	-4.722	-3.436	-4.154	-4.354	-3.095	-5.361
b_3	6.235	14.836	5.738	3.869	3.798	4.382	2.744	5.526
m	8.255	7.571	6.651	7.731	12.410	9.854	13.426	9.337
m_1	6.498	4.446	5.584	6.309	9.886	6.875	8.354	7.252
m_u	11.330	9.094	7.851	9.561	18.454	18.164	72.728	17.364

would be shifted to the right of the probit line for two or more, and the line for two or more to the right of one or more. This would accentuate the shifting of the lines due to the simple fact that it takes a bigger gap for two cars than for one and for three than for two.

The 50 percent point and its 95 percent confidence limits are estimated and appear in the last three rows of Tables 4 to 6. In Table 4, the estimate of m for SF8-1 is very small due to the position and very small slope of the probit line. The explanation

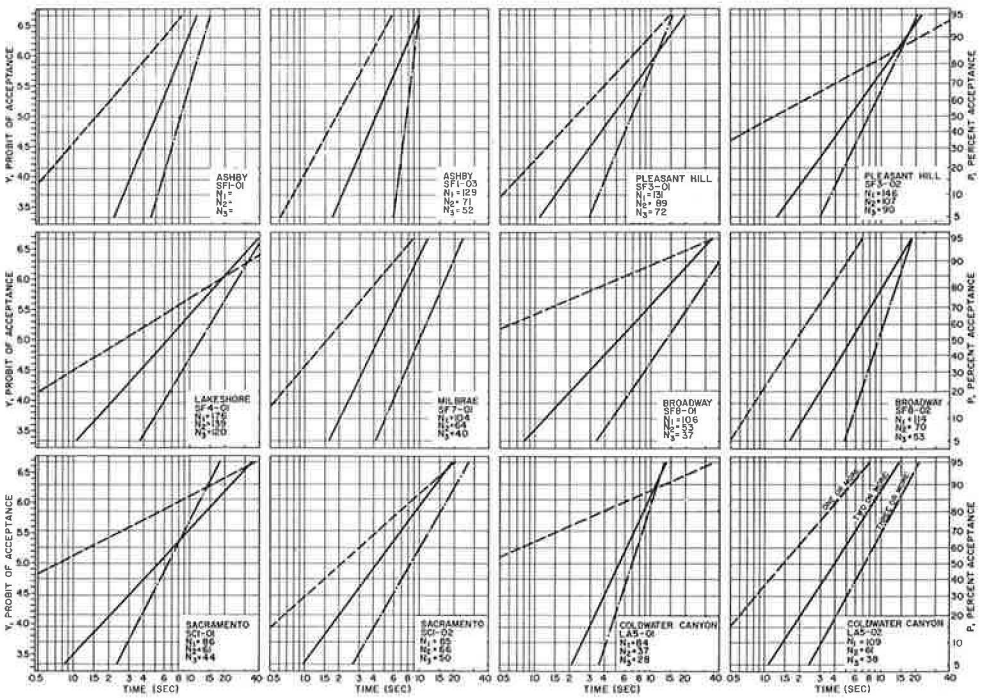


Figure 9. Gap acceptance for multiple vehicle merges.

TABLE 7

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Multiple Entries: Parallel Analyses

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
$N_1+N_2+N_3$	234	252	292	343	435	208	196	237
Sw	92.158	85.845	124.218	177.243	209.401	81.677	93.190	91.207
(1) $\bar{x}_1$	0.468	0.432	0.636	0.614	0.594	0.496	0.530	0.464
(2) $\bar{x}_1$	0.674	0.601	0.750	0.758	0.700	0.660	0.697	0.679
(3) $\bar{y}_1$	0.819	0.798	0.830	0.824	0.824	0.845	0.795	0.812
$\bar{y}_1$	0.594	0.516	0.703	0.513	0.240	0.623	0.815	0.575
$\bar{y}_2$	-0.108	-0.071	0.281	0.105	-0.264	0.113	-0.032	-0.185
$\bar{y}_3$	-0.498	-0.654	0.060	-0.189	-0.932	-0.577	-0.834	-0.781
$Sw \cdot x_1^2$	4.412	3.476	6.024	14.645	15.177	5.877	7.926	5.125
Swy_1^2	398.657	305.305	391.590	415.154	487.346	418.959	234.981	251.426
$Sw \cdot x_1 y$	15.180	14.427	20.213	30.123	32.574	17.184	12.118	16.476
χ^2_{lin}	338.595	241.512	317.583	337.517	407.910	364.094	211.783	196.815
χ^2_{par}	7.834	3.925	6.179	15.677	9.528	4.617	4.671	1.645
a_1	-1.016	-1.275	-1.431	-0.750	-1.034	-0.827	0.004	-0.917
a_2	-2.427	-2.564	-2.237	-1.454	-1.767	-1.818	-1.007	-2.368
a_3	-3.314	-3.966	-2.725	-1.883	-2.699	-3.047	-2.049	-3.390
b	3.441	4.150	3.356	2.057	2.146	2.924	1.529	3.215
$R_{1,2}$	2.571	2.045	1.738	2.199	2.194	2.183	5.253	2.828
R_1	1.874	1.576	1.306	1.492	1.578	1.503	2.538	2.031
R_U	3.801	2.774	2.395	3.443	3.250	3.390	21.184	4.288
$R_{1,3}$	4.635	4.451	2.430	3.556	5.967	5.745	22.026	5.881
R_L	3.177	3.086	1.800	2.335	3.848	3.425	7.474	3.768
R_U	7.660	7.019	3.461	6.008	10.709	11.269	238.072	10.685
m_1	1.973	2.028	2.670	2.316	3.033	1.917	0.994	1.928
m_2	5.074	4.148	4.641	5.092	6.655	4.186	5.220	5.453
m_3	9.186	9.029	6.488	8.234	18.098	11.016	21.888	11.341

for this is as presented earlier. The upper confidence limit was too large to fit into the space provided for it.

In Table 5, the χ^2 for linearity for SF1-3 is significantly large. As in other such cases this statistic is inordinately inflated by contributions from a few cases in which less than two cars rejected large gaps. The χ^2 statistic for SF7-1 again indicates misbehavior, the reasons being discussed earlier. Of special note is that the estimates for SF8-1 are much more reasonable than in Table 5. The elimination of the cases in which the gap was available to only one car evidently eliminates the heterogeneity in the data.

Table 6 gives the results of probit analyses for three or more cars per gap. The results for SF1-3 are typical of a situation encountered frequently when the gaps are not widely dispersed: the iterative procedure goes through more iterations than usual before converging, the χ^2 statistic is very small and b is very large.

Table 7 and Figure 10 show the results of probit analyses in which the probit lines for one, two and three cars per gap are constrained to be parallel under the assumption that single, double, and triple entries are equally sensitive to differences in gap size. The χ^2_{lin} for linearity has $N_1 + N_2 + N_3 - 4$ degrees of freedom and the χ^2_{par} for parallelism has 2 degrees of freedom. The χ^2_{lin} statistic for SF1-1 is clearly significant, even though none of the χ^2_{lin} statistics for the analyses on SF1-1 in Tables 4, 5 or 6 were significant. This phenomenon results from the different weights for each observation in the present analysis as compared to the analyses of Tables 4 to 6. Thus, if the b's are calculated separately using the weights of the present analysis, $b_1 = 1.811$, $b_2 = 4.278$ and $b_3 = 4.825$ result.

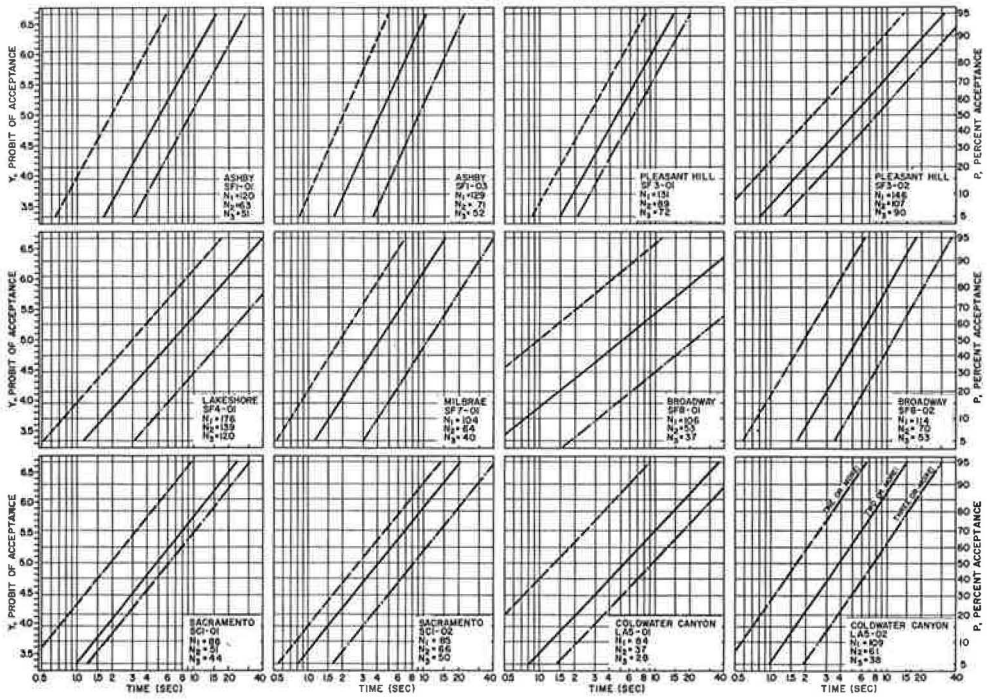


Figure 10. Gap acceptance for multiple vehicle merges (constrained or parallel).

Because of the underestimation of b_1 in the present analysis and the large number of vehicles in the "one or more" group, the contribution to the χ^2_{lin} of 278 from that group is great. The χ^2_{par} for SF1-1 is also significantly large though not extremely so. SF3-1, SF3-2, and SF4-1 show significantly large χ^2_{par} statistics. For each ramp, with the exception of SF7-1, the trend $b_1 < b_2 < b_3$ is quite pronounced and the χ^2_{par} statistics fairly large (all are significant at the 95 percent level). This fact indicates that the model considering the lines parallel is statistically inadequate. It may be concluded that double entries are more sensitive than singles, and triples than doubles, to differences in gap size. Even so, the analysis may be of use in assessing the sizes of equally effective gaps for two or more ($R_{1,2}$) and three or more ($R_{1,3}$) vehicles relative to one or more. These estimates and their 95 percent confidence limits are given in Table 7, as are the estimated 50 percent gap points for one or more (m_1), two or more (m_2) and three or more (m_3) vehicles.

EFFECT OF SPEEDS

The Ideal Merge

In the first part of this report, it was suggested that a merge must be qualified according to the terminology shown in Figure 2. Thus, the most desirable type of "gap" merge would be both "optional" and "ideal" in that it would not be made because the ramp driver had run out of acceleration lane nor would it cause turbulence in the freeway stream. The requirements of such a merge serve to document the interaction of the basic traffic elements—the driver and the vehicle, and the basic traffic characteristics—headways and vehicular speeds.

It is assumed that the average driver's normal acceleration of the merging vehicle may be represented by the following differential equation (18)

$$\frac{du}{dt} = a - bu \tag{56}$$

where u is the speed of the merging vehicle, t is time, and a and b are constants. If the merging vehicle is moving at a speed u_r at the beginning of the merge, then the limits of integration for Eq. 56 are

$$\frac{1}{-b}\int_{u_r}^{u}\frac{-b\,du}{a-bu}=\int_{0}^{t}dt$$

and the speed-time relationship is

$$\frac{\ln(a-bu)}{-b}\Bigg|_{u_r}^{u}=t$$

$$\frac{a-bu}{a-bu_r}=e^{-bt}$$

$$u=\frac{a}{b}(1-e^{-bt})+u_re^{-bt} \tag{57}$$

Since $u = dx/dt$, integration of Eq. 57 provides the equation of the time-space curve

$$x=\frac{a}{b}t-\frac{a}{b^2}(1-e^{-bt})+\frac{u_r}{b}(1-e^{-bt}) \tag{58}$$

Substitution of Eq. 57 in Eq. 56 gives the acceleration-time relationship for the merging maneuver:

$$\frac{du}{dt}=(a-bu_r)\,e^{-bt} \tag{59}$$

The units of the constants a and b^{-1} are those of acceleration and time respectively, where a is the maximum acceleration and (a/b) is the free speed in the merging area. The forms of Eqs. 56 to 59 are shown in Figure 11.

The time-space relationship of Figure 11 has been reproduced in Figure 12 along with the procedure for determining the theoretical minimum ideal gap for merging. Such a gap is composed of three time intervals: (a) a safe time headway between the merging vehicle and the freeway vehicle ahead T_r; (b) the time lost accelerating during the merging maneuver T_L; and (c) a safe time headway between the second freeway vehicle and the merging vehicle T_f. The safe headway referred to is that headway between two vehicles in a lane which will allow the following vehicle to stop safely even if the vehicle in front makes an emergency stop. If reaction times τ, braking capabilities and speeds u are assumed to be equal, then the safe headway is $(L/u) + \tau$ where L is the length of the vehicle in front.

The time necessary for the merging ramp vehicle to accelerate from a speed u_r at the beginning of the merge to attain the speed of the freeway traffic u may be obtained by solving Eq. 57 for time.

$$T_2=-\frac{1}{b}\ln\left(\frac{a-bu}{a-bu_r}\right) \tag{60}$$

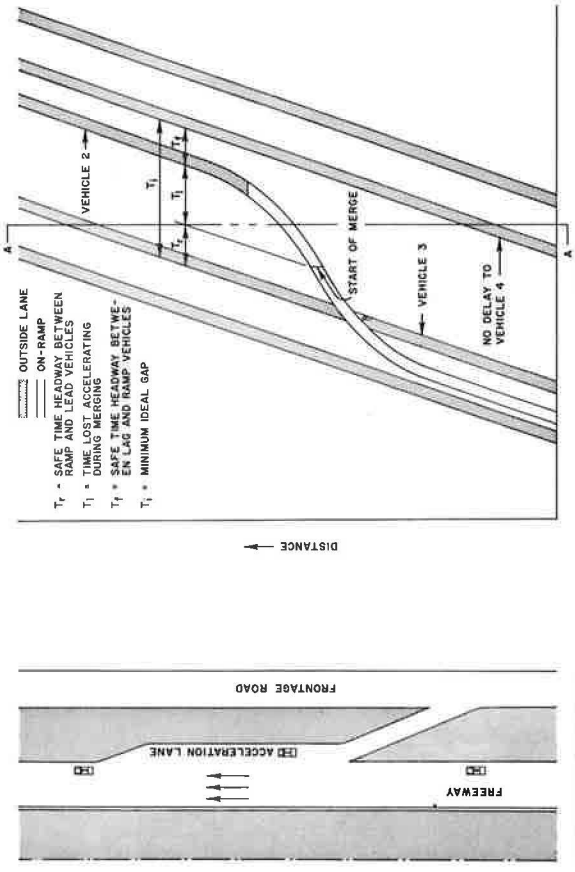


Figure 12. Theoretical minimum ideal gap for merging.

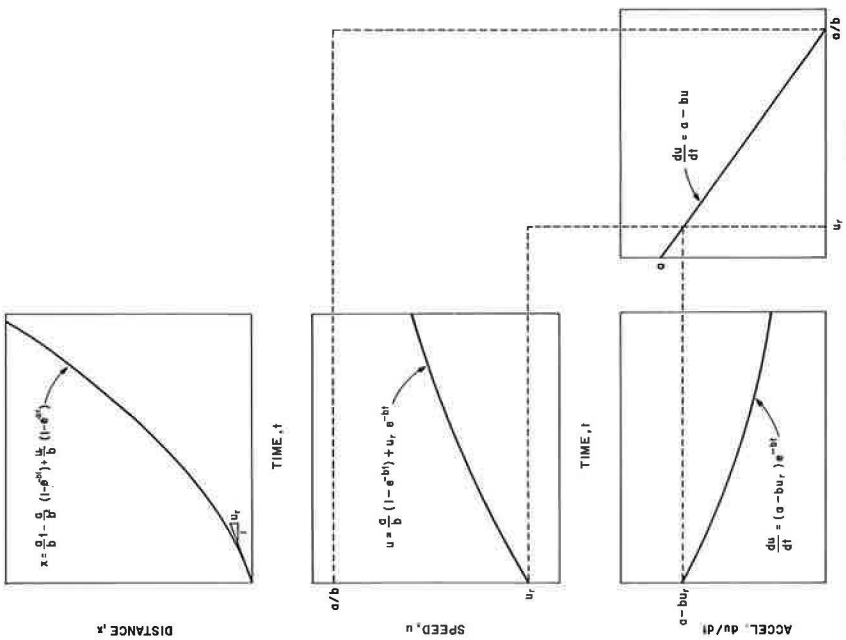


Figure 11. Theoretical speed-acceleration relationship for a merging vehicle.

Subtracting the travel time T_1 to travel the same distance covered during merging, only at a constant speed u, from Eq. 60 gives the time lost during merging T_L. The theoretical minimum ideal gap for merging ($T_r + T_f + T_L$) is

$$T = \frac{L_f + L_r}{u} + 2\tau + \frac{u + u_r}{bu} + \frac{(a/b) - u}{bu} \ln\left(\frac{a - bu}{a - bu_r}\right) \quad (61)$$

in which L_f and L_r are the lengths of the freeway and ramp vehicles. It is apparent from Eq. 61 that a merging truck would require a larger gap than a merging passenger car by virtue of its longer length L_r, and its reduced accelerating capabilities a. Similarly, if the first of the two freeway vehicles is a truck, the merging vehicle would require a larger gap because of an increased L_f in Eq. 61.

To determine representative values of minimum safe gaps, the parameters a and b in Eq. 61 must be estimated. Haight (19) suggests that some information on the parameters could be obtained from drag races which are now widely held. Knox (20) used the average result of road tests obtaining values of a = 4.8 mph/sec and a/b = 80 mph for Australian conditions.

Speed of Merging Vehicle

It is apparent from the previous section that the ramp drivers' problem in executing the merging maneuver is more than one of simply evaluating successive time headways in the freeway traffic stream until he finds a large enough gap. For example, Eq. 61 shows that the ideal gap is based on the ramp driver's reaction time τ, the vehicle characteristics L_f, L_r, a, and b, and the speeds of the mainstream vehicles and ramp vehicle u and u_r. The influence of speeds on merging should come as no surprise; we know that, for example, under conditions of forced flow on the freeway that comparatively large time headways are rejected. Theoretically, infinitely large time headways would necessarily be rejected as the concentration and movement ceased.

In delineating the effect of speeds on merging, consider a very simple model consisting of two vehicles A and B separated by a space headway s traveling at a constant velocity u on the outside lane of the freeway, and a vehicle C traveling at speed u_r on a corner intersecting the freeway at an angle δ. It follows that the gap between A and B is

$$t = s/u \quad (62)$$

However, since the driver in vehicle C is moving he views the time gap between A and B as

$$t' = \frac{s}{u - u_r \cos \delta} \quad (63)$$

where $u_r \cos \delta$ is vehicle C's speed component along the freeway. Solving for s in Eq. 63 and substituting in Eq. 62 yields

$$t = \left(\frac{u - u_r \cos \delta}{u}\right) t' \quad (64)$$

If $t' > T'$ where T' is a constant, the gap will be accepted. Therefore, the minimum acceptable gap T,

$$T = \left(\frac{u - u_r \cos \delta}{u}\right) T' \tag{65}$$

is clearly a function of the speed of the freeway traffic u and the speed of the ramp vehicle u_r.

The effect of ramp geometry on merging operation—the subject of a companion report (21)—is indicated in Eq. 65. The equation suggests that the difference between traffic operation at an ordinary intersection and that at a freeway ramp terminal is primarily in the angle at which entering traffic and through traffic converge. For an intersection (δ = 90 deg) one obtains $T = T'$ from Eq. 65; for a lane change (δ = 0 deg)

$$T = \left(\frac{u - u_r}{u}\right) T' \tag{66}$$

Whereas for an intersection, gap acceptance is independent of the speed of the merging vehicle, the speed of a vehicle merging from an entrance ramp has a profound effect on gap acceptance. This is shown in the graphs in Figure 13 for the California ramps. The dashed line pertains to ramp vehicles with speeds at the beginning of the merge greater than 20 mph; the solid line to ramp vehicles with speeds at the beginning of the merge less than 20 mph. For most of the California ramps there were virtually no slow-moving ramp vehicles hence only the >20-mph line appears. However, for the Ashby, Pleasant Hill and Sacramento ramps the effect of the speed of the merging vehicle u_r in accordance with the model exemplified by Eq. 65 is evidenced. For the higher speed a much smaller gap was needed.

Angular Velocity Model

Several researchers have hypothesized the use of angular velocity as the basis for the gap acceptance decision. The advantage of this parameter is that it takes into

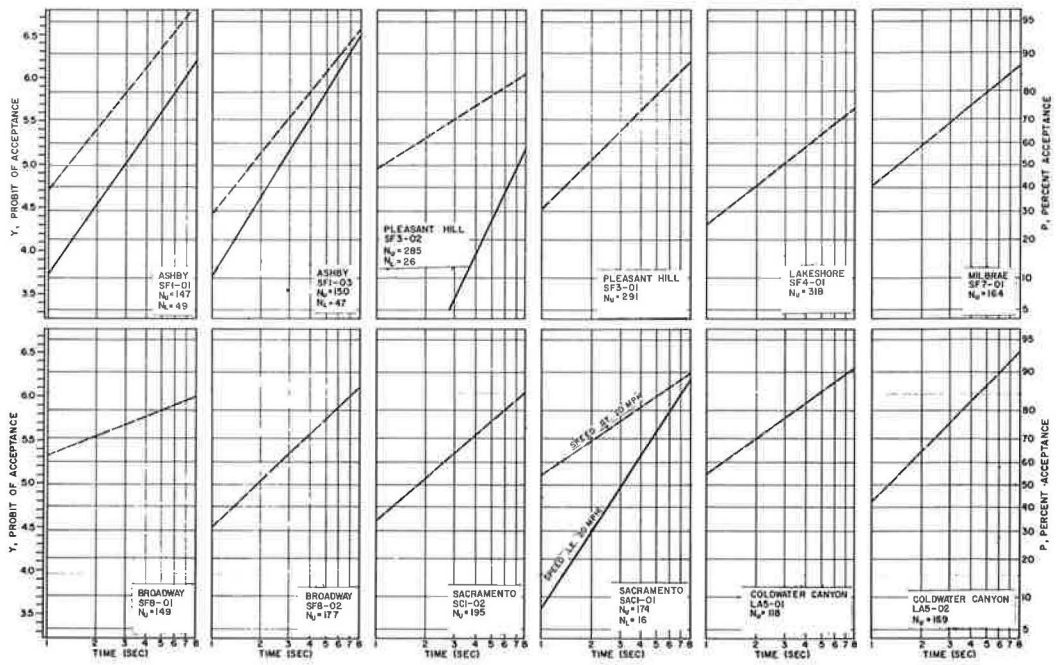


Figure 13. Gap acceptance for fast and slow merging vehicles.

consideration the distance of the approaching freeway vehicle and its speed of approach. The angular velocity model for gap acceptance is an application of the calculus of related rates. Referring to the moving coordinate system in Figure 14, we can write the equations for the relative speed of the freeway vehicle and merging vehicle and their space lag:

$$\frac{dx}{dt} = u - u_r \tag{67}$$

$$-x = w \cot \theta \tag{68}$$

Differentiating Eq. 68 with respect to time, equating it to Eq. 67 and solving for the angular velocity $\dot{\theta}$ yields

$$\dot{\theta} = \frac{w}{s^2} (u - u_r) \tag{69}$$

The basic aspect of the theory is that a ramp driver rejects a lag if he detects an angular velocity but accepts the lag if he cannot perceive any motion. Michaels and Weingarten (22) utilize this criterion to develop the equation for the minimum acceptable gap T:

$$T = s/u = \frac{w}{\dot{\theta} u^2} (u - u_r)^{1/2} \tag{70}$$

Although the direct application of Eq. 70 to determine gap acceptance times depends on the evaluation of drivers' threshold of angular velocity, the role of speed in gap acceptance is again evidenced. The relative speed $(u - u_r)$ appearing in the numerator of Eq. 70 suggests that the lower the relative speed the smaller the gap needed. The speed of the mainstream traffic u in the denominator indicates an inverse relationship. For low-speed freeway traffic, a larger gap is needed. Moreover, the fact that the latter variable is raised to a higher power than relative speed suggests it might have more effect on the gap acceptance performance. Before the effect of speeds on gap acceptance is explored further, some aspects of a two variable probit analysis need to be discussed.

Results of Two Variable Probit Analysis

It was demonstrated in the single variable probit analysis how the acceptance curve (percent acceptance as a function of the single variable gap or lag size) may be transformed into a straight line. Just as a regression analysis may involve more than one independent variable, a probit analysis may be generalized to include more than one independent variable. Based on the theory of the preceding sections, the two obvious independent variables are gap (or lag) and some speed parameter (u, u_r, or $u - u_r$). Figure 15 shows a hypothetical gap acceptance surface for a freeway speed u = 50 mph. A vertical plane parallel to the T axis yields a gap acceptance curve of the form shown in Figure 6 and discussed in the single variable analysis. It should be apparent that the equation of the surface in Figure 15 would be difficult to write—the value of the parameters of the equation from the row data even more difficult to estimate.

A probit analysis for each ramp is presented in which a combination of x_0 and u_r is considered the stimulus and r the response. In each analysis there is obtained an

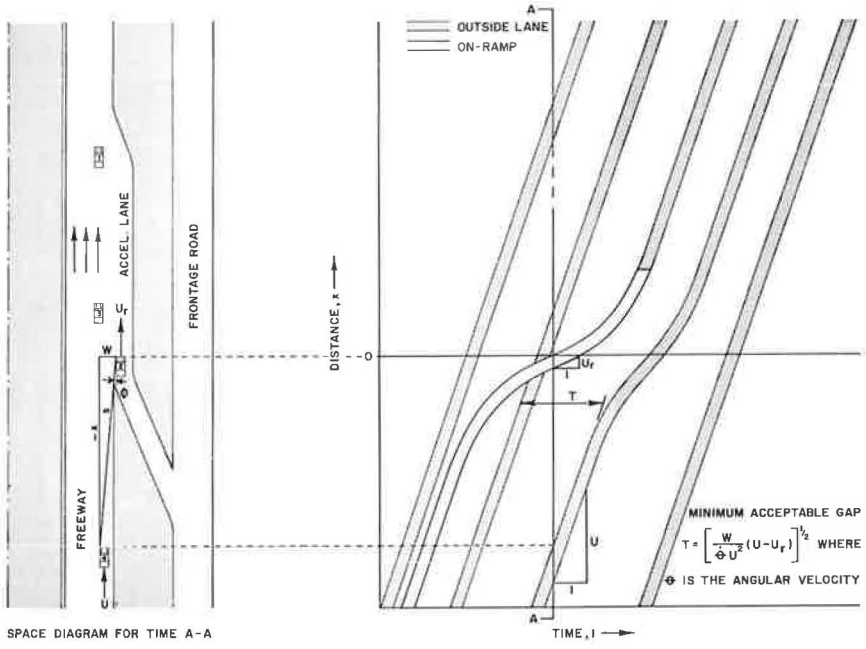


Figure 14. Theoretical minimum acceptable gap for merging.

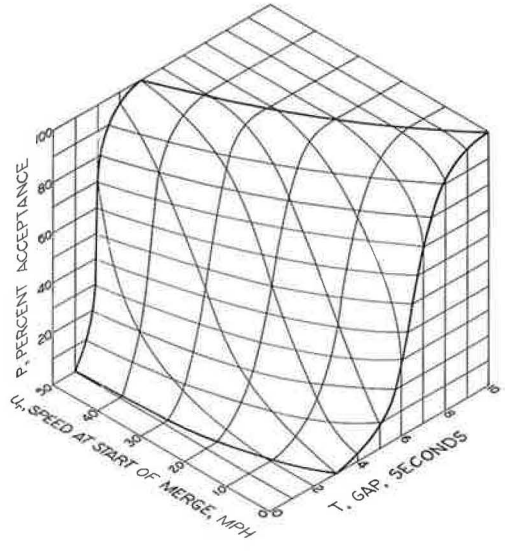


Figure 15. Gap acceptance surface for freeway speed, u = 50 mph.

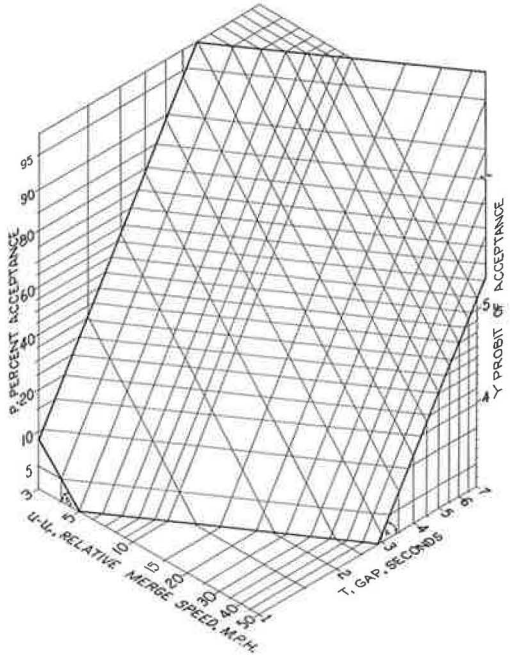


Figure 16. Acceptance plane.

TABLE 8

RESULTS OF PROBIT ANALYSES ON EIGHT FILMS

Two-Dose Analyses: $Y = a + b_1 x_0 + b_2 V_r$

Function	SF1-1	SF1-3	SF3-1	SF3-2	SF4-1	SF7-1	SF8-1	SF8-2
N	196	197	291	311	328	164	159	179
Sw	70.991	74.667	68.148	111.185	139.179	31.596	58.869	53.082
$\bar{x}_0$	0.098	0.048	0.203	0.270	0.275	0.097	0.064	0.056
$\bar{V}_r$	26.110	26.382	37.184	34.210	30.343	37.753	31.205	32.708
$\bar{y}$	0.711	0.672	0.722	0.723	0.348	0.378	0.702	0.436
Swx_0^2	13.879	13.070	8.919	19.504	17.184	2.529	8.728	6.250
SwV_r^2	7550.095	7262.986	2713.547	10108.283	4280.659	767.460	3072.899	1512.316
Swx_0V_r	-114.085	-77.616	-63.768	-35.274	-8.788	-27.331	-29.845	-17.034
Swx_0y	17.444	19.731	19.455	32.110	38.655	7.225	16.178	16.220
SwV_ry	123.334	106.030	56.143	238.825	133.231	-3.576	-13.511	68.524
Swy^2	180.575	182.233	245.760	349.495	505.946	105.377	159.074	182.510
χ^2	147.891	145.119	186.437	287.857	413.515	72.984	128.497	131.747
a	-0.498	-0.275	-3.063	-0.747	-1.362	-6.019	0.142	-2.236
b_1	1.588	1.705	2.800	1.700	2.268	4.562	1.902	2.805
b_2	0.040	0.033	0.086	0.030	0.036	0.158	0.014	0.077
$V_r = \bar{V}_r$:m	0.447	0.450	0.882	0.699	1.323	1.032	0.495	0.795
V_r=0:m	2.058	1.451	12.419	2.751	3.985	20.875	0.842	6.270

estimate of a probit plane defined by $Y = a + b_1 x_0 + b_2 u_r$. Thus for any combination of x_0 and u_r, it is possible to estimate Y and the corresponding percentage of all ramp drivers who would accept that combination. Or, going in reverse, it is possible to estimate for any given percentage acceptance the corresponding combinations of x_0 and

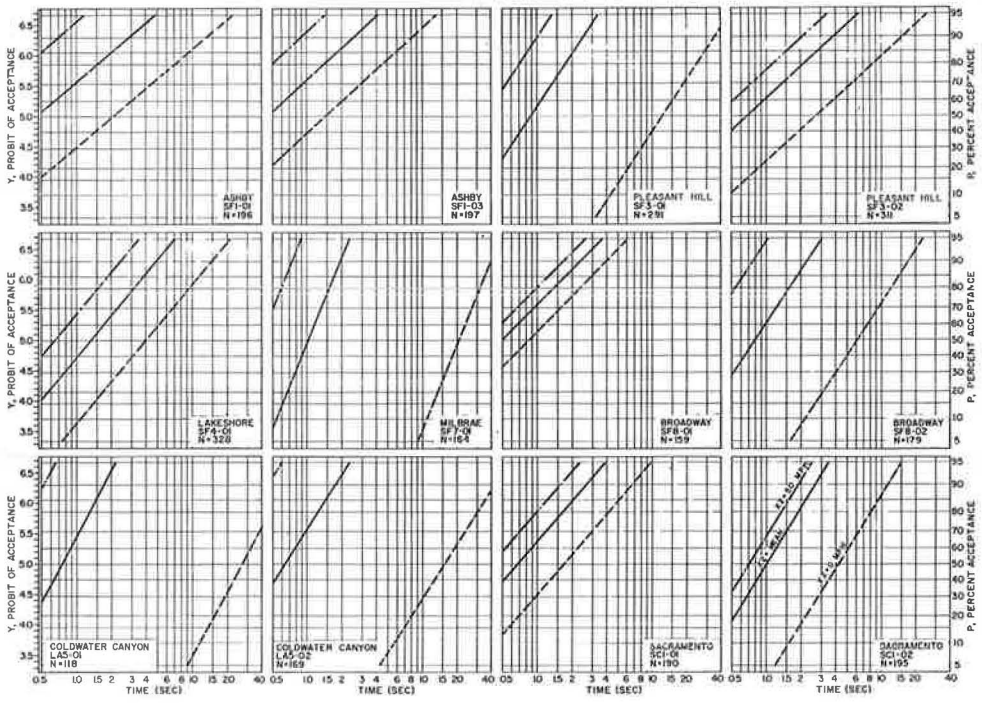


Figure 17. Effect of speed of ramp vehicle on lag acceptance.

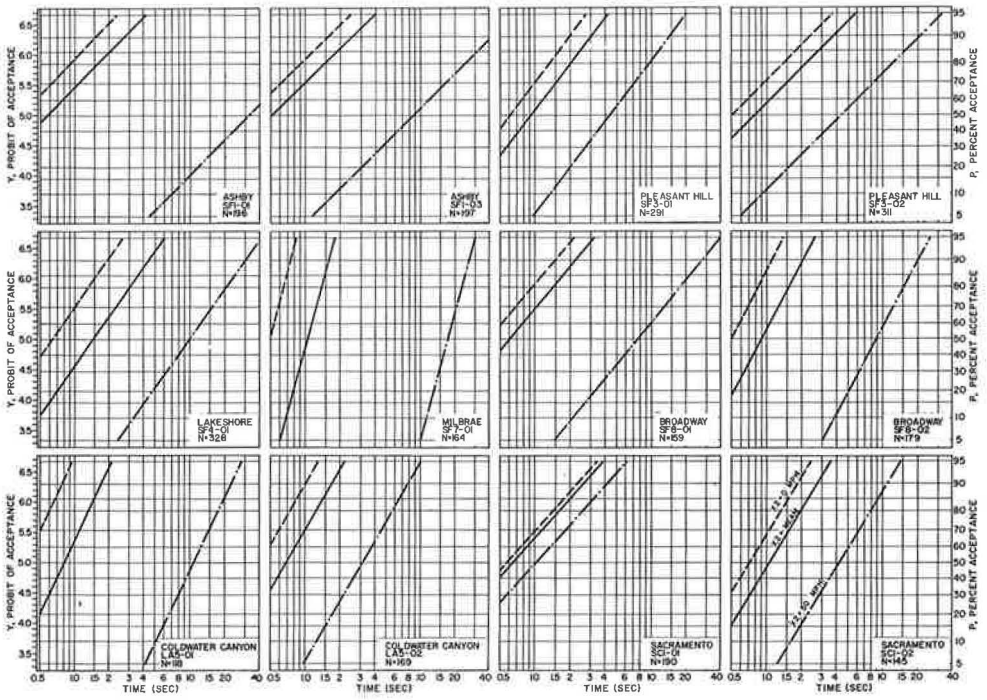


Figure 18. Effect of relative speed on lag acceptance.

u_r producing that percentage. The effect of the two variables probit analysis is to change an acceptance plane (Fig. 16). Again, Figures 15 and 16 are meant to be illustrative rather than conclusive, the actual independent variables used being the lag x_0 and the speed of the ramp vehicle u_r.

Table 8 gives the results of analyses in which x and u_r together constitute the stimulus variable. The χ^2 statistic with N - 3 degrees of freedom provides a test of linearity. The only significant χ^2 is that for SF4-1; again, the congested traffic conditions probably caused the apparent heterogeneity. Ramp SF7 shows the greatest effect of ramp speed on lag acceptance; this result supports the arguments proffered earlier explaining the apparent heterogeneity of data for this ramp. The last two rows of Table 8 give estimates of the 50 percent points for lags when $u_r = \bar{u}_r$ and $u_r = 0$. The results of the analysis summarized in Table 8 are shown in Figure 17 for the California ramps. The three lines on each graph depict from left to right a speed at the start of the merge of 50 mph, the mean speed for the study period and 0 mph.

The two variable probit analysis was repeated using the relative speed of the merging vehicle with respect to the freeway lag vehicle,

$$Y = a + b_1 x_0 + b_2 (u - u_r)$$

The results for the California ramps (Fig. 18) verify the models suggested by Eqs. 65 and 70.

SUMMARY

Theory

The first objective of this study was the development of models and useful parameters for describing the freeway merging process. This theoretical development began

with the derivation of the forms for the mean and variance of the delay to a ramp vehicle in position to merge in terms of the ramp driver's mean critical gap, freeway flow, and freeway gap availability expressed in terms of the appropriate Erlang distribution (Eqs. 27 to 34). This, a generalization of previous delay models, does not appear in the literature. The importance of this delay parameter lies in its applicability as a means to accurately forecast the capacity and operating conditions which can be expected for individual ramp-freeway junction designs. This application is the subject of a future report.

The remainder of the theoretical development concerns the treatment of the variability of critical gaps and gap acceptance among drivers. The significance of this portion of the investigation lies in the identification and application of the representative forms for both critical gap distributions and gap acceptance functions.

Characteristics

The emphasis in this report is on the collection and collation of gap acceptance characteristics. The facts that 32 ramps—chosen to reflect diverse operating, geometric, geographic and environmental conditions—were continuously filmed at 5 frames per sec for an average of an hour, and that enough data were collected to run 1344 usable gap acceptance regressions, serve to demonstrate not only the vast quantity of data involved, but the nature of the characteristics now available to interested researchers.

It is thought that the application of the individual record probit analysis has provided simple, statistically significant linear relations between gap acceptance and gap size. The idea of constraining pertinent regression lines to achieve parallelism affords the means of quantitatively comparing lags to gaps, single entry to multiple entry merges and fast to slow merging vehicles.

It is well known that the driver of a merging vehicle possesses a certain ability to adjust his merging environment. By speeding up or slowing down, he may adjust his position relative to that of a sequence of freeway vehicles, thereby enhancing his opportunity to merge. Indeed, an appreciable percentage of drivers accepted a gap before the first available gap simply by overtaking it. To account for this "dynamic" aspect of merging, lag acceptance characteristics at points along the acceleration lane were measured and summarized in the conventional graphical format used throughout this report. However, because of its applicability, these graphs are to be included in another project report establishing the effects of ramp and acceleration lane geometrics on merging operation.

Figure 13 shows that not only might the percent acceptance be related to the logarithm of the time gap for fixed freeway and ramp vehicle speeds, but that this linearity might extend to the logarithm of these speeds for a fixed critical gap. Therefore, as a representation of the effect of time and speed on gap acceptance, a surface similar to the one shown in Figure 15 seemed appropriate. To fit such a surface (with the desired log normal curve in the percent acceptance-time plane) to real data, and then test it statistically, a two-variable probit analysis was used. Under this transformation, the surface of Figure 15 becomes the plane of Figure 16. This model for the percent acceptance-time-speed data proved statistically reliable, tending to establish that the logarithms of the critical gap and some function of speed (relative or absolute) required to give any percent acceptance are linearly related.

Application

The effect of outside freeway lane volumes on gap acceptance has not been discussed. In a study of lag acceptance at urban intersection, Raff (2) wrote "it is safe to say that the main-street volume does not have an appreciable effect in the critical lag." Similarly for ramps, Wohl (23) attributed a slightly smaller percent acceptance at higher freeway volumes to sampling rather than a difference in driver behavior.

The results of this investigation tend to support the conclusions of Raff and Wohl. The graphs for individual films taken at the same location illustrate the consistency of the gap and lag acceptance curves (Fig. 7). However, since gap acceptance is a function of speed and speed is a function of volume, a relationship does exist. Because of

the form of the speed-volume relationship (the classic parabola showing two speeds for any given volume), it is evident that the volume-gap acceptance relationship is very complex. It would suffice to say that gap acceptance is independent of volume for a given level of service (free, stable, unstable, or forced flow), and that changes in gap acceptance between different levels of service may be predicted from differences in speed rather than the service volumes. From a practical point of view, since a much larger gap is needed for a freeway volume of 1800 vph under a forced flow condition (20 mph) than under a stable flow condition (40 mph), positive means should be taken to prevent the freeway from becoming congested.

An important application of the results of this report lies in the field of simulation. Computer technologists have been quite active with several digital computer programs developed for simulating freeway and interchange operation. As a design and operational tool, present programs are not sufficiently validated to warrant confidence in their ability to predict behavior or needs at freeway interchanges. While development of simulator hardware and programs has proven feasible, lack of detailed criteria on traffic stream interaction has hampered progress. While some simplifications are necessary in simulation models, simplification necessitated because of lack of knowledge of the pertinent variables reduces the model's realism. Over-simplification of the merging logic is a case in point. For example, distributions for merging gap acceptance under different geometric, traffic, and environmental conditions have been lacking.

The results of the probit analyses serve to present the pertinent gap acceptance variables in the form of distributions (the log-normal) and equations (the linear probit equation), making them extremely usable as inputs to digital simulation models. A subroutine is being written to perform this.

This project is but the first phase of a four-year research program the U.S. Bureau of Public Roads is undertaking to (a) furnish more detailed information on the effect that geometric variables have on the merging of ramp traffic, (b) develop usable distributions of traffic variables for simulation programs, and (c) develop an optimum ramp metering and merging control system. The material presented in this report should aid materially in fulfilling these ultimate objectives.

The importance of preventing freeway congestion was mentioned earlier with reference to the volume-gap acceptance relationship. Of course, the concept of ramp control as a means of deferring congestion is now a reality. It only remains for these ramp control systems to be generalized to include merging control. The design of such a merging control system will depend on the proper interpretation of merging gap acceptance characteristics.

ACKNOWLEDGMENTS

The success of this phase of the project with the national scope of its field studies is due to a large extent on the cooperation of the federal government, six state governments, and many local governmental agencies. The following partial list of personnel, their staffs and agencies is offered: California Division of Highways—Jamis E. Wilson, Karl Moskowitz and Leonard Newman; Illinois Division of Highways—Charles H. McLean; Michigan State Highway Department—Joseph Marlowe; Missouri Highway Department—James Little and James Roberts; Texas Highway Department—Dale D. Marvel, John N. Lipscomb and William V. Ward; Detroit Department of Traffic—Alger F. Malo; and New York City Department of Traffic—Edward Bonelli.

Deserving special mention is Joseph W. Hess, Acting Leader, Improved Utilization of High Speed Highways Task Group of the U.S. Bureau of Public Roads. Hess, who was instrumental in defining the scope and direction of the project in the prospectus, has helped greatly in making contacts with local agencies and selecting study sites.

The staff of the Texas Transportation Institute who worked in the collection, reduction, and presentation of the data are to be congratulated. Special recognition are due to Thomas G. Williams, Research Assistant, who assumed the responsibility for both the flight and filming aspects of the study procedure, and to Charles E. Wallace, Research Assistant, who coordinated the many activities inherent in the presentation of this report.

REFERENCES

1. Greenshields, B. D., Shapiro, D., and Erickson, E. L. Traffic Performance at Urban Street Intersections. Tech. Rept. No. 1, Bureau of Highway Traffic, Yale Univ., 1947.
2. Raff, M. S., and Hart, J. W. A Volume Warrant for Urban Stop Signs. The Eno Foundation for Highway Traffic Control, Saugatuck, Conn., 1950.
3. Blunden, W. R., Clissold, C. M., and Fisher, R. B. Distribution of Acceptance Gaps for Crossing and Turning Maneuvers. Australian Road Research Board Proc., Vol. 1, Part 1, 1962.
4. Drew, D. R. Gap Acceptance Characteristics for Ramp-Freeway Surveillance and Control. Highway Research Record 157, pp. 108-143, 1967.
5. Tanner, J. C. The Delay to Pedestrians Crossing a Road. Biometrika, Vol. 38, Nos. 3 and 4, Dec. 1951.
6. Mayne, A. J. Some Further Results in the Theory of Pedestrians and Road Traffic. Biometrika, Vol. 41, p. 375, 1954.
7. Feller, W. An Introduction to Probability Theory and Its Applications. Vol. 1, John Wiley and Sons, New York, 1950.
8. Garwood, F. An Application of the Theory of Probability to the Operation of Vehicular-Controlled Traffic Signals. Jour. Roy. Stat. Soc., Vol. 7, p. 65, 1940.
9. Tanner, J. C. A Theoretical Analysis of Delays at an Uncontrolled Intersection. Biometrika, Vol. 49, Nos. 1 and 2, pp. 163-170, 1962.
10. Evans, D. H., and Herman, R. The Highway Merging and Queueing Problem. Oper. Res., Vol. 12, 1964.
11. Major, N. G., and Buckley, D. J. Entry to a Traffic Stream. Australian Road Research Board Proc., Vol. 1, Part 1, pp. 206-228, 1962.
12. Dawson, R. F. Discusion of Gap Acceptance Characteristics for Ramp-Freeway Surveillance and Control by D. R. Drew. Highway Research Record 157, pp. 108-143, 1967.
13. Herman, R., and Weiss, G. Comments on the Highway-Crossing Problem. Oper. Res., Vol. 9, No. 6, pp. 828-840, Nov.-Dec. 1961.
14. Weiss, G., and Maradudin, A. Some Problems in Traffic Delay. Oper. Res., Vol. 10, No. 1, pp. 74-104, Jan.-Feb. 1962.
15. Finney, D. J. Probit Analysis. Cambridge Univ. Press, 1947.
16. Finney, D. J. Statistical Method in Biological Assay. Hafner Publishing Co., New York, 1964.
17. Solberg, Per, and Oppenlander, J. C. Lag and Gap Acceptances at Stop-Controlled Intersections. Highway Research Record 118, pp. 48-67, 1966.
18. Matson, T. M., Smith, W. S., and Hurd, F. W. Traffic Engineering. McGraw-Hill, New York, 1955.
19. Haight, F. A., Bisbee, E. F., and Wojcik, C. Some Mathematical Aspects of the Problem of Merging. HRB Bull. 356, pp. 1-14, 1962.
20. Knox, D. W. Merging and Weaving Operations in Traffic. Australian Road Research, Vol. 2, No. 2, Dec. 1964.
21. Wattleworth, J. A., Buhr, J. H., Drew, D. R., and Gerig, F. A. Operational Effects of Some Entrance Ramp Geometrics on Freeway Merging. Presented at the 46th Annual Meeting and published in this RECORD.
22. Michaels, R. M., and Weingarten, H. Driver Judgment in Gap Acceptance. Unpublished, April 1965.
23. Wohl, M., and Brand, D. Applications of Simulation to Highway Design. Dept. of Civil Engineering, MIT, Cambridge, Mass., 1963.

The Development and Validation of a Digital Simulation Model for Design of Freeway Diamond Interchanges

A. V. GAFARIAN, E. HAYES and W. W. MOSHER, JR., System Development Corporation, Santa Monica, California

This paper contains a brief description of the total project plans leading to a valid simulation model for design of diamond interchanges. The remainder, and by far the major portion of the paper, describes the work completed through February 1966.

In order to provide believable results applicable to the control and geometric design of diamond interchanges, the principal emphasis is on the validation problem. A description is given of experiments conducted using magnetic loop detectors as well as helicopter mounted 35 and 70-mm movie cameras for data collection. An evaluation of data reduction techniques was made. The authors' conclusions are to use a 70-mm helicopter mounted movie camera for collecting data and a semiautomatic film record reader for reducing data. The macroscopic approach the authors are using for building the total model is believed to be new for vehicular traffic simulations of this type.

•DIGITAL computer simulation shows some promise of being useful to the traffic engineer in the design and evaluation of vehicular traffic systems. This usefulness stems from two facts: (a) the phenomena involved are so complex that total mathematical analysis is virtually impossible; and (b) it is difficult, costly, and sometimes impossible to make changes and perform the evaluation in the real world; however, once a successful simulation model is built, even radical modifications to the system may be made and evaluated with relative ease.

The principal obstacle to using simulation as a tool in traffic engineering appears to be the question of realistic performance with respect to the measures of effectiveness of interest. To date, very little work has been done toward answering this question. Two basic problems are involved: the development of techniques for collecting field data, and the comparison of field data with computer output to determine if there exists a significant difference. Considerable work is presently under way in the area of data collection; however, very little is being done toward analysis of these data in conjunction with simulation studies. Such analysis is essentially a statistical problem for which some existent techniques are applicable. There are, however, some problems peculiar to vehicular traffic phenomena for which statistical techniques have not yet been developed.

The diamond interchange (Fig. 1) was selected for this study for several reasons. First, this type of interchange is frequently used on freeway systems in urban areas. Hence, any useful results developed in this study would have widespread application. Second, with good design and signalization, the diamond interchange may provide an efficient means for the interchange of traffic between a major artery and a freeway. Third, the various traffic phenomena indigenous to the diamond interchange appear as component parts in many other traffic situations. Therefore, the development of a successful model would have wide implications.

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 46th Annual Meeting.

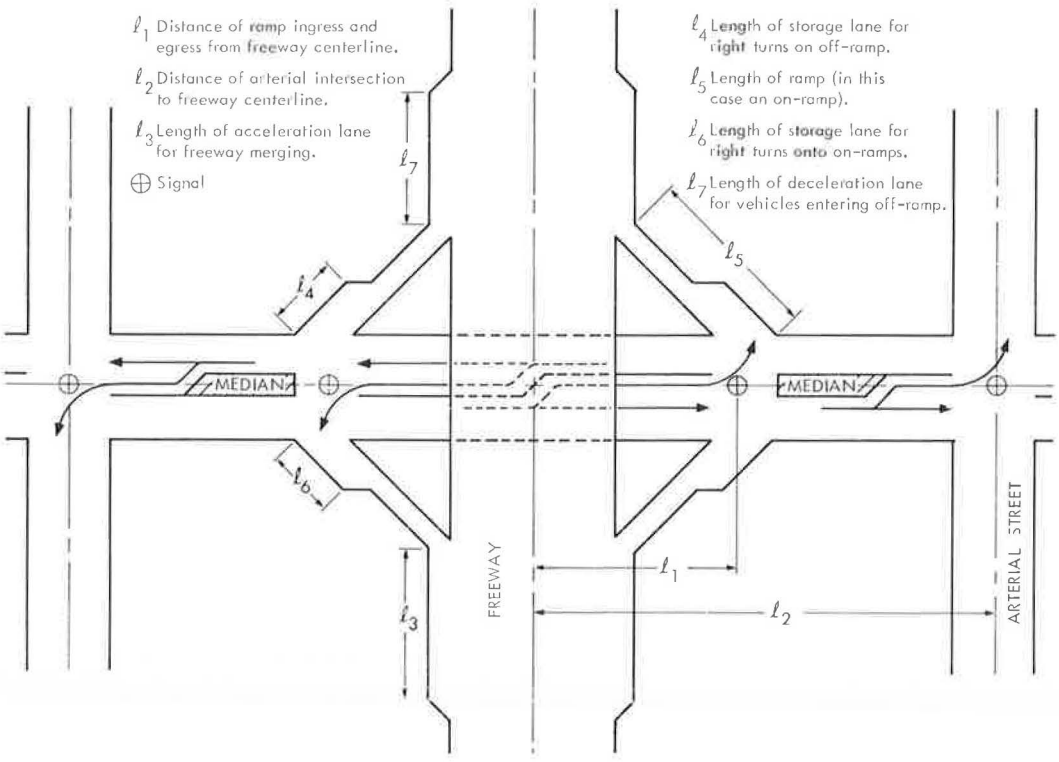


Figure 1. Diamond interchange showing some geometrical parameters.

The long-term goal of this study is to develop some needed, basic knowledge of diamond interchange operations. The specific aims are:

1. Formulation of models of the man-machine-environment interactions;
2. Simulation of these models on a large-scale digital computer, by translation of the models into computer code;
3. Validation of the computer simulation by comparing the values of a series of measurements of the operation of actual interchanges with corresponding computer simulation output;
4. Modification of the models, as suggested by the validation study, followed by repetition of the entire process;
5. Development of a computer program logic that minimizes the ratio of running time on the computer to simulated time; and
6. Performance of research on the use and design of control devices as well as on the overall geometric design of diamond interchanges.

The sixth objective is the principal one for building a flexible model of a diamond interchange. To provide believable research results, model validation becomes extremely important. If inferences about real system performance are to be obtained from the results of simulated system operation, the simulation model must be a reasonably valid representation of the real system. Some assurance of validity is provided by demonstrating that for those simulation model configurations of actual interchanges studied, the simulation produces results consistent with the known performance of these interchanges.

If the basic components of the model are properly composed, a large spectrum of possibilities is produced for investigation. Some of these possibilities are:

1. Comparative analysis of automatic, fixed time, and simple stop-sign control of the diamond interchange;
2. Design of automatic control equipment, including (a) detection placement studies, and (b) determination of control information required, such as count, velocity, following distance, as well as where and how often these data must be obtained;
3. Geometric design, such as (a) number of left- and right-turn lanes, both on the arterial street and on the off-ramps, (b) free right turns on red from extra lanes on the arterial street and the off-ramps, (c) partial cloverleaf where heavy flows conflict, and (d) placement of half diamonds every half mile instead of full diamonds every mile, i.e., left half at one location and right half at next location, upstream half at one location and downstream at next location, or diagonal half at each location; and
4. Automatic control of freeway egress and ingress through changeable message signs.

OVERVIEW

Three-Year Program

In this section a brief description of the total study is presented. The study consists of a primary effort dealing with the diamond interchange simulation and validation, and a secondary effort in which mathematical solutions will be sought to problems that not only relate to the diamond interchange simulation but are also of general interest in traffic flow research. Some theoretical work has already been completed on throughput at a saturated signalized intersection (1). The research program on the simulation and validation of the diamond interchange consists of the following four distinct efforts.

Model Construction—A complete description is formulated of the operation of a diamond interchange. The logic of the digital computer simulation model and the required input and output items of information are evolved from this description. The results are logic flow diagrams showing the overall operation of the interchange, its inputs, and its outputs.

To achieve the desired flexibility, a large number of parameters are included in the model. Some of these are (a) total cycle-length time of the fixed-time control signal system, (b) signal phasing, (c) length of the phases within a given cycle, (d) flow rates of vehicles into the system as well as their headway distribution, and (e) geometrical characteristics of the interchange, such as length of the acceleration lane for the merging of on-ramp vehicles into the freeway traffic stream, distance from the centerline of the freeway to the points of ingress and egress of the ramps at their junction with the surface street, etc. The model is also designed to permit incorporation of automatic control systems with relatively minor program modifications.

Development of the Computer Program—Once the logic flow diagrams describing the model are completed, they are coded into a computer program, and the program is debugged.

Validation—Tests are conducted to determine whether the computer simulation performs realistically. This is done by comparing the performance of selected interchanges with the performance of the simulation model when set to operate under the same conditions as the selected interchanges. It should be emphasized that the ultimate goal of the study is the generation of believable research results, and therefore, validation is essential.

Two different classes of problems are involved in the validation procedure. The first is that of making field measurements of the various inputs, outputs, and the several measures of effectiveness that are of interest in describing the performance of the interchange. These are essentially problems of designing instrumentation for data collection and devising methods for the reduction of such data. The second is concerned with determining whether or not there are significant differences between the performance of the actual interchange and the model for the same set of conditions. These are statistical problems and, owing to the peculiarities of traffic phenomena, require the development of techniques which are not available at the present time.

Research—Once a model of a diamond interchange has been developed which has realistically simulated specific situations, then some confidence can be had in the procedures used to move vehicles through the model. Thus, predictive studies can be conducted with some assurance that the conclusions reached are realistic.

Because of the extreme complexity of the total diamond interchange model, it has been divided into several smaller models, each of which will be considerably simpler to handle than the total model. Model 1, which consists of the merging of the freeway and the on-ramp, has already been formulated, coded, and debugged. Preliminary validation has also been initiated and results are presented later in this paper.

When all of the partial models of the total diamond interchange are complete and validated, they will be combined into a single whole model and total system validation will be tested. It is expected that this procedure will produce a valid total model much sooner and with fewer major difficulties than would be experienced if a total simulation were attempted from the beginning.

Validation of Simulation Model

As discussed previously, one of the principal objectives of this study is the validation of the procedures for moving the vehicles through the system. Some assurance of validity would be provided by a demonstration that, at least for some situations, the simulation produces results consistent with the known performance of the real system. To accomplish this comparison between the model and the real world, it is planned to observe the operation of at least two and perhaps three diamond interchanges in the Los Angeles area under various traffic situations. The collection of data and the various measures of effectiveness (or figures of merit) to be compared are discussed in the remainder of this section.

Collection of Data—An evaluation was made of the various techniques available for obtaining necessary field data. The two principal methods considered for ultimate use were (a) magnetic-loop vehicle detectors, and (b) aerial photography from a helicopter flying over the interchange. A third manual approach is presently being used for preliminary validation studies of version 1 of Model 1. These are discussed in the following.

Magnetic-Loop Vehicle Detectors—The proposed vehicle detector is of the loop type, where the presence of a metallic body, i.e., a vehicle within the loop, creates a phase shift of an oscillating circuit, which in turn produces a switch closure that can be monitored. The loops of these detectors would be taped to the road surface in pairs, their leading edges separated by a known distance, at various locations throughout the interchange as dictated by the design of the total model and its logic for moving vehicles through the system. The output of each vehicle detector would be sampled periodically by a multiplexor, whose output would be recorded by an instrumentation tape recorder. A timing track and synchronizing pulses would also be recorded and supplemented by verbal annotation. The tape-recorded data would be returned to the SDC computation laboratory where it would be played back and entered into the computer for reduction and processing. The information obtained for each vehicle would be presence, instantaneous velocity, length, and time of detection. The time headway between successive vehicles would also be determined. An experimental study has been completed on the feasibility of using this type of detector for these purposes (2).

Aerial Photography—Movies, taken from a helicopter flying over the interchange, were also considered as a source for collecting data on the performance and the conditions under which the system is operating. Special tape, in strips wide enough to be visible on the film, would be placed on the roadways to delineate specific geometrical sections of the interchange. The accuracy and types of information that can be extracted from the movies depend on the degree of sophistication of the equipment, i.e., the camera, lens, helicopter and helicopter mount, film, and film-reading equipment. Possible information obtainable would include vehicle presence, instantaneous velocity, length, travel time through each portion of the interchange, number of passes, and time and distance headways between successive vehicles. Some experimentation has already been performed with both 35 and 70-mm movie cameras mounted in a simple

two-axis gimbal in the helicopter. Also a survey has been completed of the equipment and data reduction techniques used by others employing aerial photography in vehicular traffic studies (3). Two automated film readers, the Information International Inc. Programmable Film Reader and the Philco automatic data reduction equipment, have been investigated to determine their feasibility for film data reduction. These studies and investigations will be reported in the near future.

Manual Observation—This procedure consists of manually counting vehicles as they pass various points in the interchange area and recording these counts on a periodic basis. It also includes a random sampling of travel times, which are determined with a stopwatch by following a vehicle between two distinctly marked points in the interchange area. The technique has been used in preliminary validation studies of Model 1 and is discussed in detail later.

It has been decided, based on the foregoing studies, to use photogrammetric techniques exclusively. The system will consist of a 70-mm moving picture camera mounted on a helicopter for collecting the data, and a conventional semiautomatic film reader for taking data off the film (automated film readers are still in their development stages for this application).

Data Analysis and Validation—The principal figure of merit used to validate the model is travel time by origin-destination. Travel times measured in the field are compared with those obtained from the computer simulation when operating under identical conditions. Other performance measures used include velocity, density, and number of passes.

The ultimate scheme used for validation in any specific situation consists of measuring, over a period of time, the arrival times and initial velocities of vehicles entering the system. These vehicles are followed and their total travel times through the system, travel times through various subsections of the system, velocities at various points, and number of passes are measured. The simulation model is then run by inserting vehicles into the system at the observed entrance times with the observed initial velocities. The program logic will then move the vehicles through the system and record the same items of information just mentioned. Next, field measurements are compared with the computer output to determine if the simulation model is behaving realistically.

The comparison of the field data with computer output must involve statistical procedures, since the flow of traffic is stochastic in nature. There are features about vehicular flow which preclude the use of existing statistical methods. Two of these are that a "steady state" may not exist in traffic flow and that observations are not independent of each other. Studies will be initiated to develop new statistical techniques or to develop methods for reducing the problem so that existing methods become applicable.

Cooperation of Governmental Agencies

The successful outcome of the validation studies depends on the support of governmental agencies of both the State of California and City of Los Angeles. Discussions were carried out with appropriate officials in both governmental bodies. The agencies have agreed to provide, at no cost to the project, the necessary crews for the installation of tape delineators as required by the computer model. The city will also provide both pilot and helicopter for the time required to obtain photographic data.

Active participation by the state has already taken place in conducting feasibility studies on the use of loop detectors in pairs for measuring traffic flow characteristics (2). The state also participated in delineating sections of an interchange with lane marking tape for photogrammetric technique evaluation study. Similarly, the city provided a pilot and helicopter for making movie-taking flights over diamond interchanges. Some of these flights were used to get 16-mm movies of traffic flow, with a hand-held camera, for qualitative use in formulating the model's computer logic. Other flights were made, using mounted 35 and 70-mm movie cameras, to determine the feasibility of getting quantitative measurements from film.

DESCRIPTION OF MODEL 1

Logic

Overall Design of Program—Since one of the goals of this simulation is to provide an easily changed model for laboratory use in the study of traffic flow on existing as well as proposed diamond interchanges, and since the stochastic nature of traffic requires either extensive replication of computer runs or one very long run for each set of parameter values, a substantial effort was undertaken to minimize the computer- to real-time ratio.

In general, computer-time to real-time ratios are inversely proportional to the programming effort, and can be reduced by either more sophisticated coding or by aggregation of vehicle movement within the model. The usual approach to traffic simulation in the past has been that of microscopic, step-by-step movement of traffic in which the location of each vehicle is determined at the end of each uniformly spaced time increment. This approach requires relatively simple program logic and coding; however, it also requires each movable component of the model (e.g., vehicles) to be checked and moved, if necessary, at relatively frequent intervals of simulated time. Time increments as small as 0.1 sec have been used for models of this type.

For this study, because of the large geographical area and correspondingly larger number of vehicles within the system at any given time, it was estimated that this approach would yield computer- to real-time ratios far in excess of unity, and the increment approach was discarded. Instead, an essentially macroscopic, significant-event simulation technique was chosen. In this approach, the model is broken down into several interconnected submodels of relatively uniform geometric configuration within which vehicles are aggregated. Individual vehicle characteristics, such as velocity, are determined only as vehicles cross the boundaries between submodels.

Figure 2 shows approximately one-quarter of a diamond interchange (Model 1), which is the subject of this study. Included are six submodels, as follows: one entrance ramp (V_{23}, V_{24} and V_{25}), the merging-weaving section (V_{26}), and a long section of uniform roadway prior to and after the weaving section (V_3 and V_{27}). The surface arterial that would normally be associated with this ramp is excluded.

To determine the time each vehicle crosses a submodel boundary, a logic was developed which considers passing, weaving, acceleration, deceleration, merging, and shock wave buildup and dissipation. Each of these characteristics is discussed in more detail later in this section on vehicle-driver performance characteristics and in the same section on inputs of vehicle-driver performance characteristics. Each time a significant event occurs, the logic changes the configuration of traffic in each submodel for the period between the significant event and the prior event.

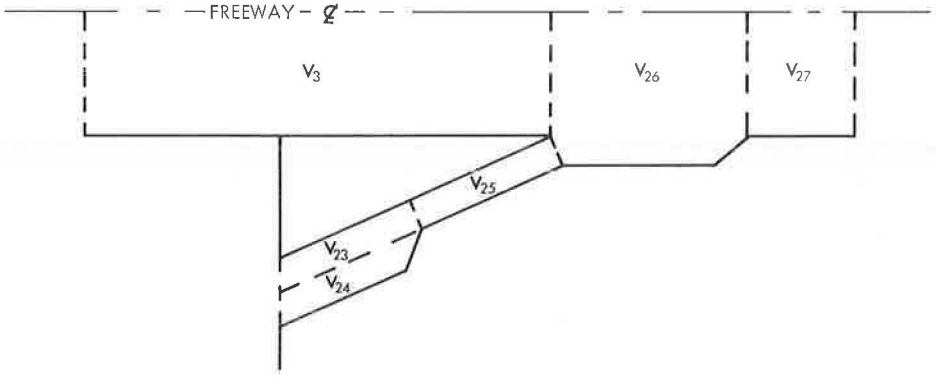


Figure 2. Diagram of Model 1.

The original design of the model defined and limited the significant events to be the phase-change times within the signal cycle. This proved unworkable due to the shortness of some of the submodels, for example, left-turn pockets and the one-lane portion of the entrance or exit ramps. For these short sections, vehicles could move through the entire submodel during a signal phase, thus causing severe logical problems in the computer determination of traffic configuration within the model at each significant event. As a result, it was decided to split each phase into several subphase intervals so that no vehicle, when moving under optimum conditions, can completely traverse any submodel during a subphase interval. The beginning of each interval is, in the present model, included with the phase-change times as the significant-event times.

With the inclusion of these subphase intervals, the model exhibits the incremental property of time-slicing simulation models; however, in actual use these intervals have been considerably longer than the time increments normally encountered in time-slicing models. The importance of interval length, or time increment, is illustrated later in this section, where a study of the effect of interval length on computer- to real-time ratios is described. This study illustrated the merit of significant-event simulation models in that substantial increases in total running time occurred as smaller subphase intervals were used. To insure ease of use of this model for any arbitrary set of diamond geometrics, the subphase interval length is variable and is determined by the computer, from the geometric data for the interchange being studied, at the time the simulation is performed.

Within each submodel, vehicle identity is maintained through the use of a vehicle storage table. This is necessary in order to assure the use of the same vehicle-driver characteristics for a given vehicle in all submodels. No attempt is made to identify the lane or longitudinal position of vehicles within the submodel.

The general procedure for moving vehicles through the model is as follows. Based on vehicle-driver characteristics, entrance time and speed, and existing density of traffic in the submodel to be entered, the desired exit time for an entering vehicle is calculated. This time is placed in a vehicle storage table and represents the earliest possible exit time for that vehicle. No additional calculations are made for a vehicle once it is in the vehicle storage table until its desired exit time becomes equal to or less than the cumulative simulated time. When this occurs, the vehicle either leaves the submodel and enters the next submodel, or it is delayed and receives a new departure time from the submodel within which it has been delayed.

As vehicles arrive at each submodel output boundary, or at the output boundary of the total model, actual traversal characteristics are recorded and in some cases stratified by origin-destinations. These traversal characteristics are used at the end of the simulation to compute various figures of merit of system performance and to compare to observed real-world data for purposes of validation. More details on validation and freeway system performance measures are discussed in this section and in the section on "Model Performance Studies."

Traffic inputs to the model are presently based on a truncated exponential distribution. In addition, a provision for inputting actual submodel entry velocity and time for each vehicle was included in order to facilitate model validation studies. Passing probability is determined from a segmented linear function of occupancy, and driver type selection is based on a normal distribution. In general, each sample of any of the stochastic variables used within the model is obtained by making an appropriate transformation of a random sample from the uniform distribution on the interval from 0 to 1. The procedure for accomplishing these transformations is included in the model in a general manner so that, as more data concerning actual distributions of driver characteristics and performance become available, the estimated distributions can be readily modified. Similarly all deterministic variables, such as length of roadway, number of lanes, signal phase and cycle length, etc., are easily changed to facilitate the study of other geometric shapes or configurations for alternative diamond interchanges. In the case of signal phase and cycle length, the possibility of automatic control and its concomitant real-time alteration of phase sequence and composition can easily be provided for in the overall logic.

Provisions for the dynamic variation of vehicle input rates and techniques for the elimination of transient effects due to beginning the simulation with no vehicles in the model have also been included in this version of the model. Stability runs were made and some results are included in the section on "Model Performance Studies." When input rates near capacity are used, the time for the figures of merit to reach a stable value approaches infinity.

Submodels—There are six submodels constituting the total model for simulating this quadrant of the diamond interchange (Fig. 2). These submodels were chosen for their geometric and logical uniformity. Within each submodel a distinct logical aggregate technique was developed which takes into account, in addition to the characteristics of that submodel, the characteristics of the preceding and following submodels, including their traffic density, vehicle composition, and behavior.

Within each submodel, analytical and stochastic techniques control traffic flow, including passing, weaving, acceleration, deceleration, exit velocity, travel time, following distance or headway, exit lane (where applicable), and exit time. For the freeway submodels, vehicles are not associated with particular lanes or locations within the submodel, but are considered to be aggregated into a statistically controlled mass within which each vehicle ultimately exits from any available exit lane.

Submodels V_3 and V_{27} are multilane sections of arbitrary length and number of lanes. Submodel V_{26} is a merging-weaving section of arbitrary length and contains one lane more than V_3 and V_{27}; V_{26} has five inputs and four outputs when simulating a four-lane freeway; V_{23} and V_{24} are both single-lane sections and constitute the individual lanes of the two-lane section of the on-ramp. These sections end at the point where it is no longer possible for two cars to be adjacent to each other. Submodel V_{25} is the one-lane section that interconnects V_{23} and V_{24} with the merging-weaving submodel, V_{26}. This submodel has two inputs and one output. Submodels V_{23} through V_{25} are all of arbitrary length and do not allow passing of vehicles within themselves. Vehicles in V_{23} and V_{24} are allowed to pass vehicles in their respective adjacent lanes, V_{24} and V_{23}. Characteristics such as grade, vertical and horizontal curvature, sight distance, etc., are not explicitly included in the model, but are implicit in the driver characteristic functions, which are distinct for each submodel.

In the case of submodel V_{26}, vehicles are entered into the vehicle storage table in order of their arrival at their corresponding input boundary until a congestion situation occurs. At this time vehicles begin to accumulate in V_{26} and speeds begin to drop at the output of V_{26}. Eventually the speed in V_{26} reduces to some minimum value and density increases to its maximum possible level (saturation density). The time required for this to occur is a function of the total input volume and the capacity at the constraining portion of the weaving-merging submodel.

When V_{26} reaches saturation density, the on-ramp begins to share with the freeway all available space for vehicles to enter V_{26} on a one-car-out-of-eight basis. In effect, this sharing procedure divides the available space of the right lane of the freeway evenly between the right-lane freeway cars and ramp cars. If the ramp traffic does not require this level of service, the excess space is reallocated to freeway vehicles. This method of operation continues until freeway plus ramp volume decreases to an extent such that congestion begins to dissipate.

Vehicle-Driver Performance Characteristics—The multiplicity of possible vehicle-driver performance characteristics and the interactions among them, of necessity, forced a selection of the few characteristics that were expected to have the greatest effect on system performance. Those selected were velocity, acceleration, deceleration, passing probabilities, final system and submodel velocities, vehicle length, and headway. This initial set of vehicle-driver performance characteristics is intended to be broad enough to aggregate some of the secondary effects such as man-machine-roadway interactions, vertical and horizontal curvature and alignments, grade and super-elevations, and traffic composition by type, purpose of trip, and experience of driver.

Sensitivity analyses of model behavior as a function of some of these characteristics were made after model validation was completed. These studies are reported in the

section on "Model Performance Studies." As additional sensitivity and validation studies are completed, it will undoubtedly become necessary to alter, refine, or add model characteristics so that more explicit inclusion of secondary traffic flow effects is provided.

Vehicle-driver types were grouped into five categories. These categories were then assigned typical free-flow (no congestion) desired velocities and accelerations. During simulation runs, vehicles attempted to achieve their desired free-flow velocity, and once there, to maintain it throughout the traversal of the model. Acceleration is generally accomplished at a rate determined by vehicle-driver type; however, in some situations, when traffic conflicts exist, lower values are used. Vehicles that cannot pass slower vehicles are slowed down to the speed of the impeding vehicle until a passing maneuver is possible. Here, as before, the driver-type acceleration is usually used to decelerate, but it can be increased or decreased if warranted by traffic conditions.

To provide for vehicle passing within the model, an algorithm containing a stochastic element was developed. This algorithm decreases the likelihood of passing as a function of the number of vehicles (occupancy) in the submodel. The detailed procedure is as follows. A minimum desired exit time is computed for each vehicle entering the submodel and is compared sequentially to the desired exit times of the vehicles already within that submodel, in inverse order of their desired exit times. If the minimum desired exit time of the entering vehicle is less than the desired exit time to which it is being compared, a passing maneuver is indicated and a passing probability, based on submodel occupancy, is used to decide whether that passing maneuver will take place.

The first failure to pass ends the process and the entering vehicle is placed in the vehicle storage table for that submodel with a desired exit time equal to that of the vehicle which could not be passed. If all desired passes are accomplished, the vehicle is placed in the table directly in front of the last vehicle passed and with its minimum desired exit time becoming its desired exit time from that submodel. The passing probability function of submodel occupancy is discussed in more detail later in this section.

Exit velocities from each submodel, and thus from the system, are analytically determined from vehicle-driver-type characteristics, traffic density, and submodel length. The basic exit velocity algorithm attempts to allow each vehicle to exit from any given submodel at the maximum desired velocity for the vehicle-driver type assigned to that vehicle. This is accomplished by computing

$$\text{est } V_f = \min [(V_o + at), \max V_f]$$

where

$\text{est } V_f$ = estimated exit velocity from the submodel,

$\max V_f$ = maximum desired free-flow exit velocity for the assigned vehicle-driver type,

V_o = actual entrance velocity to the submodel,

a = vehicle-driver-type acceleration constant for this driver in this submodel, and

t = actual time spent in the submodel.

This estimate is used as the actual exit velocity if the vehicle has not experienced any delay from its originally computed minimum desired exit time, and if the vehicle is not approaching a shock wave. If delay has been experienced, the time difference between the exiting vehicle and the immediately preceding vehicle is computed. In this case, if this time is large enough so that no potential conflict exists with the preceding vehicle, the vehicle is allowed to exit at its average submodel velocity. However, if the time difference is small the exit velocity is constrained by the exit velocity of the immediately preceding vehicle, and the exit velocity is set equal to the lower of the

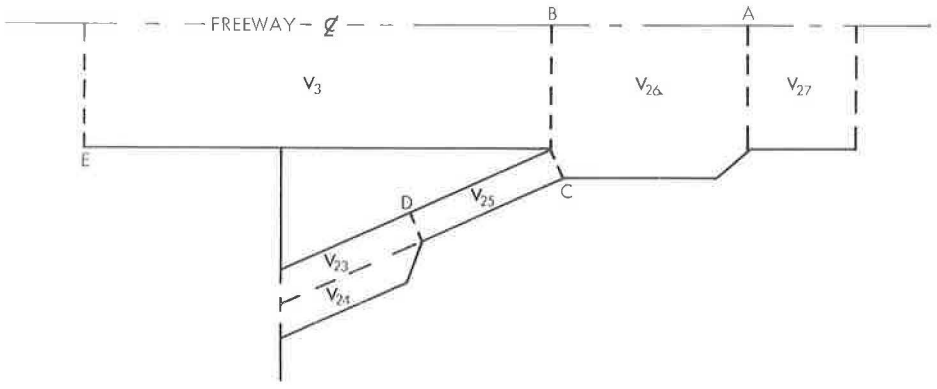


Figure 3. Diagram of Model 1 showing shock wave path.

following two values: the constraining value of the preceding vehicle, or the average submodel velocity for the exiting vehicle.

Whenever a shock wave exists near the exit boundary of a submodel, all vehicles crossing that boundary are subjected to a further exit velocity constraint. For each submodel there is a unique function from which this constraint is obtained based on the distance between the exit boundary and the shock waves at the instant the vehicle crosses the boundary. These values, which range between maximum free-flow velocity and minimum congestion velocity, are used as the exit velocities whenever they are less than any of the other constraining exit velocities.

Shock Waves and Congestion—Because the phenomena of shock wave movement and congestion buildup and dissipation are rather complex, several unique flow situations occur when such waves pass through or begin at a freeway entrance ramp. Congestion-producing shock waves occur as a result of either downstream bottleneck (resulting from some constriction either temporary, such as an accident, or permanent, such as a design fault) or from too high a total traffic flow demand at the output boundary of the ramp merging-weaving section. In both cases congestion buildup occurs in the same manner; it may not, however, occur at the same rate. On the other hand, congestion dissipation can occur in several different ways.

Figure 3 shows that a shock wave may either begin or pass through boundary A, move toward boundaries B and C, and then to boundaries D and E. The rate at which it moves in submodel V_{26} is a function of the output flow rate at A and the input flow rates at B and C. If a bottleneck exists downstream from boundary A, so that the total flow into submodel V_{27} is greater than the flow through the bottleneck, a shock wave will begin to move back from the constriction toward boundary A at a rate proportional to the difference between input rate to V_{27} and output rate at the constriction. In submodel V_{26}, the rate of shock wave movement will be somewhat lower because of the vehicle storage capacity of the extra acceleration lane. Once the shock wave has proceeded to boundaries B and C, the rates from B to E and C to D are functions of the respective volumes of the freeway and ramp and the allocation to them of the available entrances to submodel V_{26}. If the entrance ramp vehicles are assigned a substantial proportion of the use of the right-hand land of V_{27}, it is possible to have a situation in which no congestion and hence no shock wave develops in V_{25}.

Once a congestion-producing shock wave has started moving through the system, there are three sets of conditions that will cause it to disappear. These conditions are referred to as shock wave disappearances of Types 1, 2 and 3, and are described as follows.

Type 1—The condition that caused the shock wave to occur can be removed, in which case the highly congested, slow-moving queue of vehicles begins to accelerate and spread

out, thereby creating a second, but in this case congestion-dissipating, shock wave to start back through the system. If this congestion-dissipating shock wave overtakes the congestion-producing shock wave, both will simultaneously disappear. Otherwise, they will eventually leave the simulated portion of the system and normal flow conditions will exist within the model.

Type 2—The shock wave may either become stationary or return toward the causing constraint. This will occur when the input volumes become equal to, or drop below, the value allocated to their particular section of roadway as a result of the flow constraint.

Type 3—This type of disappearance is a combination of the first two. Here both the input flow and original congestion-producing constraints are changed to the extent that (a) the original congestion-producing shock wave reverses direction and returns toward its source, and (b) a congestion-dissipating shock wave begins to move from the point of the removed constraint.

For these shock wave disappearances, there are several different situations of congestion dispersal and shock wave movement which can occur as a function of the way input flow rates change. The present model, for simplicity, has been limited to Type 2 shock wave disappearance. As a result, several compromises were necessary for some of the shock wave disappearance situations. Version 2 of the model, presently being developed, will include all types of shock wave disappearance and will therefore be capable of considering all situations of congestion dispersal and shock wave movement without compromise.

Generation of Vehicle-Driver Performance Characteristics—As mentioned previously, vehicle and roadway characteristics have been separated into two groups, one of which is treated stochastically and the other deterministically. Several characteristics that ordinarily might have been treated stochastically are included in the deterministic group to simplify the implementation of version 1 of the model. Sensitivity studies will be run to determine the effect of this simplification on the overall model performance and, if indicated, some of the deterministically treated characteristics will be returned to the stochastically treated group. At present, the only stochastic elements in the model are vehicle inter-arrival times, passing, vehicle-driver type, acceleration, desired maximum velocity, and desired ramp entrance velocity.

To obtain consistent vehicle-driver performance characteristics for a given vehicle throughout the system, the vehicle is stochastically assigned a vehicle-driver type as it enters the system. This vehicle-driver type determines the values of other vehicle-driver characteristics, such as acceleration or velocity, each time they are required. Thus, in the true sense, vehicle-driver type is the stochastic variable which selects deterministic values for some of the other vehicle-driver performance characteristics.

In this model, the general procedure followed to obtain a value of a stochastic variable is to transform, by an appropriate function, a random sample from the uniform distribution in the interval from 0 to 1. These transforming functions are, in the case of vehicle-driver type, a normal distribution, and for intervehicle time gap, a truncated exponential distribution. The stochastic phenomenon of passing is accomplished by determining the probability of a successful pass from a segmented linear function of submodel occupancy. A detailed discussion of the passing probability function and the distribution function for vehicle-driver-type characteristics is presented later in this section.

Generation of Intervehicle Time Gaps—The generation of intervehicle gaps for the present model is accomplished by using a truncated exponential density function. Here the underlying assumption is that the roadways leading to the interchange are sufficiently long and uniform, so that traffic entering the model is Poisson-distributed. Truncation of the resulting exponential density function was indicated because of the occasional occurrence of arbitrarily large intervehicle gaps when the usual exponential vehicle generator algorithms were used. The long gaps are compensated for by the frequent occurrence of extremely short intervehicle gaps; however, this requires very large simulation runs before stable system performance measures can be expected. It was therefore decided to prevent the generation of extremely long intervehicle gaps by using the truncated exponential density function to generate intervehicle gaps.

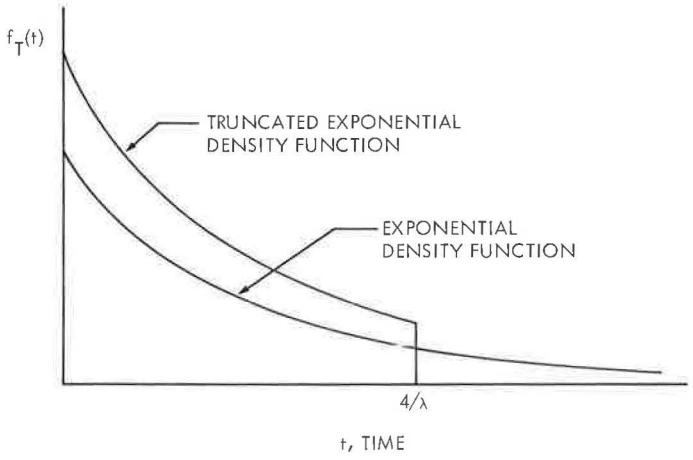


Figure 4. Density functions for generation of intervehicle gaps.

It was arbitrarily decided to truncate the exponential at the value of intervehicle gap corresponding to the expected value (mean intervehicle gap) plus three standard deviations. Thus, in the vehicle generation algorithm, vehicles are not generated with intervehicle gaps greater than this value, and thereby a situation is created in which more than the specified expected number of vehicles may be generated. To overcome this problem, it is necessary to determine the parameter values for the truncated exponential distribution in such a way that the specified expected value of intervehicle gap will result. Figure 4 shows both the truncated and non-truncated exponential density functions. It should be noted that no attempt has been made to eliminate the generation of intervehicle gaps that are impossibly short. This simplification is permissible because the logic of the model prevents any vehicles from being closer than a specified minimum time gap at any boundary of any submodel.

The exponential density function shown in Figure 4 is given by

$$f_T(t) = \lambda e^{-\lambda t}, \quad t \geq 0$$

where λ is the rate of vehicle generation per unit of time, and t is the intervehicle time gap. The cumulative distribution

$$F_T(t) = \int_0^t \lambda e^{-\lambda \xi} d\xi = 1-e^{-\lambda t}$$

is the probability of an intervehicle gap $\leq$ t. To generate intervehicle gaps with this distribution it is necessary to solve for t in the equation

$$r = F_T(t) = 1-e^{-\lambda t}$$

where r is a random sample from the uniform distribution in the interval between 0 and 1. The solution is

$$t = -\frac{1}{\lambda} \ln (1-r)$$

Since (1-r) is uniformly distributed, t can be obtained by solving $t = -1/\lambda \ln r$. For this density function the expected value (mean intervehicle gap) is

$$E[T] = \int_0^\infty \xi\lambda e^{-\lambda\xi} d\xi = \frac{1}{\lambda}$$

and the corresponding value of the standard deviation is

$$\sigma = \left[\int_0^\infty \xi^2\lambda e^{-\lambda\xi} d\xi - \left(\frac{1}{\lambda}\right)^2\right]^{1/2} = \frac{1}{\lambda}$$

The truncation value of intervehicle gap is therefore $(1/\lambda) + 3(1/\lambda) = 4/\lambda$.

As previously mentioned, this truncation affects the number of vehicles actually generated in such a way that for a given λ too many vehicles enter into the system. To overcome this effect, the value of λ used in the generating function must be adjusted so that the generating function produces the desired expected number of vehicles per hour. This adjusted value of λ' is the expected value (mean) of the truncated exponential density function.

To obtain the truncated exponential density function, it is necessary to normalize the exponential function between 0 and $4/\lambda'$. The normalizing factor is

$$\int_0^{4/\lambda'} \lambda' e^{-\lambda'\xi} d\xi = 1-e^{-4}$$

Thus the truncated exponential density function becomes

$$f_{T'}(t) = \frac{\lambda' e^{-\lambda' t}}{1-e^{-4}}, \quad 0 \le t \le \frac{4}{\lambda'}$$

The expected value (mean) of this random variable is

$$E[T'] = \int_0^{4/\lambda'} \xi \frac{\lambda' e^{-\lambda'\xi}}{1-e^{-4}} d\xi = \frac{1}{\lambda'} \frac{1-5e^{-4}}{1-e^{-4}}$$

To generate vehicles at a flow rate of λ vph, the appropriate value of λ' to use in the generating functions is

$$E[T'] = \frac{1}{\lambda} = \frac{0.924}{\lambda'} \text{ or } \lambda' = 0.924\lambda$$

Figures of Merit—Figures of merit, sometimes referred to as measures of effectiveness, are used in this model for both validation and research purposes. For validation studies, measures that are readily obtainable in supporting field studies are chosen for determination of equivalence between the simulated system and actual system and actual system performance. On the other hand, for research purposes, measures are chosen that indicate system efficiency and uniformity of traffic flow.

In designing this model several different measures were considered. Some of these have already been implemented and provision has been made for the ultimate inclusion

of the others. The following figures of merit either are implemented or can be implemented with relative ease in the present version of the simulation: (a) individual and average vehicle travel time, from origin to destination, by vehicle-driver type and by origin-destination; (b) individual and average submodel travel time by vehicle-driver type and by origin-destination; (c) proportion of vehicles leaving submodels at speeds less than their desired speeds for any given time increment by vehicle-driver type and by origin-destination; (d) number of vehicles in each submodel at any given time by vehicle-driver type and by origin-destination; (e) number of successful and unsuccessful pass maneuvers in each submodel during any arbitrary time increment by vehicle-driver type and by origin-destination; (f) time difference between minimum system or submodel travel time and actual travel time by vehicle-driver type and by origin-destination; (g) average velocity in each submodel by vehicle-driver type and by origin-destination; and (h) individual and average velocity at boundaries between each submodel by vehicle-driver type and by origin-destination.

At present, only a portion of (a) has been used to validate the model; namely, the average travel time by origin-destination. This study is reported on in the next section. The field data collected to date have been gathered using the manual statistical sampling procedures described in the previous section. Portions of figures of merit (a), (b), (d) and (h) are being used in the validation studies presently under way. For these studies, additional data are being obtained through the photographic techniques previously described.

Model Characteristics, Inputs, and Background Supporting Estimates of the Parameters

Geometrics—Several typical diamond interchanges in the Los Angeles area were studied in order to insure that the version 1, Model 1 simulation incorporated adequate storage for vehicles, adequate and reasonable maximum and minimum length for the various submodels, and that it provided a basis for estimating computer scale factors for the included measures of system performance. For the preliminary validation study, the interchange of the San Diego Freeway and Roscoe Boulevard was selected principally because a nearby tall structure provided a convenient location for performing manual data acquisition.

A 70-mm picture of the entire San Diego Freeway-Roscoe Boulevard interchange is shown in Figure 5. The southwest quarter, which is the portion of maximum traffic during the morning peak traffic period, was chosen for the present simulation. The number of lanes and the specific values of lane length for each of the submodels of Figure 3 are given in Table 1. Other factors such as curvature, alignment, grade, and lane width, which have not been used explicitly in the model, have been implicitly included in the estimates of vehicle-driver characteristics described in the next section.

The precise location of the boundaries of each submodel are, whenever possible, established by some easily identified geometric features of the roadway. Specifically, V_{23} and V_{24} are measured from the surface street curb line to the point where the total width of V_{23} and V_{24} is reduced to approximately 20 ft. At this point, referred to as boundary D, submodel V_{25} begins; it then continues to the point of the gore (boundary C) separating the freeway lanes from the auxiliary lane of the weaving-merging submodel, V_{26}. The freeway submodel V_3 also terminates at this point (boundary B). Submodel V_3 begins at the point of the gore (boundary E) separating the freeway from the off-ramp (Fig. 1). Submodel V_{26} begins at boundaries B and C and continues until the auxiliary lane is approximately 8 ft wide (boundary A). From this point to any easily identified, nearby geometric features of the freeway (e.g., bridge

TABLE 1

LANES AND VALUES OF LANE LENGTH OF MODEL 1

Submodel	Length	Lanes
3	2210	4
23	506	1
24	506	1
25	590	1
26	360	5
27	175	4

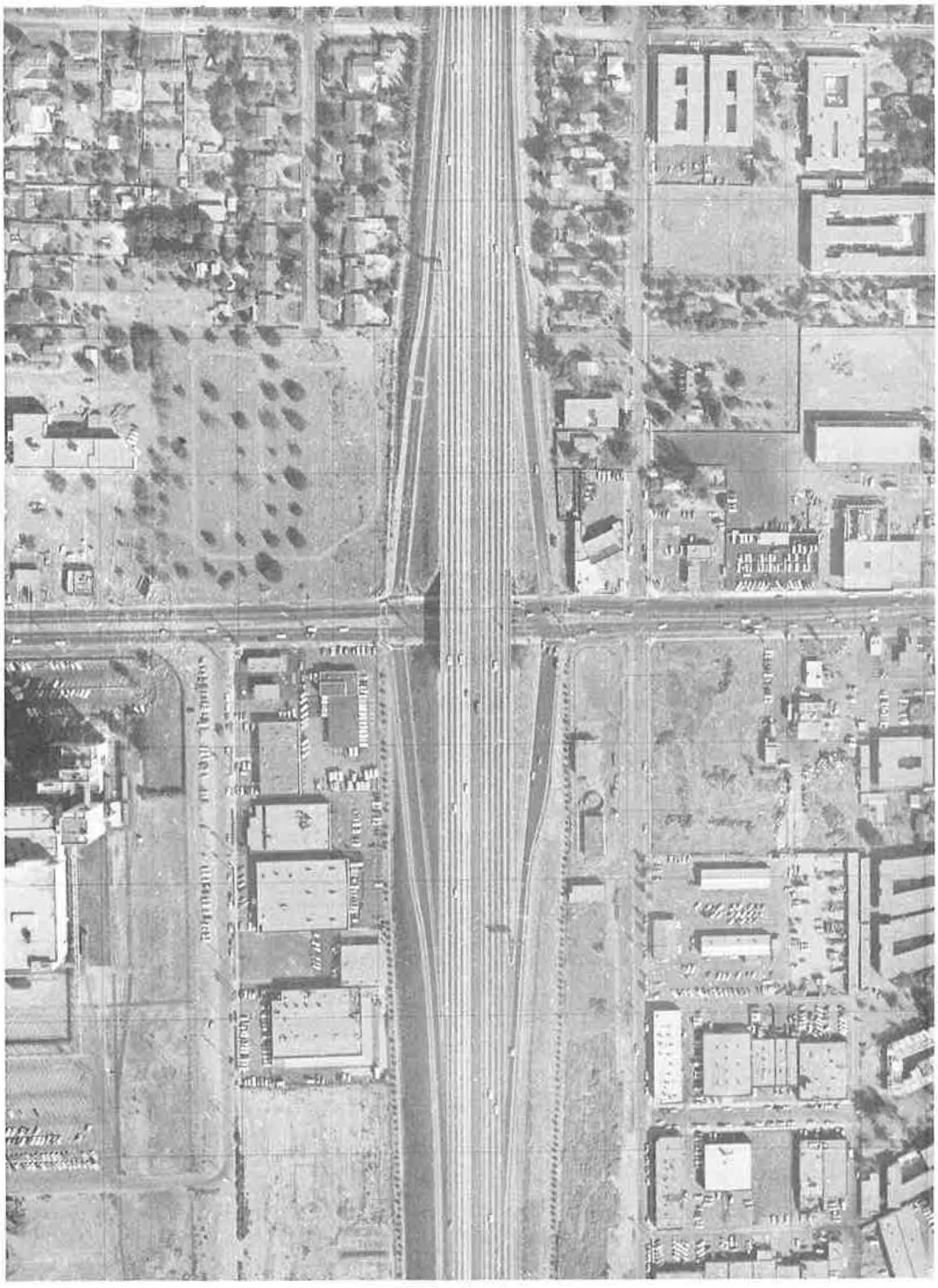


Figure 5. Roscoe Blvd.-San Diego Freeway diamond interchange.

TABLE 2
MAXIMUM DENSITY AND MAXIMUM OCCUPANCY

Submodel	Max. Occupancy	Max. Density
3	401	959
23	23	240
24	23	240
25	26	233
26	81	1188
27	31	935

abutment, light standard, sign, etc.) constitutes the length of submodel V_{27}. The choice of boundaries A and D is somewhat arbitrary; however, the rule used to locate them is based on the assumption that two vehicles will not remain side by side when the total roadway width available is reduced to 20 ft. Thus, the auxiliary lane is deemed to have ended (boundary A) when its width added to the standard freeway lane width of 12 ft produces a total width of 20 ft. Similarly, the two-lane section of on-ramp is deemed to have ended (boundary D) when the combined roadway width reduces to 20 ft.

Inputs of Vehicle-Driver Performance Characteristics—The vehicle-driver performance characteristics included in version 1, Model 1, were discussed in detail previously in this section. Specifically, they are (a) vehicle-driver type characteristics of velocity, acceleration, and deceleration; (b) passing probabilities; (c) final system and submodel velocities; (d) density; and (e) vehicle length and headway. Because it was decided to stratify drivers into five distinct vehicle-driver categories, the characteristics of desired velocity, acceleration, and deceleration are similarly stratified.

In the following discussion and for the balance of this report the terms free-flow densities, maximum density, and critical density are used as follows:

1. Free-flow densities exist when a given vehicle is able to travel at its desired velocity throughout a submodel. For vehicles with low desired velocities, free-flow operation will exit at considerably higher values of submodel density than for vehicles with high desired velocities.
2. Critical density occurs when all vehicles are traveling at approximately the desired free-flow velocity of the slowest vehicle-driver type.
3. Maximum density exists when vehicles are stalled or moving very slowly in a large group.

To simplify the model for preliminary testing, vehicle-driver performance characteristics of headway and length were chosen as constants. The minimum time-headway was set at 2 sec, and thus maximum freeway lane flow is 1800 vph. Vehicle length (22 ft) is defined as the space occupied by one vehicle when stopped in a queue, including the intervehicle gap between it and the next following vehicle. This value is the sum of an average actual vehicle length of 17 ft and an average intervehicle gap of 5 ft. From this, the value of maximum density for any submodel is given by

$$\left(\frac{M}{L}\right) 5280 \text{ veh/mi}$$

where

M = max occupancy in a given submodel = $\frac{Ln}{22}$,

L = length of submodel, and

n = number of lanes in submodel.

Table 2 gives the maximum density and the occupancy required to achieve maximum density for each submodel as computed with the foregoing equation and the values of Table 1.

It is planned to run a sensitivity analysis of these characteristics to determine their effect on the model. If it is found that uniform vehicle lengths and headways create questionable simulation results, stochastic processes will be implemented to vary these factors statistically.

TABLE 3

VEHICLE-DRIVER CHARACTERISTICS

Distribution of Vehicle-Driver Types (%)	Freeway Free-Flow Velocity (fps)	Freeway Acceleration and Deceleration (fps/sec)	On-Ramp Initial Velocity (fps)	On-Ramp Acceleration and Deceleration (fps/sec)
Slow = 7	75	3.0	6	3.0
Medium = 24	85	4.5	11	3.5
Normal = 38	95	6.0	16	4.0
Fast = 24	104	7.5	21	4.5
Very fast = 7	114	9.0	26	5.0

Using the maximum lane flow of 1800 vph with the vehicle generator described earlier in this section, an expected intervehicle gap of 2 sec with a maximum gap truncation value of 8.65 sec is obtained. Thus, no vehicle will be generated with an intervehicle gap greater than 8.65 sec.

The vehicle-driver characteristics of desired velocity, acceleration, and deceleration were each given five different values, one for each of the five vehicle-driver types of the model. These values (Table 3) were obtained, assuming a normal distribution for each characteristic, by establishing intervals for each driver type and computing the conditional mean of each such interval. These intervals were measured from the mean in standard deviation increments as follows: $(-\infty, -\frac{3}{2}\sigma)$, $(-\frac{3}{2}\sigma, -\frac{1}{2}\sigma)$, $(-\frac{1}{2}\sigma, \frac{1}{2}\sigma)$, $(\frac{1}{2}\sigma, \frac{3}{2}\sigma)$, and $(\frac{3}{2}\sigma, \infty)$. The normality assumption was experimentally justified, for the characteristic of desired velocity, from free-flow velocity data collected for 225 vehicles in the vicinity of the Roscoe Boulevard-San Diego Freeway diamond interchange. The χ^2 test for normality yielded a significance probability of 0.08. Thus, normality appears to be a reasonable assumption for desired velocity. At this time, no data have been obtained to support the normality assumption for free-flow acceleration and deceleration.

Table 4 gives the values of mean and variance used in obtaining the values in Table 3. With the exception of the experimentally obtained estimates of mean and variance for freeway free-flow velocity, all other values are estimates based on the judgment of the investigators. The mean and variance for free-flow velocity are estimates based on 225 observations. Additional data must be collected to justify or modify the remainder of the numbers.

Using the slow vehicle-driver type free-flow velocity of 75 feet per second and the minimum intervehicle gap of 2 sec, critical occupancy is

$$\frac{\text{Ln}}{(2)\,(75)}$$

TABLE 4

VALUES OF MEAN AND VARIANCE

Value	Velocities		Acceleration (Deceleration)	
	Freeway Free-Flow (fps)	Ramp Entrance (fps)	Freeway (fps/sec)	Ramp (fps/sec)
Mean	95	16	6	4
Variance	10.2	5	1.5	0.5

TABLE 5
VALUES OF MAXIMUM AND CRITICAL OCCUPANCY

Occupancy	Submodels					
	V_3	V_{23}	V_{24}	V_{25}	V_{26}	V_{27}
Maximum	401	23	23	26	81	31
Critical	58	3	3	3	12	4

where, again, L is the submodel length and n is the number of lanes in the submodel. These values, given in Table 5 together with the maximum occupancies of Table 2, are generally 15 percent of the maximum occupancy.

The passing probability function (Fig. 6), as it is presently included in the model, is based on the assumption that under very light traffic vehicles travel through each submodel of the freeway at nearly their desired speed, and passing occurs with probability near unity whenever a desired passing situation exists. Further, as occupancy increases, passing probability begins to fall off rapidly until, at critical occupancy, relatively little passing occurs. The segmented linear function, with zero and unity passing probability occurring at 0 and 100 percent occupancies, is provided with six breakpoints in the interval between 0 and 100 percent occupancy. These are at the 5, 10, 15, 20, 25 and 50 percent points. At these points the passing probabilities were arbitrarily set at 0.70, 0.40, 0.10, 0.08, 0.05 and 0.01 respectively. Because of the manner in which this function has been included in the model, it can easily be changed to any other arbitrary or observed passing probability function. It is planned to attempt to generate an actual passing probability function from the photogrammetric traffic flow data that will be collected as a subsequent task of the overall simulation study.

Submodel exit velocities are normally assumed to be equal to the freeway free-flow velocity for the vehicle-driver type assigned to that vehicle whenever traffic congestion does not interfere. Several situations can occur to prevent a vehicle from exiting at its desired free-flow velocity—shock wave formation, close proximity to a slow-moving vehicle, or the merging phenomena.

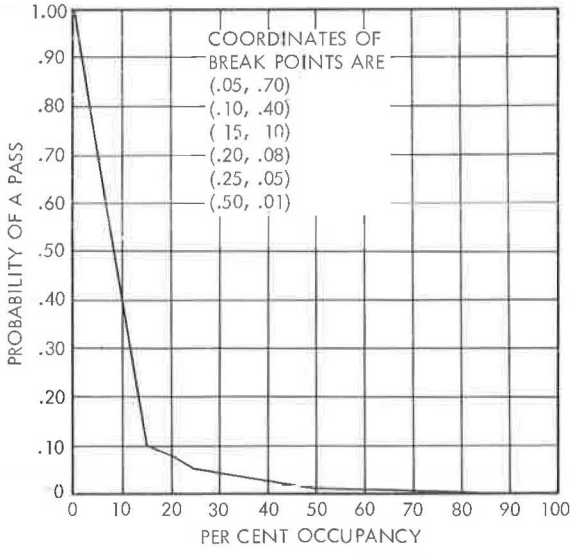


Figure 6. Probability of a pass as a function of percent occupancy.

TABLE 6

SHOCK WAVE FINAL VELOCITY CONSTRAINTS

Submodel							
V_{26}							
Occupancy of V_{26}	12	17	22	27	32	37	42
Final velocity (fps)	75	64	53	42	31	20	9
V_3							
Occupancy of V_{26}	50	55	60	65	69	73	77
Final velocity	75	64	53	42	31	21	10
V_{25}							
Occupancy of V_{26}	50	55	60	65	69	73	77
Final velocity	72	61	50	39	28	17	6
V_{23}, V_{24}							
Occupancy of V_{25}	13	15	17	19	21	23	25
Final velocity	63	53	43	33	23	13	3
Generator to V_3							
Occupancy of V_3	370	375	380	385	390	395	400
Final velocity	75	64	53	42	31	21	10
Generator to V_{23}, V_{24}							
Occupancy of 23 or 24	10	12	14	16	18	20	22
Final velocity	24	20	17	13	10	6	3

The normal freeway desired free-flow velocities are used as the exit velocity limit for any vehicle that is adequately spaced from the immediately preceding vehicle and from any shock wave that may be moving upstream. The preceding vehicle is not considered to have an effect on the vehicle of interest if the separation between the two vehicles is at least 4 sec. This value is a compromise between the 5 sec required for safe deceleration from maximum free-flow freeway velocity (114 fps) to minimum freeway free-flow velocity (75 fps), and the reasonable time-headway for vehicles traveling at the same speed (2 sec at 75 fps).

Shock wave final velocity constraints are given in Table 6. These values have been arbitrarily chosen based on: (a) length of each submodel; (b) the occupancy of the submodel at minimum spacing and slowest speed (set equal to maximum occupancy or 1 vehicle every 22 ft in each lane); (c) minimum freeway free-flow velocity (75 fps on the freeway); (d) average speed to traverse submodel V_{26} at maximum density assuming a 2-sec headway at its exit boundary, and (e) proportion of available space allocated to the various competing flows entering V_{26} and V_{25}. The minimum exit velocity for any submodel, except V_{26}, occurs when the density of the preceding submodel is close to its maximum. Since the common boundary of V_{26} and V_{27} is the point at which congestion and shock waves are produced, the exit velocity constraint of V_{26} is determined only by V_{26} occupancy. In this instance minimum exit velocity occurs at an occupancy somewhat above critical occupancy. For the present model, minimum submodel congestion exit speeds range from approximately 3 to 10 fps.

Figure 3 shows that as a shock wave reaches boundaries B and C, the allocation of the right-hand lane is split evenly between the freeway and the entrance ramp. Unless the ramp is supplying traffic at volumes above $\frac{1}{2}$ of the maximum allowable freeway lane flow rate, no shock wave will occur in V_{25}. The shock wave will, however, always move through V_3 whenever the allowed ramp plus freeway volume exceeds the maximum freeway rate. For the present values of model parameters, and for a shock wave in V_{26} resulting from a combined input from the freeway and the ramp in excess of 7200 vph, a shock wave will develop in V_{25} only if the ramp volume exceeds 900 vph. On the other hand, if a downstream constriction is causing the shock wave in V_{26}, then a V_{25} shock wave will occur for ramp volumes exceeding $\frac{1}{8}$ of the downstream flow constraint. For the cases in which no ramp shock wave occurs, the minimum submodel exit speeds provided by the model are somewhat lower than expected. Similarly, they are somewhat higher for cases in which downstream constraints reduce the boundary A exit rate below 7200 vph. This behavior is not considered significant in view of the gross nature of the shock wave algorithm as presently constituted.

In the present model, there are three distinct Type 2 shock wave disappearance and congestion dispersal situations, any one of which might prevail. For simplicity, some compromise of real behavior has been made in this part of the version 1 model. The first situation occurs if the ramp volume drops to a value such that, when added to the freeway demand, there are less than 7200 vph entering V_{26}. Here shock wave reversal takes place on the freeway and the wave moves back through V_3 toward B and then to A. The second situation is similar, except that in this case the freeway volume drops, causing shock wave reversal, and ultimately providing additional capacity for ramp traffic entering V_{26} at C. Third, the total freeway-plus-ramp flow may remain above 7200, but the split between flow on the freeway and flow on the ramp may change so that congestion can dissipate on either the freeway or the ramp, but not on both. For the parameter values of the present model, this would occur on the ramp if the ramp volume dropped below 900 vph while the freeway volume remained above 7200 minus the ramp volume; it would occur on the freeway if the ramp volume remained above 900 vph while the freeway volume dropped below 6300 vph.

As a result of limiting shock wave disappearance to the Type 2 compromise, the following unrealistic situation can occur. Congestion-dissipating shock waves, which normally occur when a constraint is suddenly removed (e.g., a wrecked vehicle is towed away), can also occur at certain times during the Type 2 disappearance situations. In particular, if the average ramp volume suddenly drops substantially below the congestion value (900 for present parameter values), the ramp will eventually become nearly empty. Excess capacity will exist at boundary A when this occurs, and as a result a congestion-dissipating shock wave should start at boundary A and move toward boundary B. Therefore, this excess capacity should not be allocated to increased freeway flow, but should be used to dissipate freeway congestion and increase freeway speed. Similarly, a decrease in average freeway volume should ultimately cause a congestion-dissipating shock wave to move down the ramp.

This lack of realism is not a problem when the model is used for the purpose of preliminary validation of the conditions which produce congested traffic flow or for validation of free-flow traffic operations. Since it would require a substantial addition of logic to the model to include congestion-dissipating shock wave properly, it was considered advisable to defer its inclusion and thus its validation to version 2 of Model 1. Version 2, presently being programmed, will be substantially more realistic in its treatment of submodel final velocity determination and will provide a complete treatment of the shock wave phenomena.

Signal Phase and Cycle Times—The present simulation model does not include any submodels which are directly influenced by signal control characteristics. However, in order to provide a method for pulsing the traffic inputs to the on-ramp submodels in a manner analogous to actual traffic situations, three on-ramp vehicle generators are provided. These generators are used, in sequence, to generate traffic under the three different signal presentation configurations used at most diamond interchange on ramps in California. Provision has been made to allow the length of time each generator is used to be set as an input parameter of the interchange studied.

In the case of the Roscoe Boulevard-San Diego Freeway interchange, these three signal phase configurations are appropriately grouped combinations of seven basic subphases (which include yellow clearance times). Each individual subphase, its duration, and the subphase grouping are shown in Figure 7. Phase 1, which is 21 sec in duration, is when westbound Roscoe Boulevard vehicles turn left to enter the southbound freeway on-ramp. Phase 2, 27 sec in duration, is when eastbound Roscoe Boulevard vehicles turn right onto the ramp during a green signal indication. During Phase 3, lasting 12 sec, eastbound Roscoe Boulevard vehicles continue to enter the ramp on a red signal indication, after first stopping. For model purposes, the yellow signal presentation subphases for Phases 1 and 2 are considered as extensions of the corresponding green phases and have been included in the corresponding phase times. The total cycle length at this interchange is represented by the sum of the three phase times—60 sec. With this implementation, it is possible to control input volume and gap distributions independently during each phase of signal operation, and thus to study

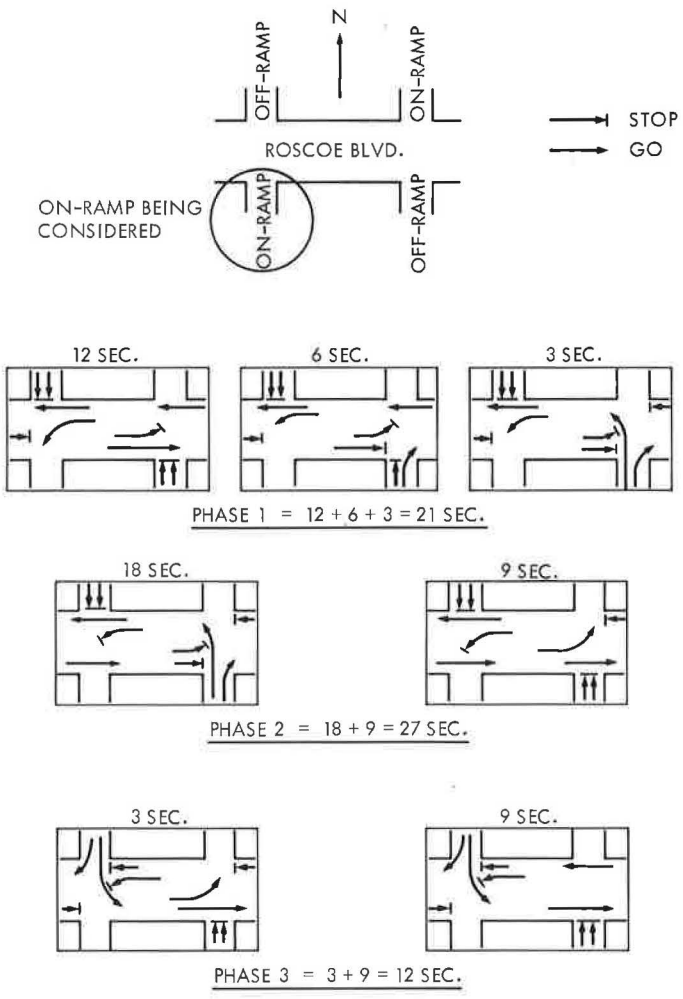


Figure 7. Phasing diagram at Roscoe Blvd.-San Diego Freeway diamond interchange, July 13, 1965.

the effect on total ramp and freeway travel time resulting from metered traffic inputs to the on-ramp.

Time Increment—Since this simulation is oriented around the macroscopic, significant-event simulation approach, and since long-time increments generally provide better computer-time to simulated-time ratios, it was decided to implement the time increment of the model on a dynamic basis. Therefore, the computer determines the value of the time increment used in each given run, based on interchange and vehicle-driver characteristics, in such a way that no vehicle moving under optimum flow conditions can pass entirely through any submodel in one time increment. For the portion of the diamond presently included in the model, the time increment is 1.5 sec.

To study the effect of time increment on the computer- to simulated-time ratio, a typical roadway configuration was simulated using a particular set of vehicle-driver and road variables. Several simulation runs were made holding all model characteristics constant except time increment, which was varied between 0.1 and 1.5 sec in increments of 0.1 sec. Figure 8 shows, for three different values of freeway volume, how the ratio of simulated time to computer time varies as a function of time increment.

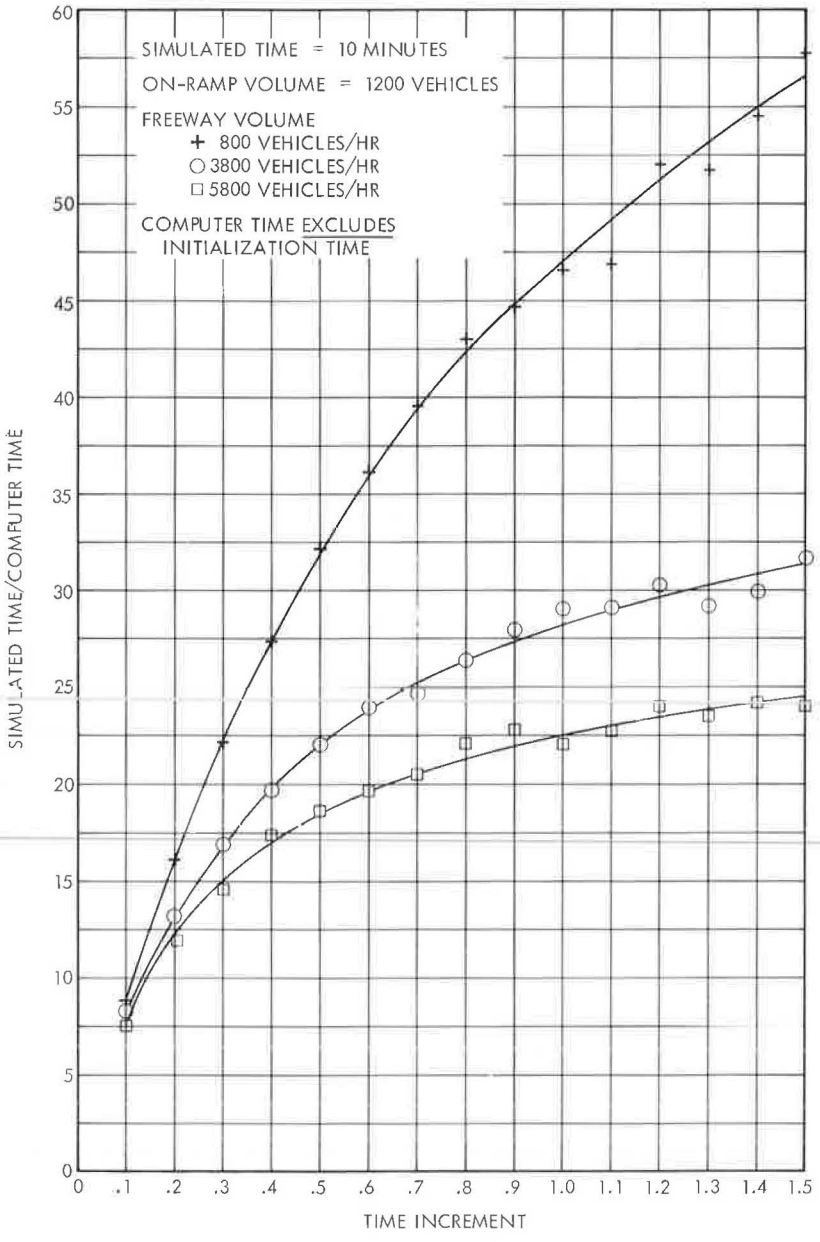


Figure 8. Ratio of simulated-time to computer-time as a function of time increment.

These ratios resulted from computer runs representing 10 min of simulation time and do not contain the effect of model initialization time, which requires approximately 5 sec of computer time. These ratios will therefore be obtained either when the computer time is substantially larger than 5 sec, or when replication runs are being made for which the model characteristics and roadway and vehicle variables remain constant, thereby enabling the initialization routine to be bypassed for all runs after the first.

Even for the very simple model being studied and for the high-volume case of 7000 vph, 3.3 times as much computer time is consumed when the smallest time increment

is used than when the largest time increment is used. Similarly, 6.7 times as much computer time is used for the low-volume case of 2000 vph.

The time ratios of Figure 8 also reflect the improvement in efficiency that results from the vehicle aggregation of the macroscopic, significant-event simulation approach. It has been estimated that aggregation can, by itself, account for a 20-fold increase in the simulated-time to computer-time ratio (4). Therefore, ratios as low as 0.4 could be expected for this model if the standard incremental approach were used. A final factor influencing these ratios is the computation of system performance characteristics. To evaluate traffic flow through the model, various measures are continuously accumulated and stored while the simulation is under way. At the completion of each run these measures are used to compute system performance for purposes of comparison of alternative geometric design, traffic flow, control, etc., as well as to establish validation of the model. A substantial amount of data reduction has been included in the runs of Figure 8.

To assure that no hidden effects on model performance occur due to the analytic determination of time increment for each simulation run, an analysis of the effect of time increment on model behavior was performed. Seventy runs (10 at each of seven values of time increment between 0.1 and 1.5) were made holding all parameter values and system characteristics constant. System travel time for each origin-destination was determined for each run.

A one-way analysis of variance test was performed to test the hypothesis that the means are identical, irrespective of time increment value. In the case of the freeway origin-destination, the significance probability is 0.124. For the on-ramp origin-destination, the significance probability is greater than 0.5. It is therefore concluded that time increment does not affect the results of the simulation.

Data Reduction

The data reduction programs convert, compute, and list variables and statistical values for use by the experimenter. Several forms of output are available under program control. These listings can be considered to be of four types.

Diagnostic—This section of the computer program was designed primarily for use in checkout of the system. The output consists of a partial decoding of the information that is maintained for each vehicle in each submodel. The printing of these records is

TIME (SEC)	VSTBL NO.	VEH COUNT	1ST ADDR	LAST ADDR
9.8	3	17	34766	34736

CORE ADDR	PREV ADDR	O.D.	INIT VEL	DRIV TYPE	SYSTEM ENTRY TIME	VSTBL ENTRY TIME	MIN XIT (SEC)	DET (SEC)	VSTBL DELAY (SEC)	DELAY BEYOND FREEFLO	NEXT ADDR
34766	34766	1	114	4	1.5	1.5	20.9	20.9	.0	.0	34772
34772	34766	1	95	2	1.0	1.0	24.3	24.3	.0	.0	34764
34764	34772	1	75	2	2.5	2.5	26.1	26.1	.0	.3	34770
34770	34764	1	85	1	1.2	1.2	27.2	27.2	.0	.0	34774
34774	34770	1	75	0	.5	.5	30.0	30.0	.0	.0	34760
34760	34774	1	85	1	4.2	4.2	30.2	30.2	.0	.0	34756
34756	34760	1	85	1	4.3	4.3	30.3	30.3	.0	.0	34754
34754	34756	1	95	2	5.0	5.0	28.3	30.3	2.0	2.0	34750
34750	34754	1	114	4	6.6	6.6	26.0	30.3	4.3	4.3	34744
34744	34750	1	95	2	7.8	7.8	31.1	31.1	.0	.0	34740
34740	34744	1	95	2	8.6	8.6	31.9	31.9	.0	.0	34762
34762	34740	1	75	0	3.0	3.0	32.5	32.5	.0	.0	34742
34742	34762	1	114	4	7.9	7.9	27.3	32.5	5.2	5.2	34752
34752	34742	1	95	2	5.6	5.6	28.9	32.5	3.6	3.6	34734
34734	34752	1	104	3	10.3	10.3	31.6	32.5	.9	.9	34746
34746	34734	1	95	2	7.0	7.0	30.3	32.5	2.2	2.2	34736
34736	34746	1	85	1	10.1	10.1	36.1	36.1	.0	.0	34736

Figure 9. Sample computer output of a diagnostic page.

under program control and may be accomplished at each time interval or at time intervals specified in a parameter table (Fig. 9).

The first line reflects the time when the examination was made. The submodel identification and the first and last computer memory addresses used by vehicles in the submodel are also printed. Examination of the first two columns reveals the address linkage scheme used in the model. This diagnostic printing also has a column of desired exit times which, by a check of ascending desired exit times against memory address linkages, assures that proper execution of the model has been accomplished. The remaining columns are printed to assist in an early evaluation of the performance of the model.

Vehicle Performance Records—Vehicle performance records can be accumulated at the end of each time increment. When such output is desired, the entry to the data reduction subroutine is made after the vehicles in the last submodel have been processed. At that time all information is available to compute and list the items associated with each submodel. Provision has been made for storing this information on tape for input to a more general data-reduction program. This two-pass approach will be used in future versions of the model if model complexity should require all of the computer memory for its program. Each page of the output is associated with a particular submodel, and each line of information corresponds to a given time increment (Fig. 10).

The time column identifies the time increment corresponding to the line of information. The signal-phase column indicates which of the three possible signal phases has been presented to vehicles entering the on-ramp of the freeway. The number of vehicles column reflects the count of vehicles in the submodel at the beginning of the time period. Percent occupancy is computed from the number of vehicles and the maximum number of vehicles that the specific submodel can contain. The columns of vehicles entered and vehicles removed are totals for the time period beginning at the time shown. Cumulative averages and totals for travel time, and number of vehicles removed are shown in the next six columns. These averages and totals are classified by vehicle origin-destination. A column of dashes is printed if an origin-destination does not

VSTBL 27

						CUMULATIVE						
						AVG VSTBL TRT		AVG SYS TRT		VEH REMOVED		# VEH
TIME	SIGNAL PHASE	# VEH	% OCUPNCY	VEH ENT	VEH REM	OD 1	OD 5,12	OD 1	OD 5,12	OD 1	OD 5, 12	PASSED
65.6	1	2	7.1	4	2	1.9	1.8	31.6	23.6	49	6	2
67.0	1	4	14.3	0	0	1.9	1.8	31.6	23.6	49	6	0
68.4	1	4	14.3	4	4	1.9	1.8	31.7	23.6	53	6	0
69.8	1	4	14.3	4	4	1.9	1.8	31.9	23.6	57	6	0
71.2	1	4	14.3	0	0	1.9	1.8	31.9	23.6	57	6	0
72.0	2	4	14.3	4	4	1.9	1.8	32.0	23.6	61	6	1
73.4	2	4	14.3	0	0	1.9	1.8	32.0	23.6	61	6	0
74.8	2	4	14.3	4	4	1.9	1.9	32.1	24.9	64	7	1
76.2	2	4	14.3	4	1	1.9	1.9	32.1	24.9	65	7	0
77.6	2	7	25.0	4	3	2.0	1.9	32.2	24.9	68	7	0
79.0	2	8	28.6	0	4	2.0	1.9	32.3	24.9	72	7	0
80.4	2	4	14.3	4	4	2.1	1.9	32.5	24.9	72	7	0
81.0	3	8	28.6	0	4	2.0	1.9	32.4	24.9	76	7	0
82.4	3	4	14.3	4	4	2.1	1.9	32.5	24.9	80	7	2
83.8	3	4	14.3	4	0	2.1	1.9	32.5	24.9	80	7	0
85.2	3	8	28.6	0	4	2.1	2.0	32.5	25.0	83	8	0
86.6	3	4	14.3	4	4	2.1	2.1	32.5	25.3	86	9	1
88.0	3	4	14.3	1	0	2.1	2.1	32.5	25.3	86	9	0
89.4	3	5	17.9	2	4	2.1	2.1	32.4	25.5	89	10	0
90.8	3	3	10.7	0	3	2.1	2.1	32.4	25.7	91	11	0
92.2	3	0	.0	0	0	2.1	2.1	32.4	25.7	91	11	0
93.6	3	0	.0	0	0	2.1	2.1	32.4	25.7	91	11	0
95.0	3	0	.0	1	0	2.1	2.1	32.4	25.7	91	11	0
96.4	3	1	3.6	1	1	2.1	2.1	32.4	26.5	91	12	0
97.8	3	1	3.6	3	1	2.1	2.1	32.4	26.5	92	12	0
99.0	4	3	10.7	1	3	2.1	2.1	32.5	27.0	94	13	0
100.4	4	1	3.6	4	1	2.1	2.1	32.5	27.0	95	13	0

Figure 10. Sample computer output of a vehicle performance page.

SYSTEM SNAPSHOT --- MOST CUMULATIVE VALUES ARE RESET EVERY 30 SECONDS

TIME	VEH REMOVED		AVG SYSTRT		-------V3-------			------V26-------			------V25-------			------V24-------			------V23-------		
(SEC)	OD1	OD5.12	OD1	OD5.12	OCPY	REM	AVTRT	OCPY	REM	AVTRT	OCPY	REM	AVTRT	OCPY	REM	AVTRT	OCPY	REM	AVTRT
30	3	2	26.5	25.4	7.9	8	22.3	4.9	6	3.6	3.8	2	7.3	8.6	0	.0	13.0	3	13.7
60	39	8	31.2	28.6	9.7	39	25.7	8.6	45	4.1	15.3	9	9.5	.0	3	13.7	8.6	9	14.3
90	82	16	33.1	30.1	8.2	45	27.2	12.3	50	5.1	19.2	8	9.8	4.3	2	14.5	26.0	7	14.7
120	114	29	33.3	31.5	6.7	38	26.8	18.5	47	5.2	15.3	14	11.1	4.3	2	15.4	8.6	11	14.6
150	152	40	33.7	32.5	11.2	34	25.6	12.3	48	5.9	19.2	9	10.6	4.3	2	16.3	17.3	8	14.0
180	202	49	33.7	32.6	10.4	50	25.7	13.5	60	5.5	.0	11	10.9	.0	1	16.3	4.3	5	15.5
210	253	55	34.0	32.7	8.9	46	27.2	9.8	53	5.2	.0	4	7.8	4.3	0	.0	17.3	4	12.5
240	291	62	33.9	32.0	11.7	45	25.7	17.2	48	4.5	11.5	9	7.6	4.3	3	13.0	8.6	9	13.2
270	336	73	34.0	32.0	10.7	51	27.5	17.2	60	5.7	19.2	9	9.6	.0	2	15.4	26.0	9	13.0
300	386	83	34.7	32.3	8.4	47	27.0	16.0	60	9.0	15.3	12	9.6	.0	0	.0	8.6	11	14.8
330	431	94	35.0	32.8	7.4	40	26.5	9.8	54	5.9	15.3	9	9.6	4.3	3	13.7	26.0	6	13.9
360	466	106	34.7	32.7	6.7	34	24.8	13.5	45	4.5	19.2	14	10.7	.0	3	13.6	8.6	12	14.3
390	502	119	34.6	32.5	9.2	30	25.8	7.4	46	4.8	11.5	11	9.5	8.6	4	13.4	17.3	5	12.5
420	538	128	34.6	32.4	8.7	42	26.3	13.5	48	4.8	7.6	11	10.6	.0	2	16.2	4.3	8	13.6
450	577	136	34.5	32.3	8.2	37	26.7	8.6	46	5.2	19.2	5	8.1	.0	3	14.3	30.4	5	13.0
480	610	147	34.5	32.2	8.9	34	26.4	12.3	45	4.7	19.2	14	10.3	4.3	1	15.3	8.6	13	15.2
510	653	160	34.5	32.4	7.9	41	26.0	7.4	56	6.0	19.2	11	9.6	4.3	1	12.5	21.7	10	13.0
540	691	174	34.4	32.0	9.7	33	24.3	1.2	52	4.7	3.8	14	9.1	.0	1	12.5	4.3	9	12.6
570	730	180	34.3	31.9	8.9	44	26.7	11.1	41	4.5	11.5	5	8.0	.0	2	16.3	34.7	5	13.2
600	773	190	34.2	32.0	8.7	43	25.2	12.3	54	4.9	11.5	12	10.8	.0	1	14.3	4.3	11	15.3
630	813	198	34.1	32.0	7.7	44	25.0	11.1	51	4.7	19.2	6	8.7	.0	3	14.3	30.4	5	13.0
660	850	209	34.1	32.0	8.9	35	26.8	9.8	49	4.7	19.2	13	10.6	.0	0	.0	8.6	13	15.7
690	891	220	34.1	32.1	7.2	40	27.2	7.4	51	5.1	23.0	9	10.6	4.3	2	12.1	21.7	8	14.1
720	926	232	34.0	32.1	5.7	33	24.4	7.4	46	4.6	23.0	13	10.9	8.6	2	13.6	4.3	11	14.7
750	958	244	34.0	32.2	10.4	28	24.5	2.4	44	5.0	15.3	12	10.7	4.3	3	13.0	26.0	7	14.3

Figure 11. Sample computer output of a system snapshot page.

VSTBL 3 VELOCITY FREQUENCY COUNTS

VEL	INITIAL	FINAL	VEL	INITIAL	FINAL	VEL	INITIAL	FINAL	VEL	INITIAL	FINAL
0	0	0	32	0	0	64	4	1	96	0	0
1	0	0	33	0	0	65	0	1	97	1	0
2	0	0	34	0	0	66	0	0	98	1	0
3	0	0	35	0	0	67	1	15	99	0	1
4	0	0	36	0	0	68	0	0	100	0	0
5	0	0	37	0	0	69	0	3	101	0	0
6	0	0	38	0	0	70	15	27	102	0	0
7	0	0	39	0	0	71	11	0	103	3	0
8	0	0	40	0	0	72	8	1	104	3	19
9	0	0	41	0	0	73	10	4	105	0	0
10	0	0	42	0	0	74	40	0	106	0	0
11	0	0	43	0	0	75	99	110	107	0	0
12	0	0	44	0	0	76	10	0	108	0	0
13	0	0	45	0	0	77	24	0	109	0	0
14	0	0	46	0	0	78	10	2	110	0	1
15	0	0	47	2	0	79	4	0	111	0	0
16	0	0	48	0	0	80	20	10	112	0	0
17	0	0	49	0	7	81	9	15	113	0	0
18	0	0	50	0	10	82	5	0	114	1	5
19	0	0	51	0	0	83	4	5	115	0	0
20	0	0	52	0	43	84	15	0	116	0	0
21	0	0	53	0	1	85	81	66	117	0	0
22	0	0	54	2	0	86	0	0	118	0	0
23	0	0	55	0	10	87	2	7	119	0	0
24	0	0	56	0	2	88	1	0	120	0	0
25	0	0	57	6	1	89	6	0	121	0	0
26	0	0	58	0	6	90	1	0	122	0	0
27	0	0	59	1	5	91	0	0	123	0	0
28	0	0	60	0	10	92	0	0	124	0	0
29	0	0	61	0	5	93	5	0	125	0	0
30	0	0	62	0	7	94	0	4	126	0	0
31	0	0	63	0	0	95	20	30	127	0	0

Figure 12. Sample computer output of a velocity frequency counts page.

VEHICLES LEAVING SYSTEM

TIME	SEQ	SYS TRAVEL TIME	DRIVER-TYPE	O.D.
19.0	1	19.9	4	12
25.2	2	24.0	4	1
29.4	3	28.7	2	1
30.8	4	29.0	2	1
32.2	5	32.0	1	1
35.0	6	32.0	1	1
35.0	7	32.0	1	1
35.0	8	31.3	2	1
36.4	9	36.4	0	1
37.8	10	31.7	4	1
37.8	11	30.5	2	1
37.8	12	25.5	3	12
37.8	13	30.3	2	1
39.0	14	32.4	4	1
40.4	15	37.7	0	1
40.4	16	35.1	2	1
40.4	17	30.6	3	1
41.8	18	35.7	2	1
41.8	19	27.5	2	12
41.8	20	30.2	3	1
41.8	21	30.6	2	1
44.6	22	31.7	4	1
44.6	23	29.7	3	1
44.6	24	34.8	1	1
44.6	25	32.0	1	1
46.0	26	29.7	2	1
46.0	27	32.4	1	1
46.0	28	32.4	1	1
46.0	29	26.3	2	5
48.0	30	33.4	1	1

Figure 13. Sample computer output showing performance characteristics of the sequence of vehicles leaving the system.

traverse a particular submodel. The count of the number of successful passing maneuvers within the submodel is also printed for each time increment. For single-lane submodels, where no passing can occur, dashes are printed.

System Snapshot—After considerable checking and experimentation, it became obvious that a less detailed output of performance records was needed. Thus, the snapshot program was devised to print selected information at the end of specified periods. If a 1-min period is chosen, one page of output will result for approximately 50 min of simulation. A total of the number of vehicles removed from the system, together with average system travel time (for both freeway and on-ramp vehicles) is accumulated, while other values are reset from line to line (Fig. 11).

Statistical Summaries—A count is maintained for every possible velocity in each submodel. At the end of each simulation run one page of velocity frequency counts is printed for each submodel. Each possible entrance or exit velocity is printed along with a count of the number of vehicles that entered and the number that exited that submodel with that velocity (Fig. 12).

Another type of output available is the sequence of vehicles leaving the system (Fig. 13). This information may be used to make comparisons, for validation studies, of simulated traffic flow with actual measurements of traffic flow taken from a freeway interchange.

In addition to the optional outputs described, each simulation run produces a single-page prologue (Fig. 14). The data in Figure 14 represent the values of the variables and model and roadway characteristics used for the corresponding simulation run. For example, vehicle generating rates, vehicle-driver type distributions, maximum free-flow velocities, accelerations, passing probabilities, submodel lengths, maximum occupancy, etc., are given. The values shown in Figure 14 are representative of the values used in the runs which have been made to produce the data presented in this report.

VEHICULAR TRAFFIC STUDY PARAMETERS

FREEWAY - DIAMOND INTERCHANGE..MODEL 1

10-01-65 -- VERSION 01 -- RUN 02

VEHICLE GENERATING	FREEWAY	POCKET LANE	CURB LANE	
RATES PER HOUR	= 3800	(GREEN) = 600	(GREEN) = 540	(RED) = 60
V23 ENTRY PROBABILITIES		= .778	= .872	= .745
VSTBL LENGTHS (FT)	V3 = 2210 V23,24 = 506	V25 = 590	V26 = 360	V27 = 175

SIGNAL PHASE TIMES (SEC)	DISTRIBUTION OF DRIVER-TYPES (%)	FWY FFVF (FT/SEC)	FWY A-FCT (FT/SEC/SEC)	ON-RAMP VO (FT/SEC)	ON-RAMP A-FCT (FT/SEC/SEC)
1 = 12					
2 = 09	SLOW = 07	75	3.0	6	3.0
3 = 18	MEDIUM = 24	85	4.5	11	3.5
4 = 09	NORMAL = 38	95	6.0	16	4.0
5 = 06	FAST = 24	104	7.5	21	4.5
6 = 06	VERY FAST = 07	114	9.0	26	5.0

FREEFLOW SYSTEM TRAVEL TIMES (SEC)

DRIVE TYPE	O.D. 1	O.D. 5.12
0	36.6	32.2
1	32.3	28.5
2	28.9	25.4
3	26.5	23.2
4	24.1	21.1

AVG VEH LENGTH + GAP (FT)	HEAD WAY (SEC)	RUN TIME (MIN)	TINC TIME (SEC)
22	2.0	25	1.5

% OCCUPANCY =	5	10	15	20	25	50	100
PROBABILITY =	.700	.400	.100	.080	.050	.010	.000

TABLES USED TO DETERMINE VELOCITIES DURING "CONGESTION"

V3 to V26		V23,24 TO V25		V25 TO V26		V26 TO V27		QUEUES TO V3		& TO V23, V24	
OCPY (CT)	VELOC (FT/SEC)	OCPY (CT)	VELOC (FT/SEC)	OCPY (CT)	VELOC (FT/SEC)	OCPY (CT)	VELOC (FT/SEC)	OCPY (CT)	VELOC (FT/SEC)	OCPY (CT)	VELOC (FT/SEC)
50	75	13	63	50	72	12	75	370	75	10	24
55	64	15	53	55	61	17	64	375	64	12	20
60	53	17	43	60	50	22	53	380	53	14	17
65	42	19	33	65	39	27	42	385	42	16	13
69	31	21	23	69	28	32	31	390	31	18	10
73	21	23	13	73	17	37	20	395	21	20	6
77	10	25	3	77	6	42	9	400	10	22	3

Figure 14. Sample computer output of a prologue page.

MODEL PERFORMANCE STUDIES

Sensitivity

There are three principal reasons for conducting sensitivity analyses of simulation models. The first is related to the independent variables in the model, which are an integral part of the logic for moving vehicles. Among such variables are average minimum time-headway, desired velocity distribution, and passing probability function of occupancy. The purpose is to determine if these variables significantly affect the measures of effectiveness of interest. Those variables which do significantly affect the measures are, in subsequent field studies, measured with great care and remain a part of the simulation model. Those variables which have no appreciable effect should be removed from the model to simplify the logic and improve the ratio of simulated time to computer time.

The second motivation for sensitivity analyses pertains to the research phase and concerns those independent variables of interest that are not an integral part of the logic for moving vehicles. In the present model these include geometric variables, such as the number of lanes in submodels, length of submodels, grade, etc., and control variables, such as cycle length, phasing schemes for fixed-time control systems, and

various automatic control disciplines. The objective in this case is to discover the effect of these variables on the measures of effectiveness to assist in the selection of design criteria for constructing and operating future interchanges or for modifying existing ones.

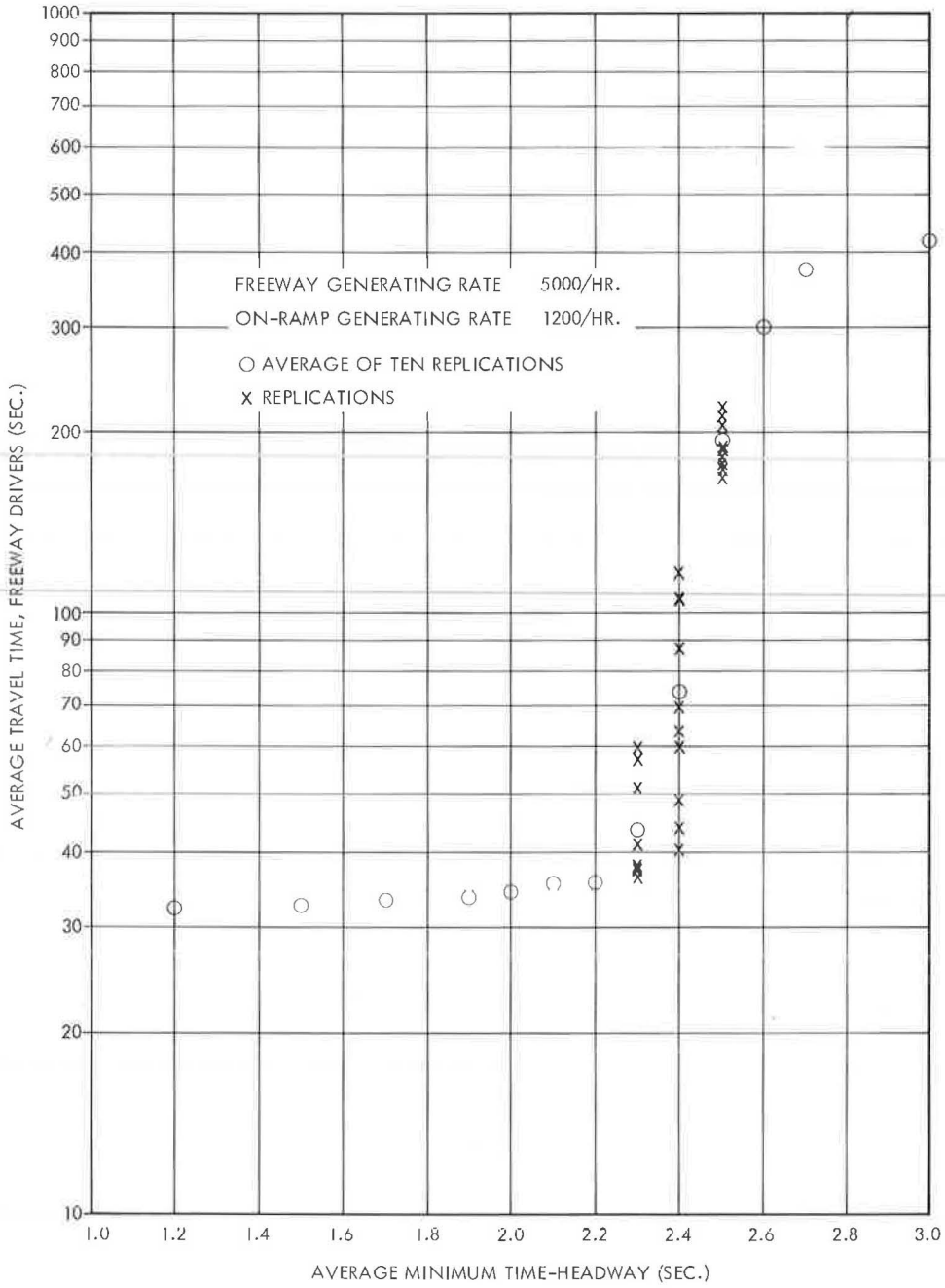


Figure 15. Average travel time of freeway drivers as a function of average minimum time-headway.

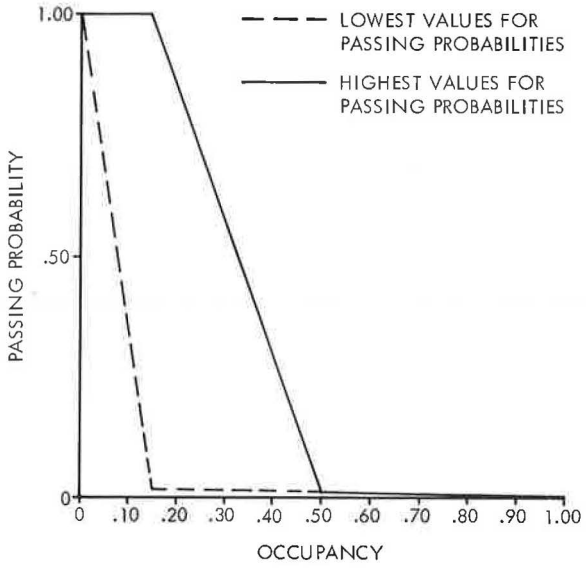


Figure 16. The two extremes of the simplified passing probability function.

The third reason is to ascertain, prior to validation, whether the model is behaving properly. If the measures of effectiveness vary unreasonably as the independent variables are changed, then either the logic of the model or the computer code, or both, are incorrect; they need close scrutiny and further debugging. Of course, sensitivity tests run for either of the first two reasons will also provide information about the reasonableness of model behavior.

The remainder of this section is devoted to discussing some of the sensitivity-run results obtained with version 1 of Model 1. These tests were conducted primarily for the first reason mentioned, namely, to determine the effect on the measures of effectiveness of certain variables that were an integral part of the logic for moving vehicles. It should be noted that in the present circumstance no sensitivity analysis in the research context may be conducted until the total interchange is simulated.

Average Minimum Time-Headway—As discussed in the previous section, the minimum time-headway for version 1 of Model 1 is set at 2 sec. Implicit in this assumption is that 2 sec is really the mean value of the minimum time-headway distribution.

To determine the sensitivity of average travel time to the selection of this independent variable, ten replications were made for each of several values of average minimum time-headway between 1.2 and 3.0 sec. Each replication was for a freeway volume of 5000 vph, an on-ramp volume of 1000 vph, and a simulation time of 2 hours. The results are shown in Figure 15 for freeway drivers. A graph with similar characteristics was obtained for on-ramp drivers.

In Figure 15, only the average of the ten replications is plotted (circles), except where the average travel time variation is large, such as at time-headways of 2.3, 2.4 and 2.5 sec. This large variation at headway values around 2.4 sec is expected, since the capacity of the freeway at an average headway of 2.4 sec is 6000 vph, which exactly equals the sum of the ramp and freeway input volumes. Thus, small fluctuations in input volume above and below the mean rate of 6000 vph, which occur because of the randomness of the traffic generator, will cause the formation and dissolution of congestion with a corresponding increase and decrease in travel times.

Figure 15 shows that there is no significant effect of minimum time-headway for values less than or equal to about 2.2 sec. Under these circumstances the input volume of 600 vph is sufficiently below capacity so that the headway constant does not impose

an appreciable influence. However, for any headway constant greater than or equal to 2.3 sec, the 6000 vph input is sufficiently close to the freeway capacity so that the headway constant drastically affects travel time by causing greater and greater delay as it is increased. Since the greatest area of interest for study occurs when the system is operating near capacity, it is concluded that extensive field measurements must be conducted under heavy flow conditions to obtain a reliable estimate of the average minimum time-headway for the model.

Passing Probability—Passing probability is a segmented linear function of submodel occupancy. In order that there may be considerable flexibility in the shape of the function, it has been provided with six breakpoints at occupancy values of 0.05, 0.10, 0.15, 0.20, 0.25 and 0.50. The passing probabilities at occupancies of 0 and 1 are 1 and 0 respectively, and are held fixed independent of the probabilities assigned to the six breakpoint occupancies.

To determine the effect of this function on the performance of the model with respect to average travel time, a simplified version of the function was employed for some computer runs. The simplification was achieved by reducing the number of breakpoints from six to two, namely those at occupancies of 0.15 and 0.50. Variation in the shape of the function was achieved by letting the value of passing probability at the 0.15 occupancy be equal to 0.02, 0.04 and 0.05 (0.05) 1.00 for a fixed probability of 0.01 at the 0.05 occupancy level. Figure 16 shows the two extremes of the passing probability as a function of occupancy.

Data runs were made for only one set of freeway and on-ramp generating rates, namely, 5121 vph for the freeway and 1159 vph for the on-ramp. These rates correspond to those measured during one of the observation periods in the validation study described

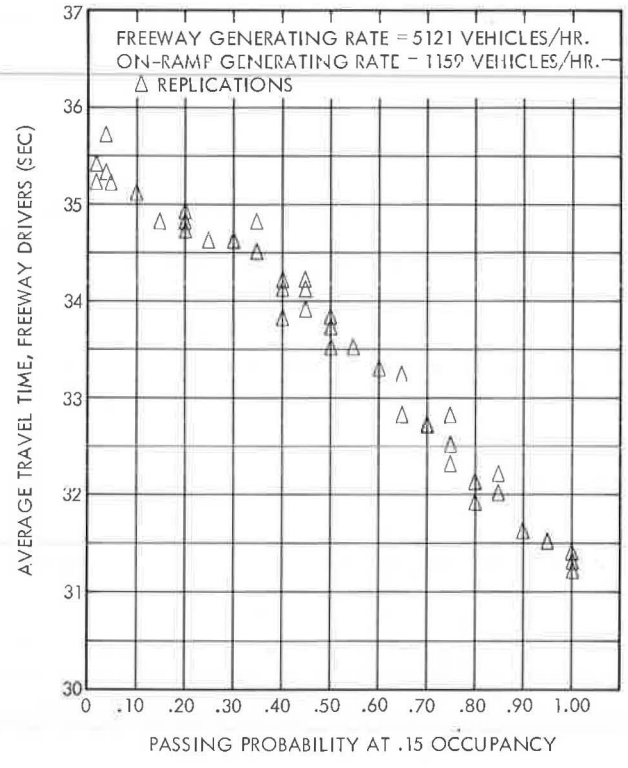


Figure 17. Effect, on average travel time of freeway drivers, of varying the simplified passing probability function at 0.15 occupancy.

later in this report. For each run, lasting 3 hours of simulated time, the average travel times for all freeway drivers and on-ramp drivers were computed. The results are shown for freeway drivers in Figure 17 and for on-ramp drivers in Figure 18. The abscissa in both figures is the value of the passing probability at an occupancy of 0.15.

It can be seen from Figures 17 and 18 that the passing probability has a significant effect on average travel time for freeway drivers, but has little or no effect on the on-ramp drivers. The probable cause for this is that the only opportunity for passing by on-ramp drivers occurs in merging submodel V_{26} and in submodel V_{27} (Fig. 2), and since the on-ramp drivers, when they arrive at the merging section, have lower velocities than do the freeway drivers, they will have correspondingly fewer passing opportunities. Since passing probability comes into play only when a potential pass arises, it follows that it is used infrequently by on-ramp drivers and therefore its value does not affect their travel times. Furthermore, V_{26} and V_{27} are relatively short submodels, offering very few potential passes.

Desired Accelerations—As discussed earlier, the vehicle-driver characteristics of desired velocity, acceleration, and deceleration were each given five different values, one for each of the five vehicle-driver types of the model. The normality assumption was experimentally justified for the characteristic of desired velocity from free-flow velocity data collected for 225 vehicles, and the values inserted into the model are based on these data. With the present manual technique for gathering data, it was not possible to make any determination of accelerations. Hence, the model values were based on the judgment of the investigators, and more accurate estimates must perforce await implementation of the ultimate data collection system.

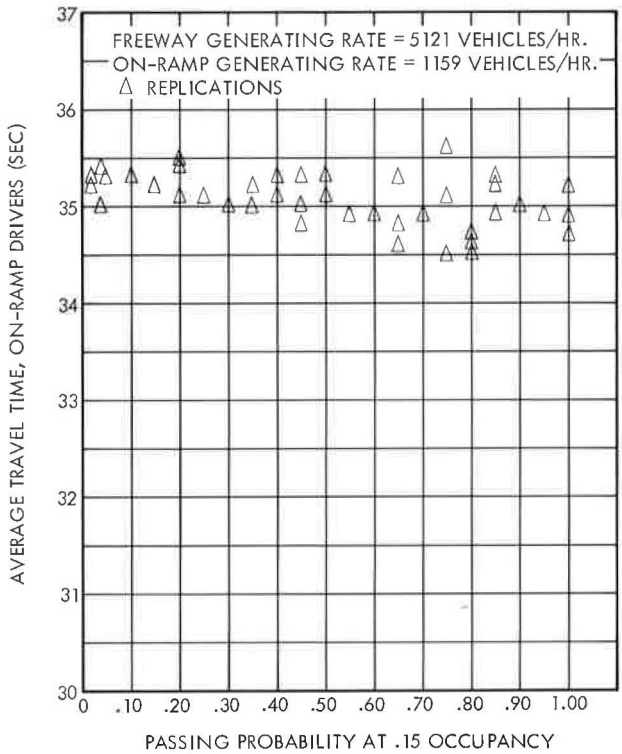


Figure 18. Effect, on average travel time of on-ramp drivers, of varying the simplified passing probability function at 0.15 occupancy.

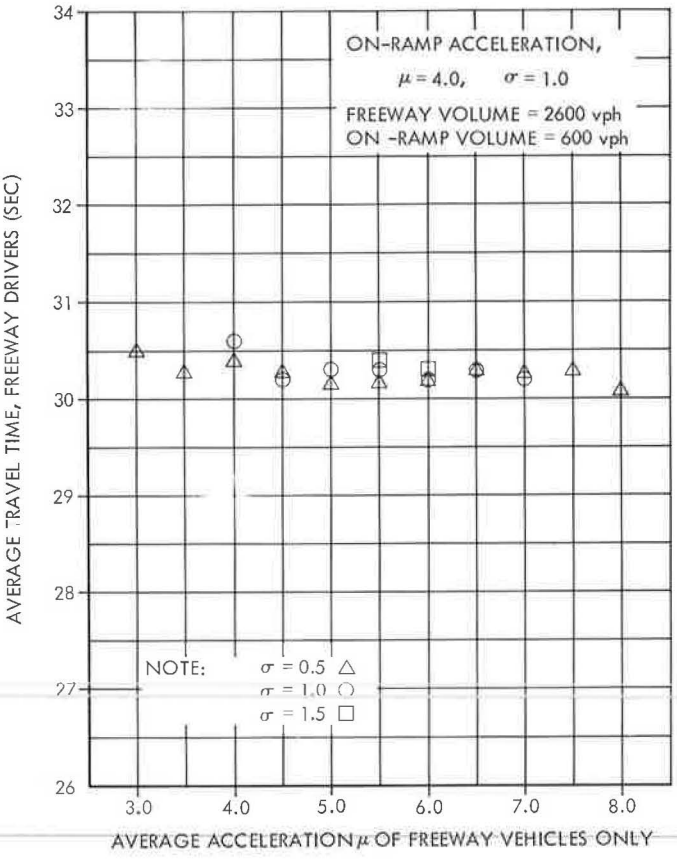


Figure 19. Effect, on average travel time of freeway drivers, of varying their desired accelerations, total volume 3200 vph.

The sensitivity tests for desired acceleration assumed that accelerations are normally distributed. Hence, mean and standard deviation completely determine the five different values for each of the five vehicle-driver types.

As in the prior two sensitivity tests, average travel time was selected as the figure of merit. Each data run covered four hours of simulated time. The results of the runs are shown in Figures 19, 20, 21 and 22. Figures 19 and 20 show the effect of changing the desired acceleration of freeway drivers while holding fixed the mean and standard deviation of on-ramp desired acceleration. Figure 19 is for a total volume of 3200 vph, while Figure 20 is for 6280 vph. The corresponding results for varying on-ramp accelerations are shown in Figures 21 and 22.

These figures show that in general the choice of acceleration has very little influence on the travel time of freeway drivers, but a substantial one for on-ramp drivers. This is not too surprising because freeway drivers do relatively little accelerating or decelerating. On-ramp drivers, on the other hand, have the primary objective of accelerating in order to merge with freeway traffic.

Start-Up

One of the principal problems facing an investigator using digital computer simulation is that of obtaining meaningful values of the model's measures of effectiveness for

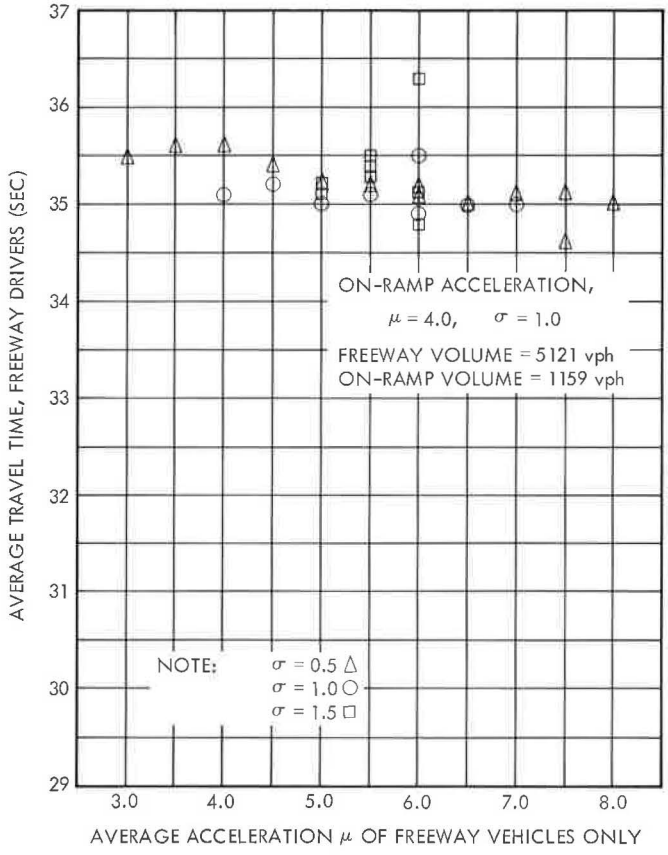


Figure 20. Effect, on average travel time of freeway drivers, of varying their desired accelerations, total volume 6280 vph.

each simulation run. This problem is discussed with respect to average travel time as the measure of effectiveness. Specifically, an attempt is made to solve the problem of determining the average travel time, say of freeway drivers, with negligible random error.

The method used was to make a computer run with a specified set of freeway and on-ramp volumes, and to record, as a function of time, the cumulative average travel time of all freeway drivers. Typical results are shown in Figure 23 for a total volume of 4000 vph. Relatively stable variation of the cumulative average is said to have occurred when the cumulative average travel times fall within a band of specified percentage deviation around a relatively stable value. The relatively stable value is defined as the midpoint of a band in which (a) the average travel times remain within the specified band from as early a time in the run as possible, (b) the total time within the band is at least 50 percent of the run, and (c) during the run at least 2100 vehicles[1] with the origin-destination of interest traverse the system. For the run of Figure 23,

[1]Assuming free-flow conditions and the asymptotic normality of average travel time, it follows, based on the 225 observations referred to earlier, that two standard deviations set equal to ½ percent of the mean travel time corresponds to 2100 cars.

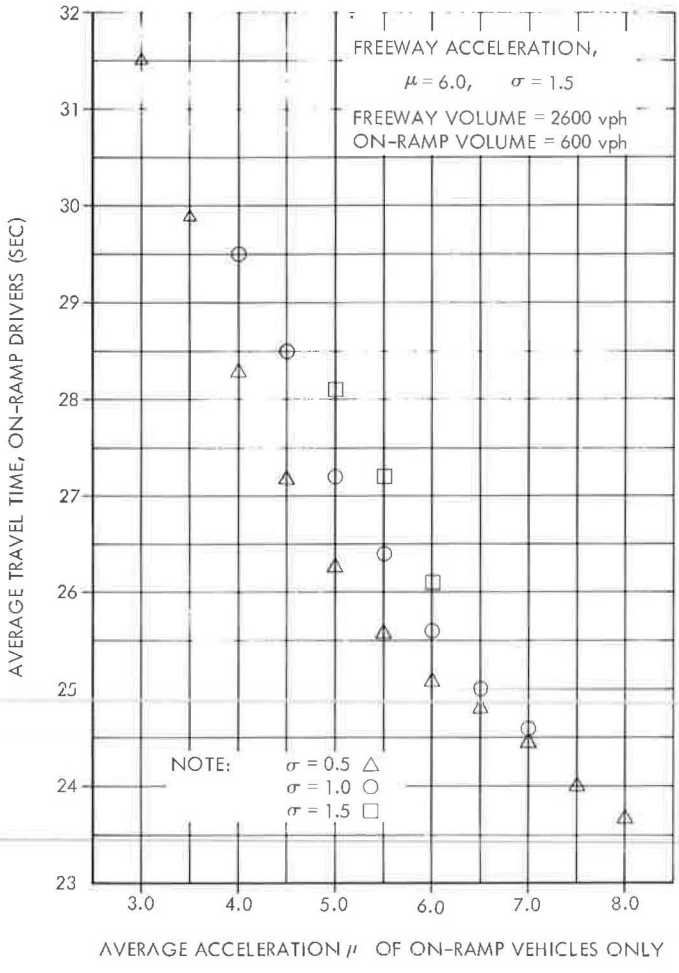


Figure 21. Effect, on average travel time of on-ramp drivers, of varying their desired accelerations, total volume 3200 vph.

the percentage deviation was chosen as ½ percent and thus the band is given by 30.65 ± 0.15. In this particular run the average travel time became relatively stable at approximately 135 sec.

The time at which relative stability is reached, as well as the relatively stable value so reached in the sense just defined, is not constant for a given freeway and on-ramp volume but varies from run to run. This is because random elements in the model, depending on the particular sequence of random numbers generated by the computer, produce different outputs. It is believed that this procedure produces an estimate of the average travel time with negligible random error.

In ten runs made at a total generating rate of 6200 vph, the time to relative stability and the corresponding band are given in Table 7. Note also that the midpoint of the band is not constant from run to run. This merely shows that the runs were not sufficiently long to reach the true value of average travel time. If costs were no consideration, each replication could have been run for, say, 20 hours of simulated time instead

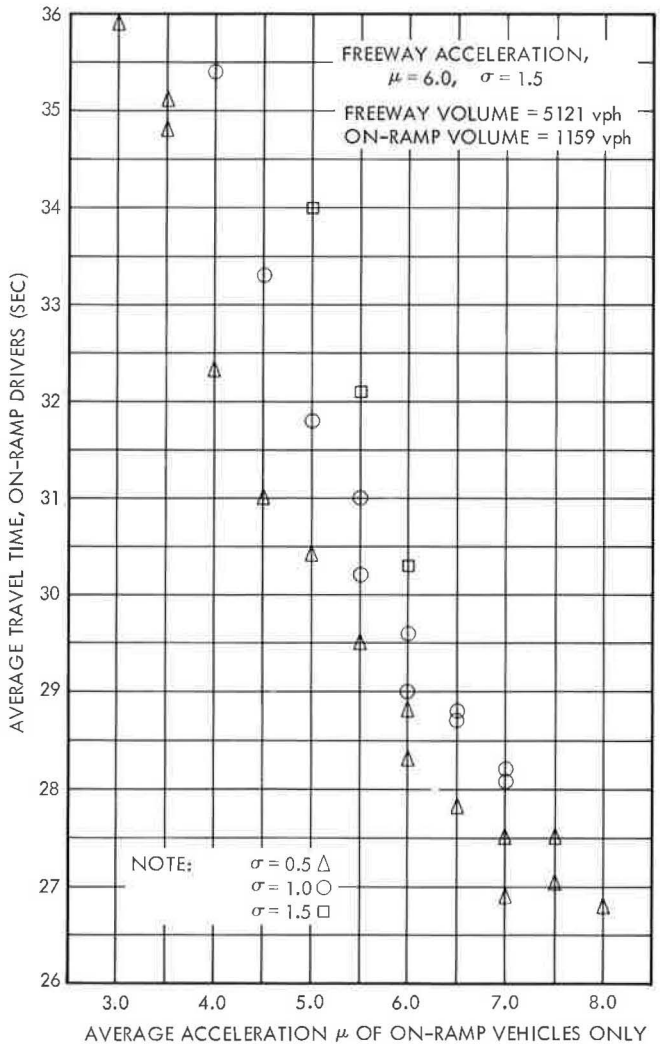


Figure 22. Effect, on average travel time of on-ramp drivers, of varying their desired accelerations, total volume 6280 vph.

of 4 hours, as was actually done, and each cumulative average might then have converged to the same mean value, namely the true equilibrium value.

The data in Table 7 reveal that from 33 to 77 min is required on any given replication for the cumulative mean value to become relatively stable. The question arises as to whether the average of cumulative averages of several replications would reach a relatively stable value much earlier than any given replication. If this average reaches relative stability sufficiently early, it is conceivable that the total simulated time for several short replications might be less than the simulated time for one replication. This, unfortunately, was not the case for the runs used to produce Table 7. As a matter of fact, the time required for the average of the cumulative averages to fall in a band which is ± ½ percent of the relative stable value was about 13 min. Thus if each of the

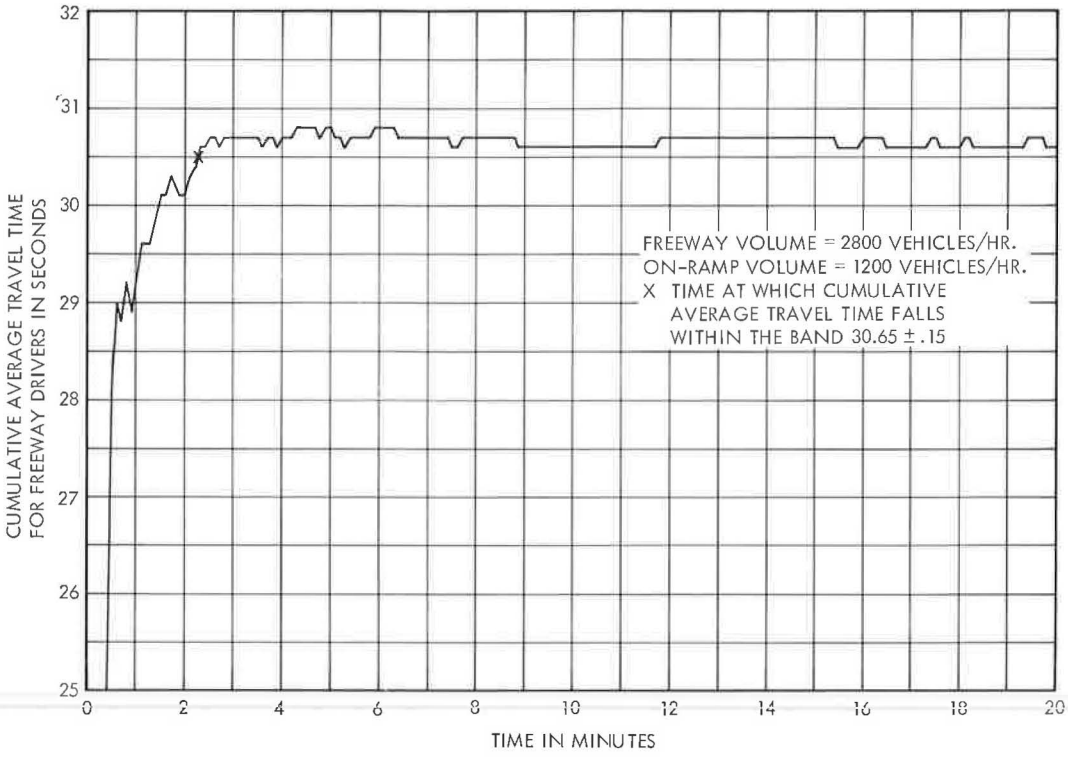


Figure 23. Cumulative average travel time of freeway drivers as a function of time.

ten replications were run for a simulated time of 13 min this would amount to a total simulated time of 130 min. Comparing this to the largest value for any replication, 77 min required for the 4th and 5th replications, it appears that performing only one replication for a sufficiently long period to reach relative stability is the cheapest way

TABLE 7

TIME TO RELATIVE STABILITY AND BAND

Run	Time to Relative Stability (min)	Band
1	64	34.75±0.15
2	57	34.35±0.15
3	33	34.45±0.15
4	77	34.55±0.15
5	77	34.35±0.15
6	53	34.65±0.15
7	35	34.35±0.15
8	49	34.75±0.15
9	40	34.45±0.15
10	65	34.85±0.15

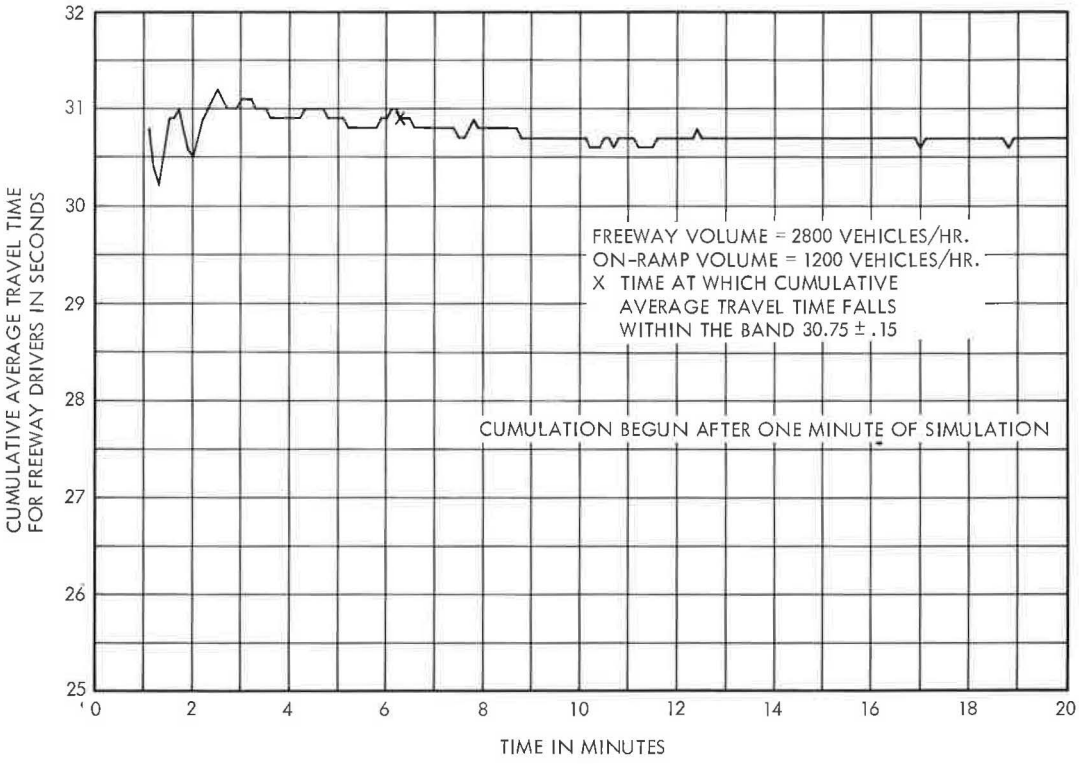


Figure 24. Cumulative average travel time of freeway drivers as a function of time. Cumulation begun after 1 min of simulation.

to get a reasonable estimate of the average travel time with negligible random error.

A possible improvement in running time for the one-replication case may be affected by accumulating travel times after the system has been in operation for a while. Intuitively, this seems like a reasonable procedure because the initial vehicles which enter the system when it is empty would possess relatively short travel times and hence would bias the cumulative average to the low side. It would appear that substantially more vehicles (or equivalently more running time) would be required to correct for this bias than if these initial vehicles were eliminated. This was not the case.

Figure 24 is the cumulative average curve for the same run corresponding to Figure 23, except that the vehicles which entered into the sum were collected after 1 min of simulation. As can be seen in this case, relative stability was achieved in approximately 6 min, whereas in Figure 23 it occurs in approximately 2 min. Several other experiments were attempted at different volumes and with the accumulation begun at various times. The results were always the same, namely, the relative stability time was always greater than or equal to the relative stability time if the accumulation had begun immediately. Thus it seems, on the basis of this rather limited study, that the best procedure for getting average travel time is to accumulate travel times, for one replication, from the beginning of the simulation and to run until the average travel time has reached relative stability.

One additional result concerning time to relative stability as a function of input volume is shown in Figure 25. Before examining this curve, the following facts should be borne in mind:

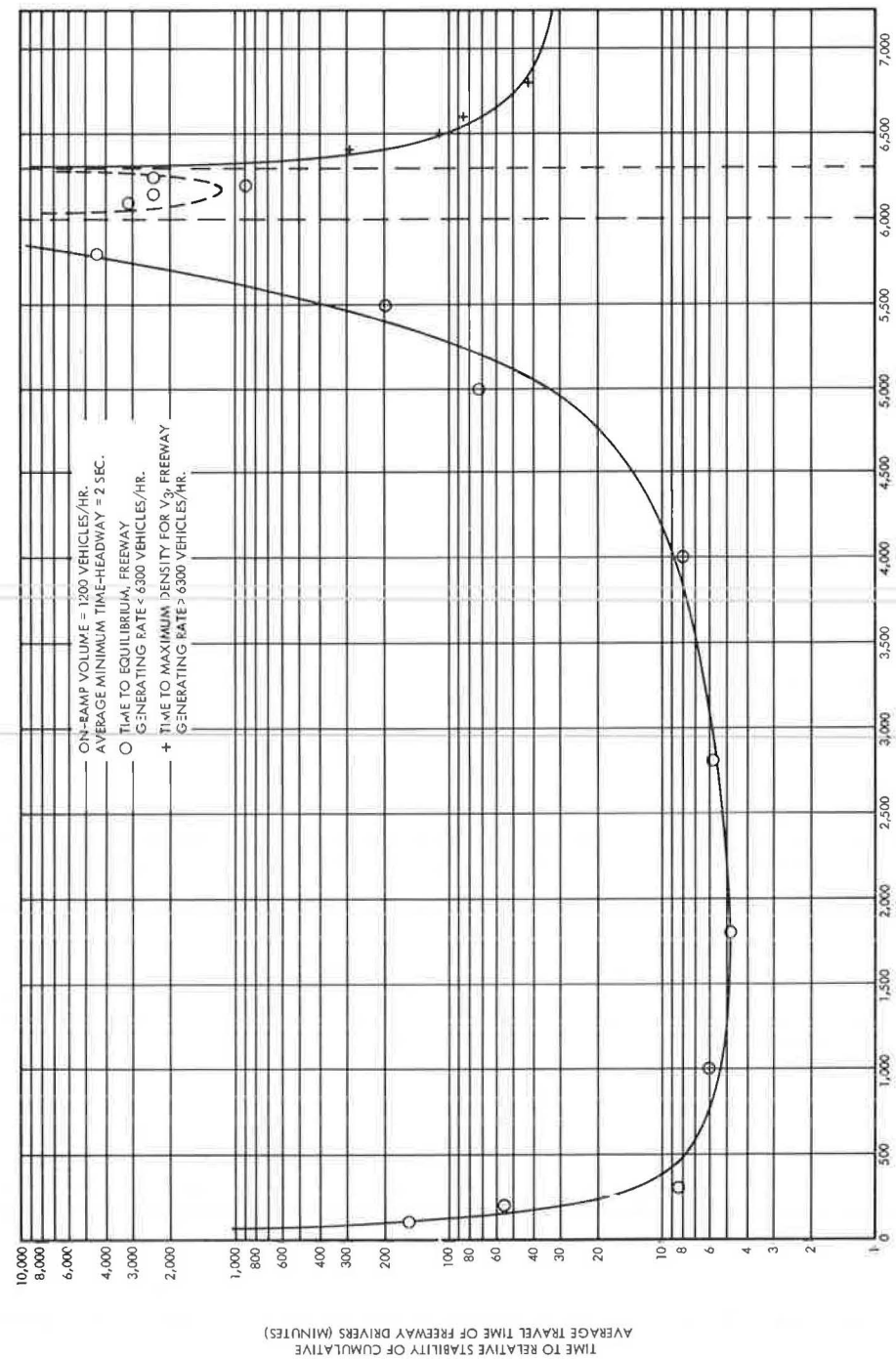


Figure 25. Time to relative stability as a function of freeway generating rate.

1. For average minimum time-headway of 2 sec, the value presently used in the model, the capacity of the freeway is 7200 vph. Hence, when the total combined demand of the freeway and on-ramp is in the vicinity of 7200 vph, the random variation in generating rate above and below 7200 vph will cause the formation and dissolution of shock waves beginning at the upstream end of V_{27}.
2. The merging-weaving logic causes the on-ramp to share with the freeway all available space for vehicles to enter V_{26} on a one-car-out-of-eight basis when V_{26} reaches saturation density. Hence on-ramp demand greater than or equal to 900 vph will restrict the freeway flow rate to 6300 vph.

The data in Figure 25 were for various freeway generating rates between 100 and 7200 vph, and a fixed on-ramp generating rate of 1200 vph. Hence, freeway generating rates between 6000 and 6300 vph will cause the on-ramp to congest; and freeway generating rates greater than 6300 vph will cause the freeway to congest. The graph of Figure 25 has three distinct regions. They are:

1. Freeway generating rates between zero and 6000 vph. It is obvious that as the generating rate approaches zero, it takes more and more time to generate enough freeway drivers to stabilize the cumulative average travel time. This clearly explains why zero is an asymptote. Similarly, a possible explanation for 6000 vph being an asymptote is that slight random variations above and below this rate cause the

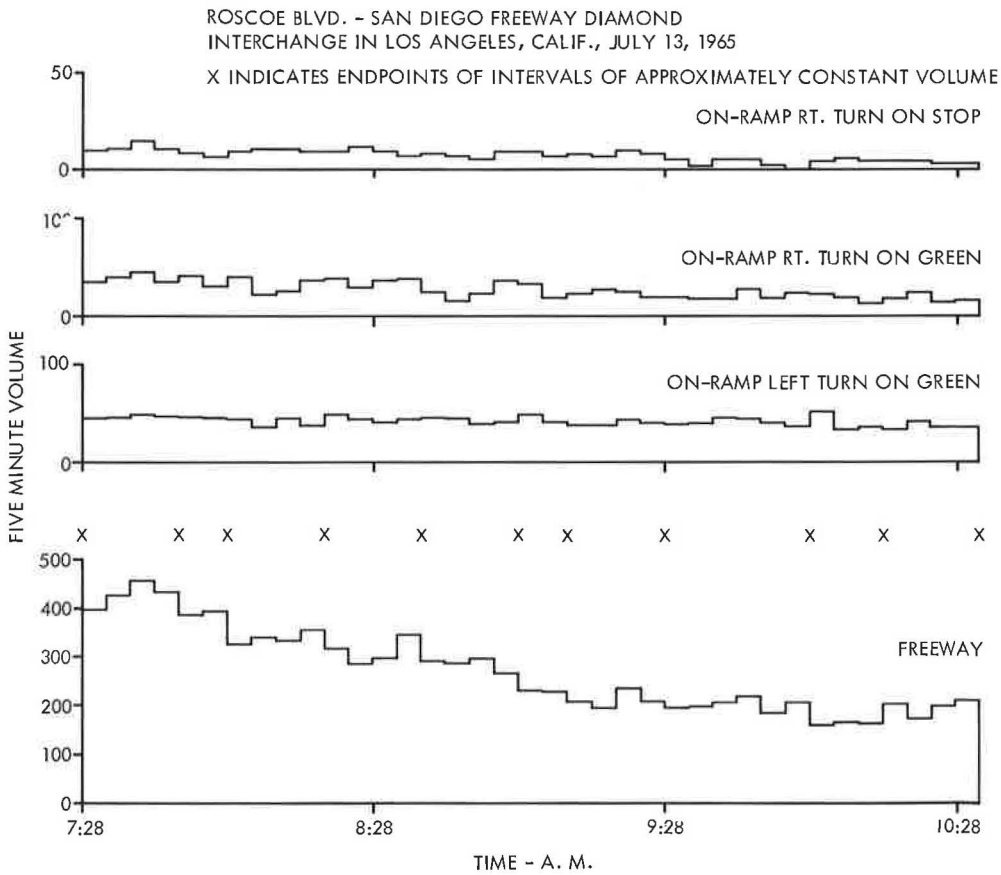


Figure 26. Five-minute volumes for three consecutive peak-period morning hours, Roscoe Blvd.-San Diego Freeway.

formation and dissolution of shock waves in V_{26}. This in turn substantially increases and decreases the travel times of freeway drivers. Thus, longer running times are required to stabilize the cumulative average travel times for generating rates in the neighborhood of 6000 vph.

2. Freeway generating rates between 6000 and 6300 vph. This curve is shown in a dashed line to indicate the paucity of data points in the neighborhood of the asymptotes and the conjectural nature of its shape. The probable reason for 6000 vph being an asymptote was just given. The probable reason for 6300 vph being a vertical asymptote is again that random variations in generating rate cause the formation and dissolution of shock waves in V_3, with consequent substantial variations in freeway travel times. This apparently requires the longer running times to stabilize the cumulative average travel time.

3. Freeway generating rates greater than 6300 vph. In this region, of course, V_3 always backs up. What is shown is the time at which V_3 reaches maximum density and not the time at which cumulative average travel time stabilizes. The latter time is much larger, and patently it would be inefficient to await its equilibrium. Rather, data should be collected only on those vehicles entering the system after the shock wave has traveled through V_3. The cumulative average travel time of these vehicles will converge with great rapidity to its equilibrium value.

One thing that should be remembered when examining Figure 25 is that only one data run was made for each point and that there would be some variation if any replications were made. Although the curve in the figure does not represent what would be obtained if the mean value of several replications were plotted, it is a reasonably close approximation to it.

Validation

The validation study for this preliminary model was accomplished by using manually collected field data. The data consisted of (a) freeway vehicle counts by the minute,

TABLE 8

RESULTS OF VALIDATION RUNS AND CORRESPONDING FIELD MEASUREMENTS

Run	Hourly Freeway Volume (f_i)	Hourly On-Ramp Volume (r_i)	Avg. of Timed Fwy. Drivers ($\bar{x}_i$)	Avg. of Timed On-Ramp Drivers ($\bar{y}_i$)	Avg. of Fwy. Driver in Simulation (ξ_i)	Avg. of On-Ramp Drivers in Simulation (η_i)	Deviations	
							$\xi_i-\bar{x}_i$	$\eta_i-\bar{y}_i$
1	5121	1158	30.8	31.2	35.0	32.4	+4.2	+1.2
2	4662	1056	31.3	29.6	33.5	31.6	+2.2	+2.0
3	4053	960	30.7	29.2	32.1	30.2	1.4	1.0
4	3732	1053	31.8	30.4	31.7	30.4	-0.1	0
5	3414	888	32.4	28.9	31.1	30.1	-1.3	+1.2
6	2748	924	30.6	30.7	30.4	29.8	-0.2	-0.9
7	2529	846	30.4	29.8	30.2	29.5	-0.2	-0.3
8	2418	768	30.7	29.8	30.1	29.6	-0.6	-0.2
9	1940	720	30.2	29.7	29.7	29.4	-0.5	-0.3
10	2312	691	32.0	30.1	29.9	29.3	-2.1	-0.8
11	1987	841	30.7	29.1	29.8	29.9	-0.9	+0.8
12	2334	888	30.0	30.3	30.0	29.8	0	-0.5
13	2625	971	30.9	30.6	30.3	30.1	-0.6	-0.5
14	2166	990	30.2	28.4	30.0	30.0	-0.2	+1.6
15	2484	965	30.1	31.0	30.2	30.2	+0.1	-0.8
16	2050	936	29.0	30.7	29.8	30.1	+0.8	-0.6
17	1896	824	30.4	28.6	29.8	30.0	-0.6	+1.4
18	1764	760	29.5	27.9	29.7	29.5	+0.2	+1.6

(b) on-ramp vehicle counts by the minute, (c) travel times through the system of freeway cars selected at random, and (d) travel times through the system of on-ramp cars selected at random. The data were collected over a period of 6 hours. The first 3 consecutive hours were in the morning peak traffic period. The second 3 consecutive hours were in the afternoon when traffic was moderately heavy. Travel times were determined (using a stopwatch) by timing a vehicle while traveling between two distinctly marked points. Vehicle counts were obtained by counting the number of cars which crossed a particular point on the freeway and on the on-ramp.

The volume counts were then plotted using a 5-min counting interval in order to smooth the data. The 6 hours of data were split into 18 time periods within which the flow of vehicles on the freeway and on-ramp remained reasonably constant. The length of these periods varied between 10 and 40 min. Figure 26 shows the flow measurements for the 3-hour morning study. Also shown are the intervals of relatively constant flow. The number of freeway cars for which travel times were obtained during each of these periods varied between 13 and 53, while the number of on-ramp cars varied between 7 and 35. Table 8 gives the results of these field measurements, which yielded f_i, r_i, $\bar{x}_i$, $\bar{y}_i$, $i = 1, \ldots, 18$, where f_i and r_i are the hourly freeway and on-ramp flow rates respectively corresponding to the i^{th} period; x_i and y_i are the averages of the freeway travel times and on-ramp travel times respectively for the i^{th} period.

Standardization of the model was accomplished by first setting its geometrical parameters equal to those of the operational system (Table 1). Next, the freeway and on-ramp vehicle generators were set to generate, by a random process, the measured field rates of traffic. Third, vehicle-driver characteristics, which in the present model implicitly include factors such as curvature, vertical and horizontal alignments, grade, and lane width, were estimated and set based on the judgment of the investigators. These are shown in Tables 3 and 4. Of these, the only vehicle-driver performance characteristics that were determined experimentally were the desired free-flow velocities for freeway drivers. These were estimated using vehicle travel times obtained with a stopwatch during very light traffic conditions when vehicles were able to pass each other with no apparent interference. Finally, a passing probability function (Fig. 6), and shock wave final velocity constraints (Table 6), were established heuristically.

Validation runs were then made, beginning with the model void of vehicles. During each run, which represented 4 hours of simulated time, the average travel times of freeway drivers and on-ramp drivers were independently accumulated. An analysis of these cumulative averages, using the technique previously described in this section shows that they become very stable and hence may be assumed to be the exact mean values for travel times with negligible random error. Thus, the outcome of the simulation runs yield the numbers ξ_i and η_i, $i = 1, 2, \ldots, 18$, where ξ_i and η_i are the exact mean travel times of simulated freeway and on-ramp drivers respectively. Table 8 gives these results together with the corresponding field measurements.

From the data in Table 8, and by letting μ_i and τ_i represent the true but unknown mean travel times of actual freeway drivers and on-ramp drivers, tests can be made to determine whether the relations $\xi_i = \mu_i$ and $\eta_i = \tau_i$, $i = 1, 2, \ldots, 18$, hold for freeway drivers and on-ramp drivers respectively. For example, for freeway drivers these tests are based on the statistics $\xi_i - \bar{x}_i$, $i = 1, 2, \ldots, 18$, and the application of a central limit theorem which implies the symmetry of their distributions. Assuming these statistics to be independent, the Wilcoxon Signed-Rank Test may be used on the numbers $\xi_i - \bar{x}_i$, $i = 1, 2, \ldots, 18$, to test $\xi_i = \mu_i$, $i = 1, 2, \ldots, 18$. In this particular case, an equal-tailed significance level of 0.61 was yielded. Similarly, for on-ramp drivers, an equal-tailed significance level of 0.21 was obtained. Based on these preliminary results it is felt that the model is behaving realistically.

CONCLUSIONS

A three-year program which will lead to a valid digital simulation model for the design of diamond interchanges has been described. The program consists of a se-

quence of four models, each with its extensive concomitant validation study. The final model of the sequence is the total diamond interchange.

Version 1 of Model 1 has been programmed, coded, debugged, and some results have been obtained. These results pertain both to the operating characteristics of the model and to its validity. Based on these results, it is felt that the present macroscopic approach to the simulation has a very high probability of developing a total model with realistic performance characteristics. Accordingly, the present version of Model 1 will be extended, using the same macroscopic notions, to a second version which will simulate more realistically the formation and dissolution of congestion. This is an extremely important characteristic to incorporate into the model since the operational characteristics of interchanges under congestion situations will be of utmost importance. Models 2, 3 and 4 will also be provided with this capability.

A brief description was given of our studies on loop detectors and on the photogrammetric methods for collecting and reducing the field data that will be required in the validation-study phase for each of the models. Based on these studies, it was concluded that the ultimate validation studies will (a) use a 70-mm movie camera mounted on a helicopter for collecting the data, and (b) reduce the data with a semiautomatic film record reader.

ACKNOWLEDGMENT

The authors wish to express their indebtedness to Robert Gray for his help in evaluating data collection and reduction equipment, Vance Griffitts for his help in coding the data reduction routines, and Allan Walker for his movie work.

REFERENCES

1. Ancker, C. J., Jr., and Gafarian, A. V. A Simple Renewal Model of Throughput at an Oversaturated Signalized Intersection. Transportation Research, Vol. 1, No. 1, pp. 57-65, May 1967.
2. Stern, S. Traffic Flow Data Acquisition Using Magnetic-Loop Vehicle Detectors. Highway Research Record 154, pp. 38-52, 1967.
3. Panel on Progress in Traffic Data Instrumentation. Proc. of Highway Conf. on the Future of Research and Development in Traffic Surveillance, Simulation, and Control, U. S. Bureau of Public Roads, pp. 13-122, 1964.
4. Gerlough, D. L. A Comparison of Techniques for Simulating the Flow of Discrete Objects. Paper presented at the National Simulation Conference, Oct. 23-25, 1958, Dallas, Texas, 17 pp.

Operational Effects of Some Entrance Ramp Geometrics on Freeway Merging

JOSEPH A. WATTLEWORTH, JOHANN H. BUHR, DONALD R. DREW, and FRANK A. GERIG, JR., Texas Transportation Institute, Texas A&M University

•A wide range of geometric design standards has been used on freeways which are now open to traffic. These different geometrics as well as different traffic patterns have resulted in a considerable range of traffic operation and safety. The designs used on early freeways were influenced by experience in street design and by lack of experience in dealing with very high traffic volumes. In many cases, poor design features were accepted and built into some early freeways because it was not known what constituted good or bad design. Since operational studies were unknown, there was little feedback of operational information to the designer, and in many cases the volumes remained low enough so that the poor design features did not cause severe operational problems. In recent years, however, much more operational experience has been gained and reported. In spite of this, the design standards in some areas have failed to keep pace with the rapidly expanding body of knowledge on the operation of freeways, and as a result some outmoded concepts of design are still being used in the design of present-day freeways. This is especially true in some primarily rural areas where some rural design philosophies are carried over into urban design situations.

On most urban freeways, traffic volumes have increased considerably and are so high that operational problems are commonplace, especially during the peak traffic periods. In fact, peak-period control of urban freeways has been necessitated to a large extent to prevent the overloading of bottleneck areas, many of which are caused by poor geometric features. The order of magnitude of the increase in volumes strived for by these projects is the same as the lost capacity due to the bottlenecks, so better design could have decreased the need for peak-period operational control and the associated costs.

Closely correlated to poor operation is low safety. Geometric features which are bottlenecks can cause shock waves (poor operation) to be generated and propagated upstream and the shock waves can cause many rear-end accidents. Also, entrance ramp designs which cause a high percentage of merging vehicles to stop before merging generally have a high accident experience as well. Many other examples showing the correlation of operations and safety could be cited.

In spite of the knowledge gained from previous studies (1, 2, 3) of entrance ramp operation, there was a need for research of a more comprehensive nature to define more clearly the operational effects of some entrance ramp geometrics, the most important of which are acceleration lane length, angle of convergence of the ramp with the freeway, and the ramp grade. One reason for this is emphasis. Despite the suggested ramp designs based on these past studies, there is still a tendency on the part of some designers to compromise from good designs because of other (sometimes very important) considerations, such as cost, right-of-way restrictions, desire for frequent access points or, in many instances, still a carryover of past design standards which may be believed to "work all right." This carryover may in most cases be due merely to a lack of adequate traffic volumes to test the operational capabilities of the designs or may be due to a lack of cognizance on the part of the designer of the operation of

Paper sponsored by Committee on Operational Effects of Geometrics and presented at the 46th Annual Meeting.

some of his designs. This report places emphasis on the correlation of design and operations which will hopefully lead to improved design on future freeways.

Another area of freeway design which is receiving attention is redesign or remodeling of existing freeways. Many of the older designs have led to poor operation and/or high accident rates, and it is realized that, along with the need for construction of new freeways, important benefits to the motoring public can be realized by modification of design on existing freeways to eliminate bottlenecks and hazards. The Los Angeles District of the California Division of Highways recently established a Freeway Operations Department (4) which is responsible for the selection of locations on the Los Angeles freeways that most need redesign or traffic control. The present study will help (a) determine which entrance ramp designs warrant remodeling; (b) determine the type of improvement which would best alleviate the problem; and (c) determine the amount of operational improvement or benefits to motorists to be expected from the remodeling.

In the area of ramp control, the data on capacity and service volumes related to geometrics should be quite helpful. There are many ways to operate ramp controls and the geometrics have never been specifically considered in any of the control theories except through the effect on traffic parameters such as capacity. This study of different types of merging areas should provide many insights into the effects of geometrics on the ramp control requirements.

Many freeway simulation programs have been developed, but their testing and calibration have been retarded by a lack of adequate field data on gap acceptance characteristics and effect of geometrics on the operational characteristics. This study will help to answer some of the need for information on the effects of geometrics on the behavior of vehicles in the freeway merging area.

OBJECTIVES

The general objective of this study was to determine the effects of entrance ramp geometrics on the traffic interaction in the freeway merging process. More specifically, the objective was to study the effects of the following specific geometric variables on entrance ramp operation: acceleration lane length, angle of convergence, and ramp grade. Most of the analyses presented in this report concentrate on the first two variables.

To accomplish this objective, it was necessary to conduct studies at many entrance ramps in order to include a fairly wide range of the geometric elements. Studies were made at 29 entrance ramps (Appendix) in Houston, Los Angeles, San Francisco, Sacramento, Chicago, Detroit, St. Louis, and New York City. The acceleration lanes of the ramps studied ranged from 240 to 1500 ft and the convergence angles ranged from 1 to 14 deg. Data from 23 entrance ramps are presented in this report. A technique involving time-lapse photography from a circling airplane was developed (5) for these studies so that a fairly long section of freeway and merging area could be filmed continuously in time. One to five rolls of film, each covering a 24-min period, were exposed at each study location.

STUDY PROCEDURES

To study the merging operation microscopically at about 40 entrance ramp locations throughout the country, a special aerial photographic study technique was developed. A 35-mm time-lapse movie camera was used with a film speed of 5 frames per second. In order to obtain a continual coverage of traffic conditions, the area of interest was circled by a Cessna 206 airplane from which the time-lapse photographs were taken. The radius of the circle was about ¼ mile and the altitude and lens were selected (a) to obtain adequate coverage of the merging area, (b) to locate the plane at the altitude of least air turbulence, and (c) to locate the airplane below any clouds. Figure 1 shows the camera mounted in the plane. The airplane made a complete circle around the study site about once every 40 sec. The rotation of the scene during analysis was not a problem because a stop-frame analysis was used.

Reference marks were placed at 200-ft intervals on the right shoulder of the freeway and on the ramp shoulder. A reference mark was placed at the physical nose of the

Figure 1. Camera mounted in aircraft.

ramp and the marked area extended 400 ft upstream of the physical nose and downstream to a point past the end of the ramp taper. Figure 2 shows a typical reference configuration at an entrance ramp.

A film speed of 5 frames per second was selected as the minimum speed which would provide the necessary accuracy of measurements. A lower film speed is more desirable from several points of view; namely, lower film cost and lower film analysis cost and a longer filming period could have been used at each ramp. In each frame the freeway scene was accompanied by a data chamber showing the time of day, the frame number, and a data slate containing information on the ramp name, film number, etc. Figure 3 shows a sequence of four frames taken at four points on the flight circles.

Because of the large amount of film obtained, analysis of the data became somewhat of a problem. During one period, three 8-hour shifts on each of three projectors were used to expedite the analysis. The data taken from the film included for each vehicle: (a) the vehicle type, (b) the location of the vehicle (ramp or freeway lane), and (c) the time the vehicle passed each reference mark. The data were recorded in a format to facilitate direct key punching on IBM cards, and all data processing was done on a large-scale computer. A companion report (5) contains further details on the procedures used in the study.

SELECTION OF STUDY LOCATIONS

The study locations were selected to meet several geometric and traffic requirements. In the selection of the ramps studied, pure merging situations were sought, i.e., locations at which the traffic operation was unaffected by upstream or downstream conditions. Merging operation at or near capacity was also desired. Local

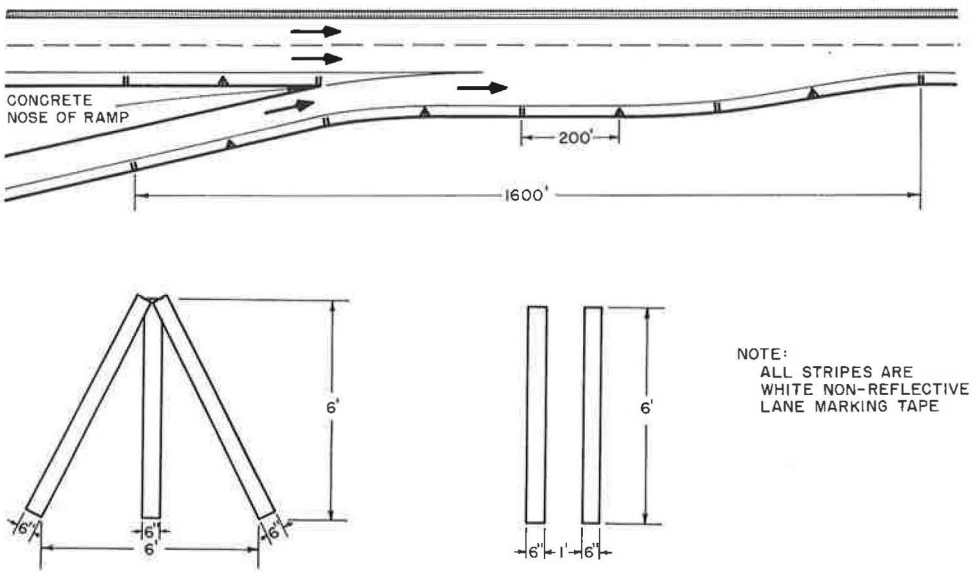


Figure 2. Typical study section.

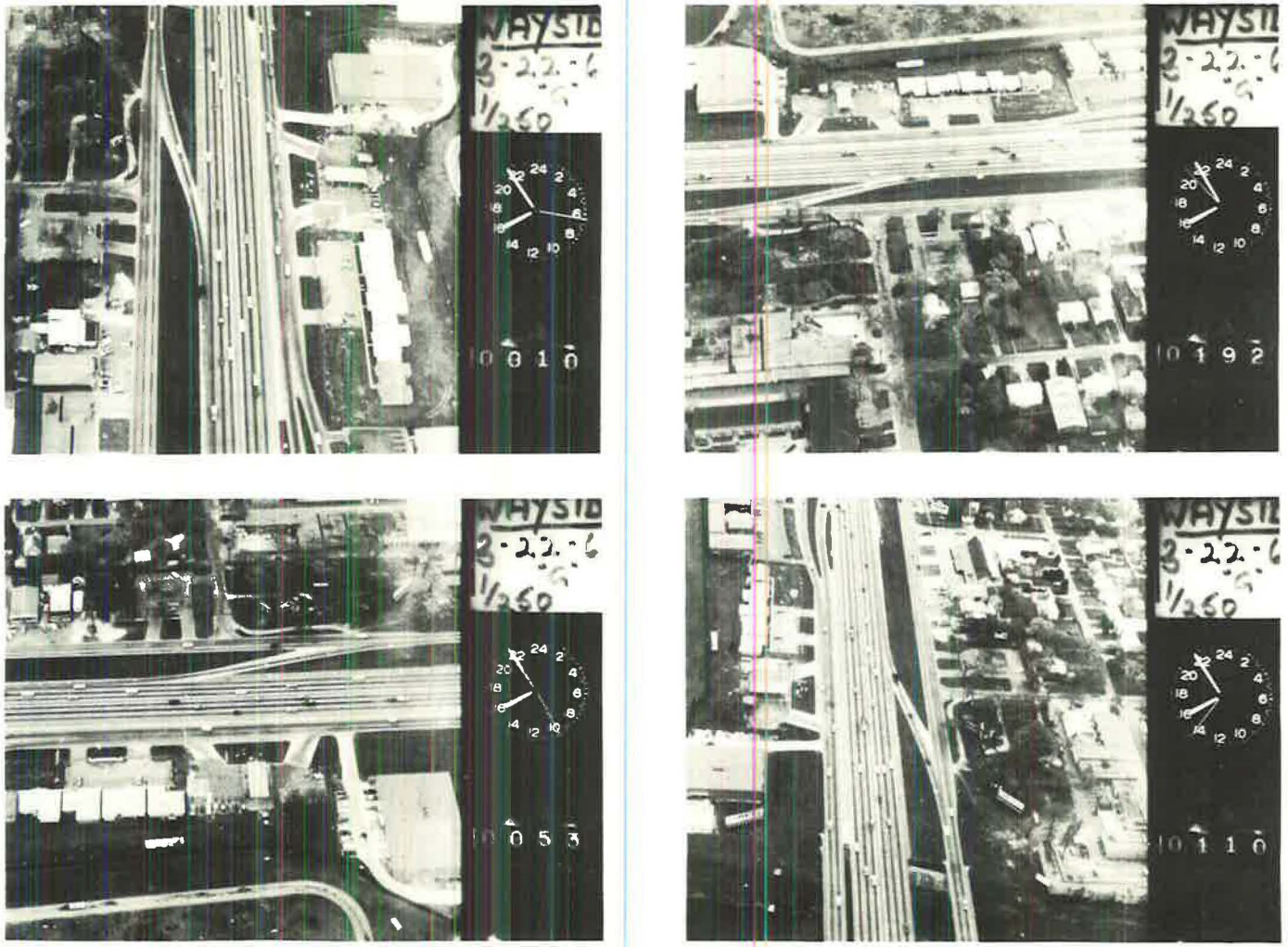


Figure 3. Typical frames from data film.

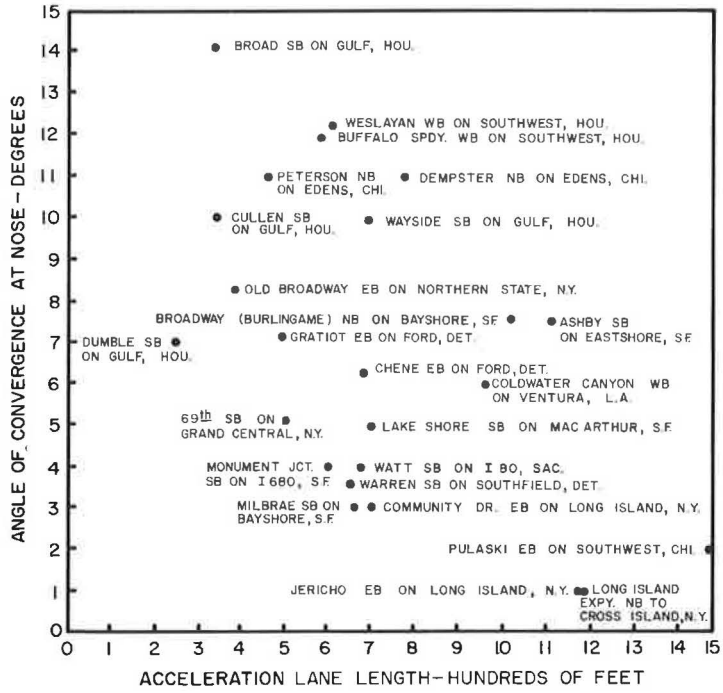


Figure 4. Summary of acceleration lane lengths and convergence angles.

authorities were extremely helpful in suggesting study sites and in providing volume and geometric data, as well as plan and profile sheets for possible study sites which were used in the final selection of sites. Before any filming was done at a location, the operation there was field checked to determine the approximate volumes and to evaluate the possible effect of any downstream restrictions. In spite of this careful checking, the operation during the film studies was not always the same as during the field checks because of changes in traffic patterns, accidents, or some other temporary situation.

One of the objectives in selecting study sites was to obtain a wide range of geometrics in the ramps studied: particularly, variations in acceleration lane length, angles of convergence, and ramp grade. Locations having straight, level freeway alignment were of primary interest, since the effect of freeway geometrics was not being studied. Figure 4 shows a geometric summary of the 24 entrance ramps for which data are included in the report. The coordinates consist of the acceleration lane length and angle of convergence. Figure 4 shows that a considerable distribution of ramp geometrics was obtained among the ramps selected. Data at several other entrance ramp locations were collected but are not presented because accidents or other downstream interference caused primarily congested merging operations.

The study period was selected on the basis of the traffic operations during the time of the field checking. The time was selected so that free flow merging operation prevailed and so that the merging volume was near capacity.

Entrance ramps included in the study are listed according to the metropolitan area in which they are located and not according to the exact suburb. For example, the Broadway northbound entrance ramp on the Bayshore Freeway is located in Burlingame, California, but is listed as being in the San Francisco area.

RAMP GEOMETRICS

The following geometric data were obtained for each of the entrance ramps which were studied:

1. Length and shape of acceleration lane,
2. Angle of convergence,
3. Ramp grade,
4. Freeway grade,
5. Length of ramp,
6. Width of acceleration lane,
7. Number of ramp and freeway lanes,
8. Freeway and ramp curvature, and
9. Curb offset at the nose.

The length of the acceleration lane, L, was taken to be the distance from the physical nose to the end of taper. Because of different offset distances of the curb at the physical nose the length of two acceleration lanes may not be directly comparable. For example, a 600-ft acceleration lane on a ramp at which the offset at the nose is 10 ft is not equivalent to a 600-ft acceleration lane on a ramp at which the offset is 2 ft. In the former case, a portion of the acceleration lane would be used by vehicles in moving laterally to the same position as would be maintained by vehicles at the nose of ramp with a 2-ft offset. Therefore, the effective length of acceleration lane on the two ramps would be different. The difficulty in defining and measuring a length of acceleration lane which would be comparable for any curb offset made it necessary to use the distance from the physical nose to the end of taper as the acceleration lane length.

Similarly, the offset distance of the curb at the physical nose has an effect on the effective angle of convergence. In order to avoid use of a misleading angle, the angle of convergence θ was measured both at the physical nose and, for the ramps with curb offsets greater than 2 ft, at a point where the left edge of the ramp (as delineated by the paint markings) was 2 ft off the edge of pavement of the freeway. Both of these angles are discussed in the analyses which follow.

Most of the geometric data were obtained from information furnished by local agencies. The form of the information supplied varied with the agencies and ranged from aerial photographs to construction drawings. The geometric data were read as accurately as possible from the plans and photographs provided, but in some cases the scale of the drawings furnished made accurate determination of the convergence angle, acceleration lane length, etc., difficult. Where such difficulties were encountered, the local agencies were asked to check the geometric data obtained from the plans.

REDUCTION AND ANALYSIS OF THE DATA

Film Reduction

The aerial, time-lapse photographs were analyzed using a vehicle-by-vehicle type of data reduction to obtain the relevant data for each vehicle on the entrance ramp or in the right lane of the freeway. Each vehicle on the ramp was traced through the merging area, and the frame number corresponding to the time the front end of the vehicle passed each reference mark (Fig. 2) was recorded. When the vehicle entered the freeway, it was treated as a freeway vehicle in the remainder of the merging area. Vehicles in the right lane of the freeway were also traced through the merging area in a similar manner.

The reference marks were normally placed 200 ft apart so that the speed of each vehicle between each pair of reference marks could be obtained. The time which each vehicle passed each reference point was known; thus, it was quite simple to obtain a time-space trajectory for each vehicle in the merging area of each ramp studied. For each roll of film analyzed, a time-space plot was made using a Cal-Comp plotter driven by an IBM 1401 computer. The time-space trajectories of vehicles on the freeway were plotted in solid lines and the trajectory of vehicles on the ramp were plotted in

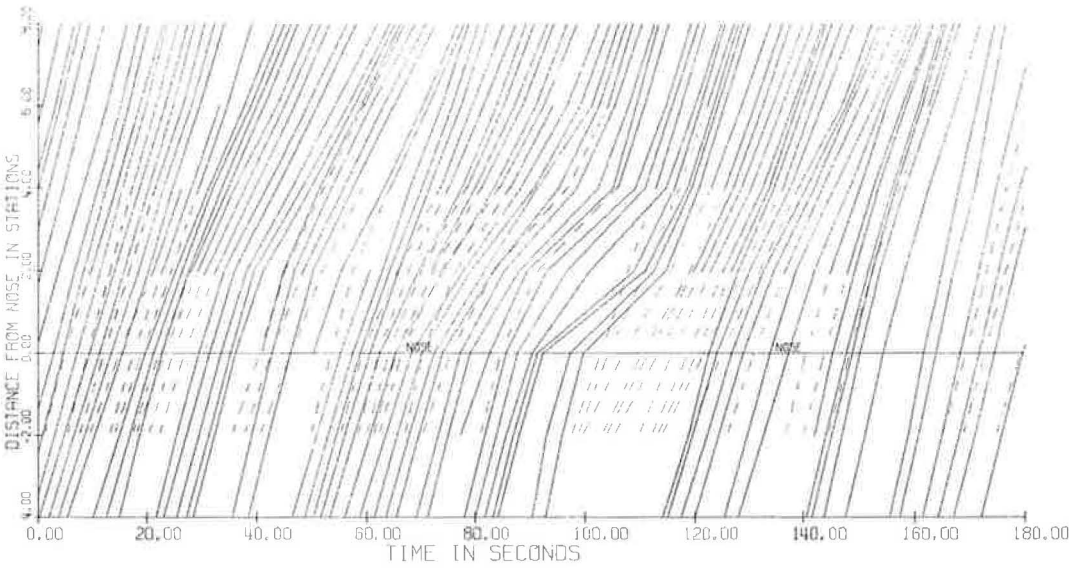


Figure 5. Time-space diagram: Broadway northbound entrance ramp, Bayshore Freeway, San Francisco. Dashed lines represent vehicles on ramp, solid lines represent vehicles on freeway.

dashed lines. The point at which a dashed line turned into a solid line on the time-space diagram indicated that a vehicle from the ramp had merged onto the freeway at that point. Figure 5 shows a sample of the time-space diagrams. Note that the only vehicles on the freeway which are included in this drawing are those on the right lane. Further details of the film reduction process are given elsewhere (5).

Analyses

Many analyses were made of various operational characteristics. Most of these were made for each of three levels of freeway operation based on freeway speeds less than 25 mph, between 25 and 40 mph, and above 40 mph. In this report the characteristics presented are limited primarily to those for which the freeway speeds were above 40 mph. The reasons for this are that the operations above 40 mph are the type intended when the freeways were designed, and that the operation at this level is more variable than in the other, since the drivers' decisions are less constrained.

Speeds of Ramp Vehicles at Ramp Nose

For each of the entrance ramps included in the study, the distribution of speeds of vehicles on the ramp over the 200-ft section upstream of the ramp nose was obtained. The speed of a vehicle on a ramp approaching the nose can be affected by several factors. Among them are (a) freeway traffic conditions, (b) ramp alignment, (c) ramp grade, (d) ramp angle of convergence, and (e) acceleration lane length. The last two factors affect the speed at the nose because they affect the gap acceptance characteristics and hence the (queuing) service rate on the ramp. At entrance ramps with minimum design standards, the merging maneuver is more difficult for the motorists and therefore the approach speeds tend to be lower than on ramps with better designs. A great deal of interference between ramp vehicles is noted on poorly designed ramps because of the relatively high frequency of very low-speed merges and the arrival of vehicles on the ramp while others are stopped waiting to merge. Considerable queuing of ramp vehicles is noted at this type of ramp.

Most of the ramps studied had a ramp alignment which would not cause a severe reduction in approach speeds. Most were diamond type ramps, cloverleaf outer connectors or direct connections, thus, the alignment was fairly good in all cases. The effect of different freeway conditions was accounted for by considering the three speed ranges on the freeway. It would have been more desirable to have grouped the data by ramp and freeway volumes, such as was done by Fukutome and Moskowitz (3), instead of by freeway speeds alone, but the sample sizes of the data did not permit such a classification.

To show the effect of ramp angle of convergence on the merging operation, the ramp speed distribution was plotted for ramps in four ranges of acceleration lane lengths (0-350 ft, 400-600 ft, 600-800 ft and 950-1200 ft). The distribution of ramp speeds was obtained only for vehicles which arrive at the ramp nose when freeway speeds are over 40 mph. Within each range of acceleration lane lengths, differences in ramp speeds are primarily caused by different angles of convergence.

Similarly, the effect of acceleration lane length on ramp speeds was obtained by examining speed distributions for ramps within three ranges of convergence angle (0-3 deg 30 min, 6-8 deg and 10-14 deg). Within each range of convergence angles, the differences in ramp speeds are primarily caused by differences in acceleration lengths.

Several analyses involving a single statistic (such as the mean of a distribution) were made by writing the value of the statistic near a point representing an entrance ramp on a special data paper. The point for each ramp was located according to the angle of convergence and acceleration lane length. The special data sheets are similar to Figure 4, with the ramp names omitted. The mean ramp speed at the nose for periods during which freeway speeds were above 40 mph was treated in this manner.

Speeds of Ramp Vehicles at Merge Point

The distributions of the speeds of ramp vehicles in the 200-ft distance in which the ramp vehicle actually merges were treated in the same manner as were the distributions of ramp speeds at the nose. For freeway speeds over 40 mph, the distributions of merging speeds were plotted vs acceleration lane length for various ranges of convergence angles, and were plotted vs convergence angles for various acceleration lane lengths. Also the mean merge speed for each ramp (for freeway speeds over 40 mph) was plotted against the angle of convergence and acceleration lane length for the ramp.

Speed Changes on the Acceleration Lane

The speed change for each ramp vehicle between the ramp nose and the point at which it merged onto the freeway was computed during times when the freeway operating speed was over 40 mph. The distribution of these speeds is an indication of the efficiency of the operation on the acceleration lane. These distributions (one for each ramp studied) are presented as functions of convergence angle and acceleration lane length in the same manner as were the distributions of speed at the ramp nose and at the merge points. The mean speed change is also shown for each ramp by acceleration lane length and convergence angle.

Relative Speed

In an ideal merge, the relative speed between a vehicle on the ramp and vehicles on the freeway is quite low at each location on the acceleration lane. A vehicle on the ramp should pass the ramp nose at a speed approximately the same as the operating speed on the freeway. It should then adjust its speed to reach the gap selected and should enter that gap at the same speed as that of the lag vehicle.

High relative speeds between vehicles generally indicate poor operation on any facility; a high relative speed between vehicles on the ramp and those on the freeway indicates poor merging operation. When freeway speeds are high (over 40 mph), a high relative speed is generally caused by the geometric design of the ramp causing low-speed operation on the ramp.

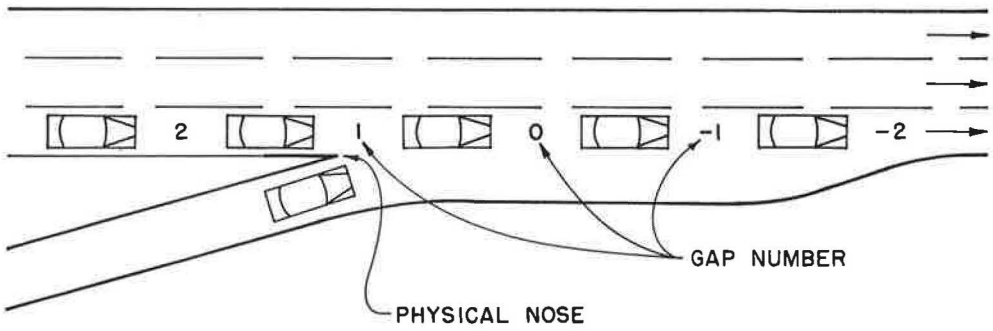


Figure 6. Number of gaps in the right lane.

The relative speed between each ramp vehicle and its lag vehicle was computed at several locations along the merging area. The first speed trap is the 200-ft length upstream of the ramp nose, and each 200-ft trap along the acceleration lane was used to compute relative speed. In addition, for each ramp vehicle the relative speed was computed at the actual trap in which the merge maneuver was completed. The distributions of these speeds are presented as functions of acceleration lane length and angle of convergence in a manner similar to the presentation of the other distributions; the means of the distributions are also shown as a function of these two geometric variables.

Accepted Gap Number

Another indication of the operational efficiency of a merging area is the gap into which the ramp vehicles merge. If the vehicles on the ramp normally select the first or second gap, the operation is better than it would be at a location at which 10 or 15 gaps pass by before the vehicle on the ramp finds an acceptable gap. The gap numbers are defined according to their locations at the instant the front of the ramp vehicle reaches the physical nose (Fig. 6). The gap that is adjacent to the ramp vehicle when it is at the nose is gap 1; the gaps are numbered from downstream to upstream.

The mean number of the accepted gap and its standard deviation were computed for each ramp and each of the three freeway speed ranges. Some of these are presented as functions of acceleration lane length and convergence angles. Gap acceptance is studied in much more detail in a companion report (6).

Acceleration Lane Use

The use of the acceleration lane was also investigated at each of the entrance ramps. For each, a distribution of encroachment points (the location relative to the physical nose at which the ramp vehicle first encroaches on the freeway) was determined. These are presented as functions of angle of convergence and acceleration lane length in order to show the effects of these geometric variables. In addition, mathematical distributions were fitted to the actual distributions.

RESULTS

Speeds of Ramp Vehicles at Ramp Nose

Ideally, the geometrics of an entrance ramp should not prevent a vehicle which is about to merge onto a freeway from attaining a speed nearly equal to the operating speed on the freeway when the vehicle passes the ramp nose. Under some freeway operating conditions, namely congestion, this is not always possible. However, when the freeway speeds are greater than 40 mph, it should be possible for vehicles on the ramp to pass the nose at high speed.

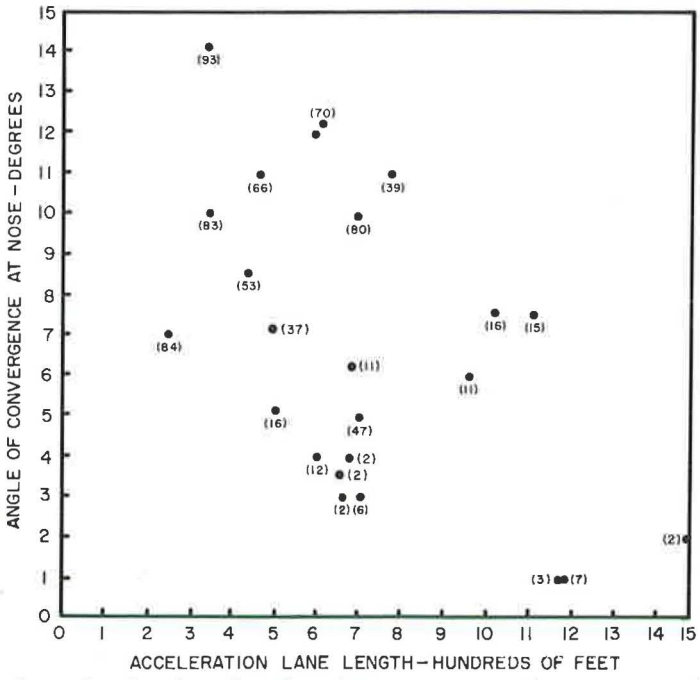


Figure 7. Effect of acceleration lane length and convergence angle on percent of ramp vehicles traveling less than 30 mph at the ramp nose.

In Figure 7, the number beside each point is the percent of ramp vehicles traveling less than 30 mph at the ramp nose when the freeway speeds are over 40 mph. Each point represents an entrance ramp and the coordinates of the point represent its acceleration lane length and angle of convergence. The effects of these geometric variables can be clearly seen in this figure. The percent of slow ramp vehicles can be seen to increase quite rapidly for ramps with angles of convergence greater than about 6 deg. The percent of ramp vehicles passing the ramp nose at speeds less than 30 mph ranges from a low of 2 for three ramps with acceleration lanes between 600 and 700 ft and with angles of convergence between about 2 to 4 deg to a high of 93 percent for a ramp with an acceleration lane of 335 ft and a 14-deg angle of convergence.

To show the effect of convergence angle on the ramp speed at the nose, the ramps studied were classified according to length and the cumulative ramp speed distributions were plotted for all ramps falling in each length class. The four length classifications which were selected are 0-350 ft, three ramps; 400-500 ft, four ramps; 600-800 ft, ten ramps; and 950-1200 ft, five ramps. Each distribution represents the speed of ramp vehicles at the ramp nose when freeway speeds were 40 mph or higher. Figure 8 shows these distributions.

Figure 8a shows the ramp speed distributions for three entrance ramps with acceleration lanes less than 350 ft. The speeds at the ramp on which $\theta = 14$ are generally about 10 mph lower than on the ramp with a smaller convergence angle of 10 deg. The lengths of these two ramps are nearly equal so that the difference in speeds can be attributed almost entirely to the difference in convergence angles. The speed distribution of the third ramp on which $\theta = 7$ deg falls between the distributions for the other two ramps. Based on convergence angle alone, it would have been expected that the speeds on this ramp would be higher than on either of the other two ramps. However, this ramp has a shorter acceleration lane and much higher ramp volumes than either of the other two ramps, and these factors account for the speeds at this location being

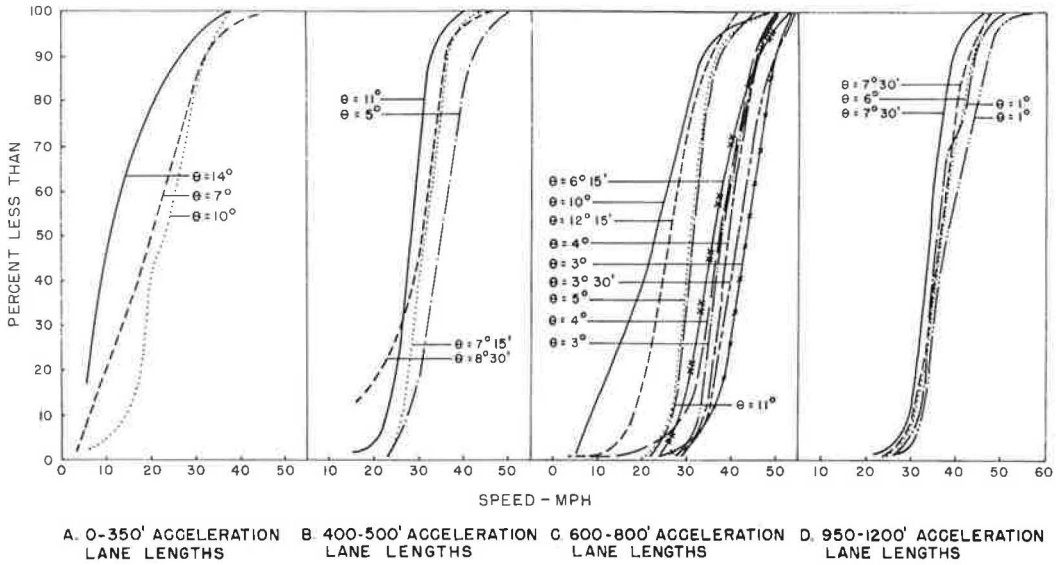


Figure 8. Effect of angle of convergence on ramp speed distribution at ramp nose.

lower than expected. That the operation at all three of these ramps is bad can clearly be seen by their speed distributions. The speed distributions in Figure 8 contain only data for times when the freeway speeds were over 40 mph.

Figure 8b shows the speed distribution for ramp vehicles at the nose of each of four entrance ramps which have acceleration lane lengths between 400 and 500 ft. The angles of convergence of these ramps range from 5 to 11 deg and the operational effects of the convergence angle can be clearly seen. The ramp with the lowest speeds is the ramp θ = 11 deg (66 percent of ramp speeds less than 30 mph) followed by the ramps with θ = 7 deg (41 percent less than 30 mph) and θ = 8 deg 30 min (50 percent less than 30 mph), while the ramp with the lowest convergence angle (5 deg) has the highest speeds, with only 16 percent of speeds below 30 mph.

Thus, for entrance ramps with short (400-500 ft) acceleration lanes, a change in convergence angle from 11 deg to 5 or 6 deg would result in about 50 percent more of the entrance ramp vehicles passing the ramp nose at speeds greater than 30 mph at times when freeway operating speeds exceed 40 mph. This change would also result in an increase in the average speed of ramp vehicles at the nose of about 8 mph and this would reduce the relative speed (difference in speed of vehicles on the freeway and on the ramp) there by about the same amount. The relative speed at the nose is more critical for entrance ramps with short acceleration lanes, since, on this type of ramp, there is little opportunity for acceleration after passing the ramp nose.

In Figure 8c, the same type of speed distributions is plotted for entrance ramps with acceleration lane lengths between 600 and 800 ft. There appear to be two clusters of data—one for the six distributions to the right of the figure and the other for the four distributions to the left. The six ramps which correspond to the higher speeds of ramp vehicles passing the nose have convergence angles ranging from 3 deg 30 min to 6 deg 15 min. The median speeds at these ramps are between 38 and 44 mph and the percent of speeds less than 30 mph ranges from 2 to 15 percent. For the other four ramps (θ = 5 deg, 10 deg, 11 deg, and 12 deg 15 min), the median speed ranges from 22 mph to 31 mph, while the percent of speeds below 30 mph is between 39 percent and 80 percent. Thus, the operation on the six ramps with the lower convergence angles is much better than on the other four ramps.

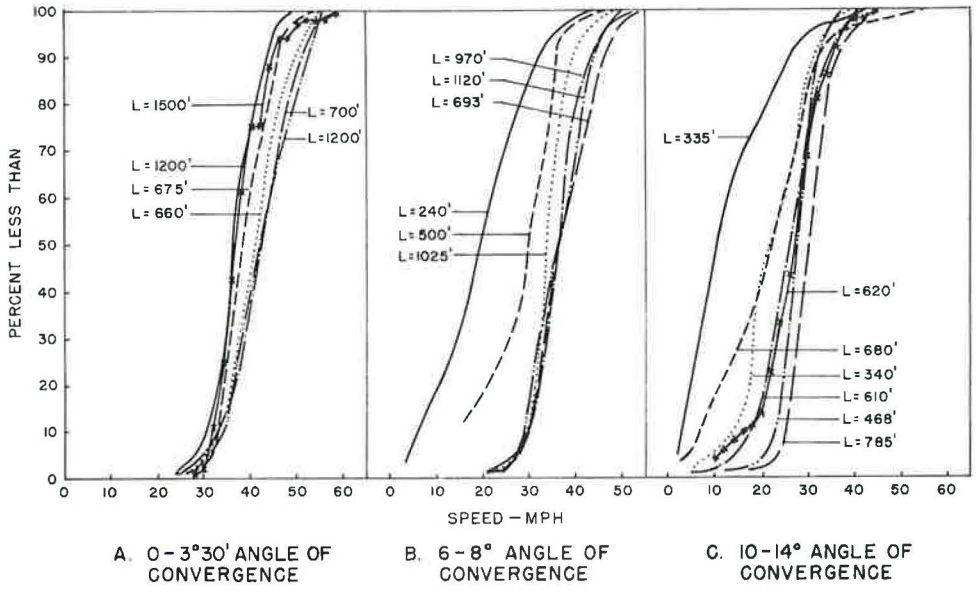


Figure 9. Effect of acceleration lane length on ramp speed distribution at ramp nose.

One of the ramps on which the speeds were relatively low had an angle of 5 deg. One would not expect to find such a large difference in operation between the ramp with a 5-deg convergence angle and the other ramps with generally lower convergence angles. The operational difference probably reflects the extremely high volume (1000-1200 vph) on the ramp with θ = 5 deg, rather than the difference in convergence angles.

Similar speed distributions for five entrance ramps with long (950 to 1200 ft) acceleration lanes are shown in Figure 8d. For these ramps, θ varies between 1 deg and 7 deg 30 min. The operation at these ramps was generally similar. On each ramp the percent of speeds less than 30 mph was less than 25 percent and the median speed ranged from 35 to 38 mph. The percent of ramp vehicles passing the ramp nose at 40 mph or faster is 10 percent and 26 percent for the two ramps with θ = 7 deg 30 min and is a minimum of 40 percent for the three other ramps (θ = 1 deg, 1 deg and 6 deg). Thus, the traffic operation was good on each of the ramps on which the acceleration lane length is at least 950 ft and the convergence angle is less than 7 deg 30 min.

In order to show the effect of acceleration lane length on the ramp speeds at the nose, the entrance ramps studied were grouped according to convergence angles with the angle categories being 0-3 deg 30 min, 6-8 deg and 10-14 deg. The speed distribution of ramp vehicles passing the ramp nose for each ramp was plotted in its proper angle group (Fig. 9). For ramps with approximately the same convergence angle, the effect of different acceleration lane lengths can be observed. All ramp speed distributions in Figure 9 represent times during which the freeway operating speed was 40 mph or higher.

Figure 9a contains the speed distributions. The acceleration lane lengths vary from 660 to 1200 ft. As might be expected of these ramps which all have low convergence angles and at least an adequate (600 ft) acceleration lane length, the speed distributions are fairly closely clustered. The median speeds range from 36 mph to 43 mph, and not more than 7 percent of the vehicles passed the nose at less than 30 mph on any of the ramps. The speeds generally increase as the acceleration lane increases with one exception—one ramp with a 1200-ft acceleration lane has the lowest speeds. The volume on this ramp during the periods studied ranged from 1200 to 1600 vehicles per hour and the total right lane merge volume was generally in excess of 2000 vph. Thus,

the very high ramp volume probably caused the ramp speeds to be lower at this location and, if the acceleration lane had been much shorter, the ramp speeds would probably be a great deal lower.

Because of a greater range of acceleration lane lengths, the effect of this geometric element can be seen quite clearly in Figure 9b. Here the distribution of speeds or ramp vehicles passing the ramp nose is shown for six ramps having a convergence angle lying between 6 and 8 deg. The acceleration lanes vary in length from 240 ft to 1120 ft. The ramp which experienced the worst operation had an acceleration lane length of 240 ft. The median speed for this ramp was 19 mph, whereas 84 percent of the ramp speeds were less than 30 mph. The distributions for the other ramps are fairly closely clustered with the median speeds ranging from 34 mph to 37 mph and from 11 to 16 percent of the speeds below 30 mph.

On one ramp the operation was surprisingly good. This ramp has a θ of 6 deg 15 min, an L of 693 ft and was a diamond interchange ramp of the direct taper type. High merging volumes (1700-2200 vph) were maintained in the right freeway lane with about 500-700 vph on the ramp. The speeds at this location are as high or higher than those of any other ramp in the 6-8 deg group, in spite of its comparatively short acceleration lane length. At this location the ramp itself is on a 4.1 percent downgrade preceding the nose for a distance of about 450 ft. In this long distance, vehicles have a good view and perspective of the freeway and are able to select a gap before reaching the physical nose and, while still on the ramp, can accelerate in order to meet the gap when it reaches the relatively short acceleration lane. This demonstrates the value of the view of the freeway which gives drivers on the ramp the ability to determine the lane in which the gap is located, and the speed of the gap before reaching the acceleration lane.

In Figure 9c, similar speed distributions are shown for five entrance ramps for which θ lies between 10 and 14 deg. The median speeds for these ramps range from 11 mph for a ramp which has an acceleration lane length of 335 ft to 31 mph for a ramp with L = 785 ft. On the ramp with L = 785 ft, 39 percent of the vehicles pass the nose at speeds of 30 mph or less, whereas for the other ramps this percent ranges from 66 percent to 93 percent, indicating a much poorer operation.

The ramp with L = 680 ft (θ = 10 deg) has a poorer operation than might be expected considering the acceleration lane length. The width of acceleration lane at this location,

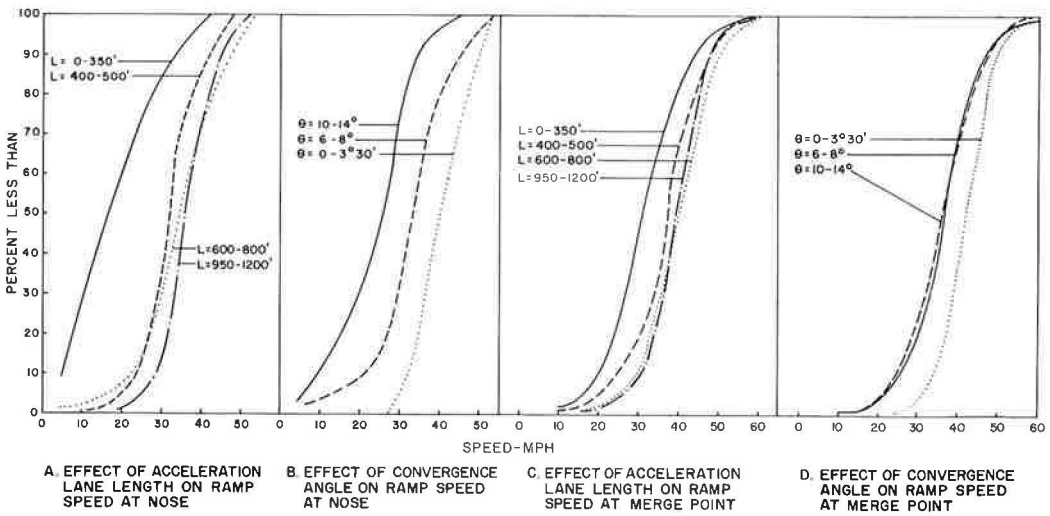


Figure 10. Effect of acceleration lane length and convergence angle on speeds of ramp vehicles at nose and at merge point.

however, is only about 9 ft and this probably contributes to the poorer operation. This location presents another interesting study of perspective. Here the freeway and ramp are both level and at the same grade. With the high convergence angle and narrow acceleration lane it is difficult for drivers on the ramp to distinguish that there is an acceleration lane which is available to them. A wider acceleration lane and a higher radius curve of transition on the right edge of the ramp and the acceleration lane (making possible a higher speed from the ramp to acceleration lane) would undoubtedly improve the operation at this ramp.

For each of the four acceleration lane length groups used in Figure 8, the average speed distribution for ramp vehicles passing the nose was plotted and is shown in Figure 10a. In this figure, the effect of acceleration lane length can be clearly seen. Speeds of vehicles on the ramps with acceleration lanes less than 350 ft are much lower than those of the other ramps, and the speeds on the ramps with acceleration lanes between 400 and 500 ft are somewhat less than the speeds for the ramps with acceleration lanes longer than 600 ft.

In Figure 10b, the average ramp speed distribution at the nose is plotted for the ramps in the three groups based on convergence angles. The speeds on the ramps with angles of convergence less than 3 deg are much higher than the speeds on the ramps with convergence angles greater than 6 deg. The lowest speeds were observed on the ramps with θ between 10 and 14 deg. The effect of the angle of convergence on ramp speeds at the nose can clearly be seen in Figure 10b.

Speeds of Ramp Vehicles at the Merge Point

The merging speed of the vehicles from the ramp during periods of free flow on the freeway has a great effect on the operation in the merging area. Normally, higher merging speeds of ramp vehicles indicate a better operation. The speed of each ramp

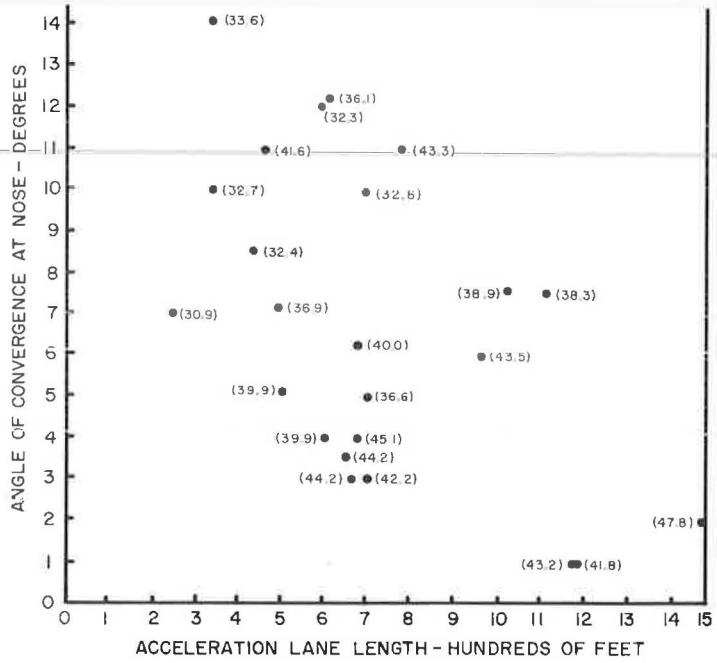


Figure 11. Effect of acceleration lane length and convergence angle on mean merge speed of ramp vehicles.

vehicle on each ramp was obtained at the 200-ft section in which the ramp vehicle first encroached on the right lane of the freeway. This was the vehicle merging speed; various analyses of the merging speeds were made and several are presented. Only periods in which the freeway operating speed exceeded 40 mph were included in these analyses.

The mean merging speed on each of the ramps included in the analyses is shown in Figure 11. The general trend is that the mean merging speeds tend to decrease as the acceleration lane lengths decrease and as the angles of convergence increase. The mean merging speed for all entrance ramps with a convergence angle less than 4 deg was about 40 mph or higher. In this group no ramp has an acceleration lane less than 600 ft and, with a convergence angle less than 4 deg, the merging speeds would be expected to exceed 40 mph when the freeway speeds were also over 40 mph.

The effect of angle of convergence on the distribution of merging speeds of ramp vehicles is shown in Figure 12. For acceleration lanes less than 350 ft in length (Fig. 12a), there seems to be little effect of convergence angle, at least over the range included (7 to 14 deg). For this range of convergence angles, it appears that the short acceleration lane length has the most important effect on merging speeds.

A somewhat divergent pattern was observed in Figure 12b. For the ramps studied with acceleration lanes between 400 and 500 ft, a distinct relationship between merge speeds and convergence angle was not apparent.

A better relationship between merge speeds and convergence angles can be seen in Figure 12c for the ramps with acceleration lane lengths between 600 and 800 ft. The two ramps with the lowest merge speeds have θ = 10 deg and θ = 12 deg 15 min. Of the five ramps which have the highest merging speeds, four have convergence angles between 3 deg 30 min and 4 deg. The fifth, however, has a large convergence angle (11 deg). It appears then that, in general, higher merge speeds can be expected at ramps which have a small convergence angle.

A similar trend was observed for the five entrance ramps with acceleration lanes between 950 and 1200 ft. The two ramps with the highest convergence angle (θ = 7 deg 30 min at both) can be seen in Figure 12d to have the lowest merge speeds. On the other three ramps, which have considerably higher merge speeds, the angles range from 1 to 6 deg. It would appear that the effect of convergence angle on merge speeds is greater on entrance ramps with longer acceleration lanes.

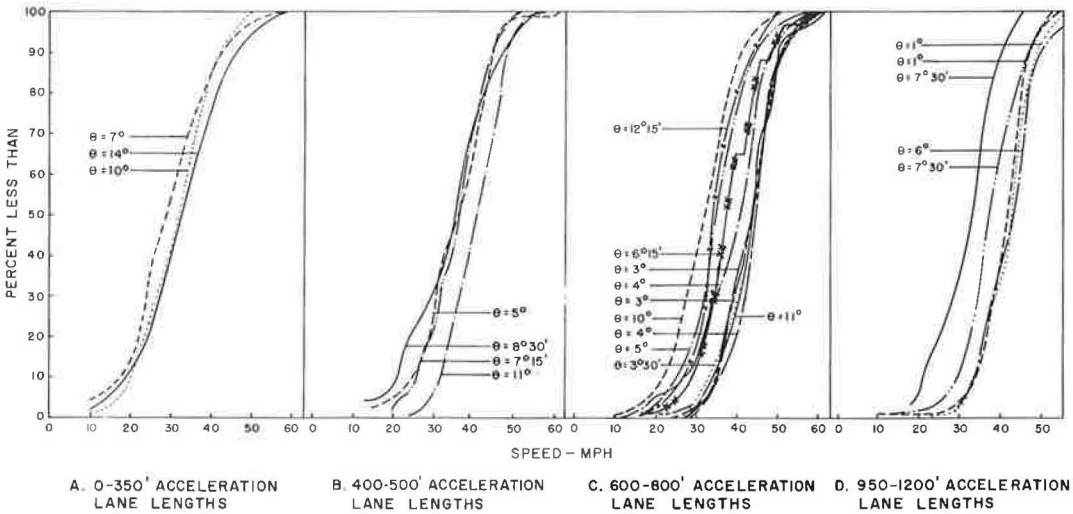


Figure 12. Effect of angle of convergence on merging speed of ramp vehicles.

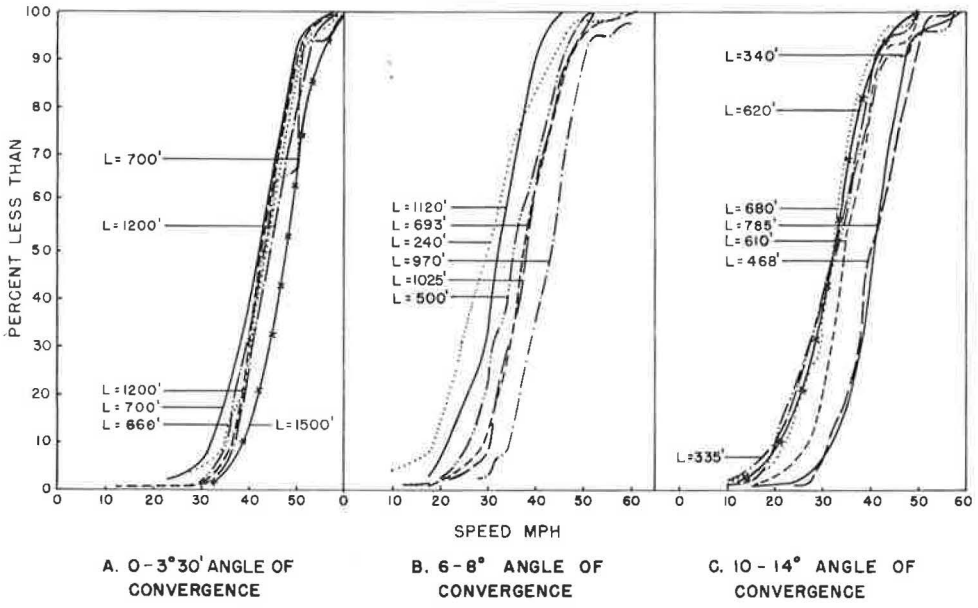


Figure 13. Effect of acceleration lane length on merging speeds of ramp vehicles.

For entrance ramps grouped by convergence angle, the effect of acceleration lane length on merge speeds can be seen in Figure 13. The distributions of merge speeds for five ramps which have convergence angles less than 3 deg 30 min is shown in Fig-

the highest merge speeds. The consistency of these distributions and the high merging speed indicate that good merging operation can be expected on ramps with a small convergence angle and at least 650 ft of acceleration lane.

The pattern is more mixed at the ramps with convergence angles between 6 and 8 deg, although a higher speed merging operation tends to be associated with ramps with longer acceleration lanes. The two ramps with the highest merging speeds have long acceleration lanes (970 and 1025 ft), while two of the three ramps with the lowest speeds have short acceleration lanes (240 and 500 ft). For ramps with convergence angles between 10 and 14 deg (Fig. 13c), no apparent patterns were observed.

When these same data which were presented in Figures 12 and 13 are averaged, the trends become much more clear. In Figure 10c and d, the average distribution of merge speeds for all ramps in each of the four acceleration lane length categories and each of the three convergence angle categories is plotted. In Figure 10c, the lowest merge speed distribution is associated with ramps with acceleration lanes less than 350 ft. Highest merge speeds are related to ramps with acceleration lanes longer than 600 ft.

Figure 10d shows that entrance ramps with convergence angles less than 3 deg can be expected to have higher merging speeds than ramps on which the angle of convergence is greater than 6 deg. Again, it can be seen that the best operation can be expected on entrance ramps with an acceleration lane length in excess of 600 ft and with a convergence angle less than 6 deg.

Speed Changes on the Acceleration Lane

The mean speed change on each acceleration lane was determined and is equal to the difference in speed of a ramp vehicle at its merge point and its speed at the nose. These

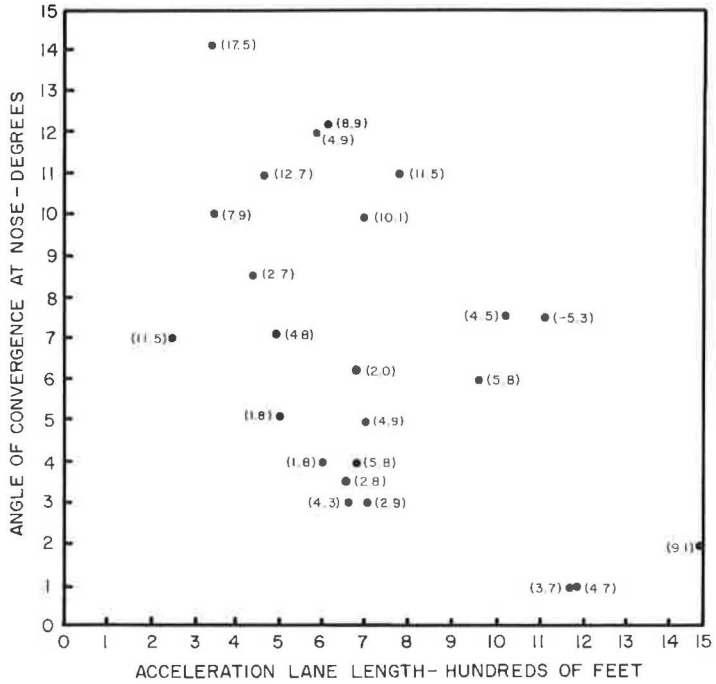


Figure 14. Effect of acceleration lane length and convergence angle on mean speed change of ramp vehicles between nose and merge point.

means are shown in Figure 14 and were somewhat surprising. Before the analysis, it was supposed that the increase in speed on the acceleration lane would be greater on the entrance ramps with the highest type of geometric design than on more poorly designed ramps. However, the opposite seemed to be the case.

For the ramps on which the acceleration lane length is greater than 450 ft and the convergence angle is less than 8 deg, none of the speed increases was greater than 6 mph. On only one of the ramps outside this group (acceleration lanes shorter than 450 ft and/or convergence angles greater than 8 deg) was the increase in speed less than 6 mph and it ranged up to 17.5 mph. This is apparently caused by the extremely low speeds at the ramp nose on the ramps with poor designs, and shows quite clearly that the design of the merging area has a strong influence on the speeds on the ramp approaching the merging area.

The effect of convergence angle on the increase in speed of vehicles on an entrance ramp between the ramp nose and the point of merge is shown in Figure 15. In each of the acceleration lane length categories, larger speed changes generally take place on entrance ramps which have larger angles of convergence. At entrance ramps with large convergence angles, vehicles on the ramp must follow a curved path to move from the ramp to the acceleration lane. The radius of this curve is a function of the convergence angle, and the safe operating speed on the curve is a function of the radius of the curve. Therefore, the speed at which vehicles can turn from the ramp to the acceleration lane can be limited by the convergence angle, so at entrance ramps with large convergence angles, the speed of ramp vehicles at the nose can be lowered because of the angle. Thus, a large increase in speed is required on the acceleration lane in order that the vehicles merge at approximately the freeway speed.

In Figure 16, it appears that for a given angle category, shorter acceleration lane length, in general, increases the amount by which vehicles increase their speed on the acceleration lane between the ramp nose and the merge point. As discussed previously,

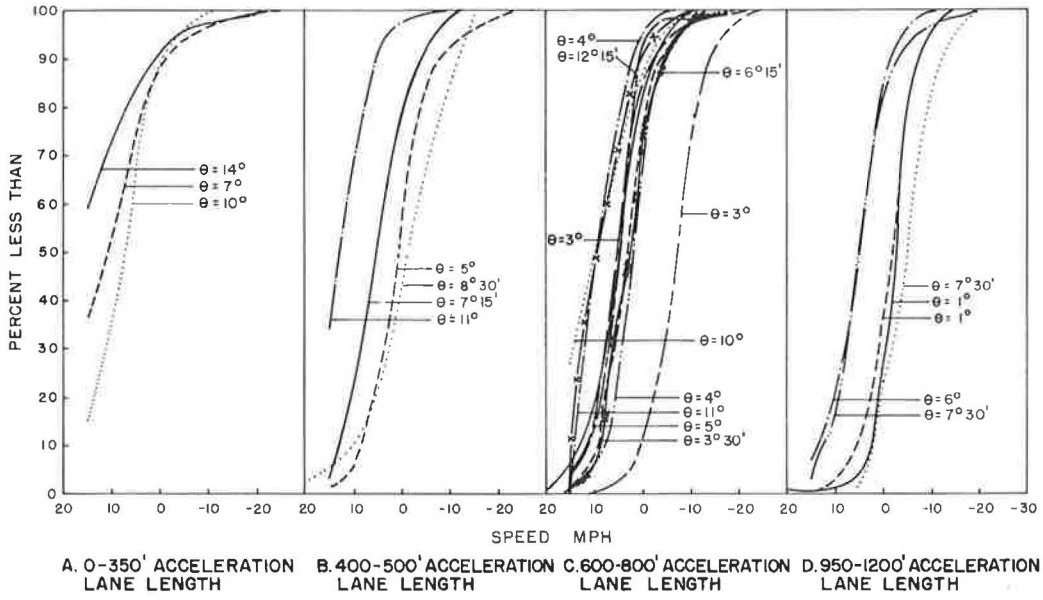


Figure 15. Effect of angle of convergence on speed changes between ramp nose and merge point for ramp vehicles.

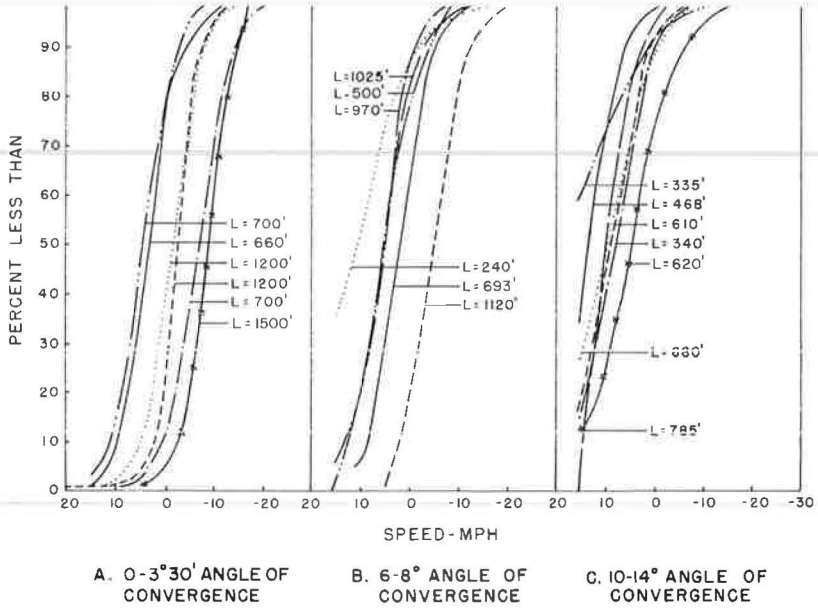


Figure 16. Effect of acceleration lane length on speed changes between ramp nose and merge point for ramp vehicles.

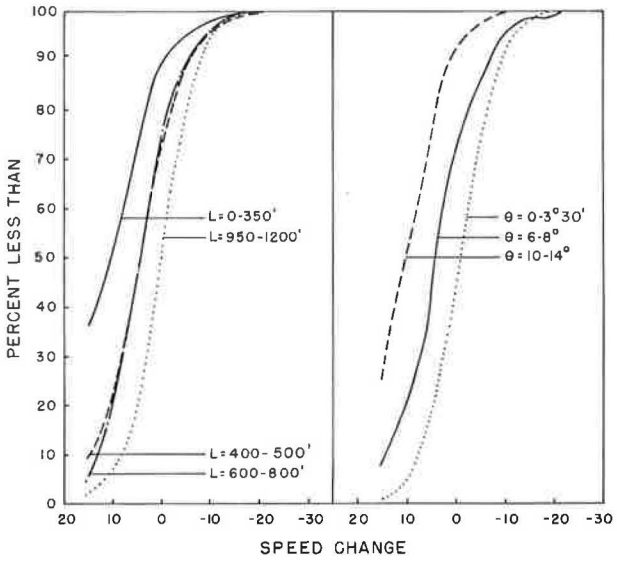


Figure 17. Effect of acceleration lane length and angle of convergence on ramp speed changes between ramp nose and merge point.

this phenomenon can probably be attributed to vehicles approaching the ramp nose at higher speeds on ramps at which the merging maneuver is easier (i.e., longer acceleration lane and lower convergence angle). Thus, on ramps with a long acceleration lane and a low convergence angle, the acceleration lane is not used to a great extent for acceleration but merely to accomplish a simple lane change. This is the type of operation for which the ramps are designed.

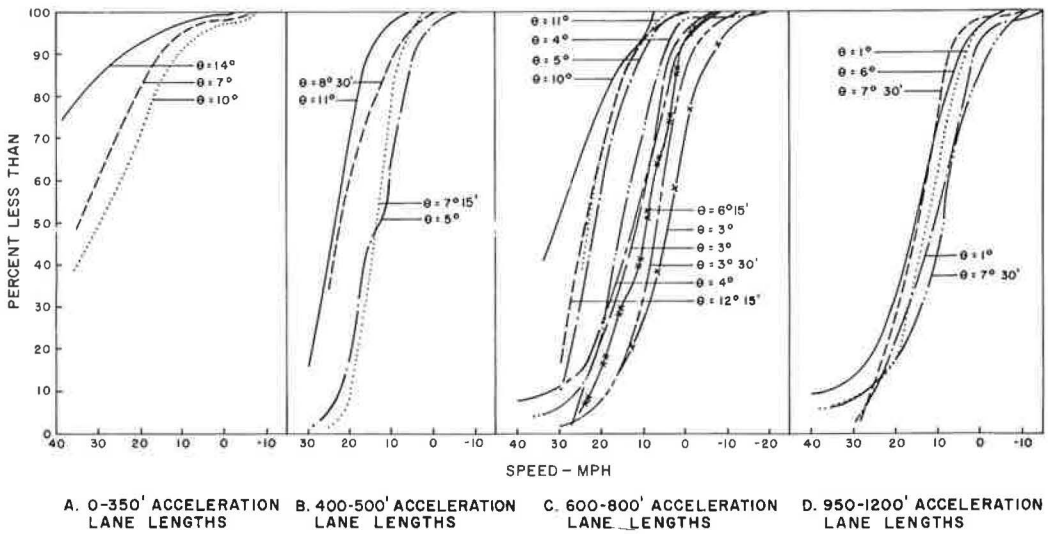


Figure 18. Effect of angle of convergence on relative speed distribution at ramp nose.

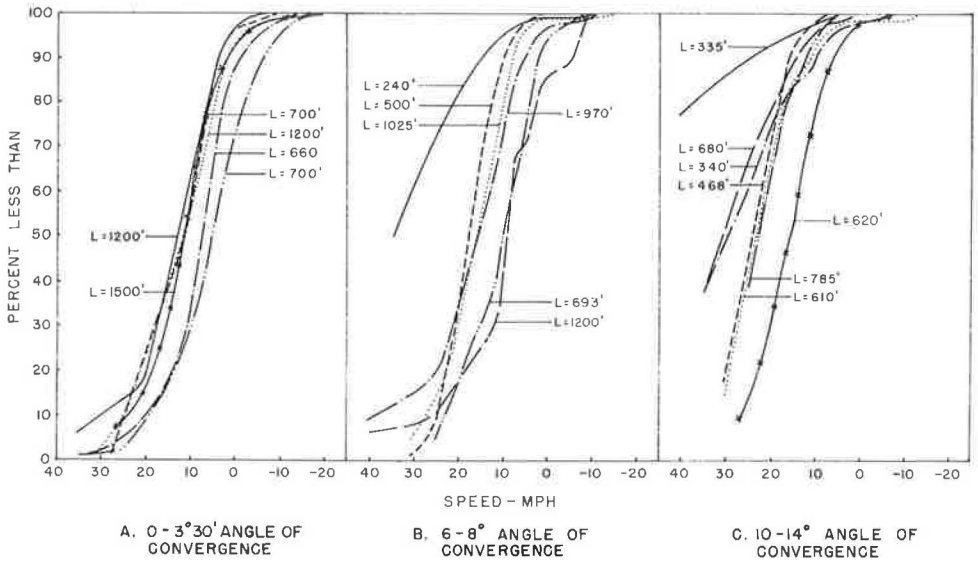


Figure 19. Effect of acceleration lane length on relative speed distribution at ramp nose.

The effects of acceleration lane length and convergence angle on the changes in speed of ramp vehicles on the acceleration lane are shown in Figure 17. Figure 17a shows that the speed changes are higher on ramps with shorter acceleration lanes, and Figure 17b shows that the speed changes are higher on ramps with large convergence

Relative Speeds at the Ramp Nose

The distribution of relative speeds (freeway lag speed minus ramp vehicle speed) at the ramp nose was plotted for each ramp grouped with ramps with similar geometric elements. To show the effect of convergence angle, the ramps were placed in four groups according to acceleration lane lengths (Fig. 18). To show the effect of acceleration lane length, the ramps were placed in three groups of convergence angles (Fig. 19).

Figure 18a shows the plots of relative speed distribution at the ramp nose for each of three ramps with acceleration lanes of less than 350 ft. The median relative speed on each of these ramps is greater than 30 mph, whereas the percent of relative speeds greater than 20 mph ranges from 73 percent to 93 percent. This indicates the poor quality, and indeed hazardous, operation at these three ramps. The relative speeds are highest for the ramp with the largest convergence angle (θ = 14 deg).

For the four ramps with acceleration lanes between 400 and 500 ft, Figure 18b shows the relative speed distributions. The two ramps with the highest convergence angles have median relative speeds greater than 20 mph, whereas the median speeds for the other ramps (θ = 5 and 7 deg) range from 9 mph to 13 mph. On the two ramps with the lowest convergence angles, from 15 percent to 18 percent of the relative speeds are greater than 20 mph, but on the other two ramps with θ = 8 deg 30 min and 11 deg these corresponding percentages are 60 percent and 72 percent.

The relative speed distributions for the ramps having acceleration lane lengths between 600 and 800 ft are shown in Figure 18c. Four ramps have relative speed distributions which are much higher than those of the other six ramps. The four ramps having the highest relative speeds have convergence angles of 5, 10, 11 deg and 12 deg 15 min and have median relative speeds ranging from 23 mph to 30 mph. On the other five ramps on which θ is between 3 deg 30 min and 4 deg, the median relative speeds were

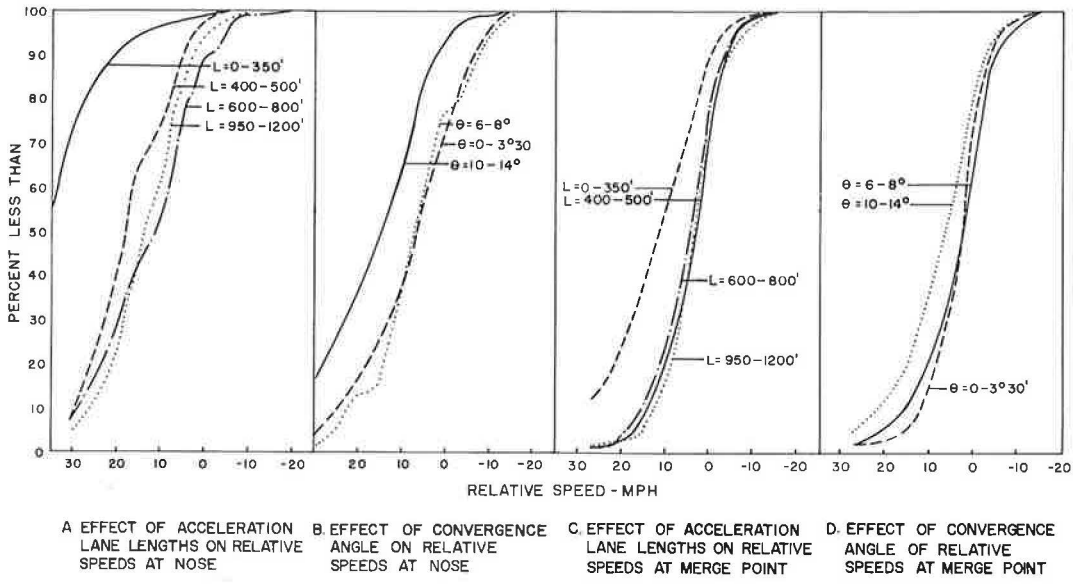


Figure 20. Effect of acceleration lane length and convergence angle on relative speeds at ramp nose and at merge point.

between 6 mph and 15 mph. (The ramp with θ = 5 deg has been previously discussed as having a ramp volume between 1000 and 1200 vph during the study which in part accounts for its high relative speeds.) Thus, the convergence angle can be seen to have an important effect on the relative speeds at the ramp nose for ramps which have acceleration lane lengths between 600 and 800 ft.

For five ramps having acceleration lane lengths between 950 and 1200 ft, the effect of convergence angles on the relative speeds is somewhat inconclusive (Fig. 18d). The two ramps with the highest relative speeds (median relative speed = 16 mph) have the high convergence angles (θ = 6 deg and 7 deg 30 min), but the ramp with the lowest relative speed (median relative speed = 9 mph) has a θ of 7 deg 30 min. The two intermediate ramps both have θ = 1 deg.

The effect on relative speeds of different acceleration lane lengths is shown in Figure 19. The ramps are again placed in three groups of convergence angles (0-3 deg 30 min, 6-8 deg and 10-14 deg). Within each group the relative speed distribution at the ramp nose is a function of acceleration lane length.

For small convergence angles (0-3 deg 30 min) shown in Figure 19a, relative speed distributions are closely grouped and no apparent relationship to acceleration lane length exists. In fact, the relative speeds seem to be higher for the ramps with the longer acceleration lane. This is probably due to the traffic patterns—higher volumes on the longer ramps.

For six ramps with convergence angles between 6 and 8 deg the effect on relative speeds of acceleration lane length is more apparent. The shortest ramp (L = 240 ft) has a median relative speed of 34 mph. The distributions for the other five ramps are fairly closely clustered. The lengths of these five ramps are 500, 693, 970, 1025 and 1120 ft, so it is somewhat surprising that the relative speed distributions are similar. The two short ramps are both at the foot of a downgrade to a depressed freeway (Edsel Ford Expressway in Detroit) and one was discussed in some detail earlier. Thus, the grade of the ramp appears to be a geometric variable which has quite a significant effect on the merging operation.

Figure 19c shows the relative speed distributions for five entrance ramps with θ between 10 and 14 deg. With the exception of the ramp with L = 680 ft, the relative

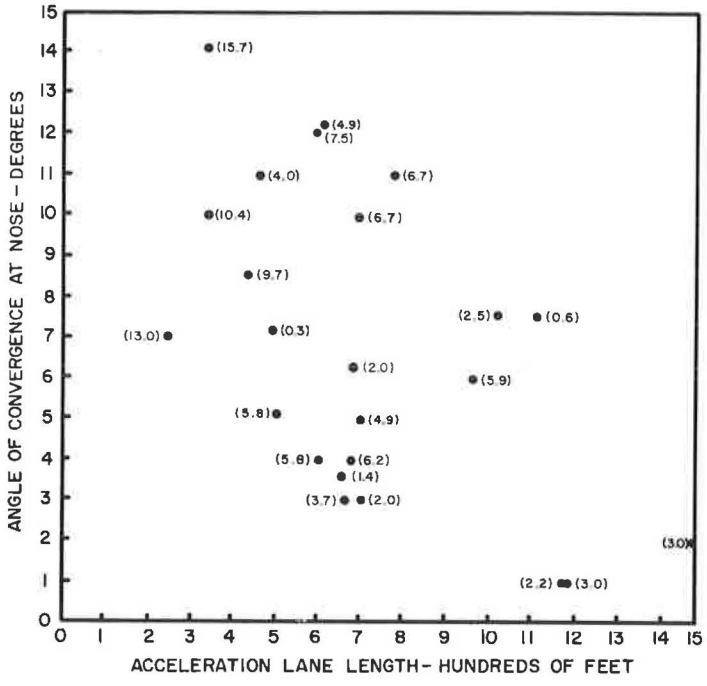


Figure 21. Effect of acceleration lane length and convergence angle on mean relative speed at merge point.

speeds decrease as L increases. For L = 335 ft, the median relative speed is greater than 40 mph, for L = 340 ft, the median is 29 mph, for L = 468 ft, the median is 24 mph, and for L = 785 ft, the median speed is 23 mph. The acceleration lane of the other ramp is 680 ft, but it is only 9 ft wide and the width probably is a major factor in the 30-mph median relative speed. The relative speeds on all of these ramps are excessive and this is probably due primarily to the high convergence angles.

The effect of acceleration lane length on the relative speed at the ramp nose can be clearly seen in Figure 20a, in which the relative speed distributions for four groups of ramps are presented. The four ramp groups are based on acceleration lane lengths of 0-350 ft, 400-500 ft, 600-800 ft and 950-1200 ft. Each distribution represents an average of the distribution for the ramps in each corresponding set of ramps in Figure 18a, b, c, and d.

The relative speeds at the ramp nose for the short ramps (0-350 ft) are excessive with the median exceeding 35 mph. The three distributions for the other three ramp groups are somewhat similar, with the relative speeds for the ramps with 400-500 ft acceleration lanes a bit higher than those for the ramps with acceleration lanes longer than 600 ft.

Figure 20b shows the effect of convergence angle on the relative speeds at the ramp nose. The three relative speed distributions represent averages for the ramps in the three angle groupings constituting Figure 19a, b, and c. The relative speed distributions for the 0-3 deg 30 min set and the 6-8 deg set are very similar. However, the distribution for the ramps with convergence angles between 10 and 14 deg indicates much higher relative speeds at these ramps. Thus, it appears that convergence angles up to 6 or 8 deg do not have a great effect on relative speeds at the ramp nose. Convergence angles greater than 8 or 10 deg do appear to have a significantly adverse effect on relative speeds at the ramp nose.

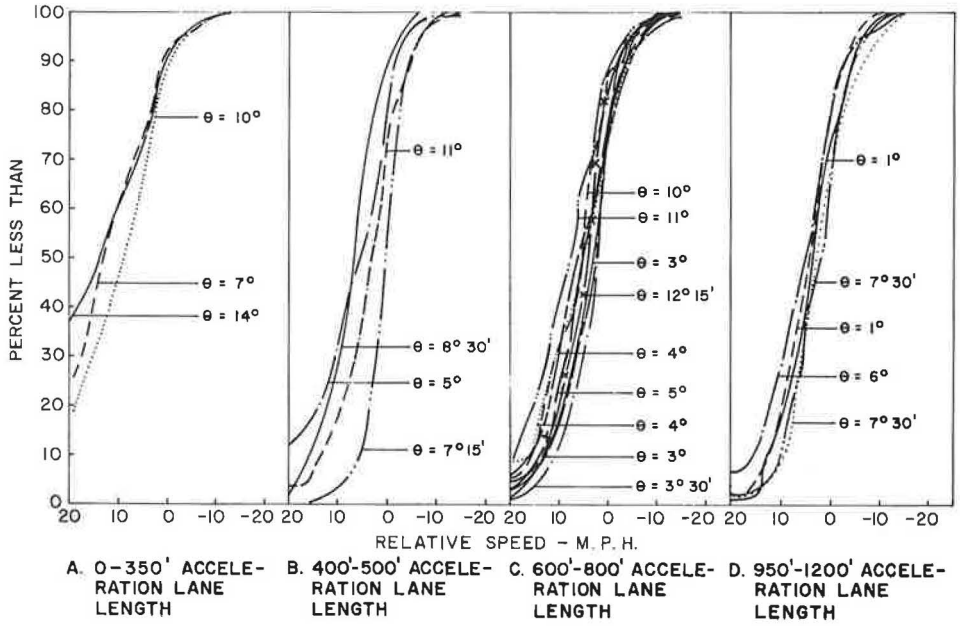


Figure 22. Effect of angle of convergence on relative merging speed.

Relative Speed at the Merge Point

The relative speed at the merge point was obtained for each vehicle for periods when the freeway operating speed was over 40 mph. For each ramp the mean relative speed (speed of freeway lag vehicle minus speed of ramp vehicle at merge point) is shown in Figure 21. While the general trend seems to be that ramps with better designs have

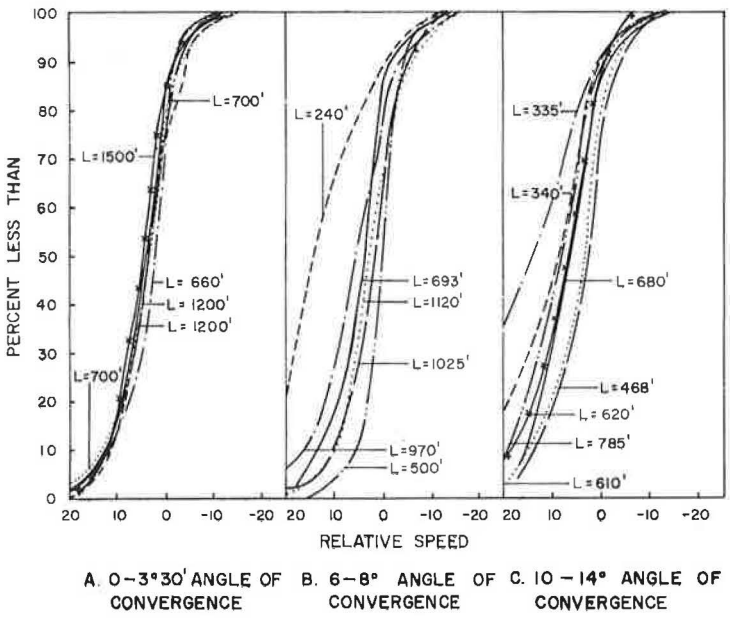


Figure 23. Effect of acceleration lane length on relative merging speeds.

lower relative merging speeds, there are many exceptions to this trend. Ramps with very short acceleration lanes and/or very large convergence angles have high relative merging speeds, but the pattern is inconclusive for ramps with less extreme geometrics.

Two ramps with relatively high convergence angles (6 deg 15 min and 7 deg) and relatively short acceleration lanes (500 and 693 ft) have very low relative merging speeds (0.3 and 2.0 mph). These entrance ramps (Chene and Gratiot eastbound on the Edsel Ford Expressway in Detroit) are both downgrade ramps to the depressed freeway. The operation at each was very good and the merging volumes were quite high. This indicates the beneficial effect of the negative ramp grade on the merging operation as discussed previously.

Figure 22 shows the effect of angle of convergence on the relative merging speeds. Generally, lower relative merging speeds are found on ramps with smaller convergence angles. The operation at the ramps with acceleration lanes between 950 and 1200 ft can be seen to be much more consistent (merging, speed distributions closely grouped) than the ramps with shorter acceleration lanes.

In Figure 23, the effect of acceleration lane length on relative merging speed is shown. For the ramps with θ less than 3 deg 30 min, a variation of acceleration lane lengths between 600 and 1200 ft had virtually no effect on the relative merging speeds (Fig. 23a).

Figure 23b shows the effect of acceleration lane length on the ramps for which θ lies between 6 and 8 deg. The ramp with the highest relative merge speeds has a 240-ft acceleration lane. On the other ramps, L ranged from 500 to 1120 ft, and the effect of acceleration lane length is somewhat unclear due to a lack of any consistent relationship between L and the relative merge speed distributions. Figure 23c, which includes the ramps for which θ is between 10 and 14 deg, presents data that show generally a decrease in relative speed for increases in acceleration lane length.

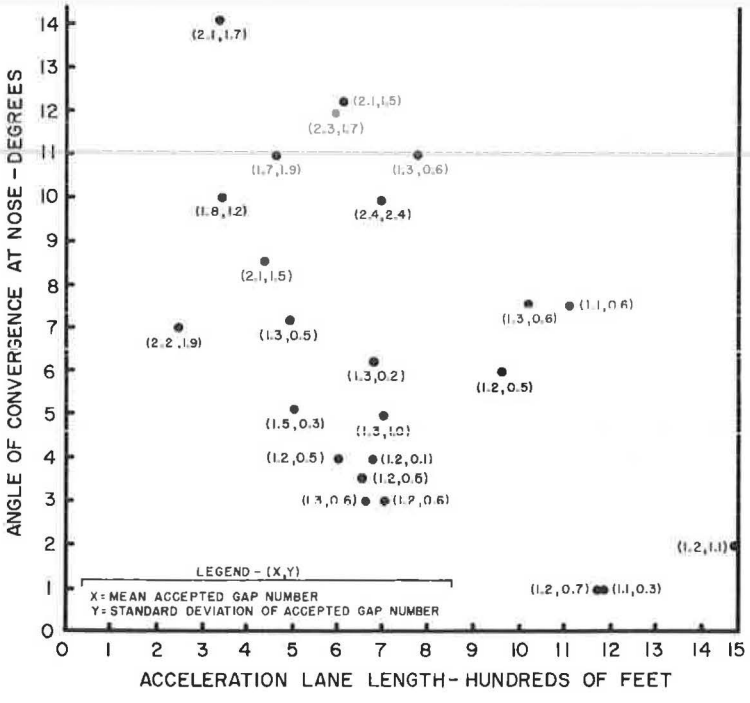


Figure 24. Mean and standard deviation of accepted gap number.

For freeway operating speed over 40 mph, the effect of acceleration lane length and angle of convergence can be seen in Figure 20c and d. The average relative merging speed for all ramps included in the study which had acceleration lane lengths below 350 ft had much higher relative speeds than did the ramps with longer acceleration lanes (Fig. 20c). The relative merge speed distributions for the other three ramp groups (based on length of acceleration lane) are very similar. Similarly, in Figure 20d it appears that convergence angles up to 8 deg have little influence on relative merging speeds. However, angles in excess of 10 deg lead to substantially higher relative speeds in the merging area. It is interesting to compare Figure 20a to 20c and Figure 20b to 20d. It is quite apparent the relative speed decreases considerably between the ramp nose and the merge points.

Accepted Gap Number

The number of the gap which was accepted by each ramp vehicle on each ramp was obtained from the films. Gap 1 for each ramp vehicle is the gap which is at the physical nose at the time the ramp vehicle arrives there. Gap 2 is behind (upstream of) gap 1 and gaps 0, -1, etc., are downstream of gap 1 (Fig. 6).

Figure 24 shows the mean and standard deviation of the accepted gap number at each ramp. Each point represents a ramp with a particular set of geometric design features; the first number in parenthesis is the mean accepted gap for that ramp, and the second number is the standard deviation of the accepted gap number. Only periods in which the freeway operating speed exceeded 40 mph were considered, so the data reflect good freeway operation at each ramp. If each ramp vehicle accepted the gap that was adjacent to the nose when the ramp vehicle reached the nose, the mean accepted gap number would be 1.0 and the standard deviation would be zero.

Figure 24 shows that ramps with low convergence angles and long acceleration lanes have associated with them a low mean and standard deviation of accepted gap number. This reflects on the ability of ramp drivers to select their gap at or before the nose on this type of ramp and indicates the consistency of operation on these ramps with good geometric designs. The consistency of behavior of ramp drivers is important in the achievement of a high degree of safety, and the ramps with low standard deviations of accepted gap number probably have a lower accident record than the ramps on which the operation is more erratic. Many ramp accidents are caused by one driver assuming that the vehicle in front will accept a particular gap (2). If this assumption proves true, all is well. If not, the following vehicle may find it necessary to decelerate rapidly (especially at ramps with large convergence angles and short acceleration lanes where a "Russian roulette" type of merging operation is observed), possibly leading to hazardous maneuvers.

The ramps which had L greater than 600 ft and θ less than 7 deg 30 min generally had good gap acceptance characteristics. With the exception of one ramp, the mean

TABLE 1

COMPARISON OF ACCEPTED GAP NUMBER FOR TWO RAMP GROUPS

Groups	Accepted Gap Number	
	Range of Means	Range of Standard Deviations
Group 1 L > 600 ft θ < 7 deg 30 ft	1.1-1.5 gaps	0.1-0.7 gaps
Group 2 L < 600 ft $\theta \geq$ 10 deg	1.7-2.4 gaps	1.2-2.4 gaps

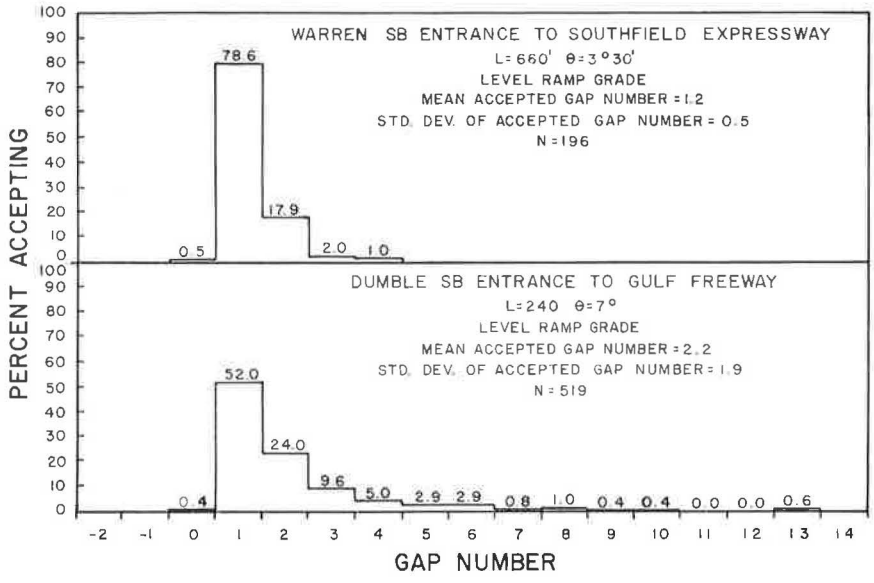


Figure 25. Effect of L and θ on frequency distributions of accepted gap number.

accepted gap number ranged from 1.1 to 1.5 and the standard deviation ranged from 0.1 to 0.7 gaps. One ramp in this group with L = 700 ft and θ = 5 deg had a standard deviation of 1.0 gaps. The flow rate on this ramp during the study period was about 1800 vph, which undoubtedly accounted for much of this variability. Even so, the means

similar group.

The ramps which have an acceleration lane of less than 600 ft and a convergence angle of 10 deg or larger exhibited poor gap acceptance characteristics. The mean accepted gap for this group was about 2.0 (a range of 1.7 to 2.4) and the standard deviation ranged from 1.2 to 2.4 gaps. This is considerably different than the similar data for the other geometric group. Table 1 summarizes the comparison of the two groups.

The effect of angle of convergence and acceleration lane length on the gap acceptance characteristics is shown in Figure 25. The frequency distributions of accepted gap numbers for two entrance ramps are also shown in this figure. The two ramps are the Dumble southbound entrance ramp on the Gulf Freeway in Houston and the Warren southbound entrance ramp on the Southfield Expressway in Detroit. Except for

TABLE 2
OPERATIONAL DATA FOR THE DUMBLE AND WARREN ENTRANCE RAMPS

Data	Dumble	Warren
Acceleration lane length	240 ft	660 ft
Angle of convergence	7 deg	3 deg 30 min
Number of films taken	5	2
Total merging volume	1200-1500 vph	1410-2030 vph
Ramp volume	400-500 vph	400-500 vph
Operating speeds	35-50 mph	40-50 mph

the length of acceleration lane and angle of convergence, the two ramps are similar geometrically. Both ramps are basically level as is the freeway at each and both ramps are of the short, "slip ramp" type. Table 2 contains much of the pertinent information for the two ramps.

In Figure 25, the distribution of accepted gaps for the Warren entrance ramp can be seen to be much "tighter" (as reflected by the smaller standard deviation) than the distribution for the Dumble entrance ramp. At the Warren entrance only 3 percent of the vehicles had not merged into gaps 0, 1, and 2, while at the Dumble ramp over 23 percent had not merged after gap 2 passed the ramp vehicle. (At a ramp similar in design to the Dumble ramp and also on the Gulf Freeway, a vehicle was observed to stop and wait for gap 35 before merging, i.e., it rejected the first 34 gaps.) These operational differences took place in spite of the substantially higher freeway volumes at the Warren entrance ramp. Thus, the better operation at the Warren entrance ramp can be attributed to the longer acceleration lane and smaller angle of convergence there.

FREQUENCY DISTRIBUTION OF MERGE POINTS

Much of the previous work in this field has involved the comparison of selected characteristics of geometry and traffic in describing acceleration lane use. This approach has certainly proved valuable in improving design of acceleration lanes, and it is used also in a later part of this paper. However, a purely mathematical description of acceleration lane use can contribute greatly to a basic understanding of the merging process. The next step will be to relate the parameters of the mathematics to the geometric and operational characteristics of the facility or facilities.

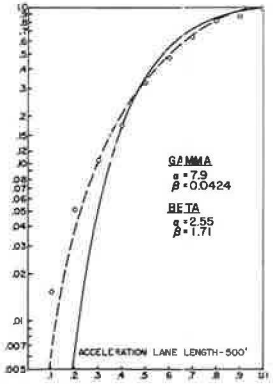
The distance from the nose to the point at which a ramp vehicle encroaches on the freeway shoulder lane is one measure of the manner in which an acceleration lane is used. It has been observed that the entry point is a result of the desire of the driver to follow a natural and easy path of entry, but that this path is affected by volumes and speeds of freeway shoulder lane traffic as well as by the geometrics of the entrance ramp and acceleration lane.

A mathematical description of entry distances must take into account not only the mean entry point, but also the dispersion of entries about this mean. Commonly, this dispersion is best described by the variance. To be useful, particularly as a tool in simulation, a usable mathematical probability distribution must be defined and shown to describe the phenomena.

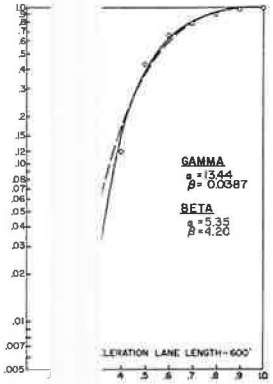
Examination of samples of data showed a frequency distribution that was not symmetrical but that had a very pronounced peak representing the activities of most drivers in entering the freeway. Generally, the entry points varied from quite near the nose to a point near the defined limit of the acceleration lane.

The Pearson Type I probability distribution (7) appears to satisfy the requirements of being nonsymmetrical, showing a pronounced peak, and having upper and lower limits imposed by the geometric design. By variation of the parameters α and β of the Pearson Type I distribution, the mean, variance, and mode can be shifted to approximately accord with the observed data on entry locations. Parameters α and β for each set of data (approximately 20 min of film) were determined by equating the normalized sample mean and variance to the theoretical mean and variance of the distribution expressed in terms of α and β. Acceptableness of the calculated curve as a valid description of the observed data was checked by the chi-square test.

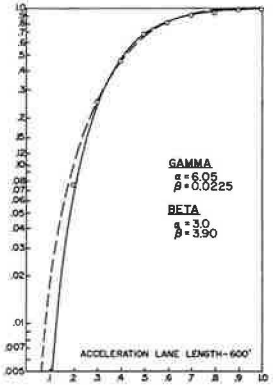
In analyzing the film, it was noted that not infrequently a vehicle would continue to travel far beyond the prescribed geometric limits of the acceleration lanes without entering the freeway shoulder lane. In some cases this meant driving over a mountable curb barrier. Consequently, since this additional travel was to greatly varying distances (although finite), the Pearson Type III distribution was investigated for its ability to describe the observed phenomena. It is nonsymmetrical, shows a pronounced peak, and has a finite lower limit. Pearson Type III distributions were determined and tested for acceptability of fit in the same manner as described for the Type I distribution.



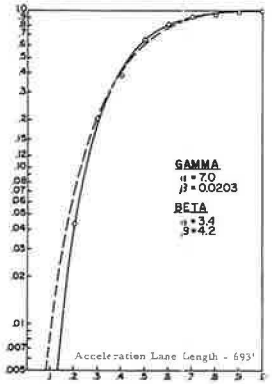
CHENE EASTBOUND ENTRANCE RAMP
EDSEL FORD EXPRESSWAY, DETROIT



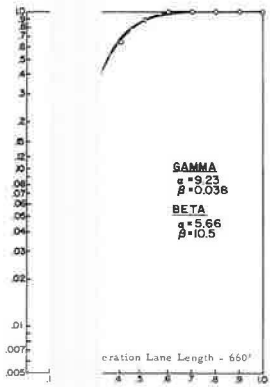
WARREN AVE
SOUTHF
OUTHBOUND ENTRANCE RAMP
EXPRESSWAY, DETROIT



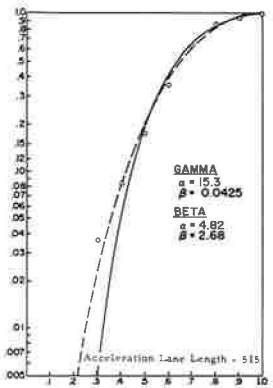
LINWOOD NORTHBOUND ENTRANCE RAMP
JOHN LODGE EXPRESSWAY, DETROIT



DEMPSTER STREET NORTHBOUND ENTRANCE RAMP
EDENS EXPRESSWAY, CHICAGO

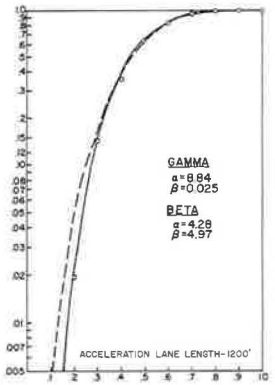


KOSTNE
EISENH
TBOUND ENTRANCE RAMP
EXPRESSWAY, CHICAGO

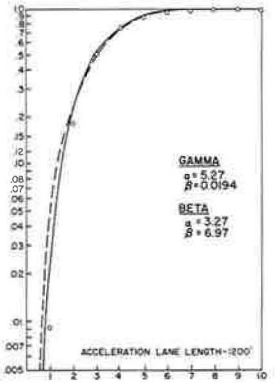


PETERSON NORTHBOUND ENTRANCE RAMP
EDENS EXPRESSWAY, CHICAGO

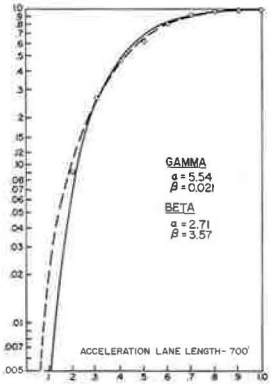
Figure 26. Freque distribution of entry points.



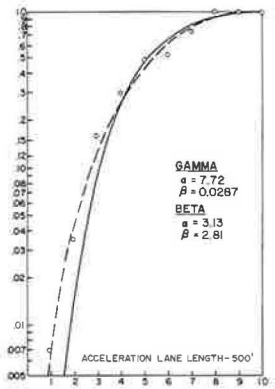
JERICHO (RT. 25) EASTBOUND ENTRANCE RAMP
LONG ISLAND EXPRESSWAY, NEW YORK



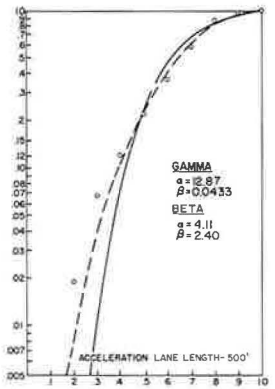
LONG ISLAND EXPRESSWAY NORTHBOUND ENTRANCE RAMP
CROSS ISLAND EXPRESSWAY, NEW YORK



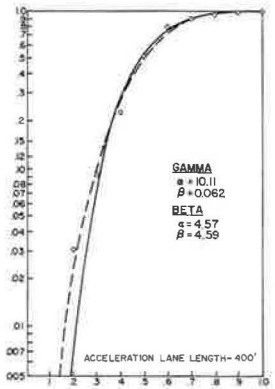
COMMUNITY DRIVE EASTBOUND ENTRANCE RAMP
LONG ISLAND EXPRESSWAY, NEW YORK



BRUSH HOLLOW ROAD EASTBOUND ENTRANCE RAMP
NORTHERN STATE PARKWAY, NEW YORK

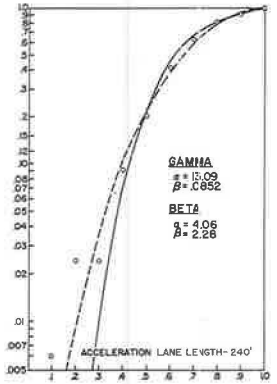


69TH ROAD (JEWEL STREET) SOUTHBOUND ENTRANCE RAMP
GRAND CENTRAL EXPRESSWAY, NEW YORK

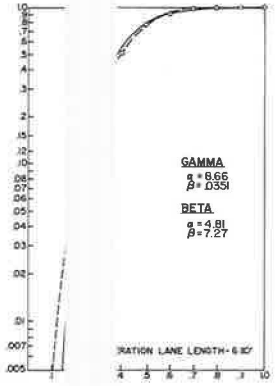


ROCKAWAY NORTHBOUND ENTRANCE RAMP
VAN WYCK EXPRESSWAY, NEW YORK

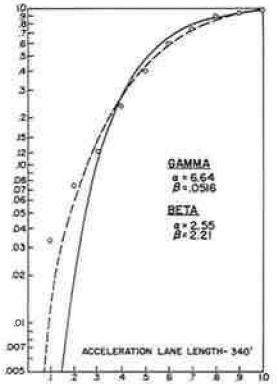
Figure 27. Frequency distribution of entry points.



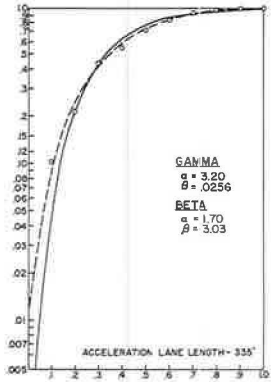
DUMBLE SOUTHBOUND ENTRANCE RAMP
GULF FREEWAY, HOUSTON



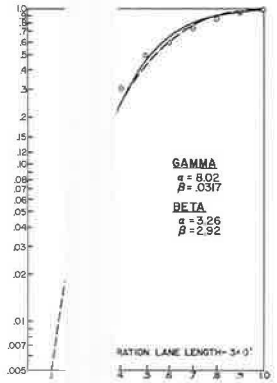
WAYSIDE S
GUL
)OUND ENTRANCE RAMP
EEWAY, HOUSTON



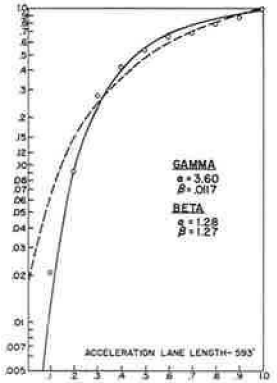
CULLEN SOUTHBOUND ENTRANCE RAMP
GULF FREEWAY, HOUSTON



BROAD STREET SOUTHBOUND ENTRANCE RAMP
GULF FREEWAY, HOUSTON

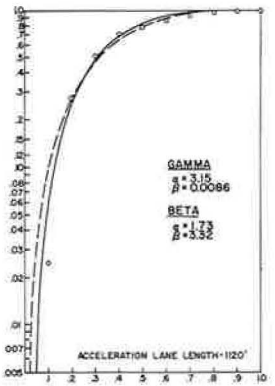


TELEPHONE
GUL
HBOUND ENTRANCE RAMP
EEWAY, HOUSTON

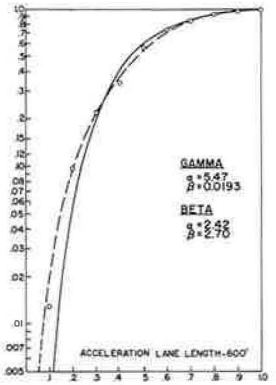


BRENTWOOD WESTBOUND ENTRANCE RAMP
DANIEL BOONE EXPRESSWAY, ST. LOUIS

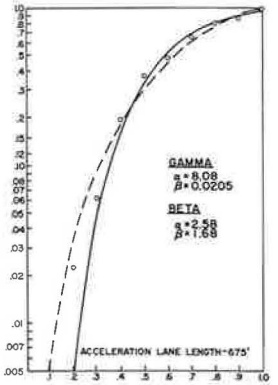
Figure 28. Freque distribution of entry points.



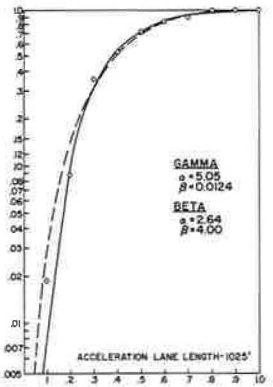
ASHBY AVENUE SOUTHBOUND ENTRANCE RAMP
EASTSHORE FREEWAY, SAN FRANCISCO



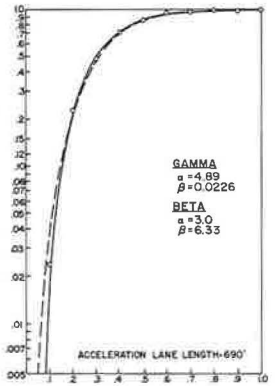
MONUMENT JUNCTION SOUTHBOUND ENTRANCE RAMP
I 680 IN PLEASANT HILL, SAN FRANCISCO



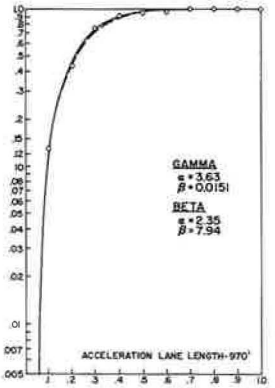
MILBRAE AVENUE SOUTHBOUND ENTRANCE RAMP
BAYSHORE FREEWAY, SAN FRANCISCO



BROADWAY NORTHBOUND ENTRANCE RAMP
BAYSHORE FREEWAY, SAN FRANCISCO



WATT AVENUE SOUTHBOUND ENTRANCE RAMP
I-80, SACRAMENTO



COLDWATER CANYON AVENUE WESTBOUND ENTRANCE RAMP
VENTURA FREEWAY, LOS ANGELES

Figure 29. Frequency distribution of entry points.

The results are summarized in the following figures. Figures 26, 27, 28 and 29 show cumulative frequencies of Pearson Type III (Gamma) in solid lines (dashed lines represent Type I distribution), plotted against normalized acceleration lane lengths. The circles are values of the raw data from which the curves were derived. Of the 24 sets of data shown in these figures, 14 (58 percent) were accepted at an 0.05 level of significance as fitting a Pearson Type III distribution and 16 (66.7 percent) were accepted as fitting a Pearson Type I distribution.

From this information, it appears feasible to describe the frequency distribution of freeway entry points mathematically. Such mathematical distributions can be used to check and validate computer simulation programs which are used to study the merging phenomena.

REGRESSION ANALYSIS OF MERGING DISTANCE

In order to delineate those factors which affect the use of acceleration lanes as described by the distance from the physical nose to the entry point, a regression analysis was carried out using several geometric and traffic characteristics as the independent variables. Using the average distance from the nose to the point of entry during each 3-minute period as a dependent variable, the independent variables which were significant at the 5 percent level were determined.

The geometric variables used were the acceleration lane length, the length of the ramp, the effective ramp grade, the angle of convergence of the ramp at the nose, and the offset of the ramp at the nose. The variables which related to traffic characteristics were the percent of trucks involved in the merge, the percent of ramp vehicles involved in the merge, the total merge volume, and the product of the 3-minute flow rate and average velocity of the freeway vehicles approaching the nose.

Notice that most of these variables have been used previously. But one of them requires some discussion, being introduced here for the first time. The new variable, relative effective ramp grade, reflects the driver's ability to see or observe the freeway as he approaches it. Ramp grade at the nose seldom describes the grade or ele-

Ramp driver behavior when approaching a freeway can be described as follows. As soon as it is possible to see the freeway and the ramp driver is close enough to begin to adjust to and utilize an acceptable freeway gap, the ramp driver will begin to judge the speed of freeway vehicles and begin to search for a gap large enough for him to merge into. If the entrance ramp approaches the freeway from below (i.e., up a positive grade), a ramp vehicle driver can begin selecting a gap and merge only when the ramp elevation approximates that of the freeway.

On the other hand, if the ramp approach is from above (i.e., down a negative grade) the driver has an earlier opportunity to begin selecting a gap and matching his speed with the freeway speed. Observation indicates that a majority of drivers begin overtly to observe freeway traffic where possible at a distance greater than 200 ft before the nose, but generally not more than 500 ft. In order to reflect the effect of observing the freeway from the ramp, the effective ramp grade expressed as a percent was defined as the algebraic difference in the elevation at the nose and the elevation on the ramp 400 ft upstream of the nose or the length of the ramp if it was shorter than 400 ft divided by horizontal distance between these two points. The relative effective ramp grade is equal to the effective ramp grade minus the freeway grade. Table 3 is a summary of the regression analyses. The test for significance was made at the 5 percent level.

The following variables were found to be significant in the results of the regression analysis:

1. Acceleration lane length—drivers tend to take a longer distance to merge on ramps with longer acceleration lanes.
2. Ramp length—a longer ramp is associated with a shorter distance to entry point. This appears to result from the ramp drivers having a greater distance to adjust to freeway conditions before reaching the acceleration lane.

TABLE 3

SUMMARY OF REGRESSION ANALYSIS

Model: $Y = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + B_4X_4 + B_5X_5 + B_6X_6 + B_7X_7 + B_8X_8 + B_9X_9$

Y = Average distance from nose to point of entry (ft)

Independent Variable, X_i	Coefficient (B_i)	Sig.
X_1 = acceleration lane length (ft)	0.128	Yes
X_2 = ramp length (ft)	- 0.087	Yes
X_3 = relative effective ramp grade (%)	+16.7	Yes
X_4 = angle of convergence (deg)	-16.9	Yes
X_5 = ramp offset at nose (ft)	18.1	Yes
X_6 = QU on freeway at nose (veh-mi/hr)	0.0026	Yes
X_7 = percent trucks in merge volume (%)		No
X_8 = percent ramp veh in merge volume		No
X_9 = merge volume (rate of flow) (veh/hr)		No

Multiple Regression Coefficient = 0.90
B_0 = 87.1 ft

3. Relative effective grade—for which an increase in relative grade is associated with an increase in distance to entry point. As the percent grade increases from minus to plus (the ramp lower than the freeway), ramp drivers cannot begin selecting gap and adjusting speed until they get quite near the nose. Hence they tend to make more of the adjustment of the acceleration lane.
4. Angle of convergence—for which an increase causes a decrease in distance to entry point. This apparently reflects a desire of drivers to continue a smooth, natural path of entry without having to perform the reverse curve that would be necessary when this angle is large and the driver went further down the acceleration lane to enter freeway. The drivers tend to stay on the path on which they are "aimed" by the ramp.
5. Ramp offset at nose—for which a larger offset distance causes an increase in distance to entry point. This is certainly to be expected, since the ramp vehicle must move a greater transverse distance.
6. The product of the 3-minute rate of flow (expressed on a vph basis) times the average speed—for which an increase causes an increase in distance to entry point.

Variables found not to be significant included the percent trucks in the merge volume, the merge volume, and the percent ramp vehicles in the merge volume.

INTERPRETATIONS AND CONCLUSIONS

From the results of these studies, it is quite apparent that the operation in the merging area is to a great extent a function of the geometrics of the entrance ramp and acceleration lane. Among the geometric variables which were found to have a pronounced effect on the merging operation are the length of acceleration lane, angle of convergence, ramp grade and width of acceleration lane. Based on these studies the following tentative conclusions are offered:

1. Based on the study of one entrance ramp with a 9-ft acceleration lane width and 23 ramps with a 12-ft acceleration lane width, it appears that the narrower acceleration lane width has an adverse effect on the merging operation.
2. On each of the ramps for which the convergence angle was 3 deg or less, the operation was extremely good. Each of these ramps had an acceleration lane at least 700 ft in length; it would appear that a desirable design standard include a low (less

than 5 or 6 deg) convergence angle and a long (at least 700 ft) acceleration lane. (For the ramps studied which had a 700-ft length, this distance was measured from the physical nose which was about 2 ft off the pavement edge. If a more desirable 10-ft offset is used, the physical nose would be moved about 150 ft back so that the desirable minimum acceleration lane length to be used with a 10-ft offset would be about 850 ft.)

3. At each of the entrance ramps which had convergence angles 10 deg or greater and/or an acceleration lane length less than 350 ft, very poor operation was noted. Based on these results and other observations, serious consideration should be given to the geometric modification or redesign of an entrance ramp if it has an angle of convergence greater than 10 deg and/or an acceleration lane length (based on a 2-ft offset at the nose) less than 600 ft for level or upgrade ramps or 500 ft for downgrade ramps. The resultant designs can lead to greatly improved merging operation and probably a substantial reduction in the number of minor accidents. If such redesign is impractical, a freeway merging control system can, in many cases, be used to improve the merging operation.

4. On the downgrade ramps studied, the operation was much better than would have been expected, based on the acceleration lane length and convergence angle only. Apparently, the better field of view on these ramps allows ramp drivers to select a gap when they are 200 ft or so upstream of the nose and to accelerate to meet the gap near the nose and merge quickly. Thus, some of the acceleration takes place in advance of the ramp nose and the need for a long acceleration lane may be somewhat diminished. Based on this conclusion, one could further conclude that better merging operation could be expected on a depressed freeway than on at-grade or elevated freeway, since downgrade entrance ramps are found on depressed freeways.

5. Variables which were found in the regression analysis to have a significant effect on the distance used by vehicles before merging are (a) acceleration lane length, (b) ramp length, (c) relative effective grade between ramp and freeway, (d) angle of convergence, (e) ramp offset at the nose, and (f) the product of merging volume and speed. Factors which tended to decrease the distance used by vehicles to merge were (a) shorter acceleration lane, (b) greater ramp length, (c) lower relative effective

product of merging volume and speed.

Further Research

This report was concentrated primarily on the effects of the angle of convergence and acceleration lane length. Several other important geometric variables (such as ramp grade) will be considered more fully in subsequent work.

In addition to the study of the effects of individual geometric elements, a mathematical model is being developed describing the effect of all pertinent geometric variables on the traffic characteristics in the merging maneuver and in particular to determine the relative importance of each of the geometric variables. The probit mathematics, which have proved so valuable in describing the gap acceptance phenomena (6), appear to hold promise in this regard.

ACKNOWLEDGMENTS

In a large project with the national scope of its field studies the success depends to a very large extent on the cooperation of local governmental agencies. Throughout the course of this project the cooperation received from all such agencies was excellent. It is not possible to acknowledge each individual who helped the project in some way so the following partial list is offered: California Division of Highways—James E. Wilson, Karl Moskowitz and Leonard Newman; Illinois Division of Highways—Charles H. McLean; City of Detroit Department of Traffic—Alger F. Malo; Michigan State Highway Department—Joseph Marlowe; New York Department of Public Works—N. C. Parsons; Long Island State Park Commission—Mr. Boyce; City of New York Department of Traffic—Edward Bonelli; Missouri Highway Department—James Little, and James Roberts; and Texas Highway Department—Dale D. Marvel, John N. Lipscomb, and William V. Ward.

Deserving special mention is Joseph W. Hess of the Bureau of Public Roads who, as the Project Contact Representative, aided the project in major ways in making some contacts with local agencies, helping select some study sites and with some preliminary formulation of the project scope and direction.

In addition, two other staff members of the Texas Transportation Institute were instrumental in much of the data collection work involved in the project. The data collection was quite extensive and the fine work of Thomas G. Williams, Research Assistant, and James Bradley, Motion Picture Production Technician, meant a great deal to the success of the project and is certainly appreciated. Also greatly appreciated is the work of W. R. McCasland in making preparations for many of the Houston studies.

REFERENCES

1. Pinnell, C., and Keese, C. J. Freeway Ramps. Texas Transportation Inst. Rept., 1960.
2. Pinnell, C. Driver Requirements in Freeway Entrance Ramp Design. Proc. Inst. of Traffic Eng., 1960.
3. Fukutome, I., and Moskowitz, K. Traffic Behavior and On-Ramp Design. HRB Bull. 235, pp. 38-72, 1960.
4. Eckhardt, J. E., and Newman, L. Freeway Operation Department Development—Interim Report. Calif. Div. of Highways, Freeway Operations Dept., District 7, 1966.
5. Buhr, J. H., Drew, D. R., Wattleworth, J. A., and Williams, T. G. A Nationwide Study of Freeway Merging Operations. Highway Research Record 202, pp. 76-122, 1967.
6. Drew, D. R., La Motte, L. R., Buhr, J. H., and Wattleworth, J. A. Gap Acceptance in the Freeway Merging Process. Presented at the 46th Annual Meeting and published in this RECORD.
7. Drew, D. R. Classification and Application of Traffic Probelms by Models. Traffic Engineering, Vol. 36, No. 2, Nov. 1965.

Appendix

STUDY OF GEOMETRICS OF RAMP STUDY LOCATIONS

Location	Ramp Name	Length and Shape of Accel. Lane P = Parallel T = Taper (ft)	Angle of Converg. at Ramp Nose (deg-min)	Angle of Converg. 2 Off Pave-ment Edge (deg-min)	Grade at Nose (%)	Freeway Direction	Ramp Type	Ramp Curvature	Curb Offset at Nose (ft)	Length of Ramp (ft)	Comments
Houston Area	Weslayan EB to Southwest Freeway	610-P	12-15	12-15	level	East	diamond	tangent	2	103	4-lane
	Buffalo Speedway WB to Southwest Freeway	620-P	12	12	level	West	diamond	tangent	2	160	4-lane free-way curve 10 right
	Dumble SB to Gulf Freeway	240-T	7	7	+0.5	East	diamond	tangent	2	170	Freeway grade past ramp +5.0%
	Wayside SB to Gulf Freeway	680-P	10	10	+0.4	East	diamond	tangent	2	380	
	Cullen SB to Gulf Freeway	340-T	10	10		East	diamond	tangent	2	180	
	Broad SB to Gulf Freeway	335-T	14	14	level	East	diamond	tangent	2	225	
Chicago Area	Pulaski EB to Southwest Expressway	1500-T	2	1-30	-0.80	East	diamond	tangent	10	900	Ramp grade is -4.7%
	Kostner EB to Eisenhower Expressway	712-T	6	6	+1.0	East	diamond	slight right	2	350	Ramp grade is -2.5%, freeway curve slight rt.
	Independence WB to Eisenhower Expressway	738-T	6	6	-0.4	West	diamond	tangent	2	432	
	Dempster NB to Edens Expressway	785-P	11	11	level	North	outer connector	slight right	2	785	
	Peterson NB to Edens Expressway	468-P	11	11	level	North	outer connector	slight right	2	1020	
Los Angeles Area	Coldwater Canyon WB to Ventura Freeway	970-P	6	6	-1.0	West	diamond	tangent	2	600	
San Francisco Area	Monument Junction SB to I080	600-T	4	4	+0.4	South	diamond	slight right	2	800	Rural, 2-lane fwy, curve slight right
	Ashby SB to East-shore Freeway	1120-T	7-30	2	level	South	trumpet directional	tangent	10	500	4-lane
	Lakeshore SB to MacArthur Freeway	700-T	5		+2.5	South	diamond	tangent	10	450	4-lane
	Milbrae SB to Bayshore Freeway	675-T	3	2	level	South	from clover-leaf CD road		6	350	
	Broadway NB to Bayshore Freeway	1025-T	7-30	2-30	level	South	trumpet directional	slight right	10	300	4-lane
Sacramento Area	Watt SB to I80	690-P	4	4	-0.2	South	diamond	slight right	2	500	2-lane freeway, curve slight right
	Warren EB to South-field Expressway	660-T	3-30	3-30	-0.4	South	diamond	tangent	2	50	Freeway grade is +3.0%
	Gratiot EB to Ford Expressway	500-T	7-15	7-15	+0.4	East	diamond	tangent	2	600	Ramp grade is -0.7%
New York City Area	Jewel SB to Grand Central Parkway	500-P	5	5	level	South	diamond	tangent	12	520	Ramp grade is -7.0%
	Rockaway NB to Van Wyck Expressway	400-T	2	2	-0.5	North	diamond	tangent	2	380	Ramp grade is -5.1%
	EB Long Island Expy. to NB Cross Island Pky.	1200-P	1	1	+0.8	North	outer connector	slight left	2	1220	Ramp grade -5.0%, fwy curve slight rt.
	Community Dr. EB to Long Island Expressway	700-P	3	3	-2.0	East	diamond	slight left	2	450	Freeway curve slight left
	Jericho EB to Long Island Expressway	1200-P	1	1	-1.0	East	outer connector	slight right	2	660	
	Broadway EB to Northern State Parkway	440-P	8-30	8-30	-1.0	East	outer connector	slight right	2	600	Ramp grade is -4.4%, free-way curve slight right
St. Louis Area	SB Brentwood to WB Daniel Boone Expressway	593-T	14-30	14-30	+0.4	West	outer connector	slight right	2	600	Ramp grade is +3.8%, 2-lane

Discussion

R. D. WORRALL, Northwestern University—In recent years, the Texas Transportation Institute has been responsible for many significant contributions to the field of freeway design. The present paper is no exception. It describes some of the initial findings of a research project which has considerable implications for both the design engineer and the traffic analyst.

The paper falls roughly into two parts: a discussion of study techniques and an analysis of the effect of three specific design variables—acceleration lane length, ramp

convergence angle, and ramp grade—on entrance ramp operations. I should like to comment breifly on part two.

The problem which the authors address is essentially one of multivariate analysis, of isolating the effect of each of a set of interdependent variables on the value of a single, dependent variable. Two alternate lines of attack suggest themselves: that of formal "experimental design" and that of inductive "statistical analysis." Both approaches, in somewhat modified format, are illustrated in the paper. The cross tabulations of merge speeds, for example, classified according to four values of acceleration lane length and five ramp convergence angles, represent an elementary form of experimental control. Similarly, the multiple regression analysis of merge distances is an example of a simple inductive analysis. In each case, the results of the analyses suggest some interesting conclusions. These conclusions should, however, be interpreted with some caution. They are based on a somewhat limited data set and on a set of analysis which are intended to be illustrative rather than statistically rigorous. For more definite conclusions, we must await the completion of the authors' research and the preparation of a final report.

The cross tabulations of merge speeds, for example, are no real substitute for a formal experimental design. They provide no control over the many external variables—external, that is, to the main analysis—which might be hypothesized to exert at least as significant an influence on the speed of a merging vehicle as do acceleration lane length or ramp convergence angle. Such factors as the ramp alignment upstream of the merge area, for example, merge visibility and short-term fluctuations in traffic flow all have been shown to significantly influence the efficiency of entrance ramp operations (8). Insofar as those variables are unidentified or uncontrolled in the authors' analysis, they represent potentially dangerous sources of statistical confusion. It is by no means apparent that ignoring them will result in an accurate identification of the desired "main effects" of acceleration lane length or ramp convergence angle.[1]

In a somewhat related vein, one may perhaps also question the diffusion of study locations. The data set developed by the authors embraces 24 different ramps in 8 different U.S. cities. Rather than adding generality to the analysis, however, such a procedure merely increases the level of residual statistical noise, and hence reduces the clarity of the results.[2]

A further point may also be noted with respect to the cross tabulation of small data samples. The distribution of 24 data points across a 4 × 5 matrix results in an extremely sparse cell composition (Table 4). In this case, it also results in a measurable degree of collinearity within the data set. A simple correlation analysis of the sample illustrated in Table 4 yields a value of $r^2 = 0.3$. Although this figure is nonsignificant statistically, it represents a sufficient degree of collinearity within the original data set to suggest that the concluded effects of acceleration lane length and convergence angle are non-independent. Any attempt to introduce a third variable into the analysis—e.g., ramp grade—would merely make matters worse. The sample size is simply too small to permit such a multidimensional analysis to be performed.[3]

Similar collinearity effects are also apparent in the regression equation summarized in the authors' Table 3. In this case, the result is an instability of the individual regression coefficients, their true values being variously under or overestimated. The

---

[1]The presence of such confounding influences is tacitly acknowledged by the authors in their explanations of several apparently illogical or unexpected experimental results.

[2]One is tempted to suggest that the authors might better have concentrated their initial field studies in one area, possibly even at one location. In the latter case, at least, a "quasi-controlled" experiment would then be possible, in which the effect of acceleration lane length could be evaluated by successively lengthening the merge area at the given ramp and observing the resultant changes in traffic behavior.

[3]Despite the small sample size, the rigor of the analyses might have been enhanced somewhat by the application of any of several simple, nonparametric statistical tests to evaluate the significance of the differences between the distributions illustrated in text Figures 8-23. One cannot help thinking that some of these differences are perhaps exaggerated in the qualitative discussion.

TABLE 4

CLASSIFICATION OF SAMPLE OF 24 RAMPS BY ACCELERATION LANE LENGTH AND RAMP CONVERGENCE ANGLE

Ramp Convergence Angle	Acceleration Lane Length (100 ft)			
	0-3.5	4.0-6.0	6.0-8.0	9.5-12.0
10-14 deg	3	1	4	0
8-10 deg	0	1	0	0
6-8 deg	1	2	0	3
3-6 deg	0	0	4	0
0-3 deg	0	0	3	2

problem here might have been avoided by subjecting the original data set to an orthogonal factor analysis, and then regressing the values of the dependent variable against the orthogonal factor scores. The result would be essentially that intended in Table 3, with the advantage of a more meaningful, orthogonal set of stable equation coefficients.[4] An example of an analysis of this type, applied to data on merging at freeway entrance ramps, is given elsewhere (9).

Finally, I would like to make two comments relating to the authors' selection of criterion variables and to their conclusions concerning the operational effects of ramp grade.

The variables selected in the paper as measures of operational efficiency are basically of two types—speed measurements made somewhere within the ramp-freeway merge area and counts of the number of gaps rejected by a ramp driver before completing his merging maneuver. These parameters certainly reflect the general level under study. As noted, for example, the speed of a merging vehicle at the ramp nose may be hypothesized to be a function of the ramp alignment upstream of the nose, as well as of the acceleration lane length and convergence angle. It would be interesting if the authors could make available some alternate measures of merge efficiency—merge volumes, merge capacity, critical flow levels, accident records, etc.—for each location, to complement their present analyses. It would also be of considerable interest to know whether the different design locations were differentially sensitive to specific combinations of ramp-mainstream merge volumes.

In their final conclusion, the authors discuss the operational effects of ramp grade, and suggest that the improved field of view at a ramp leading down onto a depressed freeway has a beneficial effect on traffic operations. This conclusion is supported empirically by a recent study at Northwestern University (8) which found that merge sight distance, expressed as a function of ramp grade and convergence angle, was the most significant single factor influencing the efficiency of ramp-freeway merge operations. Serious limitations on visibility were found to exist, particularly for merging trucks, at ramps with convergence angles of 2 deg or more and for ramps approaching the mainstream lanes either on an upgrade or from the left-hand side of the traveled way.

[4] A word of caution is in order here. These simple coefficients reflect the dimensions of the original variables; they do not reflect the comparative contribution of each variable to the value of Y unless they are first standardized. Hence the discussion following Table 3 is somewhat deceptive.

References

8. Northwestern University, Department of Civil Engineering. The Suitability of Left-Hand Ramps for Freeways. Final Report, Illinois Cooperative Research Project IHR-61, Northwestern Univ., 1966.
9. Worrall, R. D., Coults, D. W., Echterhoff-Hammerschmid, H., and Berry, D. S. Merging Behavior at Freeway Entrance Ramps: Some Elementary Empirical Considerations. Highway Research Record 157, pp. 77-107, 1967.

JAMES D. BLACKBURN, City Traffic Engineer, Raleigh, North Carolina—The research being conducted at the Texas Transportation Institute and reported in this paper is valuable. It allows a fresh look at some of the design practices which have resulted in entrance ramp geometrics providing less than desirable traffic operation.

The study measured operational characteristics including the speed of ramp vehicles at the ramp nose and merge point, the change in the speed of ramp vehicles moving along the acceleration lane, the difference in speeds between freeway and ramp vehicles at the ramp nose and merge point, and the gap number accepted by merging vehicles and the distribution of points of entrance onto the freeway. The report relates those operational characteristics to three elements of geometric design: acceleration lane length, angle of convergence, and ramp grade.

One of the more significant findings of the study was in revealing that entrance ramps having significant downgrades as they approach the ramp nose provide traffic operation superior to what would be expected, based on a ramp's angle of convergence, length of acceleration lane, and volume of traffic being handled. This superior operation was explained in a simple and logical manner, i.e., motorists approaching the freeway on a downgrade have a field of view of the freeway lanes, allowing them to select a gap for merging prior to reaching the nose and then accelerate to meet the gap immediately beyond the ramp nose. I do not believe that designers have commonly taken this aspect into consideration when designing entrance ramp elements varying from rather steep upgrades approaching the freeway to desirable downgrades commonly found on depressed freeways.

Although the report states that at some locations the large volume of ramp traffic resulted in data that were not comparable to other locations or could not be reconciled with data from other locations, no specific correlation was made between the total volume of merging traffic and operational characteristics. Poor operation occurs when the total volume of ramp traffic and traffic in the outside freeway lane approaches the capacity of that lane. Differences in that total volume from location to location may explain some of the variability of the data presented.

Common design practice has based necessary acceleration lane length on the distance required to accelerate from the ramp design speed up to the freeway design speed. I believe this approach may be inadequate. It is presumed that many of the ramps studied had acceleration lane lengths meeting the above criteria but showed less than acceptable operational characteristics, whereas others with relatively short acceleration lanes exhibited good operational characteristics.

The reported change in speed of ramp vehicles proceeding from the ramp nose to the point of merging is significant. At those locations having acceleration lanes at least 450 ft long and angles of convergence less than 8 deg, the mean speed change along the acceleration lane was never greater than 6 mph. This seems to indicate that, although long acceleration lanes are quite necessary, the term acceleration lane may be a misnomer with good design. At all those locations having shorter acceleration lanes or larger angles of convergence, the mean speed changes were greater than 6 mph and ranged upwards to 17.5 mph. Drivers approaching well-designed entrance ramps and long acceleration lanes apparently feel it is not necessary to pass the nose

at a low speed for they have considerable distance beyond the nose to accept a gap and merge. On the other hand, short acceleration lanes force the driver to approach the nose at a low speed, select his gap and be prepared to immediately merge; otherwise, he may find himself at the end of the acceleration lane and no acceptable gap available. Although ramp vehicles must certainly accelerate up to the operating speed of freeway traffic, I believe that more significant determinates of acceleration lane lengths might be ramp grades approaching the nose and volume of traffic on both the ramp and the outside freeway lane.

ADOLF D. MAY, JR., Associate Professor of Civil Engineering, Institute of Transportation and Traffic Engineering, University of California, Berkeley—The subject of this paper is one of considerable importance to design engineers, and the consequences of such analyses are essential in adopting sound design standards, designing specific ramp configurations, and evaluating existing ramps. Rather than briefly covering several parts of the paper, this discussion will concentrate on one particular aspect—that of statistical analysis and interpretation. The major theme of this discussion is to stress that this is a multivariate problem and techniques such as multiple regression analysis should be considered.

The objective of this nationwide project was to study the effects of geometric variables on entrance ramp operation. The geometric variables quantitatively examined were as follows:

1. Acceleration lane length,
2. Angle of convergence at the ramp nose,
3. Angle of convergence 2 feet off pavement edge,
6. Length of ramp,
7. Ramp grade, and
8. Ramp lane width.

The measures of effectiveness of ramp operations included:

1. Speed of ramp vehicles at ramp nose (percent traveling < 30 mph),
2. Speed of ramp vehicles at merge point (mph),
3. Speed changes on the acceleration lane (mph),
4. Relative freeway-ramp vehicle speed at the merge point (mph),
5. Mean of accepted gap number, and
6. Standard deviation of accepted gap number.

These eight geometric variables and six measures of effectiveness of ramp operations were quantitatively determined for each of the 22 selected ramps in the study.

Sample multiple regression analyses were undertaken by the discusser utilizing the measurements contained in the paper. The intent of this sample analysis was to demonstrate the application of multiple regression analysis to the data given in the paper and to present the interpretations one can draw from such an analysis. A rather unique stepwise multiple regression program was utilized on an IBM 1620 computer for this problem.

The geometric features were considered as the independent variables while each individual measure of effectiveness was considered as the dependent variable. The authors did not report on the results of any test to determine the independence of the geometric features studied. A t-test was applied to the geometric features studied by the discusser and the correlation coefficients between (a) acceleration lane length and convergence angle at nose, (b) acceleration lane length and convergence angle at the

$$Y_1 = -7.80 + 6.87X_3 + 8.71X_4$$

$$Y_2 = 43.8 - 0.715X_2$$

$$Y_3 = -0.115 + 0.954X_3$$

$$Y_4 = 11.6 - 0.00915X_1$$

$$Y_5 = 3.94 + 0.0618X_3 - 0.00037X_6 - 0.224X_8$$

$$Y_6 = 5.13 + 0.0937X_3 - 0.408X_8$$

---

where

Y_1 = speed of ramp vehicles at ramp nose (percent traveling <30 mph)

Y_2 = speed of ramp vehicles at merge point (mph)

Y_3 = speed changes on the acceleration lane (mph)

Y_4 = relative freeway-ramp vehicle speed at the merge point (mph)

Y_5 = mean of accepted gap number

Y_6 = standard deviation of accepted gap number

X_1 = acceleration lane length

X_2 = angle of convergence at the ramp nose

X_3 = angle of convergence 2 feet off pavement edge

X_4 = grade at nose

X_5 = curb offset at nose

X_6 = length of ramp

X_7 = ramp grade

X_8 = ramp lane width

Figure 30. Multiple regression equations.

2-ft offset point, and (c) the convergence angle at the nose and the convergence angle at the 2-ft offset point were significant at the 0.05 level. The authors may wish to consider these results if multiple regression analysis techniques are pursued further.

Only linear relationships were considered in this sample analysis, although the stepwise multiple regression program did permit nonlinear equations. The general equation format is

$$Y_i = A + B_1 X_1 + B_2 X_2 + B_3 X_3 + \ldots\ldots + B_N X_N$$

Inspection of several independent variable vs selected dependent variable plots indicated that linear relationships were not unreasonable but nonlinear relationships may result in improved multiple regression analysis. Again the authors may wish to consider nonlinear relationships if the multiple regression analysis techniques are pursued further.

Six multiple regression analyses were undertaken, one for each measure of effectiveness. The resulting equations which include only the significant (at the 0.05 level) independent variables are shown in Figure 30. The following interpretations were drawn from these multiple regression analyses:

1. Decreasing the angle of convergence at the 2-ft offset point and reducing the ramp grade at the nose significantly reduces the percent of ramp vehicles traveling less than 30 mph at the ramp nose.
2. Decreasing the angle of convergence at the ramp nose significantly increases the speed of ramp vehicles at the merge point.
3. Decreasing the angle of convergence at the 2-ft offset point significantly decreases the speed changes on the acceleration lane.
4. Increasing the length of the acceleration lane significantly decreases the relative freeway-ramp vehicle speeds at the merge point.
5. Decreasing the angle of convergence at the 2-ft offset point increasing the length of the ramp, and increasing the ramp lane width, significantly decreases the accepted gap number.
6. Decreasing the angle of convergence at the 2-ft offset point and increasing the ramp lane width significantly decreases the standard deviation of accepted gap number.

JOSEPH A. WATTLEWORTH, JOHANN H. BUHR, DONALD R. DREW, and FRANK A. GERIG, JR., Closure—The authors would like to express their appreciation to James D. Blackburn, Adolf D. May, and Richard D. Worrall for their fine and stimulating discussions.

It was interesting to note in these discussions the divergence in the views of the "operational world" as expressed by Blackburn and those of the "academic world" as expressed by May and Worrall. Blackburn, being an operating engineer, evaluated the results and conclusions in the framework of his experience, while the two distinguished academic representatives concentrated their discussions on the research

to the acid test of application or use by operating people. It was gratifying to note that the presently reported results and conclusions seem to agree with the experience of Blackburn.

Both May and Worrall have pointed out the fact that the paper was primarily descriptive rather than rigorously analytical, and this is certainly true. For additional analytical work on the same data set, the reader's attention is directed to the other seven reports on this project (10 to 16).

May and Worrall indicated the multivariate nature of the problem which, of course, is widely recognized. The research approach taken was that of selecting study sites at which the effects of other variables were a minimum. For example, the ramps were selected, as much as possible, to merge with freeways which were on a level, tangent section; an attempt was made to select study times to obtain traffic volumes near but not over capacity; study sites and time periods were selected so that the operation of the ramp itself was unaffected by downstream bottlenecks; of course, with an aerial photographic technique, all data were collected in good weather. This is not to say that there were no extraneous variables or that the rudimentary attempts to achieve some experimental control were entirely successful.

Experimental design is a function of several factors. Worrall discussed the experimental design from the viewpoint of the rigorous statistician. Other factors such as constraints on time and budget, frequently force the researcher to compromise in his research planning.

The concept of controlled experiments at one location rather than studying several sites has considerable merit, and to this end the Texas Transportation Institute is developing a test track to permit this type of work.

In the reported research, it would have been possible to have studied the effect of acceleration lane length at one site by varying the acceleration lane length there (if a

local operating agency could have been convinced to shorten temporarily the length of an acceleration lane so that studies could be made to determine the adverse effect of this action). Ramp grade and convergence angle could not have been so easily changed. In addition, this type of experimental design would not have fit easily within the time requirements of the project.

Worrall also expressed concern that differences in driver characteristics in different geographical areas might have increased "the level of residual statistical-noise." The authors will concede that driver characteristics might differ from one geographical area and also have a possible explanation for this. In conducting the field studies on this project, it soon became quite clear that the geometric designs also tended to be peculiar to the geographical areas. For example, in California, diamond interchanges of high type, standard designs are widely used; in Texas, most freeways are at grade and the "slip type" diamond interchange ramps are common, while in other areas the cloverleaf interchange is more usual. It could be argued quite strongly that the driving characteristics of the freeway motorist are quite a critical function of the type of geometrics on the freeways on which he learned to drive. In other words, the geographic differences in driver characteristics may be a function of the very ramps being studied.

Another interesting comment of Worrall's is that the conclusions are based on "a somewhat limited data set." This is interesting in light of the fact that the data set used and collected on this project probably represents the most complete set of data on freeway merging characteristics which has ever been assembled. To have followed some of Worrall's suggestions for increasing the sample size in each cell in the matrix and for increasing the number of variables (and hence the number of cells) would quickly have led to a study of monumental proportions. Again, the constraints of time and money provided some limitation on the sample sizes obtained. Within the budget constraint and a 20-month project duration, the project staff was successful in:

1. Developing a data acquisition system which met the requirements of (a) mobility, (b) low cost, (c) minimum interference to traffic, (d) minimum requirements for help from local agencies, and (e) necessary data;
2. Conducting studies on more than 30 entrance ramps in 8 cities from coast to coast;
3. Developing a data reduction and analysis system to obtain and process the large volume of data; and
4. Interpreting the results and preparing eight technical reports.

Both May and Worrall discuss the problem of non-independence of acceleration lane length and convergence angle and arrive at different conclusions regarding its statistical significance. Based on the geometry of merging areas and normal geometric design procedures, however, it would be surprising if the merging angle and acceleration lane length are entirely independent.

The authors would like to urge a great deal of reservation in accepting as final the results of May's regression analyses. In particular, the fact that entrance ramp grade failed to appear as a significant variable in the regression analyses should be interpreted with a great deal of caution. The authors choose to await the results of further research which is under way at the Texas Transportation Institute and which will document more completely the effects of ramp grades on merging operation, before accepting the suggestion that this geometric variable has little or no effect on merging operation.

The authors would once again like to express their thanks and gratitude to the three reviewers for their discussions of this paper. They raised many salient points which will certainly be of benefit to future research.

References

10. Buhr, J. H., Drew, D. R., Wattleworth, J. A., and Williams, T. G. A Nationwide Study of Freeway Merging Operations. Texas Transportation Institute Research Report 430-1, 1966.

11. Drew, D. R., LaMotte, L. R., Buhr, J. H., and Wattleworth, J. A. Gap Acceptance in the Freeway Merging Process. Texas Transportation Institute Research Report 430-2, 1966.
12. Drew, D. R., Buhr, J. H., and Whitson, R. H. The Determination of Merging Capacity and Its Application to Freeway Design and Control. Texas Transportation Institute Research Report 430-4, 1967.
13. Buhr, J. H. Traffic Interaction in the Freeway Merging Process. Texas Transportation Institute Research Report 430-5, 1967.
14. Drew, D. R., Meserole, T. C., and Buhr, J. H. Digital Simulation of Freeway Merging Operation. Texas Transportation Institute Research Report 430-6, 1967.
15. White, F. S. Annotated Bibliography on Gap Acceptance and Traffic Interaction in the Merging Process. Texas Transportation Institute Research Report 430-7, 1967.
16. Drew, D. R., Wattleworth, J. A., Buhr, J. H., and Williams, T. G. Gap Acceptance and Traffic Interaction in the Freeway Merging Process. Texas Transportation Institute Final Research Report, 1967.