HIGHWAY RESEARCH RECORD

Number 230

Characteristics of Traffic Flow

5 Reports

Subject Area

53 Traffic Control and Operations
54 Traffic Flow
55 Traffic Measurements

HIGHWAY RESEARCH BOARD

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

Washington, D.C., 1968

Publication 1617

Price: $2.60

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
Purdue University, Lafayette, Indiana

HIGHWAY RESEARCH BOARD STAFF

E. A. Mueller

COMMITTEE ON QUALITY OF TRAFFIC SERVICE
(As of December 31, 1967)

Johannes F. Schwar, Chairman
Ohio State University, Columbus

Jack B. Blackburn
Martin J. Bouman
Richard K. Bowler
Siegfried M. Breuning
Joseph C. Corradino
Donald R. Drew
H. A. Mike Flanakin
Stephen George, Jr.
Bruce D. Greenshields
David W. Gwynn
Edward M. Hall
C. Dwight Hixon
Hyoungkey Hong
Thomas D. Jordan
Joseph M. McDermott
Stuart R. Perkins
Charles C. Schimpeler
Ralph E. Schofer
Kenneth W. Shiatte
Vasant H. Surti
Joseph Treiterer
Robley Winfrey

COMMITTEE ON CHARACTERISTICS OF TRAFFIC FLOW
(As of December 31, 1967)

Joseph C. Oppenlander, Chairman
Purdue University, Lafayette, Indiana

Patrick J. Athol
Jack B. Blackburn
Martin J. Bouman
Kenneth W. Crowley
Olin K. Dart, Jr.
Robert F. Dawson
H. M. Edwards
John J. Haynes
Clinton L. Heimbach
James H. Kell
Russell M. Lewis
Jack C. Marcellis
Peter A. Mayer
F. William Petring
O. J. Reichelderfer
August J. Saccoccio
Charles C. Schimpeler
Joseph Seifert
William P. Sheldon
William C. Taylor
Kenneth J. Tharp
Robert J. Wheeler

Foreword

A pressing problem to those concerned with the satisfactory operation of streets and highways is the high cost and sophistication required to measure the effects of changes resulting from the application of traffic engineering measures. Researchers have been concerned with theoretical analysis, simulation and mathematical modeling as methods that might be useful in determining the adequacy and the specific effects of traffic engineering changes. The five papers in this RECORD indicate some of the extensive work that has been done in the past year. Certainly, engineers who are attempting to move traffic more efficiently and researchers who are attempting to employ more sophisticated approaches will find these papers of interest.

At the University of Connecticut, investigators used an application of probability distribution to model headways mathematically—in this instance a 2-lane, two-way highway. Testing the derived headways against two sets of data indicated good correspondence to the real-life situation although the authors caution that more substantiation may be needed for wider ranges of traffic and roadway conditions.

At the Texas Transportation Institute, simulation techniques were used to study a freeway section including exit and entrance ramps. Different modes of controlling entrance ramp traffic were evaluated. The paper discusses the computer logic and simulation techniques employed and shows a limited output of the program.

At North Carolina State University a research project attempted to develop a new way of measuring traffic volumes with data developed from a car equipped with a Drivometer. Using two of the eight Drivometer events recorded (brake applications and large steering reversals) combined with multiple-regression analysis, equations were developed that predicted traffic volumes. The authors surmise that if a suitable sample of Drivometer-equipped cars were used, volumes could be predicted within limits of statistical confidence.

A Louisiana State University professor studied left-turn signalized intersection characteristics in his effort to get suitable data for simulation modeling. Among some of the characteristics pinpointed were the observations that most left-turners stop near the center of the intersection to wait for a gap before turning, gaps shorter than two seconds are rarely accepted, and gaps longer than eight seconds are rarely rejected. There is also a 15 percent chance that left-turners will "jump the gun" to make their turns when the signal turns to green.

A Wayne State University researcher used mathematical modeling in developing a method to predict changes in a specified measure of the level of services provided by a street system under various system controls. Study of a street system having a freeway within it and attempting to maximize overall speed by various changes in the system indicated that increasing freeway capacity would achieve the greatest improvement in total system output. Use of the model could provide a method for evaluating traffic control methods and could also indicate a method for determining the effects of changes in single elements on the entire system.

Contents

The Hyperlang Probability Distribution— A Generalized Traffic Headway Model

R. F. DAWSON and L. A. CHIMINI, University of Connecticut

This research is concerned with the development of the hyperlang probability distribution as a generalized time headway model for single-lane traffic flows on two-lane, two-way roadways. The study methodology involved a process that is best described as "model evolution," and included: (a) identification of salient headway properties; (b) construction of mathematical micro-components to simulate essential headway properties; (c) integration of the micro-components into a general mathematical headway model; (d) numerical evaluation of model parameters; and (e) statistical evaluation of the model.

The proposed hyperlang headway model is a linear combination of a translated exponential function and a translated Erlang function. The exponential component of the distribution describes the free (unconstrained) headways in the traffic stream, and the Erlang component describes the constrained headways. It is a very flexible model that can decay to a simple exponential function, to an Erlang function, or to a hyper-exponential function as might be required by the traffic situation. It is likely, however, that a traffic stream will always contain both free and constrained vehicles; and that the general form of the hyperlang function will be required in order to effect an adequate description of the composite headways.

The parameters of the hyperlang function were evaluated for data sets obtained from the 1965 Highway Capacity Manual and from a 1967 Purdue University research project.

The proposed hyperlang model proved to be a sound descriptor of the reported headways for volumes ranging from about 150 vph to about 1050 vph. However, it should be substantiated and evaluated for a wide range of traffic and roadway conditions. During the conduct of any future research, careful attention should be given to proper flow rate monitoring and to proper data stratification to reflect the variations in headway characteristics that are caused by variations in traffic and roadway conditions.

•HEADWAYS, or time spacings between successive vehicles, are one of the basic characteristics essential to the description of a traffic stream. Although they are seemingly a simple aspect of traffic flow, they can be used collectively as an index of available stream capacity, or as an index of the level of stream congestion. In one of the first extensive studies of traffic headways, reported in the 1950 Highway Capacity Manual, Normann utilized the headway distribution as a basis for describing the level of service available in a traffic stream (9).

It is also apparent that at any point in the traffic lane the vehicular arrival rate is regulated by the sequence of headways between successive vehicles in the stream. In selecting a headway, a driver is simultaneously establishing a flow rate for his vehicle.

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 47th Annual Meeting.

Normally, the time-weighted mean of these individual observations is reported as the average flow-rate for the stream.

As a consequence of the importance of headways as a traffic flow descriptor, considerable effort has been expended for the purpose of developing a sound, general purpose headway model. Most of this effort has been directed toward fitting general mathematical functions to observed headways distributions. Of the many probability distributions that have been proposed, the negative-exponential function, the hyper-exponential function, the Erlang function, and the log-normal function are the better known. No one of these, however, has been found completely adequate as a general purpose headway generator.

The earliest research attempts were directed toward fitting the negative-exponential function to observed headway data (5, 6, 12). And, although this model proved to be adequate for low volume streams, it is not conceptually sound. It is based on the assumptions that successive events of concern are independent random occurrences, and that they exist over a $(0, \infty)$ range. Several studies have demonstrated that successive vehicles are not independent of each other. In an early research report, Normann showed that the relative speeds of successive vehicles were dependent on the headway between the vehicles (9). More recent research of the car-following phenomena essentially has proved intervehicular dependence (8). Of course the existence of headways over the $(0, \infty)$ range is a physical impossibility. There is a minimum headway in a traffic stream; this minimum is related to the length of the lead vehicle, to the minimum intervehicular spacing demanded by the trailing vehicle, and to the speed and acceleration of the trailing vehicle.

Some researchers have proposed a negative-exponential function with a translated axis to reflect the existence of a real minimum headway (2), but this modified model does not fully compensate for intervehicle dependence.

In an early research study Schuhl suggested the hyper-exponential model (16). Basically it is a linear combination of a negative-exponential function and a translated negative-exponential function. The translated component was proposed as a descriptor of the headways between constrained vehicles, and the simple exponential component was proposed as a descriptor of the free headways. Kell later modified Schuhl's model so that both the free and the constrained headways were represented by translated functions (11). The proposed modification was rational due to the fact that every vehicle is physically restricted from traveling at a zero headway. With this modification, Kell was able to obtain good model fits for volumes up to about 700 vph. In a recent study, however, Sword was able to fit the hyper-exponential function, as it was originally proposed by Schuhl, over the same volume range (17).

In an Australian research study Buckley proposed another version of the compound headway model, which he referred to as a semirandom model (1). He supports Schuhl's theory that a traffic stream is made up of free and constrained vehicles, and like Schuhl proposes that the negative-exponential function be used to describe the free headways. However, Buckley used a normal distribution to describe the constrained headways.

Other researchers have proposed another function. In an attempt to describe the apparent decay of randomness at higher volumes they have suggested the so-called Erlang distribution (7, 15). By varying the model parameters it is possible to simulate time spacings between events that vary from being completely random to being completely nonrandom or uniform. It is also possible to translate the axis of the Erlang function to take into account the physical constraint that prevents zero and near zero spacings. Nevertheless, there has been little success in fitting this model, except at high and low volume extremes, where the major portion of the headways are either restricted or free, respectively. The Erlang model is just not flexible enough to afford simultaneous descriptions of both free and constrained headways.

More recently Greenberg suggested the log-normal function as a headway model (4). The log-normal function, however, is merely a mathematical transformation that tends to approximate several members of the family of Erlang functions.

Scope of Study

The limited success that has been encountered in attempts to isolate a general purpose headway model can perhaps be attributed to the use of deductive research techniques. Known mathematical models have been proposed, and the parameters of these models have been determined by model fitting techniques.

A more direct study approach is proposed. It involves a methodology that is best described as model evolution, and includes the following steps:

1. Identification of headway properties that are essential for a sound, realistic description of traffic flow past a point;
2. Construction of mathematical micro-components to simulate critical headway properties;
3. Integration of the micro-components into a general mathematical headway generating function;
4. Determination of numerical values for various model parameters; and
5. Evaluation of the model.

THE PROPOSED HEADWAY MODEL

Identification of Critical Properties

Two important headway properties have been identified in previous research studies. In the first place, it has been established that there are at least two types of vehicles in a traffic stream (11, 16). For descriptive purposes, these types are referred to as free vehicles and constrained vehicles. Free vehicles are those that are not under the influence of other vehicles in the traffic stream. For the most part this condition exists when the headway from the free vehicle to preceding vehicles is of "adequate" duration, when the free vehicle is able to pass so that it does not have to modify its time-space trajectory as it approaches preceding vehicles, or when a passing vehicle has sustained a positive relative speed after the passing maneuver so that the free vehicle is still able to operate as an independent unit. Constrained vehicles, of course, are those that are under the influence of other vehicles in the stream. It has also been observed that the balance between these free and constrained vehicles varies with the flow rate of the traffic stream. As the flow rate increases, the proportion of free vehicles decreases and the proportion of constrained vehicles increases (11).

In the second place, it has been established, both by observation and rationalization, that there is a real minimum headway in a traffic stream that is related to the size and the finite velocity of the vehicle (14). If the traffic stream is a single-lane stream, and headways are measured just between successive vehicles in that stream, there must be real minimums for both the free and constrained headway distributions. If the headways are also measured between the free vehicles occupying the adjacent lane during a passing maneuver, the lower limit for the free headway distribution is zero.

Construction of Mathematical Micro-Components

Each of the micro-aspects of a headway distribution can be simulated by an appropriate mathematical function. Because of its random nature, the distribution of free vehicle headways is readily simulated by a negative-exponential distribution of the form,

$$P_f(t) = e^{-\beta_1 t}$$

where

$P_f(t)$ = the probability that a free headway is equal to or greater than t;
β_1 = the free vehicle flow rate; and
t = any time duration.

When the headways are measured on a lane-by-lane basis, without consideration for the headways of passing vehicles occupying the adjacent lane, minimum headway limits can be simulated by a boundary condition on the generator. Mathematically, the boundary is effected by translating the axis of the function so that the model takes the form,

$$P_f(t) = 1. \quad : \quad 0 \leq t \leq \delta_1$$

$$P_f(t) = e^{-\frac{(t - \delta_1)}{(\gamma_1 - \delta_1)}} \quad : \quad \delta_1 \leq t \leq \infty$$

where

$P_f(t)$ = the probability that a free headway is equal to or greater than t;
t = any time duration;
δ_1 = the minimum free headway; and
γ_1 = the average free headway.

The headways between constrained vehicles in a traffic stream tend to be nonrandom, and in some instances appear to approach uniformity. Nonrandom phenomena, ranging from phenomena that are completely random to phenomena that are completely uniform, are readily simulated by an Erlang function. Mathematically the Erlang function takes the form,

$$P_c(t) = e^{-k\beta_2 t} \sum_{x=0}^{k-1} \frac{(k\beta_2 t)^x}{x!}$$

where

$P_c(t)$ = the probability that a constrained headway is equal to or greater than t;
k = an index that indicates the degree of nonrandomness in the constrained headway distribution;
β_2 = the constrained vehicle flow rate; and
t = any time duration.

Of course the lower time bound on the constrained distribution is a real limit greater than zero. To reflect this, the Erlang function was modified to the form,

$$P_c(t) = 1. \quad : \quad 0 \leq t \leq \delta_2$$

$$P_c(t) = e^{-k\frac{(t - \delta_2)}{(\gamma_2 - \delta_2)}} \sum_{x=0}^{k-1} \frac{\left(k\frac{(t - \delta_2)}{(\gamma_2 - \delta_2)}\right)^x}{x!} \quad : \quad \delta_2 \leq t \leq \infty$$

where

$P_c(t)$ = the probability that a constrained headway is equal to or greater than t;
k = an index that indicates the degree of nonrandomness in the constrained headway distribution;

δ_2 = the minimum headway in the constrained headway distribution;
γ_2 = the average headway in the constrained distribution; and
t = any time duration.

In order to reflect the relative proportion of the total vehicles in the stream that are either free or constrained, it is only necessary to introduce linear coefficients, α_1 and α_2, before each of the component functions. The α_1 denotes the proportion of free vehicles in the traffic stream, and $\alpha_2 (\alpha_2 = 1. - \alpha_1)$ denotes the proportion of constrained vehicles in the stream.

The Complete Model

The general purpose headway model is obtained by forming a linear combination of the free and constrained components. The integrated model takes the form,

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

This model has been proposed by Buckley (1) and by Dawson (3), but it has not been researched. Dawson suggested that it be called the hyper-Erlang function. For the purpose of this paper, however, the name has been shortened to the "hyperlang" function.

The hyperlang function is a very general model; in fact it includes the negative exponential, the hyper-exponential, and the Erlang functions as special cases. This can be illustrated with schematic comparisons of the several functions.

An exponential headway simulator is depicted in Figure 1a. From a continuous queue at the entrance to a holding area, individual vehicles are randomly released. If the vehicles are metered out of the holding area at a mean flow rate of β vehicles per unit time, the resulting headways will follow an exponential distribution with a mean of $\bar{t} = 1/\beta$.

Figure 1b depicts an Erlang headway simulator. From a continuous queue at the entrance to a holding area made up of several phases in series, one vehicle at a time is allowed to enter the area. The entering vehicle goes through phase (1) where the holding times are exponentially distributed. Upon leaving phase (1), the vehicle enters phase (2), phase (3), ---, phase (k-1), and phase (k), each with a mean rate of release of $k\beta$. The distribution of times between departures of the vehicles from phase (k), and therefore from the overall holding area, are described by an Erlang distribution with a mean time,

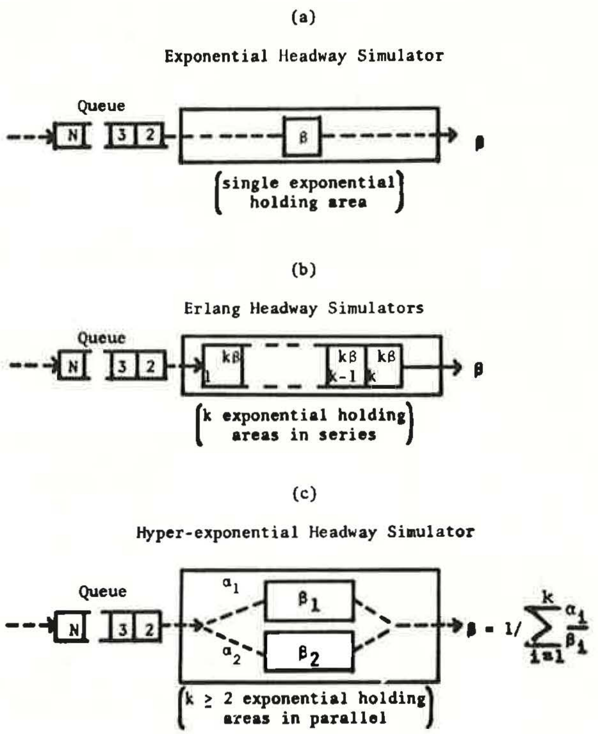


Figure 1. Schematic diagram of exponential, Erlang and hyper-exponential headway simulators.

$\bar{t} = 1/\beta$. Although the holding area is compounded, it should be considered as a single area because only one vehicle at a time is allowed in it. That is, the (n + 1) vehicle cannot enter phase (1) until the n vehicle has departed from phase (k).

Hyper-random or hyper-exponential headways can be generated by the simulator shown in Figure 1c. From a continuous queue at the entrance to a holding area made up of k channels in parallel, one vehicle at a time is allowed to enter the area. As a vehicle enters, it is assigned to one of the k channels at random, but on the average it is assigned to the i th channel α_i percent of the time. When a vehicle is held in any one of the channels, the entire holding area is considered to be occupied and no other vehicle can enter. The distribution of headways between departures from this holding area is described as a hyper-exponential distribution with a mean time between departures of $\bar{t} = 1/\beta$, where $\beta = \dfrac{1}{\sum\limits_{i=1}^{k} (\alpha_i/\beta_i)}$.

It is apparent that the single exponential holding area is a special case of both the Erlang series arrangement and the hyper-exponential parallel arrangement. It is the case in which there is only one phase in the series arrangement, and the case in which there is only one channel in the parallel arrangement. The exponential distribution is therefore a boundary between the Erlang distribution and the hyper-exponential distribution.

The hyperlang headway generator is shown in Figure 2. Close inspection shows that it is nothing more than two Erlang service channels in parallel. Of course one of these Erlang channels is the special exponential case in which there is only one holding area.

In general application it is possible for the Erlang-k channel to fade (no constrained vehicles in the stream) leaving only the simple exponential system; it is possible for the exponential channel to fade (no free vehicles in the stream) leaving only the Erlang system; and finally it is possible for the Erlang-k channel to decay to a simple exponential channel leaving a hyper-exponential system. It is likely, however, that a traffic stream will always contain both free and constrained vehicles; and that the general hyperlang function will be required in order to effect a complete description of the composite headways. The headways that are generated by this simulator will follow the proposed hyperlang function with a mean headway of $\bar{t} = 1/\beta$, where $\beta = \dfrac{2}{\sum\limits_{i=1}^{2} (\alpha_i/\beta_i)}$.

NUMERICAL EVALUATION OF MODEL PARAMETERS

Selection of Data

The parameters for the hyperlang headway model were evaluated for 1-lane flows on 2-lane, two-way roadways. Two separate analyses were conducted. One was based on

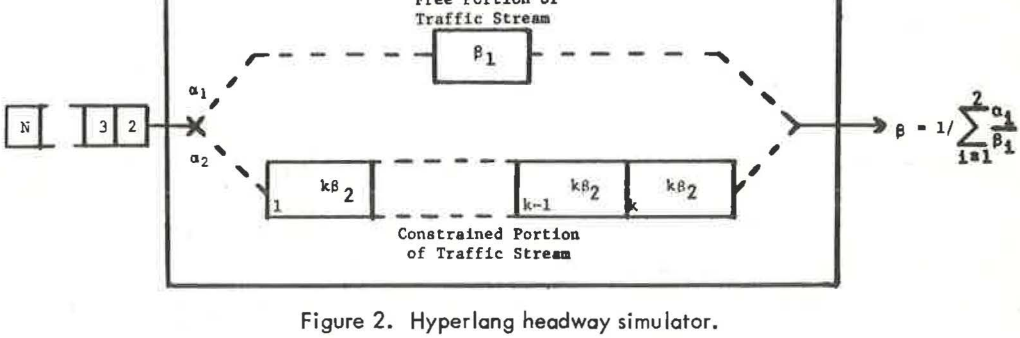


Figure 2. Hyperlang headway simulator.

the headway data for 2-lane roadways reported in the 1965 Highway Capacity Manual (10), and the other analysis was based on data from a recent Purdue University study (17). Plots of the cumulative headway distributions depicting the data sets obtained in these two studies are shown in Figures 3 and 4, respectively. The several plots obtained from

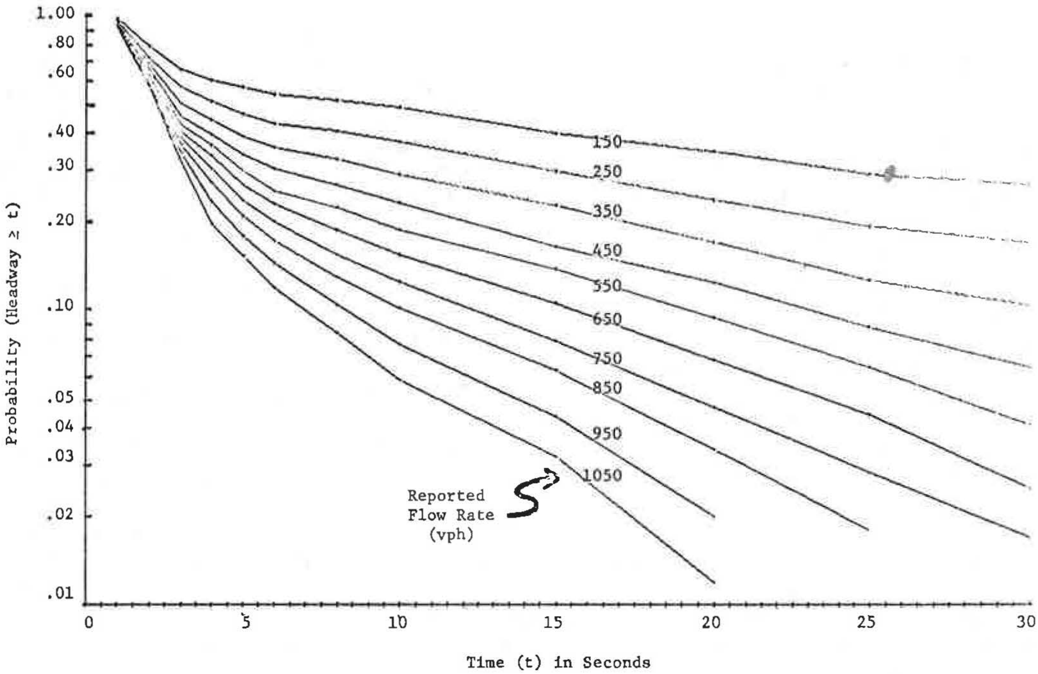


Figure 3. Observed cumulative headway distributions reported in 1965 Highway Capacity Manual.

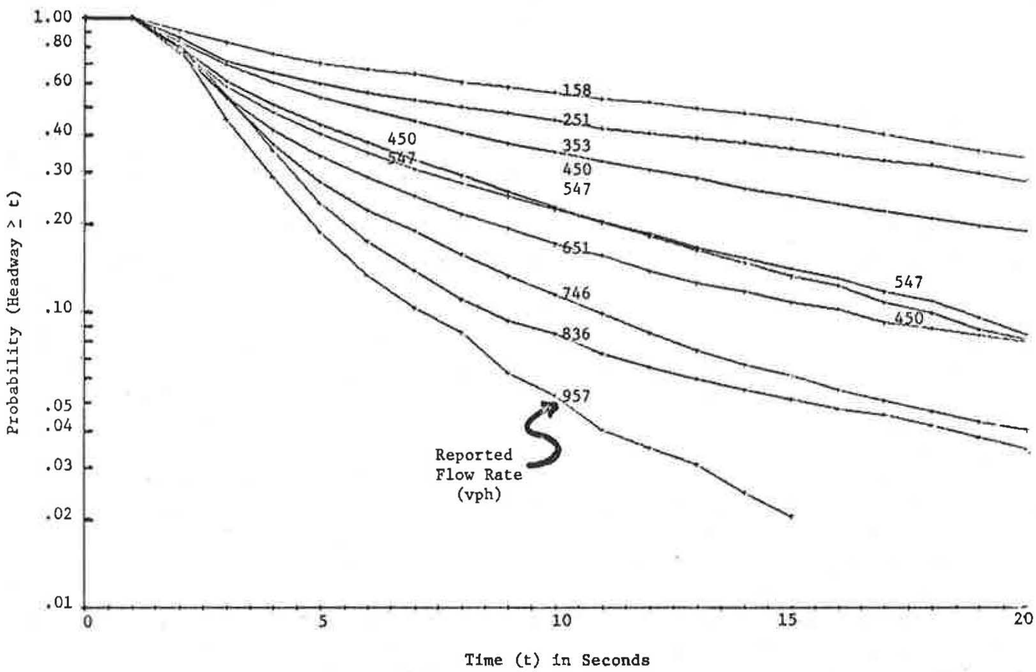


Figure 4. Observed cumulative headway distributions reported in Purdue research project.

the 1965 Highway Capacity Manual (Fig. 3) were purported to represent flow rates ranging from 150 to 1050 vph in increments of 100 vph, and in general, they tend to form a uniform family. There are some apparent irregularities in the higher volume curves, at the longer headways; but these irregularities are not as serious as they appear to be. In this region of the plots, minor irregularities are greatly exaggerated by the logarithmic probability scale.

The data obtained from the Purdue study (Fig. 4) are purported to have been collected at flow rates ranging from approximately 150 to 950 vph, in increments of approximately 100 vph; but as can be seen from the headway plots, the flow rates apparently were not monitored accurately. The 251-vph curve tends to meander from the 353-vph curve at low headways up to the 151-vph curve at longer headways. The 450, 547, and 651-vph curves are all rather closely spaced; in fact, the 450 and 547-vph curves actually cross each other. The three higher volume plots are also somewhat peculiar. The 746 and 836-vph curves tend to be quite similar, whereas the 957-vph curve tends to diverge from the family. This lack of uniformity indicates that the reported flow rates are average flow rates for periods of unsteady flow. The composite headway distribution that is formed by combining the headways from two streams with different flow rates, however, is quite different from the distribution of headways from a stream that flows uniformly at the same apparent rate. These criticisms do not invalidate the data sets; they merely emphasize the importance of properly monitoring the traffic stream while making headway measurements.

These data sets were particularly appropriate for the hyperlang evaluation studies. The hyperlang function was rationally constructed to effect a sound descriptor for traffic headways. Each of the seven function parameters has physical significance. The consistent and uniform nature of the 1965 Highway Capacity Manual data afforded the opportunity both to evaluate the parameters numerically, and to observe the relationships between the parameters and the corresponding flow rates. On the other hand the erratic nature of the Purdue University data afforded the opportunity to observe the power of the hyperlang function as a descriptive headway model. As suggested earlier these nonuniform data sets are very likely arbitrary combinations of free and constrained headways that were measured during periods of unsteady flow. The facility (or lack thereof) to describe these arbitrary distributions is an indication of the power of the hyperlang function. It was not likely, however, that the model parameters would form a consistent relationship with the reported flow rates.

Determination of Parameters

Parameter evaluation was effected in two steps. In the first step, initial estimates for the parameters of the hyperlang function were obtained using a rational subdivision technique. These initial estimates were then refined in a second step using the method of nonlinear least squares.

By definition the subdistribution of the constrained headways cannot overlap the upper region of the subdistribution of the free headways. Because of this, the upper regions of the cumulative plots of Figures 3 and 4 describe only free headways. Thus, it was possible to obtain initial estimates for the parameters of the free headway subdistribution by fitting the best translated exponential function to the appropriate portion of the headway data. The cumulative data plots (Figs. 3 and 4) proved to be of value in establishing the free headway cut-off points. These graphs are constructed with logarithmic probability scales so that the free portions of the distributions tend to plot as straight lines.

After the initial estimates for the parameters of the free headway subdistribution had been obtained, it was possible to compute and subtract this subdistribution from the overall headway distribution. The residuals that remained formed the subdistribution of constrained headways. The mean of these constrained headways was readily estimated, and initial estimates for the remaining model parameters were obtained by fitting the best Erlang function to the subdistribution. The appropriate k values for the Erlang distributions were found from graphical comparisons of the residual distributions with standardized Erlang functions. Two such comparisons are shown in Figure 5.

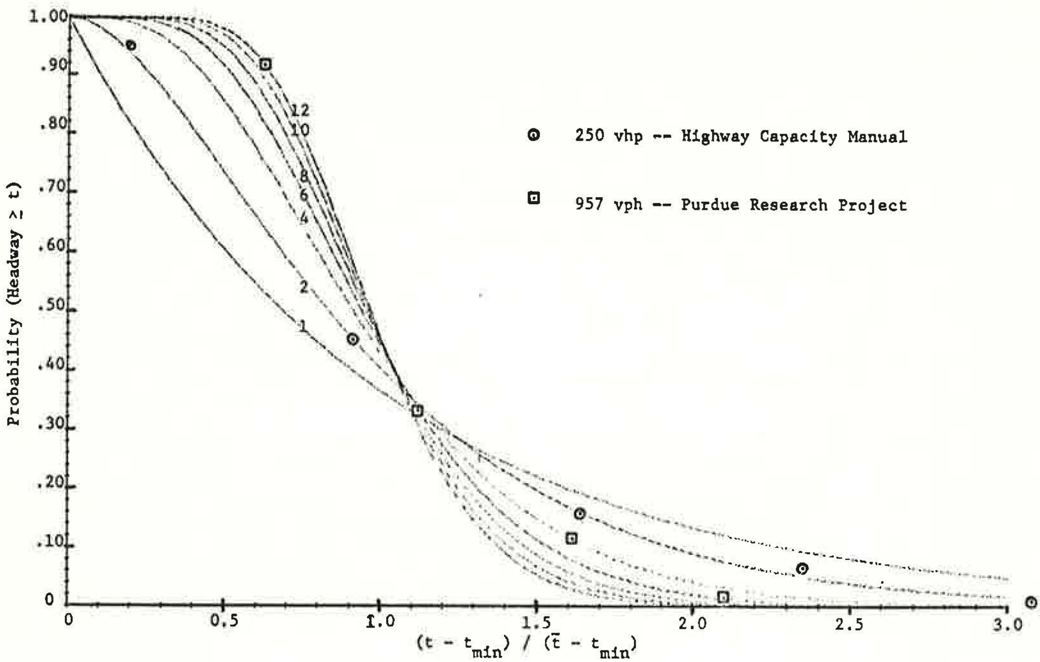


Figure 5. Comparison of residual distributions with standardized Erlang functions.

The residual distribution for the 250-vph data set from 1965 Highway Capacity Manual apparently follows the k = 2 curve; and the residual distribution for the 957-vph data set from the Purdue study was best approximated by the k = 6 curve.

The technique for obtaining the initial estimates for the parameters of the hyperlang function involved trial and measurement. Several combinations of minimum headway limits for the free and constrained distributions were tried. For each of these combinations the remaining model parameters for the free subdistribution were estimated by the method of least squares. The remaining parameters for the constrained subdistribution were estimated by computational and by graphical methods. The quality of each of the resulting hyperlang models was reported as the percentage of the total variation within the data that was removed by the model. This measurement, of course, is equivalent to R^2, the square of the multiple correlation coefficient. In each case the set of models parameters that yielded the highest R^2 value was selected for further refinement.

In the second step of the parameter evaluation process the initial parameter estimates were further refined using Marquardt's algorithm for nonlinear least-squares analysis (13). This algorithm employs the method of steepest descent and the method of Gauss to converge on a set of parameter values that tend to minimize the sum of squares of the deviations of the observed headways about the theoretical hyperlang headway function.

EVALUATION OF THE HYPERLANG MODEL

Highway Capacity Manual Study

The hyperlang models that were selected to describe the headway distributions reported in the 1965 Highway Capacity Manual are plotted as a family in Figure 6. The corresponding data sets are not superimposed on the individual curves due to their proximity; but the individual functions were graphed separately, along with the corresponding sets of raw data. The plot for the 250-vph flow rate is shown in Figure 7. The model parameters that describe these hyperlang plots are given in Table 1; all of the R^2 values fall between 0.9991 and 0.9998.

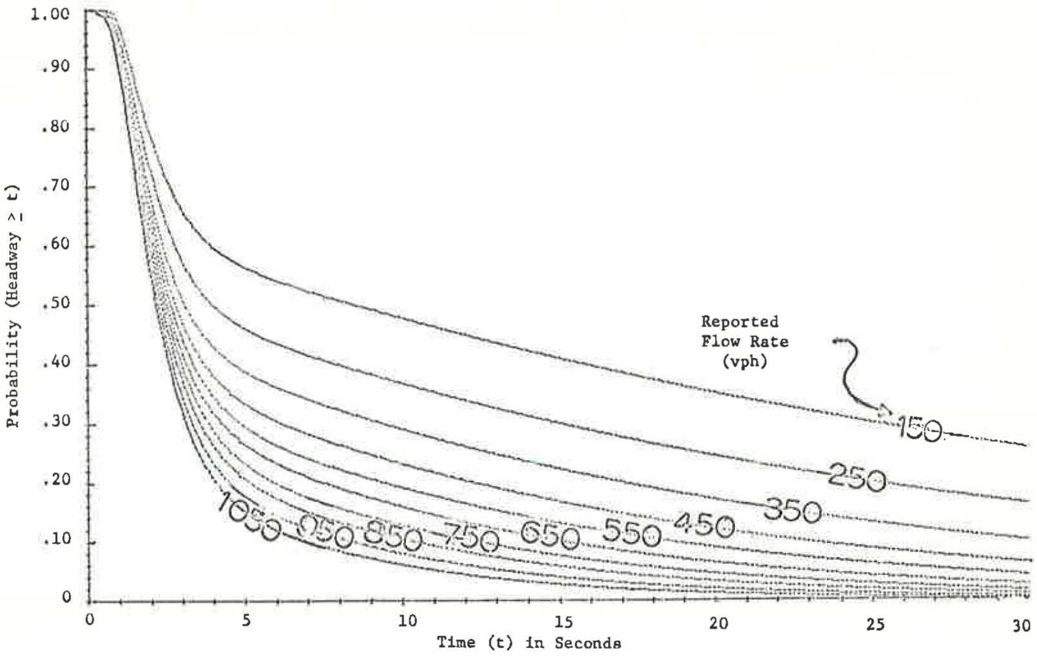


Figure 6. Hyperlang headway distributions for 1965 Highway Capacity Manual data.

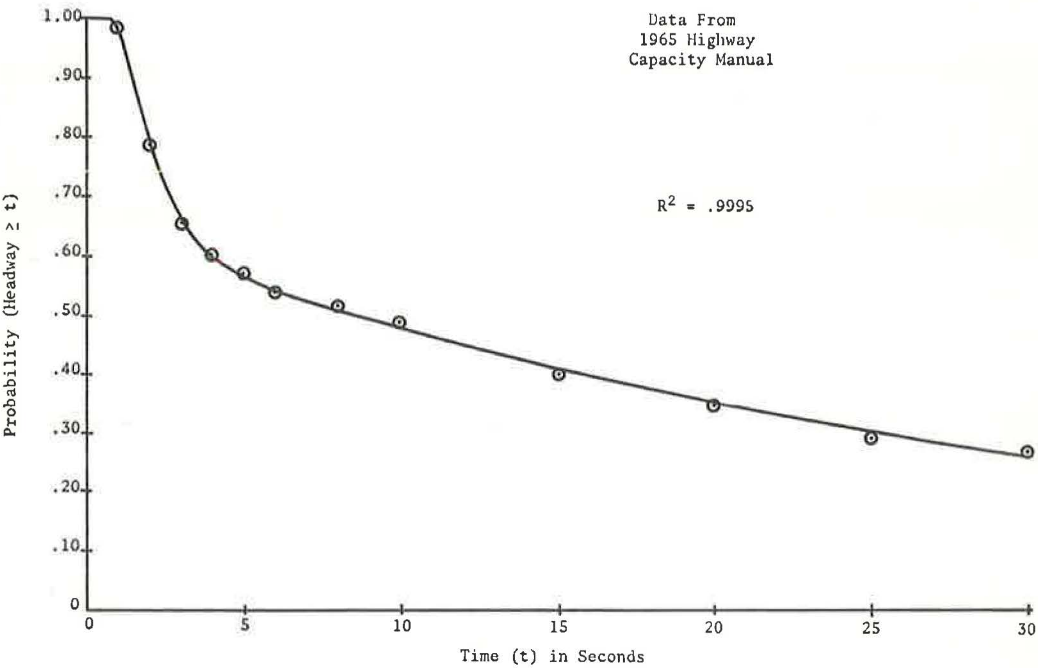


Figure 7. Hyperlang headway model volume: 250 vph.

TABLE 1

HYPERLANG MODEL PARAMETERS
(Highway Capacity Manual Data)

Monitored Flow Rate	Computed Flow Rate	R^2	α_1	γ_1	δ_1	k	α_2	γ_2	δ_2
150	166	0.9991	0.64	32.63	0.75	2	0.36	2.17	0.75
250	249	0.9995	0.55	24.62	0.75	2	0.45	2.12	0.75
350	344	0.9997	0.48	19.48	0.75	2	0.52	2.12	0.75
450	445	0.9995	0.43	15.86	0.75	2	0.57	2.14	0.70
550	528	0.9992	0.38	14.23	0.75	2	0.62	2.19	0.66
650	625	0.9995	0.36	12.02	0.75	2	0.64	2.19	0.61
750	726	0.9997	0.33	10.49	0.75	2	0.67	2.22	0.56
850	813	0.9998	0.30	9.60	0.75	2	0.70	2.22	0.55
950	926	0.9997	0.26	8.60	0.75	2	0.74	2.23	0.55
1050	1022	0.9996	0.21	8.30	0.75	2	0.79	2.25	0.55

The k indices for the Erlang components (Table 1) took on a value of 2 at all levels of flow. Before the study, however, it had been hypothesized that the k values would increase as the flow rate increased (or in other words, the constrained headways would tend to become more uniformly distributed at higher volumes). This hypothesis was based on an assumption that the increase in the number of constrained vehicles in the traffic stream would be compounded by simultaneous increases in the flow rate and in the proportion of constrained vehicles in the stream.

Although the k values did not vary as had been anticipated, the tabulated α_1 and α_2 values indicate that the proportion of constrained vehicles in the stream does increase in a direct relation with the flow rate. The proportion of constrained vehicles varied from 36 percent at a flow rate of 150 vph to 79 percent at 1050 vph. No satisfactory explanation has been found for the apparent lack of conformance with a priori hypotheses. It is possible, however, that distinctive properties in the several headway distributions were averaged out in combining data without regard for stratification to reflect variations in traffic and environmental conditions.

An interesting characteristic was observed between the minimum headways for the free vehicle distributions. The lower bounds for these distributions (δ_1) clustered in a random pattern about a central value of approximately 0.75 seconds. For this reason, δ_1, the parameter representing the minimum headway between free vehicles, was held constant at a value of 0.75 sec for all flow rates. This appears to be rational. At a flow rate of 150 vph, a time headway of 0.75 sec is indicative of an intervehicular spacing of approximately 35 ft; at 1050 vph, approximately 15 ft. Of course, these are the lower bounds for the spacings between all free vehicles.

On the other hand the minimum headway between constrained vehicles, δ_2, did not appear to be constant. This parameter varied in a consistent manner from 0.75 sec at a flow rate of approximately 150 vph to 0.55 sec at approximately 1050 vph. At the higher flow rate, the 0.55-sec headway value is indicative of an intervehicular spacing of only 7 or 8 ft. It must be remembered, however, that this is the absolute minimum spacing that will occur in the traffic stream.

A final relationship should be noted between the several headway distributions and the corresponding field-measured flow rates. When a headway distribution is specified, the flow rate is simultaneously implied. The volumes that were implied by the specific headway distributions were computed from the hyperlang parameters using the formula

$$\text{Volume} = 3600./(\alpha_1\gamma_1 + \alpha_2\gamma_2)$$

The computed flow rates (Table 1) are not in perfect agreement with flow rates that were monitored in the field, but they are not widely divergent.

Purdue University Study

The hyperlang models that were selected to describe the headway distributions reported in the Purdue study are plotted as a family in Figure 8. The corresponding data sets were not superimposed on the individual curves because of their proximity; but the individual functions were graphed separately, along with the corresponding sets of raw data. The plot for the 957-vph flow rate is shown in Figure 9.

The several headway functions in Figure 8 do not seem to constitute a very uniform family; but as can be seen in Figure 9, the hyperlang functions do fit the reported data very well. The models parameters that describe the hyperlang plots of Figure 8, and the R^2 values that are associated with each of the relationships, are given in Table 2. Again, it is interesting to note that the R^2 values are very high. The two lowest values are 0.9959 and 0.9977 but the rest of the values range between 0.9995 and 0.9998.

In general, the results obtained using the Purdue data more nearly correspond to a priori hypotheses. The k indices for the Erlang components vary directly with the flow rates. They range from ($k = 1$) at a flow rate of 158 vph, to ($k = 6$) at a flow rate of 957 vph. There also seems to be a significant relationship between these k values and the curve cluster patterns in Figure 8 and more particularly in Figure 4. The several headway distributions plotted in these two figures form four distinct clusters; and the k values for the Erlang components of the distributions in each cluster are identical.

The similarity between the headway distributions that form each of the several clusters is further highlighted by the similarity between the corresponding computed flow rates in Table 2. Tha apparent lack of a relationship between the monitored flow rates and the various model parameters tends to substantiate that the actual flow rates were not properly monitored during the field studies.

In general, the portion of free vehicles decreased and the portion of constrained vehicles increased as the flow rates increased. The α parameters of the 957-vph curve contradict this observation, but the lack of conformance within this distribution was anticipated. As shown in Figure 4, the 957-vph headway distribution diverges from the rest of the curves in the family. The divergence was apparently caused by an excessive number of long, free headways.

The minimum headway limits obtained in fitting the hyperlang function to the Purdue data do not seem to be consistent with the limits that were obtained in the analysis of the Highway Capacity Manual data. In the earlier analysis the δ_1 bounds on the free distribution were essentially constant, but in the analysis of the Purdue data the δ_1 limits tend to decrease slightly as the flow rates increase.

The limiting headways derived for the constrained vehicles (δ_2) also tend to decrease as the flow rate increases. This latter inverse relationship is in agreement with the results that were obtained in the analysis of the Highway Capacity Manual data.

TABLE 2

HYPERLANG MODEL PARAMETERS
(Purdue University Data)

Monitored Flow Rate	Computed Flow Rate	R^2	α_1	γ_1	δ_1	k	α_2	γ_2	δ_2
158	184	0.9959	0.86	22.34	0.69	1	0.14	2.88	1.65
251	219	0.9977	0.70	22.09	0.35	1	0.30	2.90	1.44
353	311	0.9996	0.61	16.75	0.74	1	0.39	3.35	1.12
450	492	0.9996	0.64	9.81	0.61	2	0.36	2.81	0.70
547	489	0.9997	0.56	11.05	0.70	2	0.44	2.73	0.90
651	567	0.9995	0.43	11.06	0.79	2	0.57	2.81	0.71
746	710	0.9996	0.40	8.35	0.88	3	0.60	2.92	0.57
836	740	0.9998	0.20	11.57	0.95	3	0.80	3.23	0.52
957	971	0.9997	0.53	4.58	1.06	6	0.47	2.71	0.72

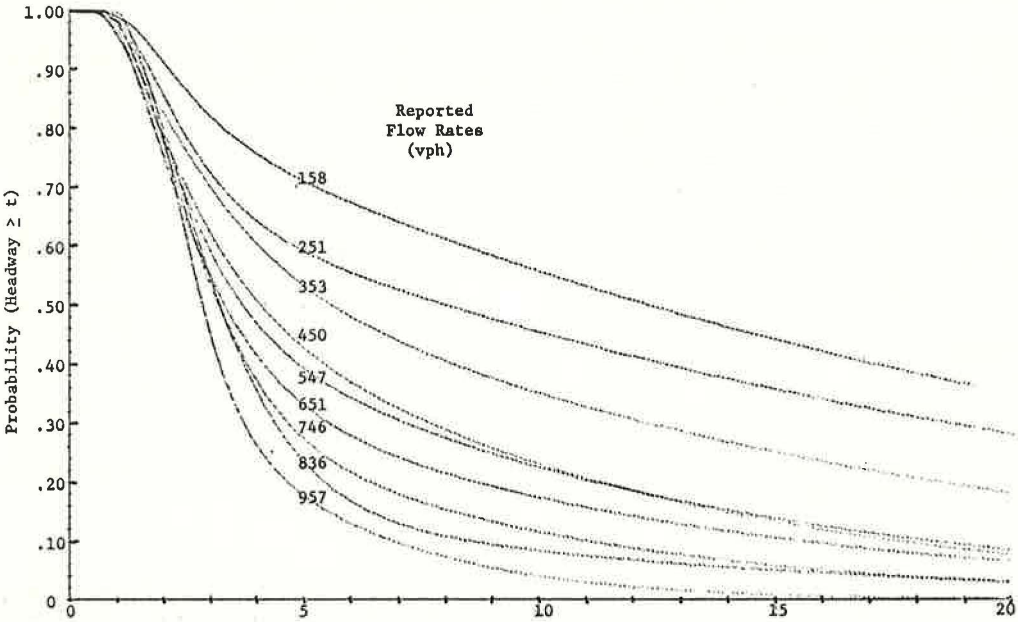


Figure 8. Hyperlang headway distributions for Purdue research project data.

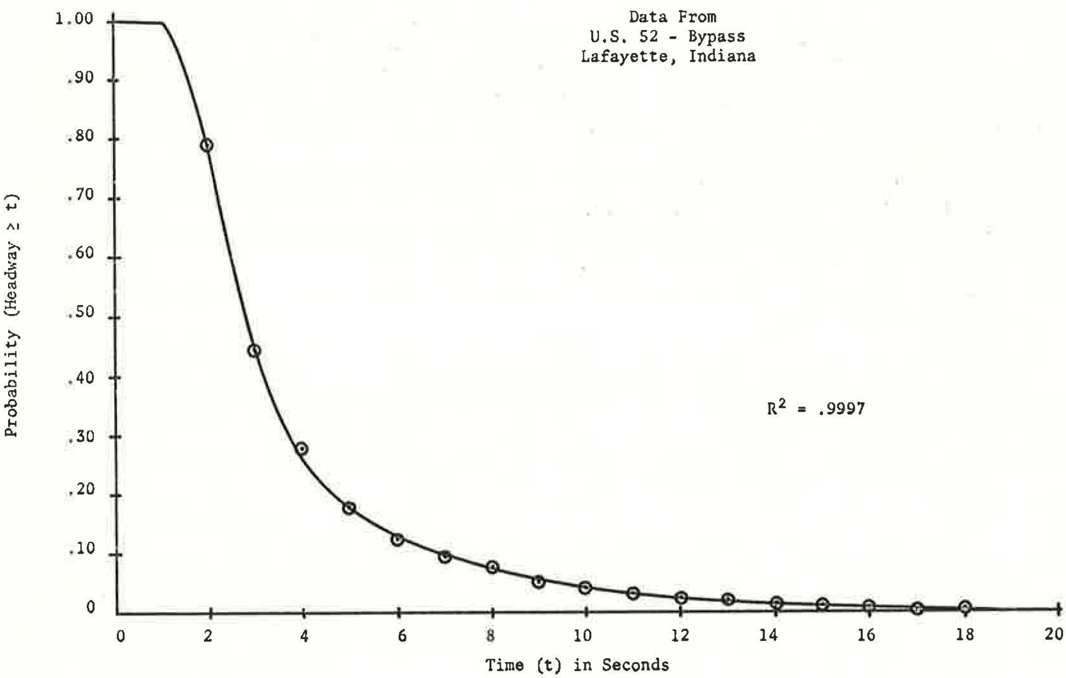


Figure 9. Hyperlang headway model volume: 957 vph.

CONCLUSIONS AND RECOMMENDATIONS

The hyperlang function is apparently a sound model for describing the headways in single-lane flows on 2-lane, two-way roadways. It is a very flexible model that can decay to a simple exponential function, to an Erlang function, or to a hyper-exponential function. It is likely, however, that a traffic stream will always contain both free and constrained vehicles; and that the general form of the hyperlang function will be required to effect an adequate description of the composite headways.

The parameters of the hyperlang model were evaluated separately for data sets obtained from the 1965 Highway Capacity Manual and from a recent Purdue University research project. In both instances the proposed model proved to be an excellent descriptor of the reported headway distributions. There is need, however, for a more comprehensive research study of the hyperlang headway model. Some of the parameter relationships that were observed in this preliminary study were not consistent for both data sets; and it appears that the differences are the result of inconsistencies in the original research data. In future studies, careful attention should be given to proper flow rate monitoring while making headway measurements. Careful attention should also be given to proper data stratification, during the analysis, to reflect variations in traffic and roadway conditions.

ACKNOWLEDGMENTS

This research was sponsored by the Highway Research Project at the University of Connecticut. The IBM 360/65, IBM 7040, and IBM 1620 computer systems that were utilized during the course of the study were partially supported by National Science Foundation Grant GP-1819.

REFERENCES

1. Buckley, D. J. Road Traffic Headway Distributions. Australian Road Research Board Proc., Vol. 1, Part 1, p. 153-187, 1962.
2. Cleveland, D. C., and Capelle, D. G. Queueing Theory Approaches. An Introduction to Traffic Flow Theory, HRB Special Report 79, p. 49-96, 1964.
3. Dawson, R. F. Discussion: Gap Acceptance Characteristics for Ramp-Freeway Surveillance and Control. Highway Research Record 157, p. 136-141, 1967.
4. Greenberg, I. A Vehicle Headway Distribution for Tunnels. Port of New York Authority (unpublished).
5. Gerlough, Daniel L. Use of Poisson Distribution in Highway Traffic. Poisson and Traffic, p. 1-58, Eno Foundation, 1955.
6. Greenshields, Bruce D., and Weida, Frank M. Statistics With Applications to Highway Traffic Analyses. Eno Foundation, 1952.
7. Haight, Frank A. Mathematical Theories of Traffic Flow. Academic Press, New York, 1960.
8. Herman, R., and Potts, R. B. Single-Lane Traffic Theory and Experiment. Theory of Traffic Flow, Elsevier Publishing Co., p. 120-146, 1961.
9. Highway Capacity Manual, U.S. Government Printing Office, Washington, D.C., 1950.
10. Highway Capacity Manual 1965. HRB Special Report 87, 1965.
11. Kell, James H. A Theory of Traffic Flow on Urban Streets. Institute of Traffic Engineers, Western Section Proc., July 1960.
12. Kinzer, J. P. Application of the Theory of Probability to Problems of Highway Traffic. Thesis, Polytechnic Instit. of Brooklyn, June 1933.
13. Marquardt, D. W. An Algorithm for Least-Squares Estimation of Non-Linear Parameters. Jour. of Society for Industrial and Applied Mathematics, June 1963.
14. May, A. D. Gap Availability Studies. Highway Research Record 72, p. 101-136, 1965.
15. Oliver, Robert M. A Traffic Counting Distribution. Operations Research, Vol. 9, No. 6, p. 802-810, Nov-Dec, 1961.
16. Schuhl, Andre. The Probability Theory Applied to Distribution of Vehicles on Two-Lane Highways. Poisson and Traffic, p. 59-75. Eno Foundation, 1955.
17. Sword, E. C. Prediction of Parameters for Schuhl's Headway Distribution. Purdue University, 1967.

A Digital Simulation Program of a Section of Freeway With Entrance and Exit Ramps

JOHANN H. BUHR, THOMAS C. MESEROLE, and DONALD R. DREW
Texas Transportation Institute, Texas A&M University

This paper describes a computer program developed for the simulation of a section of freeway, including several exit and entrance ramps. The program allows for the simulation of the traffic operation under different modes of entrance ramp control: fixed-rate metering, demand-capacity metering, gap-acceptance control and, of course, no control. The computer logic and simulation technique are discussed in detail. Limited output of the program is presented as evidence of the feasibility and realism of the simulation model.

•A penny is tossed until it comes down heads. If this happens on the first toss the player receives one dollar from the bank. If heads appears for the first time on the second toss, the player receives two dollars; on the third toss four dollars, etc., doubling each time. What should the player pay the bank for the privilege of playing this game if the game is fair?

If one has a coin handy, the simplest way to get some insight into how much a player should pay is to play, say, a thousand games and determine the average winnings per game. This seemingly unscientific approach to the Gambler's Ruin or St. Petersburg Paradox, as it is also called, represents a simple illustration of simulation.

SIMULATION

Simulation is essentially a working analogy. It involves the construction of a working model presenting similarity of properties or relationships to the real problem under study. Simulation is a technique which permits the study of a complex traffic system in the laboratory rather than in the field.

In a more general sense, simulation may be defined as a dynamic representation of some part of the real world achieved by building a computer model and moving it through time. The term, computer model, is used to denote a special kind of formal mathematical model, namely a model which is not intended to be solved analytically but rather to be simulated on an electronic computer. Thus, simulation consists of using a digital or analog computer to trace the time paths; the distinction being that the digital device counts and the analog device measures. This distinction is actually a fundamental one, being essentially the mathematical distinction between the discrete variable (digital) and the continuous variable (analog). The differences in capabilities between the digital and analog computers are manifest in the mathematical distinctions between summation and integration, or between difference equations and differential equations.

Why Simulate?

Simulation is resorted to when the system under consideration cannot be analyzed using direct or formal analytical methods. There are a few additional reasons for simulation, most of which have been found to pertain to the simulation of traffic systems.

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 47th Annual Meeting.

1. The task of laying out and operating a simulation is a good way to gather pertinent data systematically. It makes for a broad education in traffic characteristics and operation.

2. Simulation gives an intuitive feel for the traffic system being studied, and is therefore instructive.

3. Simulation of complex traffic operations may provide an indication of which variables are important and how they relate since control can be exercised over the various input parameters. This may lead to eventual successful analytic formulations.

4. In some problems, information on the probability distribution of the outcome of a process is desired, rather than only means and variances such as obtained in queueing. Where traffic interaction is involved, the Monte Carlo technique is about the only tool which can give the complete distribution.

5. A simulation can be performed to check an uncertain analytic solution.

6. Simulation is cheaper than many forms of experiment. Imagine the cost savings in simulating to find the optimum spacing of freeway interchanges.

7. Simulation gives a control over time. Real time can be compressed so that the results of a long time period can be observed in a few minutes of computer time. On the other hand, real time can be expanded and run slower than real time so that all the manifestations of the complex interactions of freeway movement can be comprehended.

8. Simulation is safe. It provides a means for studying the effect of traffic control measures on existing highways. The effect of signals, speed limits, signs, and access control all can be studied in detail without confusing or alarming drivers. Simulation offers the ability to determine in advance the effect of increased traffic flow on existing facilities. Probable congestion points and accident locations can be anticipated and changes in the physical design of the highway can be effected before the need is demonstrated through accident and congestion experience.

As a form of model, a simulation model should be compared with analysis on the one hand which involves the use of analytical, rigid, and probabilistic models and trial and error on the other, which involves devising some kind of trial solution and then taking it into actual traffic and trying it out. The relative merits of analysis, simulation, and trial and error, can best be discussed with the aid of Table 1, prepared by Goode (1).

The traffic problem has of course been attacked in the past with the tools of both analysis and trial. Simulation is actually a combination of both methods, but unlike analysis, it allows attack on the most complicated of processes. On the other hand, it does not affect traffic until the solution has been reached. Simulation is almost always midway between analysis and trial (Table 1). But as the situation being studied becomes more complex, such as in the case of traffic systems the differences between methods in terms of cost, time, etc., become more pronounced, until finally neither of the extremes can be tolerated and simulation becomes the only feasible method.

Simulation is a powerful tool and like all powerful tools it can be dangerous in the wrong hands. The increased emphasis on simulation studies and the corresponding lack of experience on the part of some people who attempt to apply the method can lead to a sort of pseudo-simulation. Pitfalls exist in simulation as in every human attempt to abstract and idealize. Some rules to follow in avoiding these pitfalls are (a) no assumption should be made before its effects are clearly defined, (b) no variables should be combined into a working system unless each one is properly explained and its relationships to the other variables are set and understood, and (c) simplification is desirable, but oversimplification can destroy the realism of the model.

For the most part, it can be said that the goals achievable by simulation in the traffic process are clear-cut and offer a profound payoff. Simulation is an ideal technique for traffic research. The simulation model is not just another means for accomplishing

TABLE 1

RELATIVE MERITS

Criterion	Analysis	Simulation	Trial
Cost	least	medium	most
Time	least	medium	most
Reproducibility	most	medium	least
Realism	least	medium	most
Generality of results	most	medium	least

what we can do today but is a tool for solving problems which cannot be solved today.

The Freeway Merging Process

One traffic problem that has generally defied analysis is the process of entrance ramp vehicles merging into a freeway stream. In the summer of 1965, the Bureau of Public Roads undertook research to furnish detailed criteria on the merging of ramp vehicles into the freeway stream. To this end, a contract, "Gap Acceptance and Traffic Interaction in the Freeway Merging Process," was awarded the Texas Transportation Institute. The general aim was the conception of relationships between the many variables associated with the interaction of vehicles traversing a ramp and merging onto a freeway so as to determine the effect of traffic and geometric characteristics on merging operation and level of service.

One objective of this project was to develop a computer program to simulate the operation in a freeway merging area. The purpose of the simulation study was primarily to show the feasibility of simulation as a tool in the study of freeway control for the eventual development of simulation programs to study optimum ramp metering and control techniques and equipment, rather than a reproduction of the minute details of vehicular behavior.

LIST OF SYMBOLS

$D_{o,i}$	Distance from the zero reference point of vehicle i
$D_{t,i}$	Tentative distance from the zero reference point of vehicle i
$D_{f,ij}$	Following distance between vehicles i and j
$D_{d,i}$	Desired following distance factor of vehicle i
$D_{a,ij}$	Acceptable minimum following distance between vehicles i and j
$D_{u,ij}$	Tentative updated following distance between vehicles i and j at the present time
$D_{t,ij}$	Tentative updated following distance between vehicles i and j during the next update
$S_{c,i}$	Current speed of vehicle i
$S_{d,i}$	Desired speed of vehicle i
$S_{w,i}$	Lane changing speed of vehicle i
$A_{s,i}$	Smoothing acceleration of vehicle i
$A_{x,i}$	Maximum acceleration of vehicle i
$A_{b,i}$	Normal acceleration of vehicle i
$B_{s,i}$	Smoothing deceleration of vehicle i
$B_{x,i}$	Maximum deceleration of vehicle i
$B_{b,i}$	Normal deceleration of vehicle i
$B_{n,i}$	Minimum deceleration of vehicle i
$L_{c,i}$	Calculated lag time for vehicle i
$L_{a,i}$	Minimum acceptable lag time for vehicle i
$V_{s,i}$	Length of vehicle i
T	Time increment between updates

THE SIMULATION PROGRAM

The Basic Model

The program allows for the simulation of a freeway segment with a number of entrance and/or exit ramps, having a geometric configuration that can be changed at will. The number of exit ramps is limited to 2, whereas the number of entrance ramps plus freeway lanes is limited to 6. Thus, for example, if the freeway segment to be simulated has 3 lanes, then the number of entrance ramps is limited to 3.

The program initially places a number of vehicles on the freeway segment and then proceeds to process each vehicle in the system according to the programmed flow logic, starting with the vehicle most distant from the beginning of the simulated section, regardless of the lane it is in. During simulation, new vehicles are generated on the freeway lanes and entrance ramps according to a Poisson distribution of arrivals.

Each vehicle is assigned a number of characteristics such as length, current speed, desired speed, distance from the zero reference point (beginning of the simulated section), etc., and the program, using a periodic scan technique, updates these characteristics for each vehicle during the scan interval. The program therefore has available, at any time, all the information concerning each vehicle in the system. The general logic organization is shown in Figure 1. Essentially, each vehicle attempts to travel at its desired speed, subject to certain restrictions. If the vehicle being processed is traveling at its desired speed and a safe headway exists between it and the vehicle ahead, it is simply allowed to proceed down the freeway. If it encounters a slower vehicle ahead, it will, under certain conditions of following distance, attempt to change lanes, first to the left and then to the right. If unsuccessful, the vehicle will decelerate.

The program was written in the Fortran IV computer language and is generally run on an IBM 7094 Model I computer. However, it can be run on other computers with a Fortran IV compiler. The simulation program consists of 1 monitor routine and 16 subroutines. Each subroutine is completely modular so that any logic changes in any subroutine will not affect the remainder of the program.

The program also provides the option of simulating the merging operation under various modes of entrance ramp control: (a) no control, (b) fixed-time metering, (c) demand-capacity metering, and (d) gap-acceptance control.

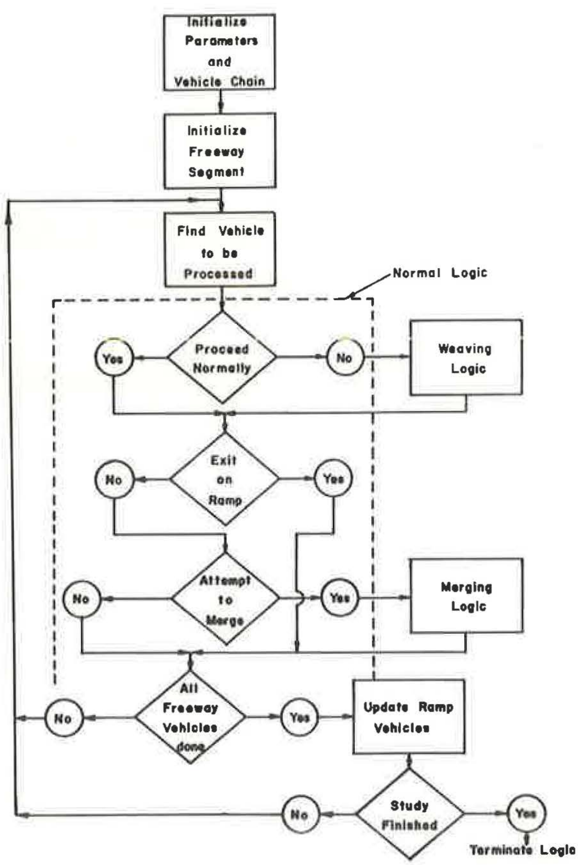


Figure 1. Logic flow diagram of the system.

Input Parameters

Inherent in the formulation of a simulation model is the determination of the significant input and output variables. Inputs may generally be divided into four categories: geometrics, traffic characteristics, driver policy and vehicle performance. A fifth category of input serves to control the program directly.

The inputs to this simulation program are listed below by category. Each input parameter has a built-in default value which is used if no other value is specified. The default values are shown in parentheses behind each input parameter.

(1) Geometric Characteristics
- (a) Length of freeway section (6000 ft)
- (b) Number of through lanes (3)
- (c) Number of entrance ramps (1)
- (d) Number of exit ramps (1)
- (e) Location of entrance ramps (4000 ft)
- (f) Location of exit ramps (1000 ft)
- (g) Length of each on-ramp (500 ft)
- (h) Length of each acceleration lane (300 ft)
- (i) Location of ramp signal
- (j) Grade (0)
- (k) Start of grade (4000 ft)
- (l) End of grade (5000 ft)

(2) Traffic Characteristics
- (a) Lane volumes in vehicles per hour*

  Default: Lane 1 = 1430
  Lane 2 = 1730
  Lane 3 = 1840
  Lane 4 = 600
  Lane 5 = 0
  Lane 6 = 0
- (b) Proportion of vehicles on each lane that exit on first ramp

  Default: Lane 1 = .30
  Lane 2 = .10
  Lane 3 = .05
  Lane 4 = 0
  Lane 5 = 0
  Lane 6 = 0
- (c) Proportion of vehicles on each lane that exit on second exit ramp

  Default: Lane 1 = .80
  Lane 2 = .15
  Lane 3 = .05
  Lane 4 = .20
  Lane 5 = 0
  Lane 6 = 0
- (d) Proportion of commercial vehicles in each lane

  Default: Lane 1 = .06
  Lane 2 = .03
  Lane 3 = .01
  Lane 4 = .03
  Lane 5 = 0
  Lane 6 = 0

(3) Driver Policy
- (a) Average acceleration (3 ft/sec^2)
- (b) Average deceleration (6 ft/sec^2)
- (c) Minimum deceleration (2 ft/sec^2)
- (d) Maximum speed (60 mph)
- (e) Crowding factor (0.7)
- (f) Gap acceptance characteristics, specified by two probit equation coefficients (0.5, 2.0)
- (g) Location of exit decision stations (710 and 0)

*The freeway shoulder lane is designated "lane 1." The adjacent freeway lanes and then the entrance ramps are numbered consecutively.

(4) Vehicle Performance
  (a) Maximum acceleration (11 ft/sec^2)
  (b) Maximum deceleration (15 ft/sec^2)

(5) Program Control
  (a) Scan interval (1 sec)
  (b) Analysis time (4 min)
  (c) Warm-up time (1 min)
  (d) Length of warm-up section (500 ft)
  (e) Print option (no)
  (f) Plot option (no)
  (g) Number of plot stations (9)
  (h) Size of plot increments (200 ft)
  (i) Location of last plot station (4300 ft)
  (j) Number of check stations (12)
  (k) Location of each check station

Most of the input parameters given above are self-explanatory. The purpose and use of those that are not immediately evident, will become clear later.

Internal Bookkeeping

The internal bookkeeping procedure used in representing the flow of vehicles within the computer and in keeping track of the characteristics of each unit is perhaps the most complex aspect of simulation requiring the highest degree of programming skill. It is of vital importance to the efficient and successful operation of the program. To implement any practical simulation program, the method of bookkeeping must keep the core storage requirements at a minimum and lend itself to fast sequential processing.

In this simulation program, each vehicle is assigned a subscript number between 1 and 500, allowing no other vehicle to have the same subscript. Each vehicle characteristic is stored in an array so that any particular characteristic, such as current speed for example, of any particular vehicle, can be found by addressing the appropriate array with the subscript number assigned to the vehicle of which the characteristic is desired. If the simulation is allowed to operate over a long period of time, a considerable quantity of data is collected. To conserve storage in the computer, only the characteristics of the vehicles presently in the system are stored in memory. The characteristics of a vehicle which has passed the end of the study section are no longer stored. This is accomplished by reassigning the subscript number of the vehicle using a chaining technique.

The chaining technique is a method of logically organizing the relative position of a vehicle to all other vehicles in the system. There is assigned to each vehicle a value in each of two characteristic arrays which contain the subscript numbers of the vehicles immediately ahead of and behind it and in line with its movement. The arrays of these characteristics are named LAST and NEXT, respectively. This gives the processing program access to the characteristics of the vehicles between which a vehicle being processed is situated and allows determination of those characteristics of the vehicles which must be changed.

TABLE 2

ARRAY ORGANIZATION OF A CHAIN

Subscript Number	Last	Next
2	1	4
4	2	8
8	4	10
10	8	23
23	10	49
49	23	75
75	49	76

Since each vehicle has characteristics in arrays LAST and NEXT, the processing program has an overall picture of the placement of vehicles in the system. These two arrays can be thought of as being a chain with each link being one subscript of a vehicle. For example a chain array might appear as in Table 2.

A chain from the arrays NEXT and LAST containing the subscripts of vehicles shows the organization of one lane of traffic. From Table 2, the lane represented by this chain would physically consist of vehicle 1, followed consecutively by vehicles 2, 4, 8, 10, 23, 49, 75, and 76.

At the zero reference point a location called the generation pool is constructed for each lane of the system, always containing one vehicle already assigned its characteristics. When the logic determines that a vehicle will enter the freeway, a new vehicle and its characteristics are generated. The vehicle in the generation pool of that lane enters the freeway, and the newly generated vehicle is placed into the generation pool and remains there until another vehicle is generated. This method of always having one vehicle in each generation pool allows the chain associated with each lane to have an end link at the zero reference point, facilitating processing.

The vehicle most distant from the zero reference point in each lane has a value in the LAST array as does every other vehicle in the system. Since this vehicle does not have a vehicle between it and the end of the freeway segment, a fictitious vehicle is placed outside the freeway segment to represent the vehicle ahead of the most distant vehicles. Only one fictitious vehicle is generated for all lanes and is assigned a complete set of characteristics and a unique subscript number. This allows all chains to begin with the same unique vehicle logically terminating the chains. As a vehicle passes through the freeway segment its characteristics are stored, but once it passes off the end of the freeway segment its characteristics are no longer saved and its subscript number is removed from the chain of that lane. The vehicle directly behind the vehicle just processed off in the same lane considers the fictitious vehicle as the vehicle now ahead of it.

As a vehicle is introduced into a generation pool, it is assigned a subscript number. To prevent it from accidentally acquiring the subscript of some vehicle already in the system, the new vehicle obtains its subscript number from a free links pool. Initially one chain is formed which contains all the subscript numbers available for the system linked together using the NEXT and LAST arrays. When the system is initially set up, the subscript numbers not assigned to vehicles placed on the freeway segment, form the free links pool. This pool subsequently contains every available subscript number not being used in the system. As a vehicle passes off the end of the freeway segment its subscript number is placed in the free links pool, and when a vehicle enters a generation pool it obtains its subscript number from the free links pool.

The purpose of the chaining logic is to permit an organized handling of vehicle characteristics while using a method which is easily adapted for digital computation. In the digital computer this process reduces the number of internal storage locations required. If this process were not used and specific locations were required for the characteristics of each vehicle, both in and out of the system, the number of vehicles that can pass through the system would be quite limited. Using the chaining technique, the restriction imposed on the system is that not more than 500 vehicles be on the freeway segment at any one time.

Initial Setup

The first step after reading in all the input parameters, is the geometric arrangement of the simulated section in the computer and the placement of vehicles on the freeway lanes and ramps.

Initially, one vehicle is placed in each freeway lane at the end of the study section and one vehicle on each entrance ramp, 10 ft upstream of the ramp nose. Each of these vehicles, as well as subsequent vehicles, is assigned a desired speed and its current speed is set equal to its desired speed. These speeds are pseudorandom numbers generated so as to be normally distributed with parameters:

$$\text{Mean} = 0.85 \times \text{maximum freeway speed}$$
$$\text{Standard deviation} = 0.07 \times \text{mean} + 1.375$$

Any speed generated to be higher than the maximum freeway speed, is set equal to the maximum freeway speed.

Based on their assigned speeds, vehicles are then placed at a distance behind the previous vehicle so as to correspond to an exponential speed-density curve. This relationship, first proposed by Greenberg (2), is given by

$$u = c \ln (k_j/k)$$

where

u = speed,
k = density,
k_j = jam density = 175 veh/mile, and
c = constant of proportionality = 35 mph.

Each vehicle is further assigned a "desired following distance factor" which is simply a random number, uniformly distributed between 0.2 and 1.0 and a length of either 16 or 32 ft based on whether it is a passenger car or truck. The decision as to the type of vehicle is made by comparing a uniformly distributed random number to the proportion of commercial vehicles specified as an input parameter. Similarly, each vehicle is tagged as to whether it desires to exit or not, and if so, which ramp it will exit on.

During simulation, new vehicles are generated from a Poisson distribution of arrivals by comparing a uniformly distributed random number to the volume in vehicles per scan interval, the volume in each lane being an input parameter. Vehicles so generated are assigned a current speed, which equals its desired speed, and a length, in the same manner as during initial setup, but it is placed at the beginning of the simulated section.

The program runs for a period, specified as the "warm up time" input parameter, before any vehicle data are collected and stored.

Simulation Logic

The simulation logic for stepping vehicles through the system, once the inputs are known, may be divided into three classifications: (a) flow logic for unimpeded vehicles, (b) car-following logic for platooned vehicles, and (c) maneuvering logic for vehicles executing maneuvers involving more than a single stream of traffic. In reality, all drivers of vehicles within the roadway system are continually and simultaneously making decisions and modifying their behavior. In the course of a simulation, the classification of most vehicles—unimpeded, following or maneuvering—will change many times. The computer, however, can make only one simple logical choice at a time. To control all the occurrences at any given instant, it must process all decisions sequentially. In other words, it must process each decision for every vehicle, for each vehicle in every lane, and for each lane within the system. It must do this in accordance with a prescribed sequence for each instant of time to be considered.

The program is organized into independent logical divisions with one monitor division to direct the control among the other divisions. By using separate divisions, experimentation can be carried out in one division without changing any other logic.

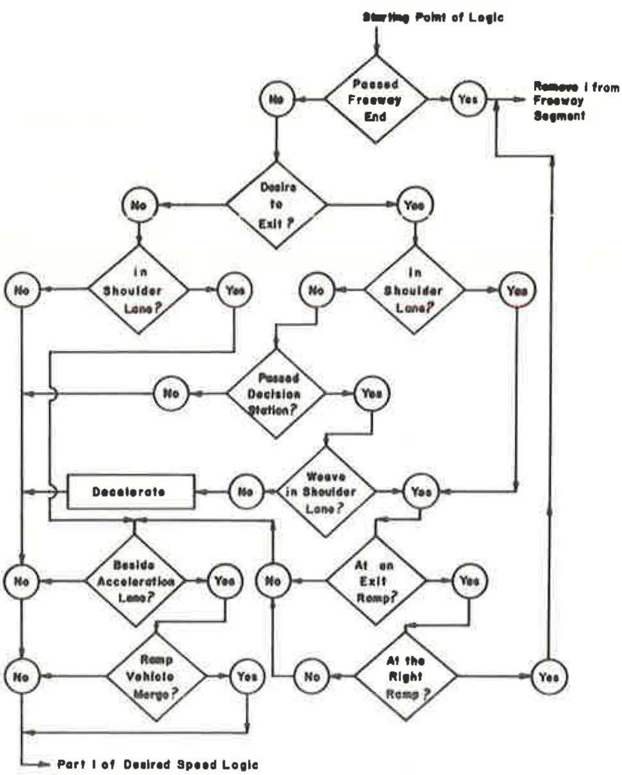


Figure 2. Distance logic flow diagram.

The monitor division incorporates the tasks of initializing all parameters including the chains, initializing the freeway segment for the first time period, and handling the normal or through flow of traffic. This section determines which of the vehicles on the freeway segment is the most distant nonprocessed vehicle and perform a series of tests on it. The tests are in the form of the following questions.

1. Will this vehicle travel past the end of the freeway segment during the next scan interval?
2. Does this vehicle exit on a ramp?
3. Does a vehicle merge in front of this vehicle from the acceleration lane of an entrance ramp?
4. If this vehicle is not in the shoulder lane and desires to exit soon, can it weave into the shoulder lane?
5. Is this vehicle traveling as fast as it desires?

Once the monitor division has operated on all freeway vehicles and has updated them, the ramp vehicles are updated.

Normal Flow—The normal flow section determines the behavior of vehicles in the same lane and is divided into two segments: the distance logic and the speed logic. A flow diagram of the distance logic is shown in Figure 2.

The distance logic stems from the fact that certain points on the freeway segment are of particular interest in the processing of vehicles. These are the end of the freeway segment, the beginning of the exit ramps, the exit decision stations, and the area next to the acceleration lanes. The number of these points and their distances from the zero reference point are input parameters which can be varied to give the scheme flexibility. This section determines whether a vehicle has passed one of the points during a given time period by calculating a tentative distance $D_{t,i}$.

$$D_{t,i} = D_{o,i} + (S_{c,i})\,T$$

$D_{t,i}$ represents the position of vehicle i if no speed change occurs. If this tentative distance is greater than the distance to one of the significant points, and the original distance $D_{o,i}$ of vehicle i is less, vehicle i will pass this point for the first time. Depending on which point it is, different algorithms are employed.

Vehicle i has passed the end of the freeway segment if its $D_{t,i}$ is greater than the distance to the end of the freeway segment. A sequence is then initiated to remove i from the chain of that lane and place its subscript into the free links pool. The exit decision stations are located upstream of the exit ramps so as to give a vehicle desiring to exit time to weave into the shoulder lane. After an exiting vehicle has passed such a station, it will attempt to weave until it is successful, even after it has passed the desired ramp. When a vehicle is in the shoulder lane and it desires to exit, its $D_{t,i}$ and $D_{o,i}$ are compared with the distance to each exit ramp on the freeway segment to determine if this vehicle is now at its desired exit ramp, and if so, it is processed out of the system.

The distance logic is based on the assumption that there is no change in the speed of a vehicle during a time interval. This is, of course, not necessarily true. After the position of a vehicle on the freeway segment has been compared with all the important points and this vehicle does not leave the freeway, control is transferred to the first part of the three-part speed logic.

The desired speed logic is broken down into three parts based on the relationship between the vehicle being processed and the vehicle ahead of it in the same lane. A following distance and an acceptable following distance are calculated for a pair of vehicles, i and j, where i is the vehicle being processed and j is its lead vehicle. The following distances are given by

$$D_{f,ij} = D_{o,j} - D_{o,i} - V_{s,j}$$

$$D_{a,ij} = V_{s,i} + S_{c,i} + \left[\frac{0.5\,(S_{c,i} - S_{c,j})^2}{B_{b,i}} + D_{d,i}\,(S_{c,i})\right] K$$

where

K = 0 if $S_{c,i} \leq S_{c,i}$, and
K = crowding factor if $S_{c,i} > S_{c,j}$.

The factor, K, allows for a smaller acceptable following distance between two vehicles if the lead vehicle is traveling faster than the following vehicle.

The three parts of the speed logic are shown by the flow charts in Figures 3, 4, and 5. The first part of the speed logic deals with the case where the following distance is less than acceptable, the second part with the case where the following distance is exactly equal to the acceptable following distance and the third part with the case where the following distance is bigger than the acceptable following distance.

If the following distance is less than acceptable (Fig. 3) but i is traveling slower than j, then the logic is terminated by simply updating vehicle i at its current speed. However, if i is faster than j it will decelerate by either the normal deceleration, $B_{b,j}$, or the maximum deceleration, $B_{x,i}$, depending on how close i is to j.

If the following distance between i and j is equal to the acceptable following distance (Fig. 4) and i is traveling slower than j, then, if maximum acceleration would put it above its desired speed, it simply accelerates to its desired speed and the logic is terminated. Otherwise, it accelerates by its maximum acceleration and then attempts to change lanes. On the other hand, if i is traveling faster than j, then it either decelerates by the maximum amount or it decelerates to the same speed as j. If, in the latter case, its speed is now less than desirable, it will attempt to change lanes, otherwise, it will decelerate further to its desired speed.

The last part of the speed logic (Fig. 5) deals with the case where the following distance is greater than the minimum acceptable following distance. If, under this condition, i is traveling slower than j, it is accelerated by an amount $A_{s,i}$ given by

$$A_{s,i} = A_{x,i}\,(S_{d,i}/S_{c,i}) - A_{x,i}$$

Therefore, if it is already traveling at its desired speed, its speed will not change, but otherwise, it will adjust its speed so as to approach its desired speed asymptotically.

However, if under these conditions, vehicle i is traveling faster than vehicle j, a tentative following distance is calculated to inspect what will happen if i does not change its speed. The tentative following distance is given by

$$D_{u,ij} = D_{o,j} - D_{o,i} - (S_{c,i})T - V_{s,j}$$

If this tentative following distance is less than or equal to the minimum acceptable following distance, vehicle i decelerates by the normal deceleration $B_{b,i}$. Otherwise, the distance between i and j will remain safe after this scan interval and the vehicle is processed

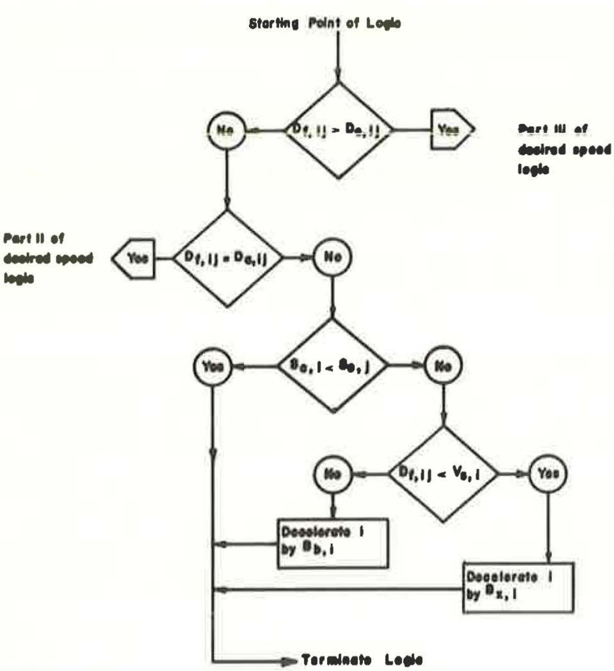


Figure 3. Speed logic flow diagram—Part I.

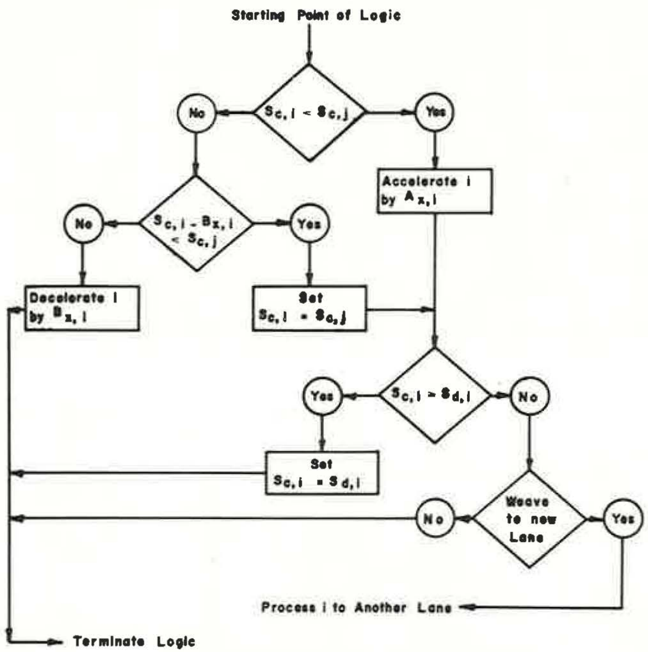


Figure 4. Speed logic flow diagram—Part II.

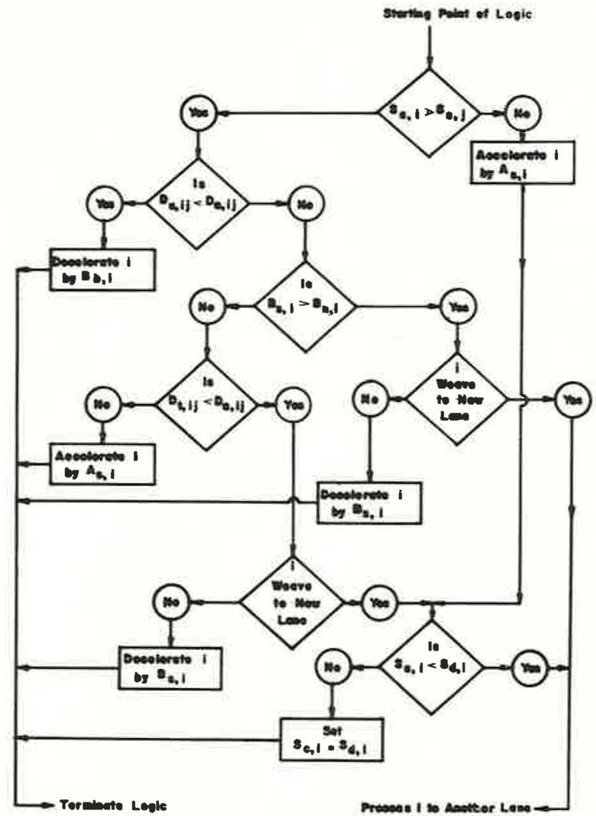


Figure 5. Speed logic flow diagram—Part III.

so as to eliminate erratic movement, by using a smoothing deceleration as expressed by

$$B_{s,i} = \frac{(S_{c,j} - S_{c,i})^2}{2(D_{u,ij} - D_{a,ij})}$$

This gives the deceleration that i must take for it to be moving at the same velocity as j when the following distance between them equals the acceptable following distance. If this smoothing deceleration is greater than the minimum deceleration $B_{n,i}$, vehicle i will try to change lanes. If vehicle i cannot change lanes, it is decelerated by $B_{s,i}$. If, on the other hand, $B_{s,i}$ is less than $B_{n,i}$, the possibility of accelerating vehicle i is explored and a new tentative following distance, being the projected following distance two scan intervals hence, is calculated, assuming $S_{c,i}$ and $S_{c,j}$ constant during the following time period.

$$D_{t,ij} = D_{o,j} + (S_{c,j})T - D_{o,i} - 2(S_{c,i})T - V_{s,j}$$

If this projected following distance is less than the minimum acceptable following distance, i will either decelerate by the amount $B_{s,i}$ or will change to another lane, if its speed is less than desirable or will increase its speed to its desired speed. On the other hand, if the projected following distance is greater than the minimum acceptable, vehicle i will accelerate by $A_{s,i}$.

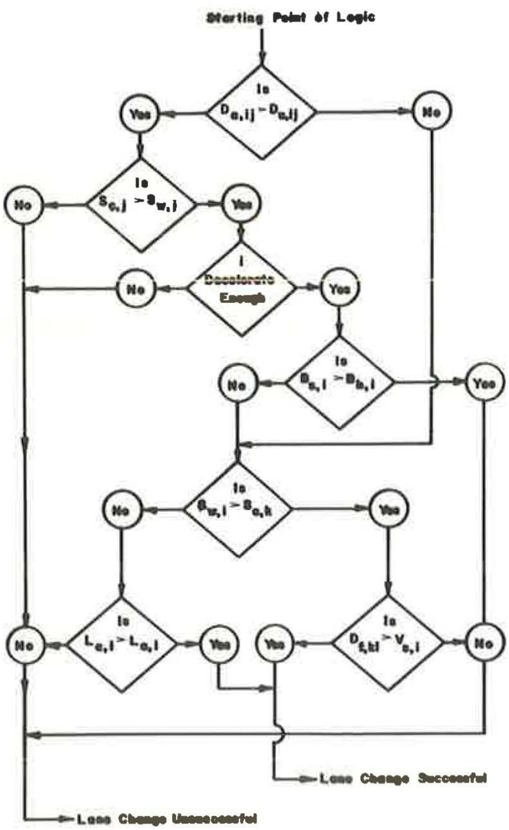


Figure 6. Lane change logic flow diagram.

Lane Changing—When it is determined that a vehicle desires to change to another lane, the possibility of changing to the left is explored and then to the right. If a vehicle is on either the shoulder lane or the median lane, only one direction is tried. If all lane change attempts are unsuccessful, the vehicle remains in its present lane. If successful, the vehicle is removed from its old lane to the new lane by rearranging the appropriate chains. After a lane has been chosen into which an attempted entry will be made, the changing vehicle is accelerated by the maximum amount $A_{x,i}$. This new speed is called the lane changing speed and is designated $S_{w,i}$. The vehicle will change lanes if it will not come hazardously close to the vehicles ahead of or behind it in the new lane. If the gap is inadequate the attempt is unsuccessful. The lane changing logic is diagrammed in Figure 6, assuming that vehicle i attempts to change lanes into a space between vehicles j and k.

First, the relationship between the lane changing vehicle i and the lead vehicle, j, in the new lane is investigated by calculating a tentative following distance $D_{u,ij}$ and a minimum acceptable following distance $D_{a,ij}$.

$$D_{u,ij} = D_{o,j} - D_{o,i} - V_{s,j} - (S_{w,i})T$$

$$D_{a,ij} = V_{s,i} + [S_{w,i} + D_{d,i}(S_{w,i})]K$$

where

$K = 0$ if $S_{w,i} \le S_{c,j}$, and
$K =$ crowding factor if $S_{w,i} > S_{c,j}$.

If the following distance is bigger than the minimum acceptable following distance, the relationship between the lane changing vehicle and the following vehicle, k, in the new lane will be investigated. However, if the following distance is less than acceptable then the relative speed between i and j is investigated. If i is traveling faster than j, the lane change attempt is unsuccessful but if i is slower, a test is made to see if i can slow down so that a safe distance will exist between it and vehicle j, if i uses its maximum deceleration. If i cannot decelerate fast enough, the lane change attempt is unsuccessful, otherwise, a smoothing deceleration, $B_{s,i}$, is calculated to see if i is closing in on j too fast.

$$B_{s,i} = \frac{(S_{c,j} - S_{w,i})^2}{2(D_{o,j} - D_{o,i} - (S_{w,i})\,T)}$$

If this deceleration that would be required is greater than the normal deceleration, vehicle i does not change lanes, otherwise, it is assumed that the relationship between i and j is safe and the relationship between i and k is investigated.

If vehicle i is traveling faster than vehicle k and there is at least a vehicle length between them, a lane change occurs. Otherwise, the attempt is unsuccessful.

If vehicle i is slower than k, the time relationship between vehicles i and k is investigated. First, a minimum acceptable lag time, $L_{a,i}$, is generated as a random variable, normally distributed with a mean of 0.5 sec and a standard deviation of 0.1 sec. However, if the generated variable is less than 0.1 sec, it is set equal to 0.5 sec. This minimum acceptable lag time, is compared to a lag time, $L_{c,i}$, defined as the time required by vehicle k to come within the minimum safe distance of i and is given by

$$L_{c,i} = \frac{(D_{o,i} - D_{o,k} - V_{s,i} - D_{a,ik})}{(S_{w,i} - S_{c,k})}$$

If this lag time is less than acceptable, the lane change does not occur. Otherwise, vehicle i changes lanes and the chains are appropriately rearranged.

Ramp Discipline—The behavior of vehicles on an entrance ramp is treated in two parts: the movement of vehicles on the ramp and the merging of vehicles onto the freeway segment from the ramp.

The movement of vehicles on a ramp is handled similarly to the vehicles on the freeway segment, except that ramp vehicles cannot change lanes while on the ramp. The ramp vehicle closest to the end of the ramp is processed as a special case because this vehicle must stop at the end if it is unable to merge. This vehicle is gradually decelerated as it approaches the end of the acceleration lane. When it reaches the end, it stops. The remaining ramp vehicles are processed similarly to the freeway vehicles by calculating an acceptable following distance, given by

$$D_{a,ij} = V_{s,i}/2 + (S_{c,i}/2)\,K$$

where

$K = 0$ if $S_{c,i} \le A_{b,i}\,T$, and
$K = 1$ if $S_{c,i} > A_{b,i}\,T$.

This acceptable following distance is shorter than those of through vehicles since ramp vehicles generally crowd closer together when attempting to merge. The procedure determines whether a vehicle will accelerate or decelerate by comparing $D_{a,ij}$ with its following distance during the last time period, the present time period, and the next time period. Again smooth processing is the key criterion and when a vehicle approaches another vehicle it will do so gradually.

Gap Acceptance—The merging subroutine processes vehicles from the acceleration lane onto the shoulder lane of the freeway segment. Vehicles are processed starting with the one closest to the end of the acceleration lane, each vehicle will attempt to merge into the gap adjacent to its current position. The space relationships between the merging vehicle, the leader and the follower are first investigated and if these are found acceptable, the time gap is compared to an acceptable gap for the merging vehicle to determine if the merge will take place.

The criteria used to determine if the physical limitations will preclude a merge, depends on the relative speeds of the vehicles involved. If the speed of the leader is greater than the speed of the ramp vehicle, the acceptable distance between them is one vehicle length. The acceptable distance is reduced by one-half if the leader merged in the same scan interval. If the speed of the leader is less than or equal to the speed of the ramp vehicle, the acceptable distance between them is one vehicle length for each 10 mph of speed of the ramp vehicle.

The acceptable distance between the ramp vehicle and the follower also depends on the relative speed of the two vehicles. If the ramp vehicle is the faster of the two, the acceptable distance is one vehicle length. Otherwise, the acceptable distance is one vehicle length for each 10-mph ramp speed.

If any of the acceptable distances are less than the actual distances, the ramp vehicle does not merge. For vehicles which are forced to stop at the end of the acceleration lane, an "impatience factor" is introduced. This causes the acceptable following distances to be reduced by 5 percent for each scan interval that it waits, to a minimum of 60 percent of the original acceptable following distances.

After it has been determined that the physical conditions will permit a merge, an acceptable gap is generated for the ramp vehicle. Acceptable gaps are generated as random variables, distributed as a lognormal distribution. This is done by generating a normally distributed pseudorandom number, finding its probit value by means of a table look-up and comparing this probit value to the value given by the probit equation specified by the input parameters, solved for the available freeway gap (3).

The Output

The normal output of a simulation system is a table of average values of quantities such as travel times and volumes processed. This gives an overall view of the operation of a system, but it does not give any information about individual vehicle movement. A number of small mistakes can be present in the logic without being detected. A table of values indicating vehicle movement is cumbersome for a large study and, at best, is difficult to interpret. Part of the output of this simulation is such a table of average values but a more complete picture of the operation was developed as part of this research by displaying the movement of vehicles in a graphical form by means of a time-space diagram.

The abscissa of the time-space diagram represents time, the ordinate represents distance from the zero reference point, and one continuous line represents the movement of an individual vehicle in time and space. The plot therefore gives the position of each vehicle on the freeway segment for the entire period of the study plus a picture of relative speeds and positions. It therefore serves not only to give an overall view of the operation, but is also of tremendous value in debugging the program by displaying the occurrences of logic errors.

Figure 7 is an example of part of the time-space output of the simulation of an uncontrolled merging situation. It shows vehicles in the right-hand freeway lane plotted in continuous lines while vehicles on the ramp are represented by dashed lines. Lines that suddenly terminate, indicate that the vehicle has changed lanes. Different lanes and the ramp movement can be plotted using continuous lines and dashed lines or by using different colored lines to distinguish among them. A plot showing a large number of lanes can become quite cluttered and will lose its effectiveness.

Each curve of a plot is made up of a number of points with a line drawn between them. Each point represents the position and time of a vehicle in the system. Since the digital

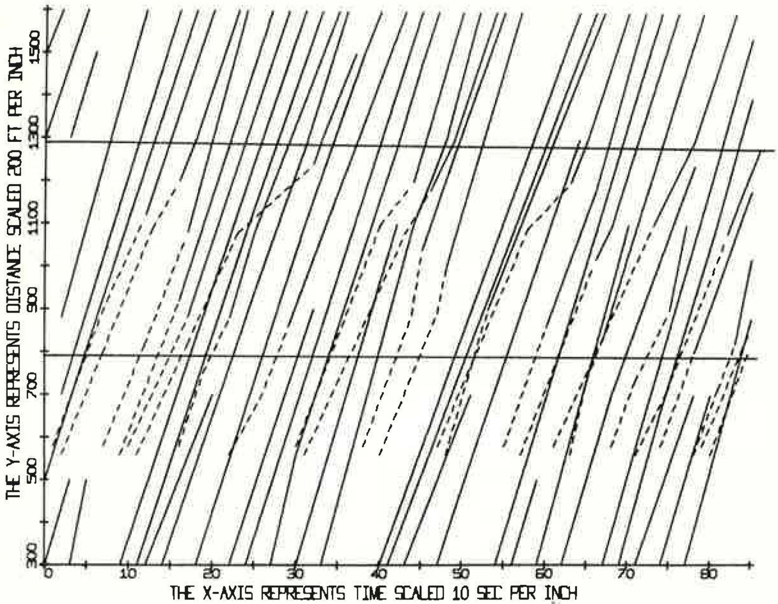


Figure 7. Time-space plot of uncontrolled ramp.

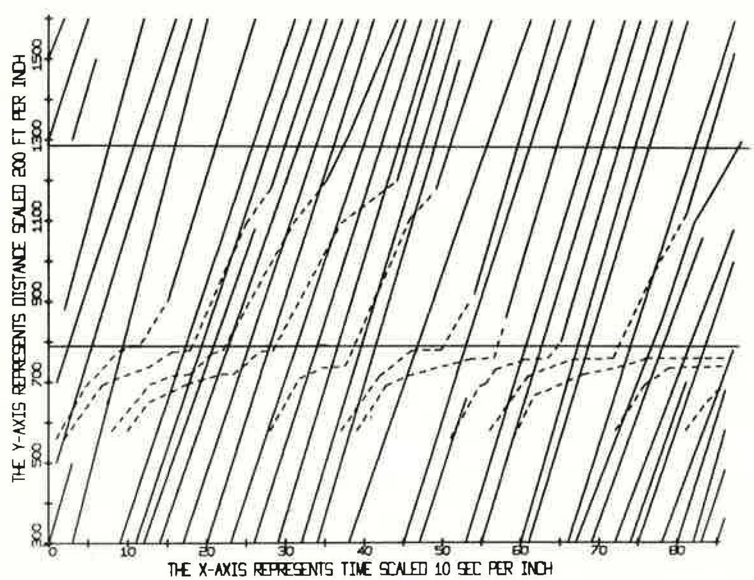


Figure 8. Time-space plot of controlled ramp.

computer program makes a periodic scan, it decides when data are to be saved and writes these items on magnetic tape, one record per block, indicating the name of the vehicle, its position on the freeway segment, and the time. After the study has been completed the data on the tape are sorted and rearranged into the proper form for plotting the time-space diagram.

Figure 8 is an example of part of the time-space output of a ramp under simulated gap-acceptance control. The behavior of vehicles on the freeway and on the ramp as

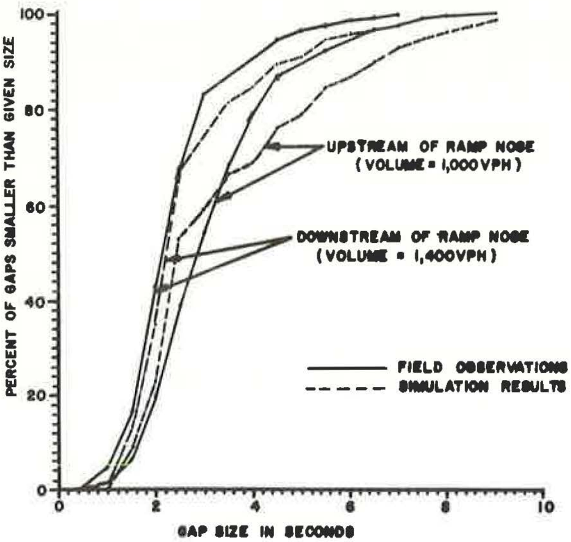


Figure 9. Gap distributions yielded by simulation program compared to field data.

demonstrated by these plots seems to represent real life conditions with a fair degree of realism.

A third type of output provided by this simulation is the time-space characteristics of each vehicle in tabular form, punched into computer cards. This output is in the same form as the input used on an extensive library of analysis programs developed by the Texas Transportation Institute and used for the analysis of the merging process (4).

Model Calibration and Validation

If programmed properly, the realism of the computer output is a function only of the realism of the system model and the inputs to the model. A simulation model is essentially a hypothesis and, therefore, must be tested before it can be accepted as fact. Such tests include its feasibility, realism and validity.

The feasibility and realism of this simulation study is illustrated by the time-space diagrams. Little work has been done on the calibration and refinement of the various models and assumptions that constitute the overall simulation program. Such studies are now in progress and indications are that the model can be satisfactorily calibrated.

Although it was not the intent of this simulation study to reproduce all the minute details associated with vehicular behavior, it is essential that the results be compared with known real world responses to the same inputs. This was done by running the simulation program using the geometrics of the outbound Cullen entrance ramp on the Gulf Freeway in Houston. A single 20-min period was simulated, using as input parameters the volume levels and gap acceptance characteristics observed at this ramp over a 20-min study period. Further analysis of the computer output revealed that simulated drivers generally behaved in a more uniform manner than real drivers. They generally maintained higher and more uniform speeds while following each other at greater distances. However, their speed-spacing relationships as revealed by a comparison of time headway distributions showed an encouraging agreement between simulated and actual conditions (Fig. 9). Through the calibration of the various models that constitute the simulation, any particular set of data for any particular ramp can probably be duplicated quite closely. However, the study was designed to simulate the operation in any merging area, with the result that the calibration and validation of this simulation program will involve a large number of runs under greatly varying conditions with associated compromises in adjustments.

Theoretically, the model must duplicate the characteristics which the highway engineer uses as design criteria, or the characteristics that the traffic engineer uses as operational criteria. As is usually the case, however, universally useful design and operational criteria cannot be precisely defined and each design application requires the selection of suitable criteria based on engineering judgment. One might adopt a microscopic philosophy in which attempts would be made to duplicate, in the computer, the specific details of field samples of moderate time length. Otherwise a macroscopic approach might be utilized in which computer runs seek to reproduce gross statistical properties of field samples accumulated over long periods of time. Since traffic is a stochastic process, valid arguments can be raised in support of either approach. In this study, greater emphasis is being put on the macroscopic results. The calibration, validation and further refinement of the model is presently in progress.

ACKNOWLEDGMENT

The simulation program described in this paper was developed under a research contract sponsored by the U.S. Department of Transportation, Federal Highway Administration, Bureau of Public Roads.

REFERENCES

1. Goode, H. H. Computers, Simulation, and Traffic. Unpublished.
2. Greenberg, H. An Analysis of Traffic Flow. Operations Research, Vol. 7, 1959.
3. Drew, D. R., LaMotte, L.R., Buhr, J. H., and Wattleworth, J. A. Gap Acceptance in the Freeway Merging Process. TTI Research Report 430-2, Dec. 1966.
4. Buhr, J. H., Drew, D. R., Wattleworth, J. A., and Williams, T. G. A Nationwide Study of Freeway Merging Operations. TTI Research Report 430-1, Dec. 1966.

Traffic Volume Measurement Using Drivometer Events

J. E. CLARK and P. D. CRIBBINS, North Carolina State University

An investigation to attempt to develop a technique for measuring the driving hazard of one highway relative to another began in 1966. Data for this study were collected with a Drivometer-equipped test car operating on six functionally different highways. This paper is an outgrowth of the driving hazard study and was specifically intended to explore any correlation that might exist between the recorded Drivometer events and traffic volumes on a given facility.

Data were collected from a driving schedule which utilized a Latin square experimental design to minimize bias between drivers, times, and facilities. Eight Drivometer events, each recorded on a 0.1-mi basis in one direction on an urban arterial street, were selected as independent variables to be compared with 15-min traffic volumes on an urban arterial street in Raleigh, N.C. These Drivometer events included total travel time, change in speed, running time, small steering reversals, large steering reversals, brake applications, accelerator applications, and changes in direction of travel.

For the facility under investigation, the mean number of brake applications correlated most closely with traffic volumes. However, none of the simple correlations was high enough to assume complete linear dependence of traffic volume on any single Drivometer event. By combining the two most significant events (brake applications and large steering reversals) and using a multiple regression analysis, equations were developed to predict traffic volumes on the investigated facility.

Although specific findings are obviously applicable only to the study site, it can be concluded that, by measuring a suitable sample of drivers with a Drivometer-equipped vehicle and recording their brake applications and large steering reversals, traffic volumes can be computed with statistical confidence.

•BECAUSE of the magnitude of travel on the nation's streets and highways, new techniques for measuring traffic volumes are needed to supplement present manual and mechanical methods. This investigation was undertaken to develop a new approach to traffic volume measurement consisting of a sampling procedure using Greenshields-Platt Drivometer events.

During the summer of 1966, a research project (ERD-110-67-6) was initiated by personnel of the Highway Research Program in the Department of Civil Engineering, North Carolina State University. The long-range objective was to develop a technique for measuring the driving hazard of one highway relative to that of another. Data were collected by means of a test car equipped with a Drivometer and operating on six functionally different highways. The Drivometer was designed to furnish an objective visual measure on a time and/or distance basis of driver performance, vehicle motion and the traffic and highway environment through which the driver is traveling.

The research reported herein is an extension of the investigation conducted within the scope of ERD-110-67-6 (hereafter referred to as Project 67-6). Specific objectives of

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 47th Annual Meeting

this additional investigation are (a) to correlate traffic characteristics collected by the Drivometer-equipped vehicle with traffic volumes on a given facility, (b) to develop prediction equations for traffic volumes using Drivometer events, and (c) to test the hypothesis that the frequency of the Drivometer events varies with volume on a given facility.

Drivometer events can be used to establish "norms" for different kinds of highways and to check the effect of improvements in the physical design of the highway on traffic flow. There are other applications for the Drivometer, but for most of these it would also be desirable to have an estimate of traffic volumes on the given facility.

SELECTION OF STUDY AREA

After establishing site criteria, the study site was selected from the highway segments composing the test route used in Project 67-6. Each section of the entire test route chosen for Project 67-6 belongs to a particular functional classification. Care was taken to insure that each section was fairly homogeneous and had some uniform physical characteristics such as number of lanes, lateral spacing of opposing lanes, and type of access control. The highway segments investigated are given in Table 1.

Westbound Hillsboro Street was selected as the study site for this research. Hillsboro has a relatively high ADT (16,000) and a volume sampling program indicated that there was a low variation in volumes along the 3.08-mile segment. There was also a high peaking in the traffic volumes for morning and afternoon periods.

Hillsboro Street can be further described as a high volume urban arterial with high-density roadside development and no access-control or median. Most of westbound Hillsboro consisted of two traffic lanes except for 0.05 mi consisting of three lanes and 0.545 mi consisting of one lane. There are 23 traffic signals on the study segment with two different timing cycles.

The maximum degree of horizontal curvature is 3.2 deg and the maximum grade is 4 percent. The posted speed limit is 25 mph except for the last 0.49 miles on the west end of the study site where the posted speed limit changes to 35 mph. Curb parking is permitted along 1.89 mi of the 3.08-mi segment.

TRAFFIC VOLUME DATA COLLECTION

General Procedure

The westbound traffic volume counts along Hillsboro Street during the interval of the study were recorded with portable, penumatic-actuated machine counters. All counters

TABLE 1

TEST ROUTE SEGMENTS[a]

Highway Segments	Route	Description	Volume (ADT)	Length (mi)
A	US 64, Western Blvd.	Four-lane divided urban arterial at-grade expressway with some control of access	18,000	3.76
B	US 1 Bus., Hillsboro Street	Four-lane undivided urban arterial with high-density roadside development	16,000	3.08
C	US 1 Bypass, Beltline	Four-lane divided freeway with full control of access	12,000	8.96
D	US 70	Four-lane divided rural primary trunk route with moderate roadside development	13,500	5.83
E	S.R. 1002	Two-lane rural secondary collector	2,800	6.92
F	NC 54	Two-lane rural primary trunk route	3,600	6.39

[a] D.L. Heimbach, P.D. Cribbins and J.W. Horn. 1966. Developing traffic flow indices for the detection of high accident potential highways in North Carolina. Unpublished.

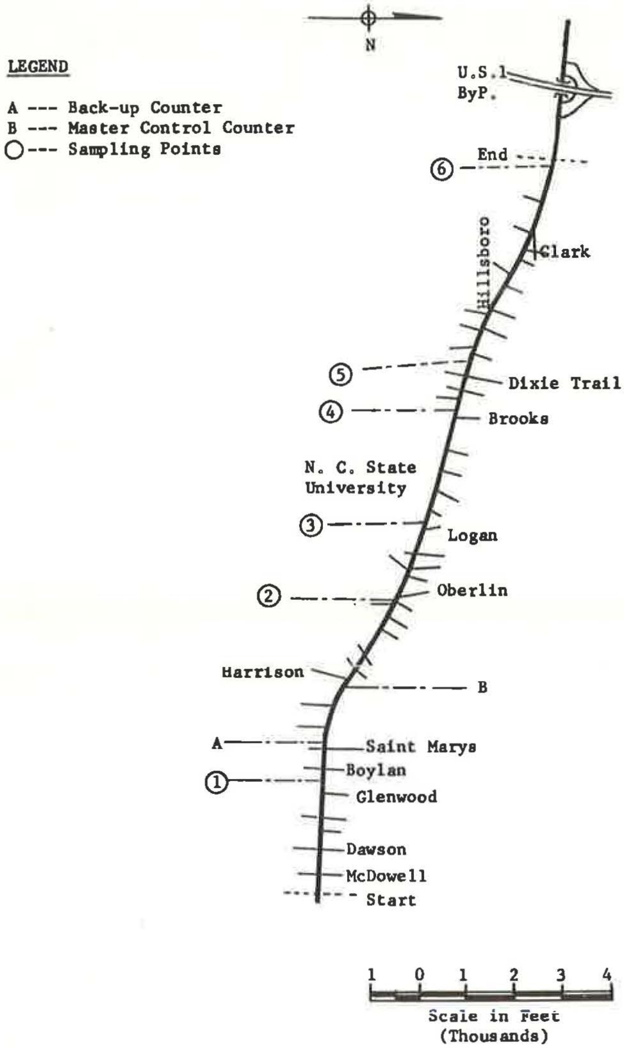


Figure 1. Hillsboro Street study section.

recorded two actuations as one vehicle except the K-Hill counters which recorded axles. Because the percentage of trucks on Hillsboro Street (4%) was so low, no correction was applied to the recorded volumes for multi-axle vehicles. Figure 1 shows the test section along Hillsboro Street with the location of major intersections and sampling points for traffic counters.

The master control counter and sampling counter printed cumulative totals on a paper tape roll every 15 minutes and reset to zero on the hour. The back-up counter furnished a continuous traffic count recorded on a circular chart. The counter also recorded and displayed a cumulative total traffic count on a dial. Traffic volumes recorded by the master control counter concurrent with the test runs are plotted in Figure 2. The 15-min volumes were measured, beginning on the half and hour marks, for a duration of 15 min.

Traffic Volume Profile

Traffic volume profiles were plotted for an evaluation of the variability in traffic volumes along the study section because the Drivometer events were so recorded. The six sampling points (Fig. 1) were located near major street intersections where a change in traffic volumes on Hillsboro Street might be expected to occur. A 24-hr sample with 15-min totals was taken at each of the six sample points.

Volume profiles for 0630 and 1700 hour time periods are shown in Figure 3. The 0630 hour profile is relatively flat but the 1700 hour profile shows a peak at sample point 4, which was located just west of Brooks Avenue, where three lanes of traffic were counted. The peak can be accounted for by the left-turn lane for traffic turning onto Dan Allen Drive. These profiles are characteristic of an arterial street without control of access where there are intersecting feeder streets.

DRIVOMETER EVENTS DATA COLLECTION

Description of Equipment

Drivometer events measured concurrently with the previously mentioned traffic volume counts were needed to accomplish the objectives of this research. These additional data were also collected during the summer of 1966 by personnel of Project 67-6.

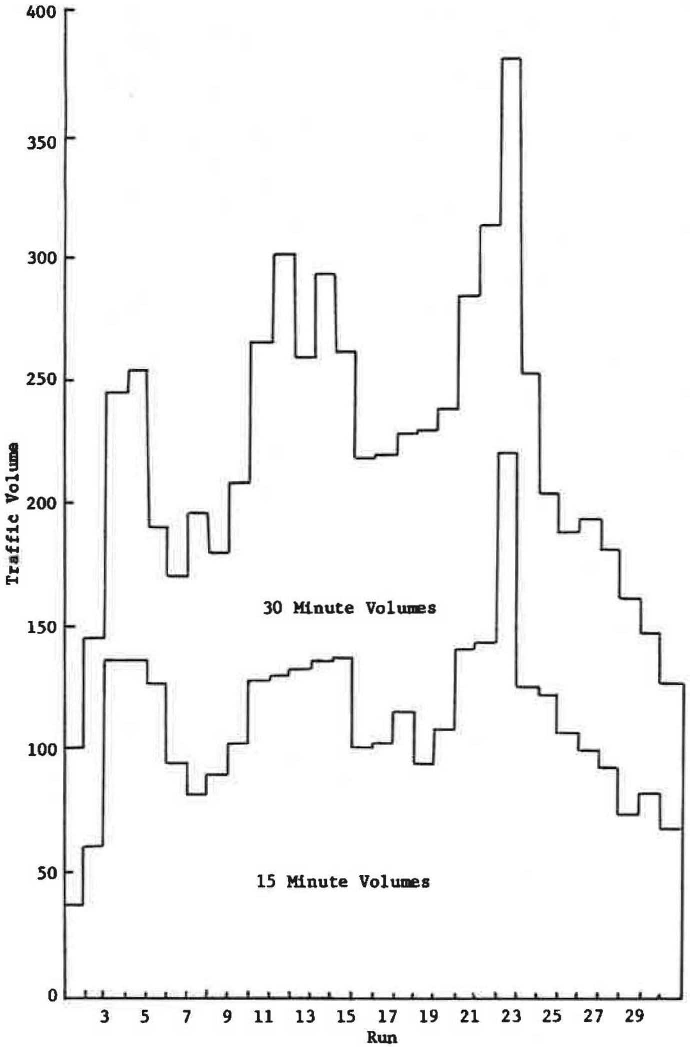


Figure 2. Master control volume counts.

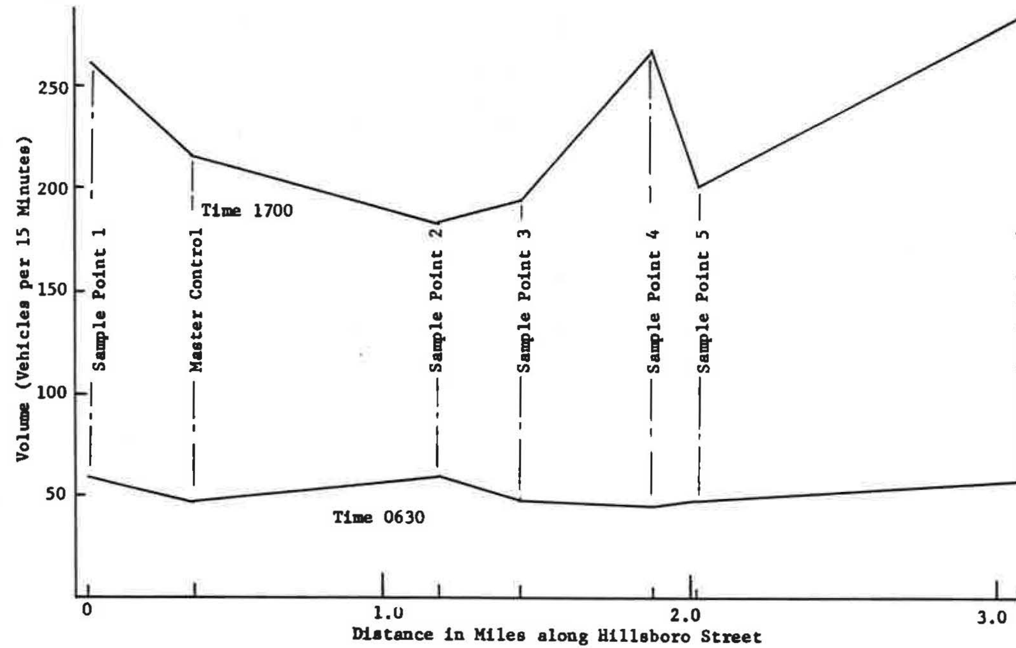


Figure 3. Volume profile for 15-min volume counts.

TABLE 2

TYPICAL LATIN SQUARE DRIVING SCHEDULE[a]

Road Section	Run (days)					
	1	2	3	4	5	6
A	0630 (1)	0900 (2)	0830 (3)	0800 (1)	0730 (2)	0700 (3)
B	0700 (2)	0630 (3)	0900 (1)	0830 (2)	0800 (3)	0730 (1)
C	0730 (3)	0700 (1)	0630 (2)	0900 (3)	0830 (1)	0800 (2)
D	0800 (1)	0730 (2)	0700 (3)	0630 (1)	0900 (2)	0830 (3)
E	0830 (2)	0800 (3)	0730 (1)	0700 (2)	0630 (3)	0900 (1)
F	0900 (3)	0830 (1)	0800 (2)	0730 (3)	0700 (1)	0630 (2)

[a]See Table 1, footnote a.

The Drivometer can be defined as an electro-mechanical device which was designed to furnish on a time and/or distance basis an objective visual measure of driver performance, vehicle motion and the traffic and highway environment through which the driver is traveling. The Drivometer and Traffic Events Recorder at North Carolina State University is the third unit to be constructed incorporating these basic ideas. The component parts of the unit consist of the following:

1. A speed change unit which is composed of an auxiliary speedometer with photocell switches to give the number of speed changes in 4-mph increments and the distance traveled in 0.01-mi intervals. This unit was mounted on the steering column and connected to the speedometer cable by means of a "tee."
2. A steering wheel switch which was arranged to actuate a counter whenever the steering wheel was reversed from any position.
3. A modified airplane gyroscopic compass was arranged to furnish a count for each $2\frac{1}{2}$ deg of turning in any direction. This arrangement permitted changes in direction of travel of the test vehicle to be read directly from a counter in $2\frac{1}{2}$ deg units for any distance.
4. An accelerator switch which was attached to one of the engine bolts and was actuated by the small rod which operates the carburetor. This switch registered through a counter all of the up and down movements of approximately $\frac{1}{4}$ in. or more of the accelerator at any speed.
5. A brake switch which was connected to the brake pedal and registered all applications of the brake.
6. A hand-held keyboard which enabled an observer sitting in the car to count all traffic events such as a parking or turning vehicle in the traffic stream. No such events were recorded by the observer during this research.

The overall function of the Drivometer Events Recorder is to record driver actions, vehicle motions and traffic and/or highway events in digital form. These actions, motions and events were all recorded on a bank of 25 counters which was photographed by a 16-mm camera at 0.1 mi and 1-min intervals. The dials and recording camera were all housed in a metal recorder box which was mounted in the rear of a 1965 custom model Ford in which a portion of the rear seat had been removed. The test vehicle was equipped with standard factory equipment except for an automatic transmission. There were no markings or external attachments to distinguish the test vehicle from any other vehicle on the road.

Data Collection for the Present Research

Research performed by the personnel on Project 67-6 yielded data from the six different highway sections (Table 1). The basic experimental design (driving schedule) which was used for the collection of data for the first 6 half-hour periods (6:30 A.M. to 9:30 P.M.) is given in Table 2. Runs were made on Mondays through Fridays with holidays excluded in order to provide reasonably homogeneous conditions especially with respect to the within-day patterns. The entry in each section of highway actually took less than 30 min to drive; therefore, some dead time existed between the sections. Five different drivers participated in the tests. In any one run through the six sections, three drivers were assigned to the test vehicle and drove according to the rotation schedule specified by the experimental design. In Table 2, each cell of the 6 x 6 square is the time when the test run on the given section was to be started; the three drivers assigned to the square are identified as (1), (2), and (3). The drivers rotated driving assignments as shown with each driver driving two sections on each run. The basic square was repeated to cover the additional time periods of 0930-1230, 1230-1530, 1530-1830, and 1830-2130. Each of the five squares was repeated for runs in the opposite direction which made a total of ten Latin Squares. Five different drivers with three drivers assigned to each square gave ten possible combinations of three which were assigned according to the pattern in Table 3. The numbers in parentheses are driver identification numbers.

TABLE 3

DRIVER ASSIGNMENTS[a]

Time Period	Direction I	Direction II
0630-1930	(1), (2), (3)	(1), (2), (4)
0930-1230	(1), (4), (5)	(3), (4), (5)
1230-1530	(2), (3), (4)	(1), (3), (4)
1530-1830	(2), (4), (5)	(1), (2), (5)
1830-2130	(1), (3), (5)	(2), (3), (5)

[a]See Table 1, footnote a.

The preceding discussion was necessary to describe the basic experimental design of Project 67-6 from which a portion of the collected data was used for this study. Only Drivometer events collected for the westbound direction of Hillsboro Street were used. The resulting schedule of test runs by time of day, day of week, and driver is given in Table 4. The total number of westbound runs was 30 and each of the five drivers drove exactly six times. The runs were distributed by days of the week with seven runs each on Monday and Tuesday, five each on Wednesday and Thursday, and six runs on Friday.

The data collection began on command from the observer as the test vehicle passed the marked point which indicated the beginning of the test segment. The 16-mm recording camera and Traffic Events Recorder were turned on and the test vehicle moved with the traffic stream or at the posted speed limit in the absence of traffic on the test section. After completion of the test runs, the 16-mm film was processed and placed on a film editor. Measurements per 0.1 mile of the total travel time, running time, small steering reversals, large steering reversals, changes of direction, brake applications, accelerator applications and speed changes were read from the 16-mm film and manually transferred to a data sheet.

TABLE 4

SCHEDULE OF WESTBOUND TEST RUNS ON HILLSBORO STREET

Time	Mon.	Tue.	Wed.	Thur.	Fri.
0630			(3)		
0700		(2)			
0730		(1)			
0800	(3)				
0830					(2)
0900				(1)	
0930		(5)			
1000	(4)				
1030	(1)				
1100					(5)
1130				(4)	
1200			(1)		
1230			(4)		
1300		(3)			
1330		(2)			
1400	(4)				
1430					(3)
1500				(2)	
1530		(5)			
1600	(4)				
1630	(2)				
1700					(5)
1730				(4)	
1800			(2)		
1830	(5)				
1900					(3)
1930					(1)
2000				(5)	
2030			(3)		
2100		(1)			

METHODOLOGY AND ANALYSIS

Methodology

The following eight Drivometer events, recorded on a 0.1-mi basis, were chosen as the independent variables to be used in the analysis:

1. Total travel time or time required to travel the study section (in sec).
2. The change in speed of the vehicle measured in half-units for each 2-mph change in speed.
3. The running time or time that the vehicle is in motion on the study section (in sec).
4. The number of small steering reversals where each unit represents the magnitude of a $^3/_8$-in. change in the steering wheel.

5. The number of large steering reversals where each unit represents the magnitude of a 1-in. change in the steering wheel.

6. The number of brake applications where each unit represents one depression of the brake.

7. The number of accelerator applications where each unit represents a ¼ in. up and down movement of the accelerator.

8. The number of changes in direction where each unit represents a 2.5-deg change in direction of the test vehicle.

During a test run on Hillsboro Street approximately thirty 0.1-mi samples of each Drivometer event were recorded. The means and standard deviations of the Drivometer events for each of the test runs were calculated by an IBM 360 program and a mathematical relationship between the 15 and 30-min traffic volumes, measured by the master control counter concurrent with the test runs, and each Drivometer event was investigated. The method of least squares was used to determine the existence of any linear relationship of the following form:

$$Y = A + BX$$

where

Y = dependent variable (traffic volume),
A = constant or Y-intercept,
B = coefficient or slope, and
X = independent variable (Drivometer event).

The following data were also computed and printed out:

r^2 = coefficient of determination, the dependence of Y on X,
r = correlation coefficient, and
σ_Y = standard error Y.

When the r^2 and σ_Y values for each Drivometer event were known, it could be determined if further analysis were possible. If these values gave some indication of linear dependence, then further analysis could be attempted using a multiple regression analysis. With a multiple regression analysis, the effect of each Drivometer event would be measured and the insignificant events eliminated through a step-wise elimination procedure. Only the most significant events would remain in the final equation. The multiple regression equation was of the following form:

$$Y = B_O + B_1X_1 + \ldots + B_kX_k$$

where

Y = dependent variable (traffic volume),
B_O = constant,
B_{1-k} = correlation coefficients,
X_{1-k} = independent variables (Drivometer events), and
k = total number of independent variables.

The following data were also computed and printed out:

R^2 = same as r^2 except for multiple linear regression, and
σ_Y = standard error Y.

The R^2 and σ_Y values were necessary so that some degree of confidence could be placed in the final equation.

After calculation of the multiple regression equation, the results could be checked against the actual traffic volumes measured concurrently with the test runs.

Analysis of Drivers

Five male drivers for the test vehicle were used and were rotated according to the experimental design (Tables 2 and 3). These drivers were all graduate research assistants with similar educational backgrounds and driving experience.

An analysis of variance of selected Drivometer events was calculated on a distance basis for the total 360 half-hour runs on the test route segments in Table 1. An F test was used and significance was determined at the 0.05 α level. The results are given in Table 5 and indicate that there is no statistical difference among the five drivers used in the tests and the Drivometer events, speed changes and brake applications. At the 0.05 α level, there is a difference among the five drivers and the number of changes in direction. The driver schedule (Table 4) shows that each of the five drivers drove exactly six times.

Analysis of Drivometer Events

For computational ease, the mean values of the Drivometer events were multiplied by the constants in Table 6. These adjusted means were used throughout the analysis.

The first step in the analysis was the preparation of a simple linear regression using traffic volumes as the dependent variables and Drivometer events as the independent variables. The regression analysis was performed by computer and the results are given in Table 7. Plots of selected scatter diagrams are shown in Figures 4 and 5.

According to the r^2 and σ_y values (Table 7) several of the variables indicate that some linear dependence is present but none of the r^2's are high enough to dictate complete linear dependence of traffic volumes on any one of the Drivometer events. The computed value of r^2 explains what proportion of the variation in the traffic volumes is due to the regression of the traffic volumes on the Drivometer events.

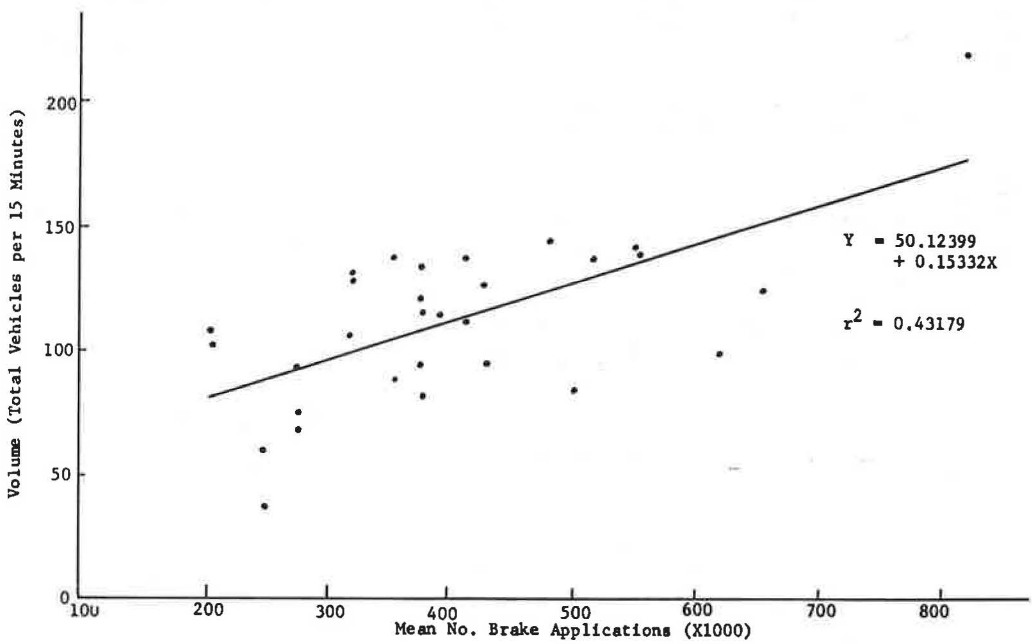


Figure 4. Fifteen-minute volumes vs mean number brake applications.

TABLE 5

ANALYSIS OF VARIANCE RESULTS OF SELECTED DRIVOMETER EVENTS ON DISTANCE BASIS FOR THE TOTAL 360 HALF-HOUR RUNS[a]

Source of Variance	Degrees of Freedom	Speed Changes			Change of Direction			F Value Table (α = 0.05)	Brake Applications		
		Sum of Squares	Mean Sum of Squares	F Value[a] Calculated	Sum of Squares	Mean Sum of Squares	F Value[a] Calculated		Sum of Squares	Mean Sum of Squares	F Value[a] Calculated
Highway direction	1	54.55	54.55	3.17	3.40	3.40	0.66	3.89	0.001	0.001	0.01
3-Hr time period	4	180.34	45.09	2.62	58.02	14.50	2.80	2.41	5.060	1.265	5.74
Highway direction × time period	4	44.25	11.06	0.64	25.20	6.30	1.22	2.41	0.709	0.177	0.80
Highway facility	5	55678.49	11135.70	647.95	10763.30	2152.66	415.71	2.26	731.542	146.308	664.21
Highway facility × Direction	5	14.53	2.91	0.17	30.59	6.12	1.18	2.26	0.257	0.051	0.23
Time period × high facility	20	701.28	35.06	2.04	176.58	8.83	1.71	1.62	19.959	0.998	4.53
Direction × time period × highway facility	20	568.33	28.42	1.65	177.33	8.87	1.71	1.62	3.200	0.160	0.72
Runs within Latin squares	50	730.66	14.61	0.85	372.89	7.46	1.44	1.46	8.979	0.180	0.81
Half-hour time periods within squares	50	995.15	19.90	1.16	279.99	5.60	1.08	1.46	9.798	0.196	0.89
Drivers (adjusted)	4	51.09	12.77	0.74	52.51	13.13	2.53	2.41	1.158	0.290	1.31
Residual	196	3368.55	17.186		1014.94	5.18			43.171	0.220	
Total	359	62387.22			12954.75				823.835		

[a]Brothers, B.T., 1967. An Investigation of Drivometer Measurements as Related to the Accident Potential of a Traffic Facility. Unpublished master's thesis, Department of Civil Engineering, North Carolina State University.

Note: F values that are statistically significant are underlined.

TABLE 6

MULTIPLIERS FOR DATA MEANS

Drivometer Event	Multiplier
Total travel time	10
Small steering rev.'s	10
Large steering rev.'s	100
Brake applic.	1000
Changes in direct.	100
Accel. applic.	100
Speed changes	100
Running time	10

TABLE 7

SIMPLE LINEAR REGRESSION VALUES OF THE 15-MINUTE VOLUMES ON EACH DRIVOMETER EVENT

Independent Variable	A	B	r^2	σ_y
X_1 - total travel time	- 96.16678	1.23433	0.38397	26.83188
X_2 - small steer. rev.	30.42558	0.96440	0.24943	29.61734
X_3 - large steer. rev.	4.30017	0.27169	0.30867	28.42453
X_4 - brake applic.	50.12399	0.15332	0.43179	25.76994
X_5 - changes in direction	35.11278	0.41749	0.32046	28.18109
X_6 - accel. applic.	86.30275	0.43559	0.06326	33.08714
X_7 - speed changes	45.41776	0.16600	0.22492	30.09709
X_8 - running time	-192.78383	2.02294	0.39291	26.63651

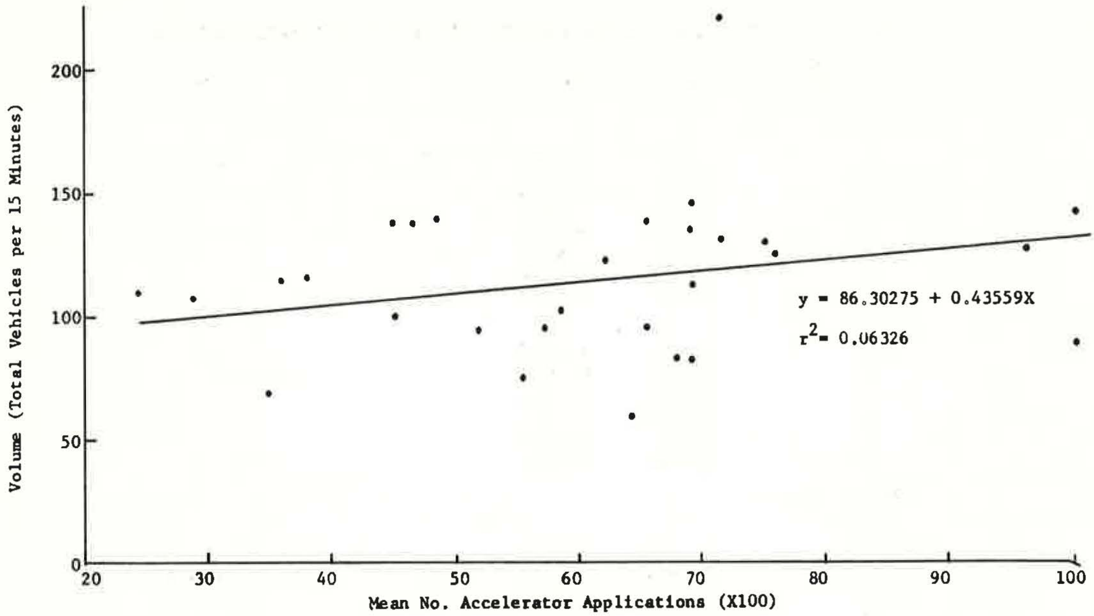


Figure 5. Fifteen-minute volumes vs mean number accelerator applications.

In a stepwise multiple linear regression, all of the independent variables may be forced into the equation and later removed on the basis of a significance test. MULTRGSN 1[1] was used for the multiple regression analysis. In this particular program, an F test value must be chosen to enter or delete a variable. The F value is selected by deciding what α value and degrees of freedom to use (1, p. 125, defines degrees of freedom as "the total number of variates minus the number of independent relationships existing among them"). An α value of 0.05 is usually selected for use in engineering work. This means that the probability of rejecting a coefficient when it should not be rejected is 5 percent. The degrees of freedom can be determined when the number of observations and variables is known. In this study, 30 observations and 9 variables were used. The degrees of freedom in the numerator is 1 and corresponds to the number of degrees of variance among the means. The degrees of freedom in the denominator is 21 which is the number of observations minus the number of variables. An F value of 4.32 was taken from F tables and specified on the input control cards as the value to enter and delete a variable.

In the first step, the 8 independent variables were forced into the equation with the stepwise results (Table 8) for the 15-min volumes. The student's t of each of these variables was computed, and when squared ($t^2 = F$) would eliminate the variable if it did not reach the prescribed F value. This procedure continues until all the variables which fail to reach the prescribed F value are eliminated.

Of the 8 independent variables only two remained in the final results for the 15-min volumes: (a) X_3 = large steering reversals, and (b) X_4 = brake applications.

By observing the R^2 and σ_y values throughout the steps, it can be seen that little change takes place in these variables. This indicates that the eliminated variables have

[1]Performs general multiple regression calculation. MULTGRSN 1. N.C. State University Computer Center, Bulletin 65-7. Revised for IBM 360 system by John Graham, Program Assistant, N.C. State University Computer Center, Raleigh. The MULTGRSN 1 program has been adapted from GUIDE General Program Library writeup 11.3.001, Stepwise Multiple Linear Regression Analysis on the IBM 1410, by W. D. Stevens of Skelly Oil Company, Tulsa, Okla.

TABLE 8

STEPWISE MULTIPLE LINEAR REGRESSION VALUES FOR 15-MINUTE VOLUMES

Independent Variable	B_k	Student t	R^2	σ_4	B_0
Step 1					
X_1	0.82364	0.86776	0.5867	25.3781	-15.14836
X_2	-0.50226	-0.72837			
X_3	0.23896	1.36915			
X_4	0.11740	1.72715			
X_5	0.16409	1.15393			
X_6	0.02127	0.86136			
X_7	-0.11576	-1.08578			
X_8	-0.70839	-0.50228			
Step 2					
X_1	0.45225	0.77317			
X_2	-0.52567	-0.77738			
X_3	0.24206	1.41195			
X_4	0.10476	1.68801			
X_5	0.14633	1.08092			
X_6	0.02022	0.83602			
X_7	-0.09652	-0.98701			
Step 3 - Step 5 (omitted)					
Step 6					
X_3	0.14818	2.02319	0.5421	24.0075	-11.60750
X_4	0.09198	2.20724			
X_5	0.14941	1.17273			
Step 7					
X_3	0.16027	2.19511	0.5178	24.1737	0.18364
X_4	0.11922	3.42242			

TABLE 9

COMPUTED VS MEASURED 15-MINUTE TRAFFIC VOLUMES ON WESTBOUND HILLSBORO STREET

Test Run	Measured Volume	Computed Volume[a]	Deviation	Percentage
0630	38	77	-39	-102.5
0700	61	72	-11	- 18.0
0730	137	103	34	24.8
0800	137	107	30	21.8
0830	128	117	11	8.6
0900	95	110	-15	15.8
0930	83	113	-30	- 36.1
1000	90	94	- 4	- 4.4
1030	103	76	27	26.2
1100	129	108	21	16.3
1130	131	96	35	26.7
1200	134	106	28	20.9
1230	137	132	5	3.6
1300	138	139	- 1	- 0.7
1330	112	115	- 3	- 2.7
1400	113	109	4	3.5
1430	116	110	6	5.2
1500	95	117	-22	-23.2
1530	109	105	4	3.7
1600	142	135	7	4.9
1630	145	127	17	11.7
1700	221	194	27	12.2
1730	127	139	-12	- 9.5
1800	123	100	23	18.7
1830	108	118	-10	- 9.2
1900	100	146	-46	-46.0
1930	94	89	5	5.3
2000	75	111	-36	-48.0
2030	84	122	-38	-45.1
2100	69	86	-17	-24.6
Total	3374	3373		

[a]Rounded off to nearest whole number.

a negligible effect on the dependent variables. Using these two Drivometer events, approximately 52 percent of the regression can be explained. From Table 8, the prediction equation for 15-min traffic volumes can be written as follows:

$$Y = 0.18364 + 0.16027X_3 + 0.11922X_4$$

To check this equation, the values of the Drivometer events were substituted into the equation and volumes were computed for each of the 30 test runs. Comparisons between the measured and computed volumes (Table 9) indicate that 12 of the 30 computed volumes are within ±10 percent of the actual or measured volumes. The deviation column indicates that the computed volumes are within $2\sigma_Y$ of the measured volumes on 100 percent of the test runs and within σ_Y on 63 percent of the runs. The magnitude of σ_Y is ±24 vehicles. The best results were obtained during the time periods from 1230 to 1830 hours. This can perhaps be explained by the fact that parking maneuvers, pedestrian movements and vehicle turning movements were more stable during this period. The early morning and late afternoon time periods produced the largest deviations. The slight differences between the computed and measured volume totals are due to rounding error.

After comparing the measured and computed results, it appears that this volume equation is significant. All results checked within the limits of prediction work. Caution should be exercised in the application of this equation in that it is unique for the westbound direction of Hillsboro Street, and for the test vehicle and equipment used. This volume equation obviously should not be applied under any other conditions without further analysis.

SUMMARY

Findings

1. The frequency of the Drivometer events varied with traffic volumes on the study section. Within the range of traffic volumes recorded, the frequency of the Drivometer events increased with increasing traffic volumes.
2. For the urban arterial facility used in this study, the best correlation with traffic volumes was the mean number of brake applications. Each of the Drivometer events except accelerator applications had a better correlation than speed changes with the traffic volumes. However, none of the simple correlations were high enough to assume complete linear dependence of traffic volume on any single Drivometer event.
3. By combining the two significant Drivometer events (large steering reversals, and brake applications) and using a multiple regression analysis, an equation was derived to predict or statistically measure the traffic volumes on the given facility ($Y = 0.18364 + 0.16027X_3 + 0.11922X_4$).

All the findings of this research were based upon the linear relationship of each Drivometer event and traffic volumes measured concurrently with the collection of the Drivometer data.

Conclusions

The first two objectives of this investigation were to correlate traffic characteristics collected by the Drivometer-equipped vehicle and traffic volumes on a given facility, and to develop a prediction equation for traffic volumes using a Drivometer-equipped vehicle traveling on the given facility. It can be concluded that these two objectives were met and that a prediction or measurement of traffic volumes on the study section of the urban arterial facility can be made using the equation developed in this study. By measuring the number of brake applications and large steering reversals, the traffic volume on the study section of Hillsboro Street can be calculated with statistical confidence.

The third objective was to test the hypothesis that the frequency of the Drivometer events varies with volume on a given facility. It can be concluded that the frequency of each of the Drivometer events does indeed, on the average, vary with traffic volumes on the study section and within the range of traffic volumes measured.

REFERENCES

1. Adler, H. L., and Roessler, E. B. Introduction to Probability and Statistics. W. H. Greenman and Co., 1960.
2. Baerwald, J. (ed.). Traffic Engineering Handbook. Institute of Traffic Engineers, 1965.
3. Department of Commerce, Bureau of Census. Statistical Abstract of the United States. U.S. Government Printing Office, 1966.
4. Greenshields, B. D. Quality of Traffic Transmission. HRB Proc., Vol. 34, p. 508-522, 1955.
5. Greenshields, B. D. The Quality of Traffic Flow. In B. D. Greenshields, H. P. George, N. S. Guerin, M. R. Palmer, and R. T. Underwood (eds.), Quality and Theory of Traffic Flow, A Symposium, Vol. 34, p. 1-40, Bureau of Highway Traffic, Yale University, 1961.
6. Herman, R., Montroll, E. W., Potts, R. B., and Rothery, R. W. Traffic Dynamics: Analysis of Stability in Car Following. Operations Research, Vol 7, p. 86-105, 1959.
7. Jones, T. R., and Potts, R. B. The Measurement of Acceleration Noise—A Traffic Parameter. Operations Research, Vol. 10, p. 745-763, 1962.
8. Montroll, E. W., and Potts, R. B. Car Following and Acceleration Noise. HRB Special Report 79, p. 39-47, 1964.
9. Surti, V. H., and Gervais, E. F. Peak Period Comfort and Service Evaluation of an Urban Freeway and an Alternate Surface Street. National Proving Ground for Freeway Surveillance, Control and Electronic Traffic Aids, Detroit, p. 37-45, 1966.
10. Wardrop, J. G., and Charlesworth, G. A Method of Estimating Speed and Flow of Traffic from a Moving Vehicle. Proc. Instn. Civ. Engineers, Part II, Vol. 3, 1954.
11. Wine, R. L. Statistics for Scientists and Engineers. Prentice-Hall, 1964.
12. Wohl, M. Vehicle Speeds and Volumes Using Sonne Stereo Continuous Strip Photography. Traffic Engineering, Vol. 29, No. 4, p. 14-19, 1959.

Left-Turn Characteristics at Signalized Intersections on Four-Lane Arterial Streets

OLIN K. DART, JR., Louisiana State University

The purpose of this research was to determine left-turn characteristics that would be suitable for defining a left-turn component of a signalized intersection computer simulation model.

The primary characteristic studied was the left-turn gap acceptance distribution and its variation with types of left turns (non-stop or turn from waiting position) and types of gaps in the opposing traffic stream. Studies were conducted at 6 intersections in 4 large Texas cities using multiple-event recorder and 16-mm time-lapse filming techniques to develop data. Over 1000 left-turning vehicles were observed for gap acceptance and other left-turn characteristics.

Gap acceptance distributions were developed on rectangular coordinates, probability-logarithmic coordinates and as regression lines on probit-log time coordinates. Chi-square tests show that there are differences in gap acceptance distributions for moving vs stopped position turns and between distributions for different types of gaps. Comparisons with more recent data from California were very favorable while other research for different left-turn situations was shown to be quite different.

•THE research reported in this paper was the first phase of a project which developed a digital computer model for the simulation of traffic operating through a signalized intersection (12).

In preparing to formulate a model it was found that the literature adequately defined most intersection characteristics but was practically devoid of data relating to left-turn characteristics. Therefore, it was necessary to conduct enough studies to enable these characteristics to be defined and used in the simulation model. This research was necessarily limited to isolated signalized intersections on 4-lane arterial streets, since funding and time would not permit a larger scope of study. Intersection approaches studied included both those with and without separate left-turn lanes, but none of the signals were equipped with separate left-turning phases.

The principal characteristics were the acceptable gap distributions applicable to drivers of vehicles turning left from the inside lane of one approach across the two opposing lanes of traffic. Other left-turn characteristics evaluated include:

1. The percentage of drivers that "jump-the-gun" at the beginning of a green phase;
2. The location within the intersection area that left-turn drivers position their vehicles while evaluating gaps in the opposing traffic stream;
3. The frequency of left turns made after the end of the green phase; and
4. The change in accepted gap requirements with delay.

PREVIOUS RESEARCH

When a driver prepares to make a left turn at a signalized intersection, he will generally encounter an opposing traffic stream which may consist of one, two, or three

Paper sponsored by Committee on Characteristics of Traffic Flow and presented at the 47th Annual Meeting.

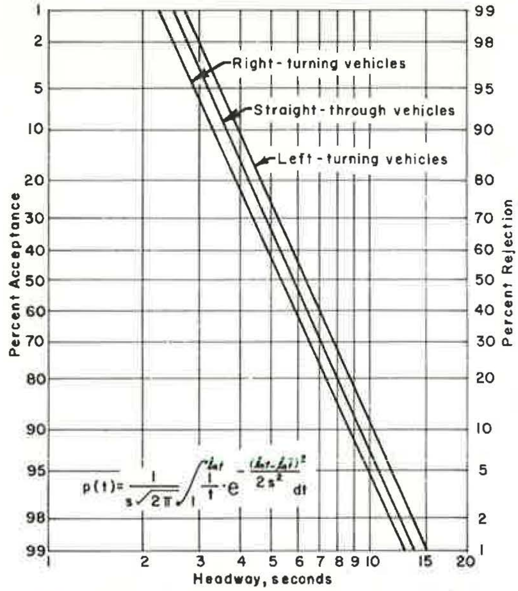


Figure 1. Gap acceptance distributions for vehicles passing a stop sign (2).

lanes of traffic depending on the location. The left-turning driver has to evaluate the gap sizes in this opposing stream and select an opening which he considers large enough for him to cross through the stream safely into the cross street. The gap is known as the acceptable gap for this particular maneuver.

Very little earlier research in determining acceptable gap distributions was directly applicable to the type of maneuver described. Raff (1) and Bissell (2) have provided data for vehicles leaving a cross street under stop sign control, but they do not apply to a 4-lane street left turn under signal control. Bissell's distributions are shown in Figure 1 and were used in Kell's simulation work (3). (Note the typical log-normal form for these distributions.)

More recent studies by Solberg and Oppenlander (4) provided gap and lag-acceptance data for minor street vehicles crossing or entering main street traffic streams from a stopped position. In addition, they utilized the technique of probit analysis which permitted a statistical evaluation of differences in gap-acceptance times. In general, they concluded that there was general agreement among the three methods investigated.

Research reported by Wagner (5) deals with the same type of gap and lag acceptances at a single intersection, but he investigated the effects of certain factors on driver decisions. These factors included vehicle type, pressure of traffic demand, direction of movements through the intersection, sequence of gap formation, and conditions on the opposing side-street approach.

A number of student theses at the Yale Bureau of Traffic have dealt with left turns but only three have dealt specifically with gap-acceptance characteristics. These were limited studies dealing with 2-lane, two-way traffic streams, each conducted for a single intersection in the city of New Haven. Kaiser (6) found for a sample of 158 drivers of left-turning vehicles at an unsignalized intersection that the smallest gap accepted was 3.75 sec and the largest gap rejected was 4.75 sec. Clark (7), at another unsignalized intersection, found a critical gap size (size where there are an equal number of rejected gap sizes larger than this size and accepted gap sizes smaller than this size) of 3.2 sec as compared with 4.2 sec in Kaiser's study; the difference is probably attributable to the geometric differences and location of the two intersections. Noblitt (8) found gap acceptances for left-turning truck combinations to be 1.4 to 1.8 times as large as the required gap for cars, and 1.2 to 1.5 times as large as the required gap for single-unit trucks.

Based on 500 field observations on 2-lane two-way streets, Kell (9) found the left-turn gap-acceptance distribution given in Table 1. Again, he pointed out the log-normal nature of these data.

The most recent field studies evaluating left-turn gap and lag acceptances were conducted by the Traffic Research

TABLE 1

GAP ACCEPTANCES BY LEFT-TURN VEHICLES ON TWO-LANE, TWO-WAY STREETS[a]

Gap Size (sec)	Cumulative % Accepting	Gap Size (sec)	Cumulative % Accepting
1.0	0	4.5-5.0	94.7
1.0-1.5	1.4	5.0-5.5	96.4
1.5-2.0	10.2	5.5-6.0	97.9
2.0-2.5	18.3	6.0-6.5	98.2
2.5-3.0	31.3	6.5-7.0	98.5
3.0-3.5	50.0	7.0-7.5	99.3
3.5-4.0	64.6	7.5-8.0	99.4
4.0-4.5	85.3	8.0+	100.0

[a]After Kell (9).

TABLE 2

SUMMARY OF STUDY SITE CHARACTERISTICS

CHARACTERISTIC	INTERSECTION NUMBER 1	2	3	4	5	6
CITY	AUSTIN	FT. WORTH	HOUSTON	HOUSTON	HOUSTON	WACO
STREET NAMES (STREET STUDIED LISTED FIRST)	24th. ST. AT LAMAR	UNIVERSITY AV. AT E-W FREEWAY FRONTAGE RD.	SHEPHERD ST. AT W. GRAY	WASHINGTON ST. AT SAWYER	FRANKLIN ST. AT CAROLINE	18th. ST. AT WACO DRIVE
INTERSECTION TYPE	4 – WAY	TEE	4 - WAY	TEE	TEE	4 - WAY
SIGNALIZATION	FULL - ACTUATED	UNSIGNALIZED	FIXED - TIME 62 & 75 SEC. CYCLE LENGTHS (50-50 SPLIT)	FIXED - TIME 70 SEC. CYCLE (75-25 SPLIT)	UNSIGNALIZED	FIXED - TIME 62 SEC. CYCLE (40-60 SPLIT)
ON STUDY APPROACH (1/2 OF FULL STREET WIDTH)						
NO. OF LANES & APPROACH STUDIED	2 - BOTH	3 - NORTHBOUND	2 - SOUTHBOUND	2 - EASTBOUND	2 - WESTBOUND	2 - BOTH
CHANNELIZATION	NONE	MEDIAN OPENING	NONE	NONE	PAINTED - LEFT-TURN SLOT	LEFT - TURN SLOTS OUTLINED BY 6" RAISED BUTTONS
PARKING	NONE	NONE	NONE	YES - PROHIBITED 1-WAY AT PEAK	NONE	NONE
MAXIMUM 1-WAY VOLUME, VPH	700	1250	1050	2000	1170	800
% LEFT TURNS	5 - 12	10 - 30	5 - 20	5 - 10	15 - 40	10 - 25
STUDY TYPE	MOTION PICTURES ON GREEN PHASE	20-PEN RECORDER & LIMITED TIME-LAPSE PICTURE	1-SEC. TIME-LAPSE PICTURES	1-SEC. TIME-LAPSE PICTURES	1-SEC. TIME-LAPSE PICTURES	20-PEN RECORDER & 1-SEC. TIME-LAPSE PICTURES
HOURS STUDIED	7-8 AM, 5-6 PM	8 AM - 4:30 PM	2:30 - 5:15 PM	3-6 PM	7-9 AM	11 AM - 6 PM

Corporation (10). They were interested in validating probability tables used in left-turn decision components of signalized intersection simulation models. Three intersections in the Berkeley-Oakland area of California were studied with left-turn gap and lag-acceptance distributions being developed for streets that were 40 to 46 ft wide.

INTERSECTION FIELD STUDIES

Variables

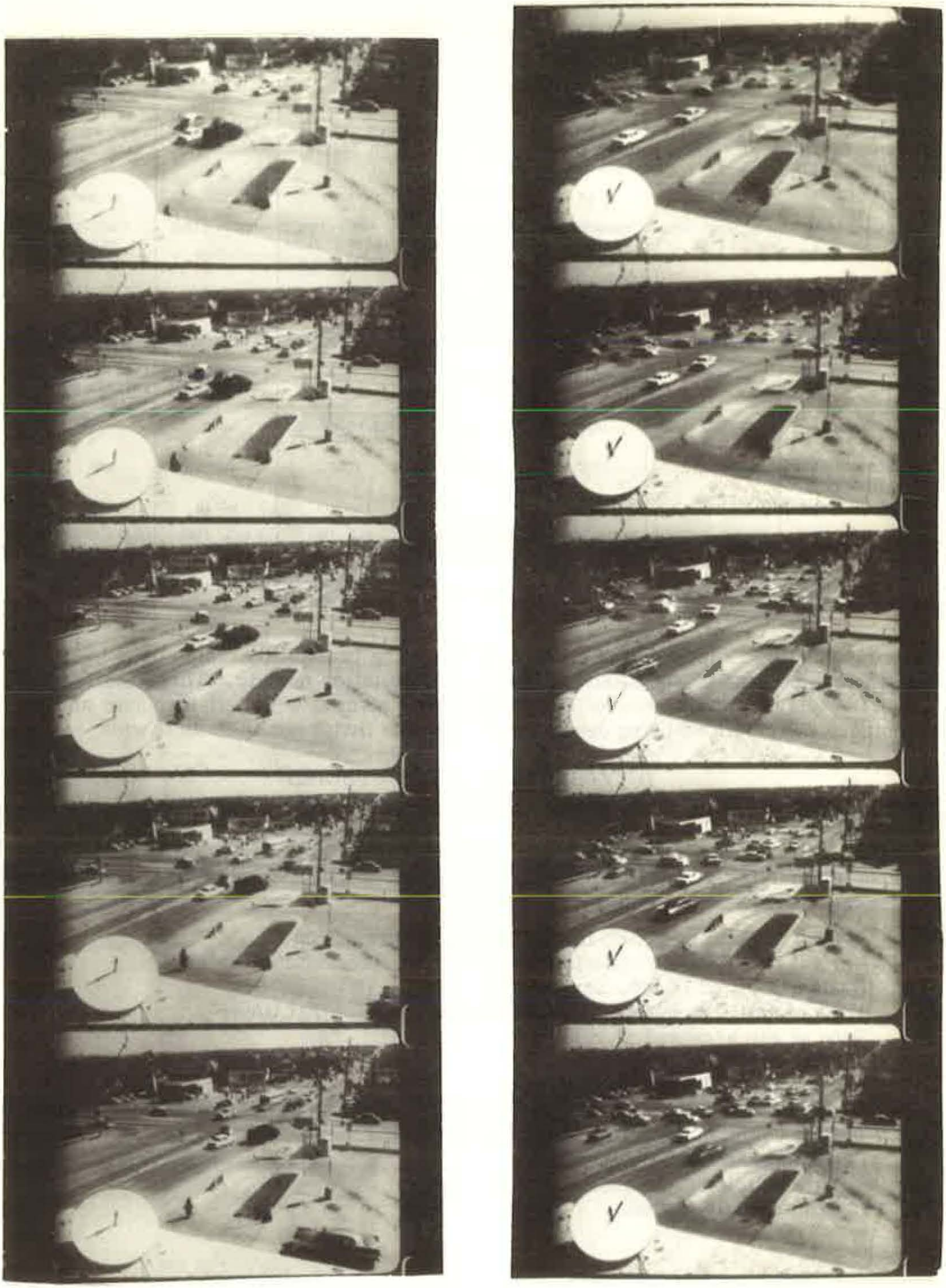
To evaluate gap acceptance of left-turning drivers and to provide other data for the purpose of verifying simulation output, it was decided that the following variables had to be measured in any field studies conducted:

1. Arrival time of each left-turn vehicle at intersection;
2. Arrival time of each opposing vehicle in the intersection center;
3. Waiting position of left-turn vehicle;
4. Time of actual turn maneuver and signal phase at that time;
5. Type of gap accepted or rejected;
6. Vehicle type;
7. Sex and approximate age of driver;
8. Arrival and depature time of each vehicle from intersection approach "system";
9. Traffic volumes operating in both directions;
10. Percentage of turns in each lane;
11. Signal cycle length and phasing split; and
12. Length of study period.

Site Selection

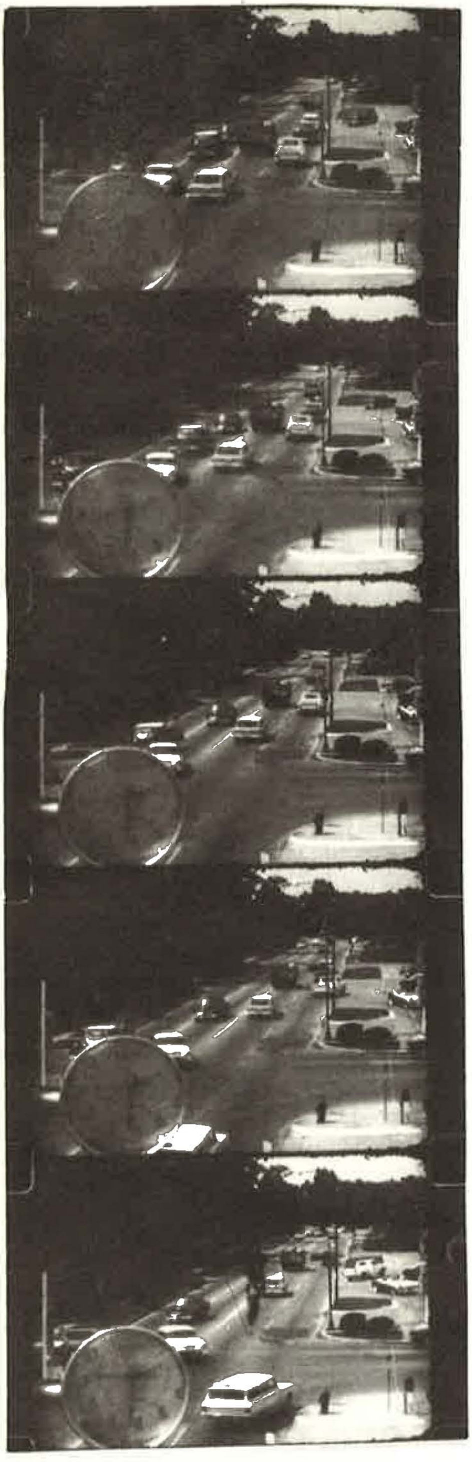
It is difficult to locate many intersections with all the desirable characteristics that one would prefer for study purposes. All of the intersections finally selected were not ideal but did provide a good range of traffic volume and percentage of turns.

On the basis of preliminary investigations, six intersections were selected for detail studies. They were located in the following cities: Austin, 1; Ft. Worth, 1; Houston, 3; and Waco, 1. Specific characteristics of each study site are summarized in Table 2.

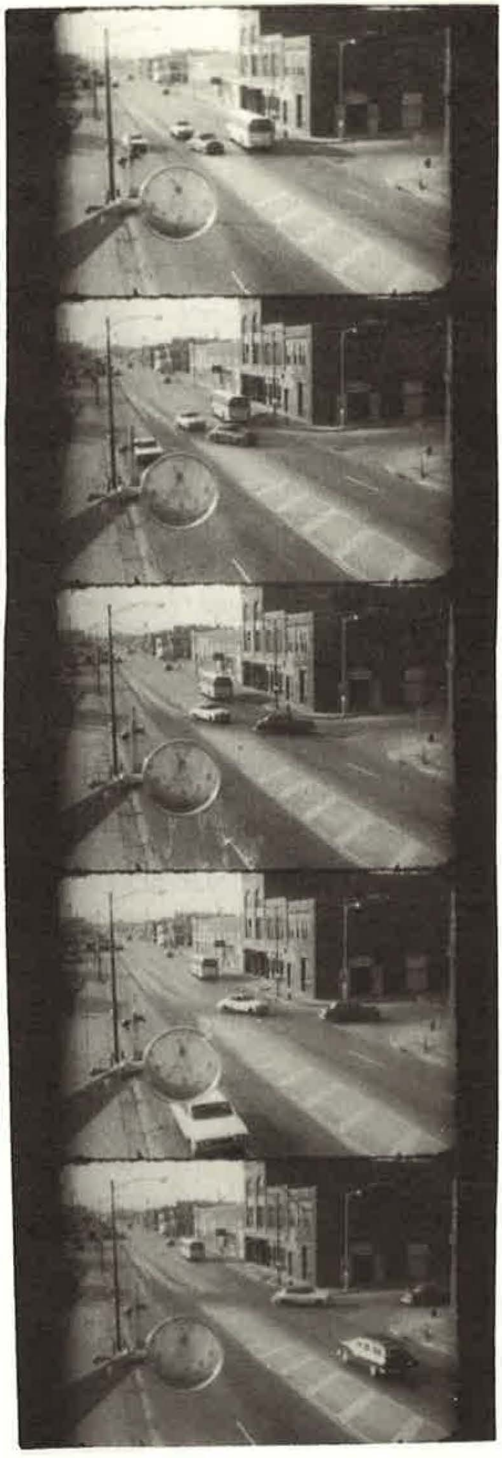


(a) (b)

Figure 2. Typical time-lapse film strips, 18th at Waco Drive: (a) "jumping-the-gun" and (b) clearing on the red phase.



(a)



(b)

Figure 3. Typical time-lapse film strips, Houston: (a) turning on amber, yielding to opposing right turn and (b) gap acceptance.

Although principally interested in 4-way signalized intersections, two of the locations studied were unsignalized T-intersections with relatively high volumes (Table 2). This was permitted for the following reasons:

1. Gap acceptance for left turns from a stopped position at an unsignalized intersection should be the same as those operating on the green phase of a signalized intersection; and
2. A relatively high percentage of left turns at this type of intersection are made without stopping and therefore gap sizes required for this different maneuver could be evaluated and applied to the comparable maneuver at a signalized intersection.

Study Methods Considered

Accepted methods for obtaining the required data include the use of a multiple-event recorder connected to road tube actuated air switches and hand switches, 16-mm motion-picture photography, 16-mm time-lapse photography, or combinations of these methods. Each of these techniques was used for at least two hours of study at one location. The details of each study method can be obtained elswhere (12).

On the basis of these studies, it was decided that the most satisfactory and economical procedure, from both field study and data analysis time standpoints, was the time-lapse photography technique. This procedure was used for all three Houston studies (Table 2).

A 16-mm camera and tripod were set up on either a nearby roof-top or a platform truck within 250 ft of the intersection. The single frame button on the camera was actuated at 1-sec intervals through a special electrical timer-solenoid circuit.

This technique required only one-tenth as much film as the motion picture technique and still provided gap measurements to an accuracy of ±0.5 sec. A filming technique is more desirable than the 20-pen recorder method, since a complete picture of the intersection system is available at any instant and complete analysis of a situation can be obtained by running and rerunning the film through a projector. Typical film strips obtained by this method are shown in Figures 2 and 3.

Data on left-turn operation obtained from time-lapse films were transferred to punch cards and a Fortran program was written for an IBM 709 digital computer to analyze these data.

There were four types of gaps considered in this analysis of 4-lane street studies (Fig. 4). There are two types of lane gaps, i.e., gaps formed by two successive inside-lane vehicles or two successive outside-lane vehicles. In addition there are two types of offset gaps, i.e., gaps formed by a vehicle in one lane trailed by a vehicle in the adjacent lane.

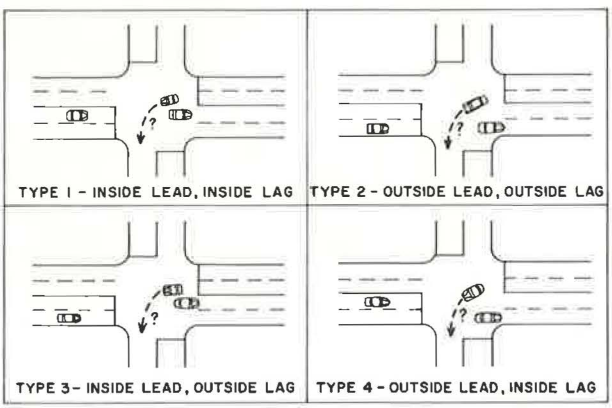


Figure 4. Opposing traffic stream gap types confronting left-turn vehicles.

Other studies (4, 5, 6, 10) have distinguished between "gaps" and "lags" in developing acceptance distributions. The difference in these two terms is shown by Figure 3b. The first left-turn vehicle accepts a gap formed by the bus and the trailing station wagon. The second left-turn vehicle accepts a lag or the time elapsed between its arrival at the center of the intersection and the arrival of the station wagon at the same location (in its own lane). In this research, there was generally no distinction between lags and gaps except for turns made without stopping. The drivers of these vehicles accept lags rather than gaps.

RESULTS

The principal objective of the field studies was to determine the distributions of acceptable gaps for left-turning vehicles on 4-lane arterial streets. This included variations for moving and stopped vehicles and different types of opposing gaps.

Other characteristics of left turns were also noted such as the waiting or starting position in the intersection area, the percentage of left-turning vehicles that jump-the-gun at the start of the green phase, the frequency of left turns clearing the intersection after the end of the green phase, the percentage of left-turn drivers accepting gaps smaller than the largest gap previously rejected, and the variation of gap acceptance with driver characteristics.

Turn Types and Starting Positions

Aside from left turns that jump-the-gun, most left turns at a signalized intersection under moderate to heavy traffic conditions are made from a stopped or waiting position in the intersection proper.

Other left turns may be made while the vehicles are slowly moving (or creeping) into the intersection proper; still others are made by vehicles that never slow much below maximum permissible turn speed. Under relatively heavy traffic conditions, these vehicles follow on the "heels" of other left-turn vehicles; under relatively light traffic conditions, the majority of the turns are likely to be nonstopping turns.

Analysis of data for 174 turns at the most typical intersection studied, West Gray at Shepherd in Houston, showed 126, or 72.4 percent, were made from a stopped position. Of these 126 turns, 75, or 59.6 percent, were waiting in a position near the center of the intersection; 41, or 32.5 percent, were waiting in a position about half-way into the intersection from the stop line; and the remaining 10 (9 of which jumped-the-gun) were made from the stop line.

At the unsignalized intersection of Franklin and Caroline Streets in Houston, data for 690 left turns showed 447, or 64.9 percent, were moving-type turns.

Left-Turn Gap Acceptance

There are four possible gap types to be found in the opposing traffic stream (Fig. 4). In each of the three principal studies, gap acceptance distributions were determined for each gap type. For each particular gap size (nearest 0.5 sec), the percent of gaps accepted is defined by the number of these gaps accepted divided by the total number of these gaps accepted and rejected. Turns that jumped-the-gun, forced their turn, or completed their turn after all opposing traffic had stopped were not considered to have accepted a normal gap.

Turns From a Stopped Position—The largest sample of turns from a stopped position came from the intersection of Franklin and Caroline Streets in Houston. A total of 232 accepted gaps were distributed as to type as follows: 51, Type 1; 66, Type 2; 55, Type 3; and 60, Type 4. These 232 turns also rejected a total of 809 gaps. The resulting cumulative distribution of acceptable gaps by type is shown in Figure 5, where there is considerable scatter of data points.

When unsatisfactory turns (forced, unopposed, etc.) were eliminated from the 126 vehicle sample at the West Gray and

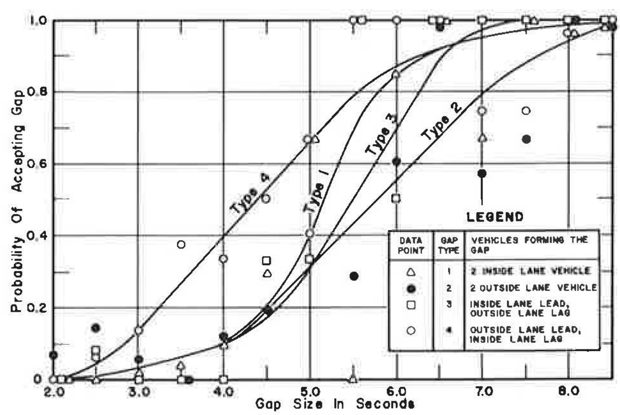


Figure 5. Probability of left-turn vehicle accepting gap from stopped position at unsignalized, channelized T-intersection; Franklin at Caroline, Houston.

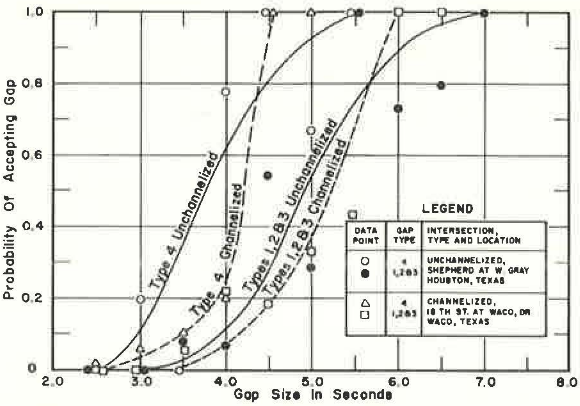


Figure 6. Probability of left-turn vehicle accepting gap from stopped position at signalized 4-way intersections.

Shepherd location, there were only 75 turns left. Accepted gaps distributed as to type included: 14, Type 1; 21, Type 2; 11, Type 3; and 30, Type 4. These sample sizes were too small to provide satisfactory cumulative curves, but there did appear to be a distinct difference between Type 4 gaps and the other 3 combined (Fig. 6).

At the Waco location, only those turns that could be verified by time-lapse films in the same manner as in the other two studies were used for further comparison. With a high percentage of turns being made without opposition or jumping-the-gun, only 35 of 232 turns could be analyzed, providing data on 16 Type 4 acceptable gaps as compared with 19 gaps of Types 1, 2, and 3. The cumulative distributions are also shown in Figure 6.

Turns From a Moving Position—The gap accepted by a moving left-turn vehicle is not defined by two opposing vehicles but by the difference in arrival times (at the intersection center) of the turning vehicle and the nearest opposing vehicle. (Some authors designate such a gap as a lag.) Therefore, there were only two types of gaps to be evaluated; one with the nearest opposing vehicle in the outside lane (Type 2) and the other with the nearest opposing vehicle in the inside lane (Type 1).

The 447 moving left-turns filmed at Franklin and Caroline Streets in Houston were classified as either creeping or moving at normal turn speed as they began their turn. An analysis of gap acceptance distributions for both conditions showed no significant difference between the two for both Type 1 and Type 2 gaps. The data were then separated into two groups by type providing 185 Type 1 and 262 Type 2 gaps.

Since a moving turn does not reject a gap, it was necessary to check the first gap rejected by each of the 243 left turns that stopped at this location before making the turn. Analysis provided 129 Type 1 and 114 Type 2 gaps that were rejected. The cumulative gap acceptance distributions provided the two curves shown in Figure 7. As would be expected for a given gap size, a lower percentage of drivers will accept the gap if it is formed by an opposing vehicle in the the outside lane, as compared with one in the inside lane; it takes longer to clear the outside lane than the inside lane when making the left turn.

Left Turns That Jump-the-Gun—It was observed that many drivers whose left-turn vehicle is at the head of the inside lane or left-turn channelization queue when the signal changes to a green indication will jump-the-gun, i.e., make the turn before the opposing traffic enters the intersection. In some states this maneuver is outlawed; in others it is encouraged—but it definitely does occur at the intersections studied in Texas cities.

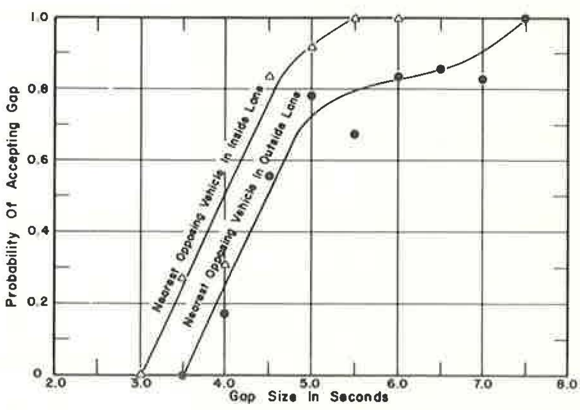


Figure 7. Probability of left-turn vehicle accepting gap without stopping at unsignalized, channelized T-intersection; Franklin at Caroline, Houston.

To determine the probability of its occurrence, the number of signal cycles at which a left turn was

TABLE 3

PROBABILITY OF LEFT-TURN DRIVER'S JUMPING-THE-GUN

Intersection Location	Approach Type	No. of Signal Cycles: Total	With L. T. at Queue Head	With LT's Jumping-the-Gun	Probability of Occurrence
Houston	Nonchannelized	130	35	6	0.171
Austin 1	Nonchannelized	65	12	2	0.167
Austin 2	Nonchannelized	65	8	0	0.000
Total	Nonchannelized	260	55	8	0.145
Waco 1	Channelized	127	79	8	0.101
Waco 2	Channelized	127	86	16	0.186
Total	Channelized	254	165	24	0.145

observed at the head of the inside lane or channelized queue was divided into the number of cycles when such a vehicle jumped-the-gun. The results for 5 signalized approaches are given in Table 3. The same percentage occurred for the total observations of each type of intersection.

Change of Required Acceptable Gap—As hypothesized, there were a number of drivers of left-turn vehicles who accepted a gap that was smaller than the largest one previously rejected. In general the difference was only about 0.5 to 1.0 sec. The frequency of such a change in gap requirements and consequent probability of occurrence is given in Table 4. Although there was an attempt to relate this occurrence with vehicle waiting time and opposing traffic stream density, no relationship was apparent.

Turns Made on Yellow and Red Phases—Many left turns were completed after the signal phase had changed to yellow or even red. Observation of 260 cycles of operation on 3 nonchannelized approaches in Houston and Austin found 58 left-turn vehicles still waiting to turn when the green phase ended. All 58 of these turns were then completed, 40 percent of them on the red phase.

At these and other locations, more than one vehicle turned after the end of the green phase. As many as four turned on one occasion at the Waco location, the last one almost colliding with cross street vehicles that had started legally (Fig. 2b).

ANALYSIS OF GAP DISTRIBUTIONS

The distribution curves in Figures 5, 6, and 7 were fitted by eye to the highly scattered data after attempts to fit various mathematical distributions were unsuccessful. However, this type of data usually takes a log-normal form and can be represented by straight lines on probability X log paper as in Figure 1. The other statistical treatment of this type of data involves the use of probit analysis as described by Finney (11). This technique is used successfully in biological assay work where it is necessary to analyze the relationship between insecticide concentrations (or dosages) and its effectiveness in percent of insects killed at each concentration. The typical cumulative distribution of percent killed vs insecticide concentration takes the form of the nonlinear normal

TABLE 4

PROBABILITY OF LEFT-TURN DRIVER'S REDUCING ACCEPTABLE GAP

Study Location and Approach Type	No. of Left Turns: Total	Accepting Gap Smaller Than Largest Rejected	Probability of Occurrence
Houston, nonchannelized	119	14	0.117
Waco, channelized	241	22	0.091
Grand total	360	36	0.100

TABLE 5

PROBIT ANALYSIS REGRESSION EQUATIONS

Location and Turn Type	Gap Types	Regression Equation	df	χ^2
Franklin at Caroline, Houston—turns from stopped position (channelized)	1	Y = -1.265 + 8.937X	4	0.604
	2	Y = 1.795 + 4.749X	4	2.687
	3	Y = 0.469 + 7.609X	1	1.299
	4	Y = 0.455 + 6.985X	4	0.604
	All	Y = 1.007 + 5.931X	5	12.752[a]
W. Gray at Shepherd, Houston—turns from stopped position (unchannelized)	1-3	Y = -4.044 + 13.342X	3	0.867
	4	Y = -0.747 + 11.018X	2	1.208
	All	Y = -2.142 + 11.094X	4	0.803
18th St. at Waco Dr., Waco—turns from stopped position (channelized)	1-3	Y = -5.241 + 15.368X	2	1.690
	4	Y = -2.310 + 14.075X	2	8.500[a]
	All	Y = -3.840 + 13.662X	2	2.157
Franklin at Caroline, Houston—turns without stopping (channelized)	1	Y = -8.027 + 21.874X	3	0.142
	2	Y = -1.979 + 10.807X	4	2.240
	All	Y = -2.638 + 11.795X	4	1.245

[a]Significant at 5 percent level.

sigmoid curve. By utilizing the probit transformation, the relationship is expressed in linear form.

Solberg and Oppenlander (4) applied this technique to lag and gap acceptances at stop-controlled intersections. They showed that the probit of the gap acceptance percentage, Y, was related to the logarithm of the time gap, X, by the straight-line equation

$$Y = a + bX$$

They also determined the parameters of the tolerance distribution, its mean and variance, as well as the median gap and lag acceptances. The probit transformation is given by the linear equation

$$Y = 5 + \frac{1}{\sigma}(X - \mu)$$

where Y is the probit of acceptance percentages and X is the log of the time gap. The median is estimated from values of X when Y = 5 (when percent acceptance = 50 percent), which is the median effect dosage in the biological assay work.

Probit Analyses

The probit transformation was applied to the gap acceptance data and produced the linear regression equations in Table 5 and Figures 8 through 12. Table 5 includes the results of testing the goodness-of-fit of these equations to the experimental data. The data were grouped in 1-sec intervals (1.26-2.25, 2.26-3.25, etc.) to eliminate some of the inconsistencies. This shows that most of the experimental data can be represented adequately by the straight-line probit relationships.

To ascertain if there were differences between types of gaps and similar data at different locations further evaluations were made using the χ^2 test.

TABLE 6

CHI SQUARE TEST OF GAP DISTRIBUTION DIFFERENCES

Intersection	Gap Types Compared	df	χ^2
Franklin at Caroline, stopped turns, channelized	Type 1 vs Type 4	5	14.09[a]
	Type 2 vs Type 4	4	5.81
	Type 3 vs Type 4	2	0.47
	Type 1 vs Type 2	3	3.36
Franklin at Caroline, moving turns	Type 1 vs Type 2	2	6.59[a]
	All moving vs all stopped	5	11.28[a]
W. Gray at Shepherd, unchannelized	Types 1-3 vs Type 4	1	28.40[b]
	All vs all stopped turns at Franklin-Caroline	5	3.94
18th St. at Waco Dr., channelized	All vs all stopped turns at Franklin-Caroline	5	11.24[a]

[a]Significant at 5 percent level.
[b]Significant at 1 percent level.

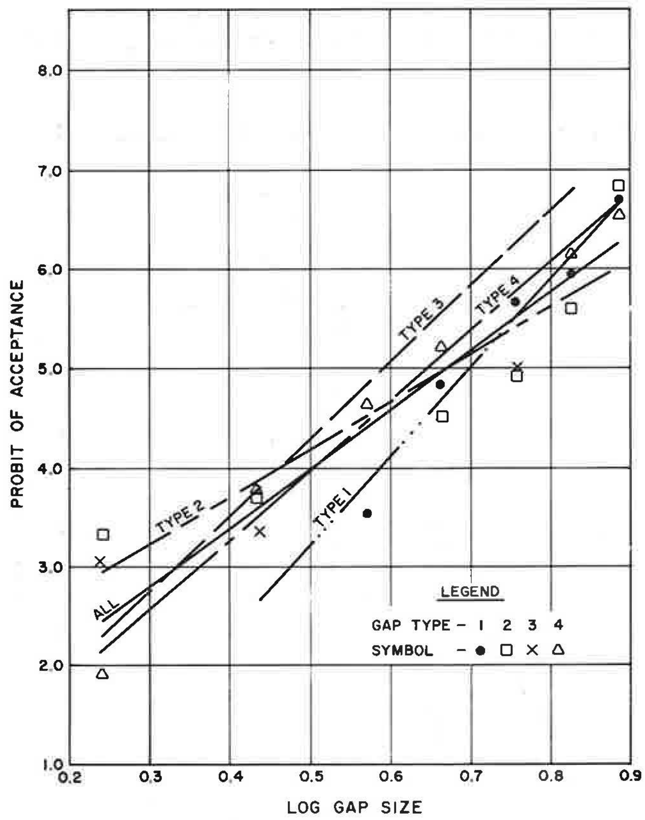


Figure 8. Probit regression lines by gap types, Franklin at Caroline, turns from stopped position (Houston).

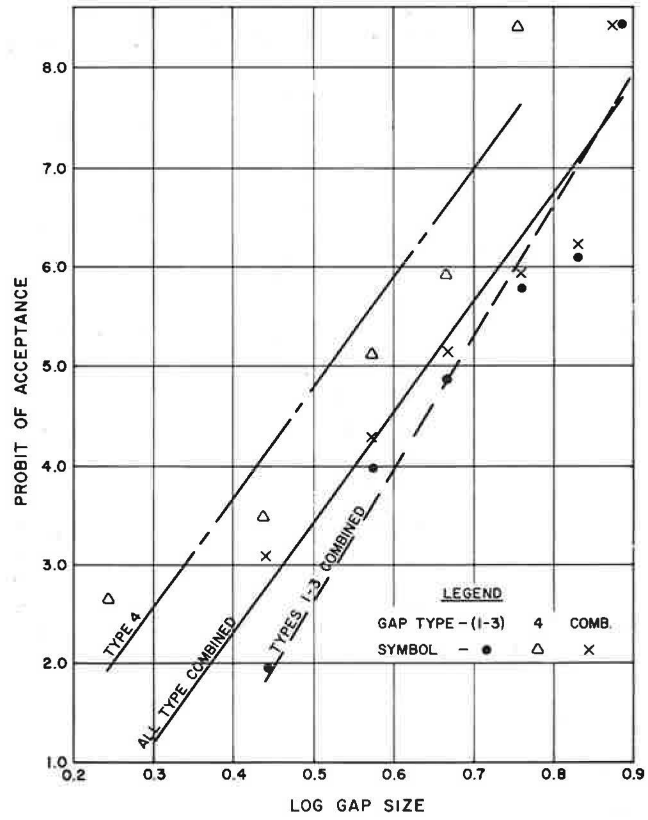


Figure 9. Probit regression lines by gap types, W. Gray at Shepherd, unchannelized (Houston).

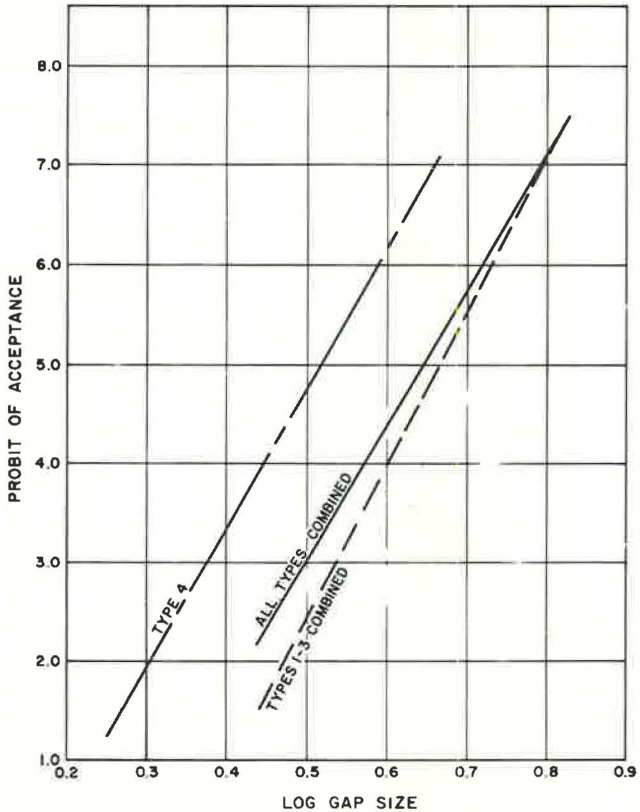


Figure 10. Probit regression lines by gap types, 18th St. at Waco Drive, channelized (Waco).

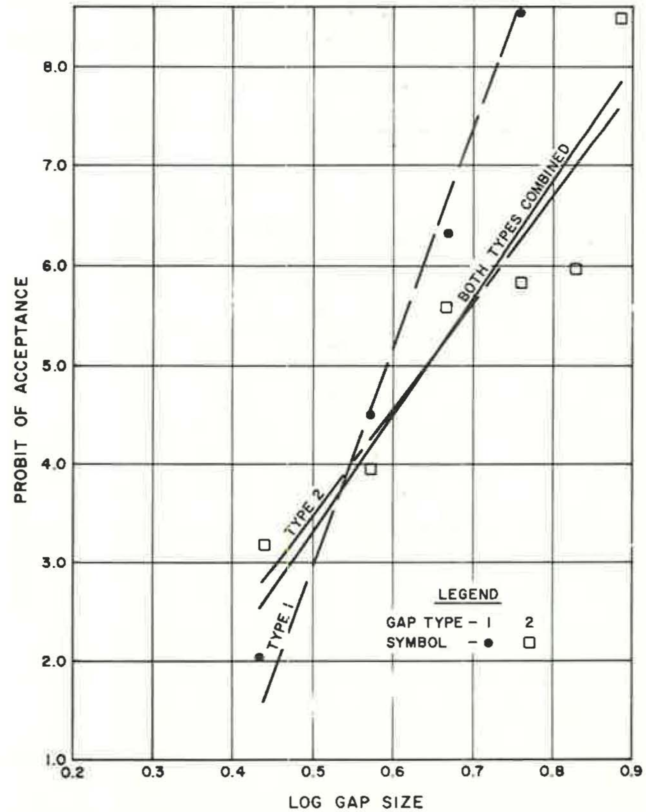


Figure 11. Probit regression lines by gap types, Franklin at Caroline, moving turns (Houston).

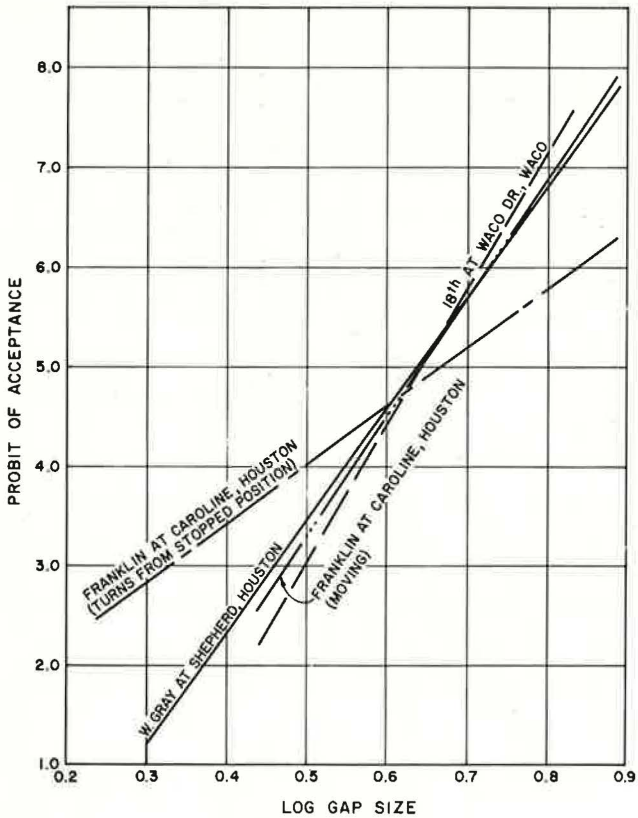


Figure 12. Probit regression lines, all gap types combined.

Observed data for one group were tested against the regression line of the group being used as the expected or theoretical data. The results are given in Table 6. There is a significant difference between Type 1 and Type 4 gap acceptance distributions at the Franklin and Caroline intersection in Houston and between acceptance of Types 1, 2, and 3 gaps as a group vs Type 4 gap acceptance at the Waco intersection.

For moving turns there is also a significant difference between the gap (lag) acceptance distributions for Types 1 and 2 gaps (lags). Furthermore there is a significant difference between the gap acceptance distribution for turns made without stopping as compared with the distribution for turns made from a stopped position.

When all gap types are combined into one distribution, the somewhat limited data of the Waco intersection produce a distribution that is different from that obtained at the Houston intersection. (Both intersections are channelized but the Houston one is unsignalized.)

Comparisons With Other Research

It is interesting how the distributions developed in this study compare with those of other researchers. Figure 13 compares two distributions with combined (all gap types) data from this study with the combined left-turn gap acceptance data from the validation study of Traffic Research Corporation (10) and the left-turn gap acceptance from a cross street developed by Bissell (2) and Solberg and Oppenlander (4).

The gap acceptance distributions obtained from the intersection of Franklin and Caroline compare quite favorably with those obtained by Traffic Research Corporation. For

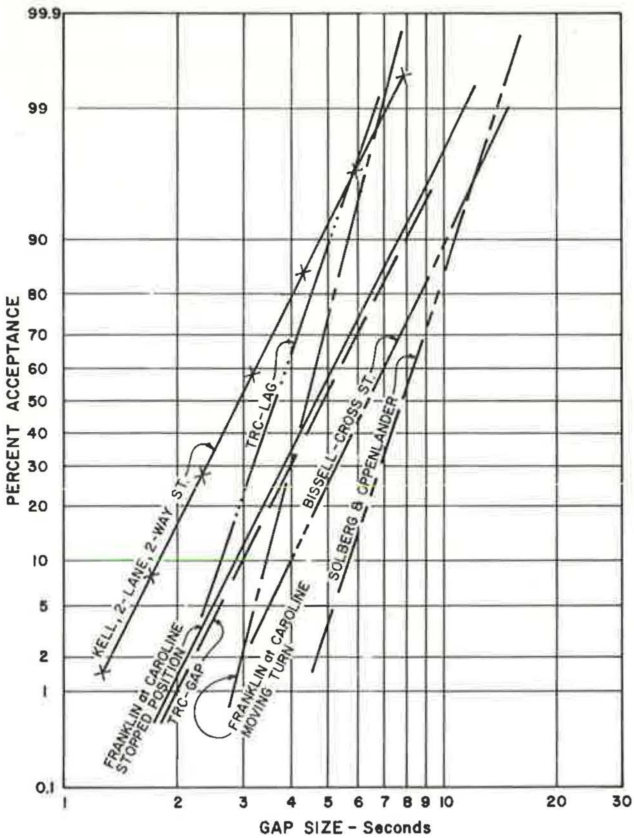


Figure 13. Gap acceptance comparison.

turns from a stopped position, the distribution is almost identical with that obtained for the combined California data. The moving-turn gap distribution is similar to the TRC acceptable lag distribution but has a steeper slope.

As would be expected, there is a definite difference between distributions obtained for left turns from cross streets under stop sign control and the left turns across opposing traffic streams as reported in this paper. Another comparison with Kell's data (9) shows that left turns across a single-lane opposing stream require smaller gaps than turns across a two-lane stream (Fig. 13).

CONCLUSIONS

On the basis of the research reported in this paper the following conclusions seem to be in order. Although these are based on a limited number of studies, it is felt that the results are representative of similar types of signalized intersections.

1. A time-lapse 16-mm filming technique that utilizes 1-sec intervals was found to be the most satisfactory and economical method for studying gap-acceptance characteristics of left turns at a signalized intersection.

2. More than one gap-acceptance distribution or one critical value is necessary to cover the existing range of left-turn types on a 4-lane street: (a) moving-turn gap or lag-acceptance distributions are significantly different from distributions for turns made from a stopped position; (b) shorter gaps are required for most moving turns, if the next opposing vehicle is in the inside lane as compared with one in the outside lane; and (c) for turns made from a stopped position, there is evidence that Type 4 gaps will more likely be accepted than any of the other three gap types of the same size.

3. There was no appreciable difference between gap acceptance requirements for left turns and other turn characteristics on a channelized approach as compared with those on an unchannelized approach. In general, it is very rare that a gap smaller than 2 sec will be accepted or that one larger than 8 sec will be rejected.

4. With the exception of turns that jump-the-gun or turn without stopping almost every left turn is made from a waiting position within the intersection area ahead of the stop line. There are likely to be twice as many turns made from a position (front of vehicle) at the center of the intersection than from a position about half-way between the stop line and the intersection center.

5. There are a significant number of drivers who will jump-the-gun in turning left at the beginning of a green phase, and this feature should be incorporated in any left-turn model. The probability of this occurring appears to be about 0.15 for all left-turn vehicles at the head of a queue when the signal first displays the green phase.

6. Although 10 percent of left-turning drivers accepted gaps smaller than the largest gap rejected, this research did not establish any significant relationships with driver delay or opposing traffic stream characteristics to explain this phenomenon. The difference in gap sizes being generally less than 1 sec makes this result of little value to a simulation model.

7. A large number of left turns will be completed after all opposing traffic has passed and many will turn even after the green phase has ended. This fact makes it difficult to collect gap acceptance data under light to moderate traffic conditions.

ACKNOWLEDGMENTS

This paper reports results obtained from part of the research conducted by the author for his doctoral dissertation at Texas A&M University. He wishes to express his appreciation to the Institute of Traffic Engineers who sponsored the field studies and data analyses through their Technical Development and Research Fund. The author is further indebted to the Texas Transportation Institute at Texas A&M University for the provision of traffic measuring equipment and data analysis facilities. This research was also greatly aided by the helpful counsel of Charles J. Keese and other staff members of the Texas Transportation Institute.

This research was further made possible through the cooperation of traffic engineering personnel of the Texas Highway Department and several Texas cities.

REFERENCES

1. Raff, M. S. A Volume Warrant for Urban Stop Signs. Eno Foundation, 1950.
2. Bissell, H. H. Traffic Gap Acceptance from a Stop Sign. Unpublished Graduate Research Report, ITTE, Univ. of California, Berkeley, 1960.
3. Kell, James H. Intersection Delay Obtained by Simulating Traffic on a Computer. Highway Research Record 15, p. 73-97, 1963.
4. Solberg, Per, and Oppenlander, J. C. Lag and Gap Acceptances at Stop-Controlled Intersections. Highway Research Record 118, p. 48-67, 1965.
5. Wagner, Frederick A., Jr. An Evaluation of Fundamental Driver Decisions and Reactions at an Intersection. Highway Research Record 118, p. 68-84, 1965.
6. Kaiser, Fred J., Jr. Left Turn Gap Acceptance at an Unsignalized Intersection. Thesis manuscript, Bureau of Highway Traffic, Yale Univ., May 1951.
7. Clark, Thomas R. The Left Turn, A Criteria (sic) for Intersection Control. Thesis manuscript, Bureau of Highway Traffic, Yale Univ., May 1960.
8. Noblitt, Jack L. A Study of Gaps Associated with a Left-Turning Truck. Thesis manuscript, Yale Univ., May 1959.
9. Kell, James H. Data obtained by ITTE, University of Calif., Aug. 10, 1961.
10. Traffic Research Corporation. Final Report, Validation of Traffic Simulation Model Components. Prepared for U.S. Bureau of Public Roads, Dec. 1966.
11. Finney, D. J. Probit Analysis. Cambridge Univ. Press, 1962.
12. Dart, Olin K., Jr. Development of Factual Warrants for Left Turn Channelization Through Digital Computer Simulation. Ph. D. dissertation, Texas A&M University, Jan. 1966.

Optimization of Traffic Flow Splits

WILLIAM C. TAYLOR, Wayne State University

The objective of this study is the development of a method for predicting changes in a specified measure of the level of service provided by a street system under various system controls. The procedure consists of the development of a maximization model, an analysis of model sensitivity, and the comparison of this model output to the output from models presently in use.

The model is an application of the technique of maximization under constraints. The objective function selected for maximization is the number of vehicle-miles per hour produced within the designated street system. The variables studied include street capacities, flow characteristics, one-way streets, freeway capacities, and ramp capabilities.

The model is used to illustrate the effectiveness of various possible control systems when evaluated by increases in the objective function. The results of this analysis indicated that increasing the capacity of a freeway has the greatest effect on the total system output.

Although the development of the technique for measuring system output is the primary objective of this study, the application of these results in the decision-making process is the ultimate goal. There are two major areas where use of this model type would add significantly to the field of traffic assignment and control. First, it provides an objective measure for evaluating control methods, and second, it provides a method for determining the system effect of changes in an individual element.

The results of the model were used to illustrate the decision-making process for three selected criteria: (a) a minimum acceptable average velocity; (b) a maximum acceptable number of stored vehicles; and (c) a maximum output per dollar expenditure.

•THE street system in major metropolitan areas provides one means by which the transport of people and goods takes place. At the present time, it is the primary system for moving people in all cities and the exclusive system in many. The extent to which this system can be made to operate in an efficient manner will affect the entire well-being of the city.

The rapid growth in the number and use of automobiles led to the introduction of a new element in the street system. This element, the freeway or expressway, is designed for the exclusive purpose of moving vehicular traffic in a more efficient manner than is possible on the other elements of the street system. However, the addition of a freeway to an existing street network does not always result in the greatest possible increase in operating efficiency of the system. In fact, this study illustrates that unless certain constraints are placed on the driver's choice of alternative routes, the increased efficiency will be less than maximum.

True maximization of the efficiency of a system can be obtained only if the new element is integrated into the system, not superimposed on existing streets. The total

Paper sponsored by Committee on Quality of Traffic Service and presented at the 47th Annual Meeting.

concept of this integration into the street system involves many phases. These phases can be considered in the broad classifications of planning, design, construction, and operation. However, as is true in many complex systems, these categories are not mutually exclusive. Even though this study is directed toward the last of these categories—operation, the results are dependent on the interrelationships with the other categories.

The operation of a system of streets involves several phases:

1. A decision must be made as to the desired level of service to be provided by the system;
2. A measure of the extent to which this level of service is being accomplished must be specified;
3. A technique for evaluating this measure must be derived;
4. A method for predicting changes in these values which result from the application of alternate controls must be developed; and
5. The selection criteria to be used in the decision-making process must be specified.

This study is directed toward the fourth phase—the development of a method for predicting changes in the specified measure of the level of service provided by the application of alternate controls. The fulfillment of this phase cannot be completely isolated from the remaining phases. For this reason, values and relationships for the other phases are assumed where it is necessary for the development of the study.

The development of a technique for predicting changes depends on an objective measure by which quantitative measures can be evaluated and applied to an existing system of streets. The procedure used to accomplish the objective consists in the development of a mathematical model of traffic flow, an analysis of the effect of changes in various parameters of the model on this flow, and the comparison of the actual efficiency to the optimum efficiency for a particular example.

The literature was reviewed to determine the extent to which this approach had been explored and to learn where this study could make a significant contribution. The principal finding was that the analytic approach had not been applied to the system evaluation of the application of alternate control devices. Techniques using models have been developed to maximize the flow on single elements of a system (1), but this concept has not been extended to include the adjoining surface street system.

SCOPE

The technique developed in this study is based on a corridor analysis. The choice between alternative paths, whether free or forced, is limited to streets in close proximity to the freeway. The development of this method of evaluation encompasses:

1. The development of a theoretical model which yields universally applicable results; that is, the output of the model is in the form of parameters that are not unique to any one site, but are common to all traffic corridors.
2. The establishment of descriptive measures and tolerances on the necessary input for the model.
3. The testing of the model.

The study is limited to an analysis of the results of the model for one objective function only; however, the technique is applicable for other functions as well. Three applications of the results of the model in the decision-making process are described.

MODEL DEFINITION

The term "model" is defined as a representation of reality that attempts to order and describe some aspect of it. This broad definition of the term covers many areas which are not commonly thought of as models. The field of traffic engineering contains many examples of models. Route maps, intersection flow diagrams, accident spot maps, and capacity charts are all examples of models in common use. Each of these models is designed to illustrate either pictorially or mathematically, an event which has occurred or can occur. Although these models are not physical representations of real world situations, their meaning is quite clear.

The models vary from a physical representation at a different scale, as in the route maps, to a completely abstract mathematical formulation. The type of model chosen to depict an event is dependent on the aspect of the event selected. In this study, the form of the model was dictated by the desired output; the distribution of traffic on alternative routes in a corridor. The term "traffic assignment" is commonly used to describe this distribution process.

Traffic assignment models have been constructed with various criterion measures. Currently, the most widely accepted are those based on the minimum time assignment: each driver is assigned the route with the shortest travel time from his origin to his destination. These models have been summarized in a publication dealing with traffic engineering (2). The more sophisticated models include capacity restraints that reassign trips by an iterative process until each trip has been assigned to the route which minimizes the individual travel time. The total travel time on the systems is then expressed, Σt_{min}, where t_{min} is the minimum time required for each vehicle to reach its destination.

The minimization of total travel time in a corridor is not equivalent to the minimization of the total travel time for each vehicle in that corridor. The former is a system optimization; the latter is an example of summing optimized subsystems. The model developed in this study and the application illustrated will amplify this difference. Examples are solved to illustrate the difference in present traffic assignment models and $_{min}\Sigma t$ which is the objective function of the optimization model.

Figure 1 shows the logic used in the development of the optimization model. The equations and mathematical expressions are for the objective function chosen as an example. The logic would remain the same for other selected objective functions, but the equations would vary.

An intrinsic element in the development of this flow chart is a series of definitions. Throughout this paper, the following definitions are used in the analysis and description of the model.

Optimization model—The model assignment which results in the maximum total vehicle-miles per hour.

Equal time model—The model assignment which results in all trips from the origin to the destination being equal.

Corridor—The system of streets serving the origin and destination.

Element—Any section of the corridor between points of access or egress.

Output or throughput—The number of vehicle-miles per hour produced within the corridor.

Demand—The number of vehicles per hour wanting access through the corridor; this is considered to be a steady-state condition.

Volume and Flow—The number of vehicles per hour using the element; this is also considered a steady-state condition.

Velocity—The average velocity through the element, i.e., element length divided by time to traverse the element.

An initial assumption is that of nonsegregation. It is assumed that increased output is equally desirable on all possible routes; that is, output is independent of the route or routes on which it is achieved. This type of decision is not so clear-cut in actual practice. Often, the efficiency of flow on the freeways is considered to be more important than the flow on the arterials. Likewise, there is often a choice between serving the suburban commuter and the central city resident within the corridor. This is not a simple question to resolve, but any resolution could be incorporated into the model without changing the basic model structure.

An additional assumption is that controls or additional facilities will be added at any time the demand for facilities exceeds the combined capacity of the elements. The flow chart (Fig. 1) illustrates this assumption in the final check $\sum_{i=1}^{N} f_i = F$. This assumption

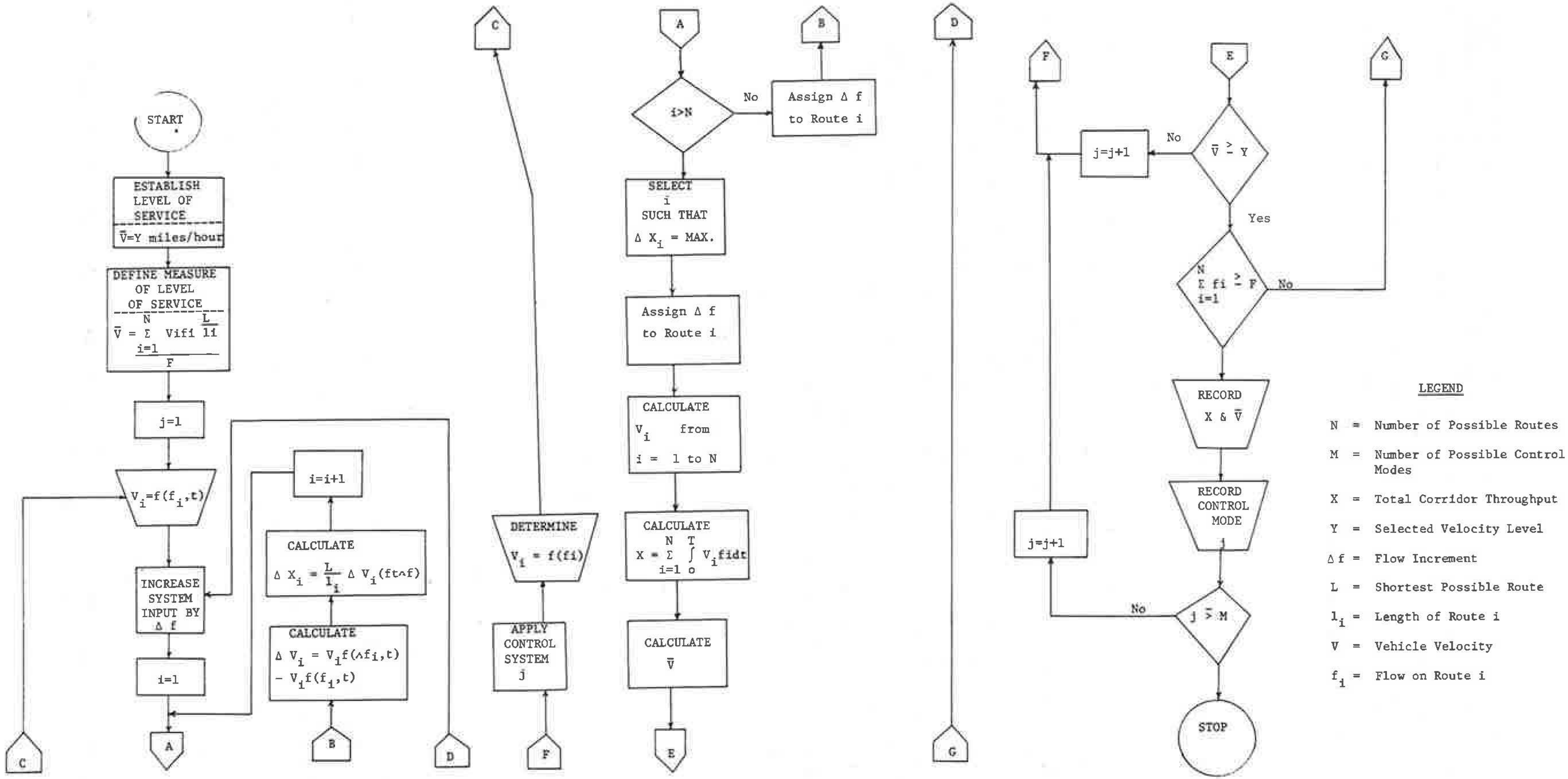


Figure 1. Flow chart for optimization model.

was made solely as an expedient to model application. The development of the model would be identical for the situation in which the demand was allowed to exceed the capacity, the only difference being that a velocity-flow relationship for each of the elements would have to be specified for this condition.

The previous two assumptions are basic to the philosophy of the model. In addition, the following assumptions are made in the examples used for illustrative purposes:

Each element of the system is of equal length.

$$\frac{L}{l_i} = 1 \text{ for } 1 \leq i \leq N$$

Demand flow is constant over the study period.

$$F(t) = K \text{ for } 0 \leq t \leq T$$

There is one location at each decision point where traffic can be switched to either the freeway or the arterial route without additional delay.

The unit of measure is the ability of a given system of alternative streets to perform the function of carrying a given number of trips from origin to destination. The most efficient volume split is the one which accomplishes this function with the minimum total travel time. The total travel time is defined as the composite individual travel times summed over the total corridor demand.

A dimension for expressing this unit is in terms of vehicle-miles per hour. For a given demand on a given system, the split which produces the greatest number of vehicle-miles per unit of time weighted to include route length is the same split that will produce the minimum total travel time. This weighting of the route lengths is accomplished by including a $\frac{L}{l_i}$ factor (Fig. 1). This unit is more readily expressed in terms of measurable variables, and thus, was chosen for the model.

MODEL MATHEMATICS

If the number of vehicle-miles produced on street i is denoted X_i, the necessary equations to express the desired model output can be written:

$$X_i = \int_0^T v_i(f_i, t) f_i(t) d_t \qquad \text{where} \quad 0 \leq f_i \leq F; \; 0 \leq v_i$$

$$\sum_{i=1}^{N} \frac{L}{l} X_i \text{ to be maximized}$$

The terms of the equations are defined in the following way:

X_i = vehicle-miles per time unit on street i;
$v_i(f_i, t)$ = the travel speed on street i as a function of the flow (f_i) and time (t);
$f_i(t)$ = the flow of traffic on street i expressed as a function of t;
T = total time under study;
N = the total number of alternative streets through the corridor;
F = total corridor demand;
L = shortest distance from the origin to the destination; and
l_i = distance from the origin to the destination on route i.

The maximization is subject to the constraint that $F = \int_{i=1}^{N} f_i(t)dt$, be equal to the total demand on the system. This particular model type is one which lends itself to solution through the technique of maximization under constraints.

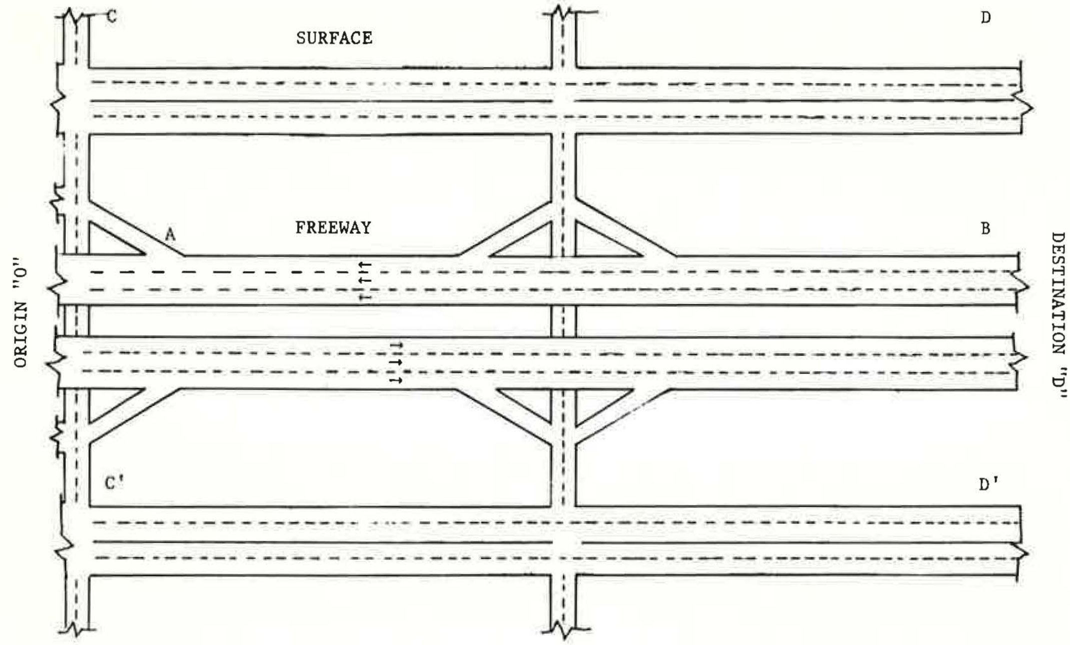


Figure 2. Schematic diagram of the corridor.

A complete description of the mathematics involved in the solution of this problem type can be found in many textbooks in the field of mathematics and operations research (3).

The degree of difficulty encountered in the optimization of a model of this type depends on three factors: (a) the form of the term $v_i(f_i, t)$; (b) the number of alternative paths; and (c) the number of constraints imposed on the solution.

A schematic diagram of a simple corridor is shown in Figure 2 with the origin referred to as O and the destination as D. The two alternative paths are labeled A-B and C-D. Route A-B is a four-lane freeway with a maximum speed of 60 mph and a velocity-flow relationship described by the curve in Figure 3. Routes C-D are 4-lane urban arterials with a maximum speed of 40 mph and the hypothetical relationship between speed and flow shown in Figure 3.

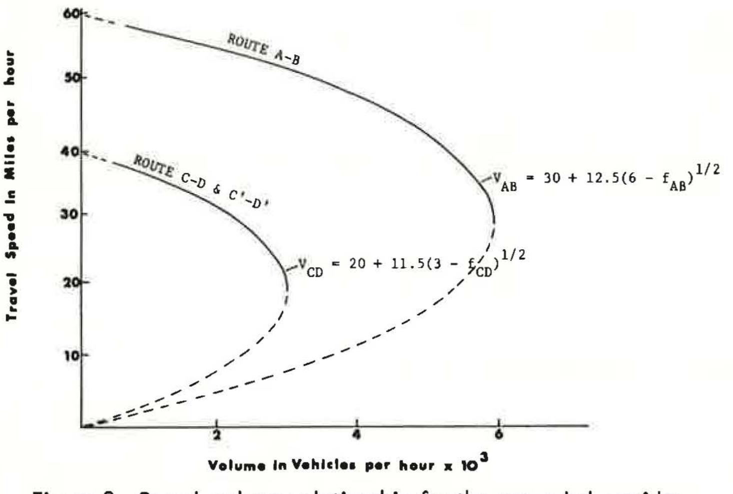


Figure 3. Speed-volume relationship for the expanded corridor.

The technique for determining the volume split between A-B and A'B' which results in the maximum vehicle-miles of travel per time unit is known as maximization under constraint. The equation as written previously, $X_i = \int_0^T v_i(f_i, t) f_i(t) d_t$, can now be written with the proper values of $v_i(f_i, t)$ substituted.

$$X_{AB} = \int_0^T \left[30 + 12.5(f_sAB - fAB)^{1/2}\right] fAB\,dt$$

$$X_{CD} = \int_0^+ \left[20 + 11.5(f_sCD - fCD)^{1/2}\right] fCD\,dt$$

The expressions $\left[30 + 12.5(f_sAB - fAB)^{1/2}\right]$ and $\left[20 + 11.5(f_sCD - fCD)^{1/2}\right]$ are the equations of the velocity-volume relationship (Fig. 3). The terms X_{AB} and X_{CD} are the respective throughput of vehicle-miles per time unit on the routes A-B and C-D. The total number of vehicle-miles produced within the corridor is $X_T = X_{AB} + 2X_{CD}$.

This is the objective function which is to be maximized. For the example, a value of 8000 vph is chosen as the demand. This subjects the solution to the constraint that $f_{AB} + 2f_{CD}$ must equal 8000 vph.

The problem is solved in the manner described elsewhere (3).

$$X_T = \int_0^T \left[30 + 12.5(f_sAB - f_{AB})^{1/2}\right] f_{AB}dt + 2\int_0^T \left[20 + 11.5(f_sCD - f_{CD})^{1/2}\right] f_{CD}dt$$

for T = 1 hour

$$X_T = 30f_{AB} + 12.5f_{AB}(6 - f_{AB})^{1/2} + 40f_{CD} + 23f_{CD}(3 - f_{CD})^{1/2}$$

Rewritten with constraint added:

$$Q = T\left[30f_{AB} + 12.5f_{AB}(6 - f_{AB})^{1/2} + 40f_{CD} + 23f_{CD}(3 - f_{CD})^{1/2} + \lambda(f_{AB} + 2f_{CD} - 8)\right]$$

$$\frac{\partial Q}{\partial f_{AB}} = T\left[30 - \frac{12.5f_{AB}}{2}(6 - f_{AB})^{1/2} + (6 - f_{AB})^{1/2}12.5 + \lambda\right] = 0$$

$$\frac{\partial Q}{\partial f_{CD}} = T\left[40 - \frac{23f_{CD}}{2}(3 - f_{CD})^{1/2} + (3 - f_{CD})^{1/2}23 + 2\lambda\right] = 0$$

$$\frac{\partial Q}{\partial \lambda} = f_{AB} + 2f_{CD} - 8 = 0$$

Solving the three simultaneous equations:

$$f_{CD} = 4000 - \frac{f_{AB}}{2}$$

$$f_{AB} = 4470 \text{ vph} \qquad f_{CD} = 1765 \text{ vph}$$

These results indicate the maximum number of vehicle-miles per hour is obtained when 1765 vph use each arterial, and 4470 vph use the freeway. The total output of the system per hour is expressed as X.

$$X = 1\left[30 \times 4470 + 12.5 \times 4470\ (1.53)^{1/2}\right] + 40 \times 1765 + 23 \times 1765\ (1.235)^{1/2}$$

$$= 319,100 \text{ vehicle-miles per hour}$$

In a similar manner, the average speed of the vehicles on the two alternatives can be found by substituting the proper values in the equation for (V_i).

$$V_{AB} = 30 + 12.5\ (6 - f_{AB})^{1/2} = 45.5 \text{ mph}$$

$$V_{CD} = 20 + 11.5\ (3 - f_{CD})^{1/2} = 32.8 \text{ mph}$$

It is evident that the optimum split is not the one which equalizes speeds on the two routes. In fact, to obtain the highest efficiency a 13-mph speed differential should be maintained between the freeway and the arterial. The output of an equal speed solution can be determined by varying the preceding solution. To accomplish this, the two expressions for V_i must be set equal, and the resultant volume split determined.

$$V_{AB} = V_{CD}$$

$$30 + 12.5\ (6 - f_{AB})^{1/2} = 20 + 11.5\ (3 - f_{CD})^{1/2}$$

$$f_{AB} + 2f_{CD} = 8000$$

$$f_{AB} = 5760 \qquad f_{CD} = 1120$$

The resultant speed on each of the paths is then 35.8 mph. This would have the effect of reducing the speed on the freeway from 45.5 to 35.8 mph and raising the speed on the arterial from 32.8 to 35.8 mph. The total output of the system for this volume split can be determined by substituting into the equation for X.

$$X = 1[(30 + 12.5 \times 0.46)\ 5760 + (20 + 11.5 \times 1.37)\ 1120]$$

$$= 286,000 \text{ vehicle-miles per hour}$$

The actual difference between the two solutions for this example is slightly less than 12 percent. The example was solved for total demands of 6000, 7000, 9000 and 10,000,

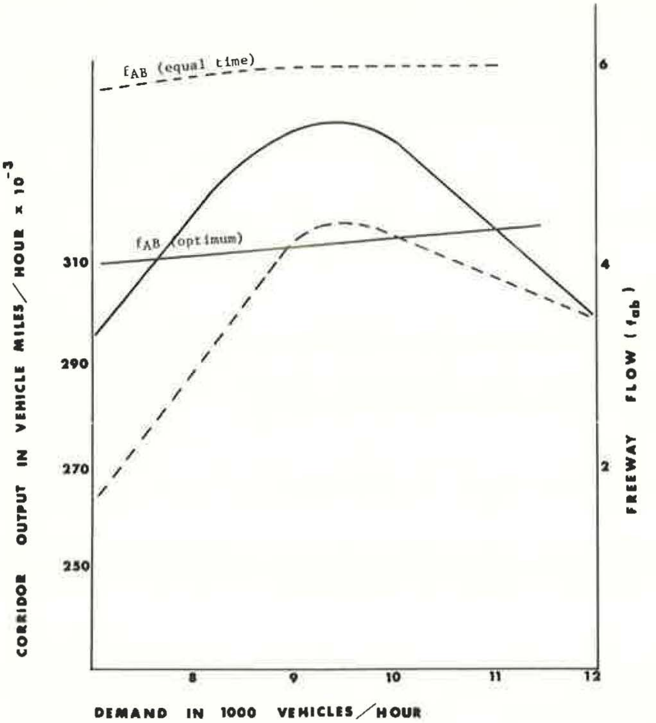


Figure 4. Comparison of optimum and equal time model solutions.

as well as for the case of 8000. By plotting the model output versus the total demand, the effect of traffic growth on the system can be shown. Figure 4 shows the results. The straight lines represent volumes on the alternate paths for the various demands, and the curved lines plot the total throughput in vehicle-miles per hour for the corridor. The shape of this graph corresponds to what would be reasonably expected. The output of the system is less than maximum for low volumes due to insufficient demand and is less than maximum for high volumes due to congestion. For this example the maximum output is reached when the demand is 9200 vph.

Figure 4 also illustrates the difference between this model output and the output obtained with the volume split which results in equal velocity on each route. The curved lines represent the corridor output measured in vehicle-miles per hour obtained at the various demand levels. The dashed line represents the output from the maximization model; the solid lines, the output from the equal time solution.

MODEL RESULTS

The preceding example illustrates the mathematics of the model solution and indicates the type of output which can be derived. The analysis and comparison of these outputs for different sets of operating conditions provides the information necessary for the determination of the most efficient operation of the corridor elements.

The selection of the mode of operation within the corridor can be based on a comparison of the output for two conditions: (a) the benefits to be derived from forced assignment (such as ramp metering) and (b) the benefits to be derived from improvements made to the arterial street system.

TABLE 1

FLOW AND VELOCITY CHARACTERISTICS OF STREETS A-B, C-D, AND C'-D' FOR OPTIMUM OPERATION AT VARIOUS DEMAND LEVELS

Demand (Q)	Flow (f)		Velocity (v)	
	A-B	C-D & C'-D'	A-B	C-D & C'-D'
6,000	3,880	1,060	50.0	35.2
7,000	3,960	1,520	47.8	34.0
8,000	4,050	1,980	47.5	31.6
9,000	4,120	2,440	47.1	28.6
10,000	4,200	2,900	46.8	23.6

TABLE 2

FLOW AND VELOCITY CHARACTERISTICS OF STREETS A-B, C-D, AND C'-D' FOR EQUAL TRAVEL TIME OPERATION AT VARIOUS DEMAND LEVELS

Demand (Q)	Flow (f)		Velocity (v)	
	A-B	C-D & C'-D'	A-B	C-D & C'-D'
6,000	5,500	250	39.0	39.0
7,000	5,670	665	37.0	37.0
8,000	5,800	1,100	35.8	35.8
9,000	5,860	1,570	34.7	34.7
10,000	5,900	2,050	31.1	31.1

Using the previously described technique, the values in Table 1 were determined. In all cases, the optimum flow results when the velocity on the freeway is kept higher than on the arterial streets. This means that some method must be found to prohibit drivers from selecting the fastest route if the optimum output is to be achieved.

The value of metering flow onto the freeway, and thus forcing the optimum distribution, can be determined by comparing these results with characteristics associated with an equal speed solution. Table 2 gives the values associated with the equal travel time solution.

The results of this model can be used to illustrate further the advantage of optimizing the output. The average velocity of vehicles passing through the corridor can be found by dividing the total vehicle-miles per hour by the total flow. When this is done, the optimum solution can be compared with the equal travel time solution and the percentage increase in velocity calculated (Table 3). Figure 4 illustrates this difference between the two solutions for various demand levels; the solid lines represent the optimum solution.

Table 3 indicates that the percent increase in efficiency becomes lower as the system approaches capacity. This has to be true since there can be no diversion of traffic if all elements of the system are operating in the congested range.

ARTERIAL STREET IMPROVEMENTS

The second category of model applicability is the determination of the effect of arterial street improvements on corridor flow. If the traffic flow is assigned to the various elements of the system in the optimum manner, then it is apparent that a change in the operational characteristics of any element of the system will change this assignment. The optimum output for various operational changes can be determined by the model.

The Wisconsin Avenue Study (4) illustrates the range of possible changes in the operation of a surface arterial street. The results of this study, and the information available in the Highway Capacity Manual (5) can be used to determine the limits of

TABLE 3

INCREASE IN THE AVERAGE VELOCITY ACHIEVED BY OPTIMUM OPERATION

Demand (Q)	Avg. Velocity Optimum	Avg. Velocity Equal Time	Percent Increase in Velocity
6,000	44.8	39.0	14.9
7,000	41.8	37.0	13.0
8,000	39.7	35.8	10.9
9,000	37.1	34.7	6.9
10,000	33.3	31.1	6.7
11,000	30.2	29.1	3.8

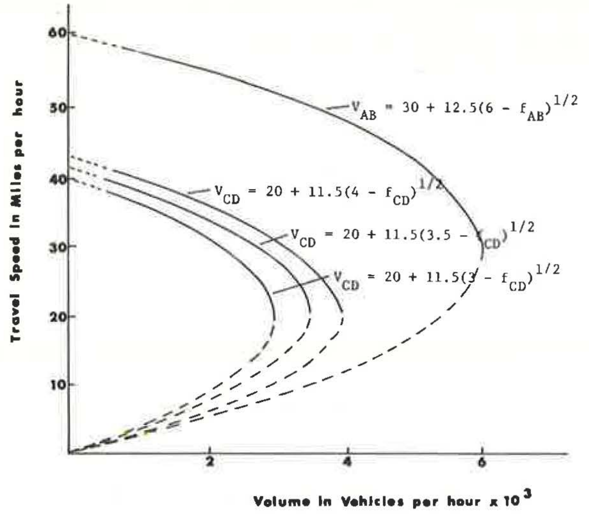


Figure 5. Speed-volume relationships for various capacity levels.

possible flow improvements on arterials. The model can then be used to determine the required traffic flow changes to optimize the operation of the corridor.

The same corridor used in the previous example is used to illustrate the effect of increased capacity. The schematic diagram of this corridor is shown in Figure 2. The assumed volume-speed relationship used follows the same form as in Figure 3 and used in the preceding analysis, but the capacity of the surface street is increased.

The equations describing the curves are shown in Figure 5. The difference between the three curves describe the flow on the surface streets associated with an increase in the street capacity from the previous value of 3000 vph to 3500 and 4000 vph. This represents about a 16 percent increase in the capacity for the first case, and a 33 percent increase for the second. These increases are high, but the Wisconsin Avenue Study has shown that capacity increases of this order can be effected by street improvements. Interpolation of the results will provide an indication of the benefits to be derived from smaller changes in the capacity of the arterials.

Figure 6 plots the optimum corridor output versus volume for each of three capacities. Two conclusions can be drawn by comparing the three curves and the nonoptimized equal-time output curve. First, maximum hourly output of the corridor increases substantially with an increase in the capacity of the arterials. The increase from a capacity of 12,000 to 14,000 vph results in an increase in the maximum output from 352,000

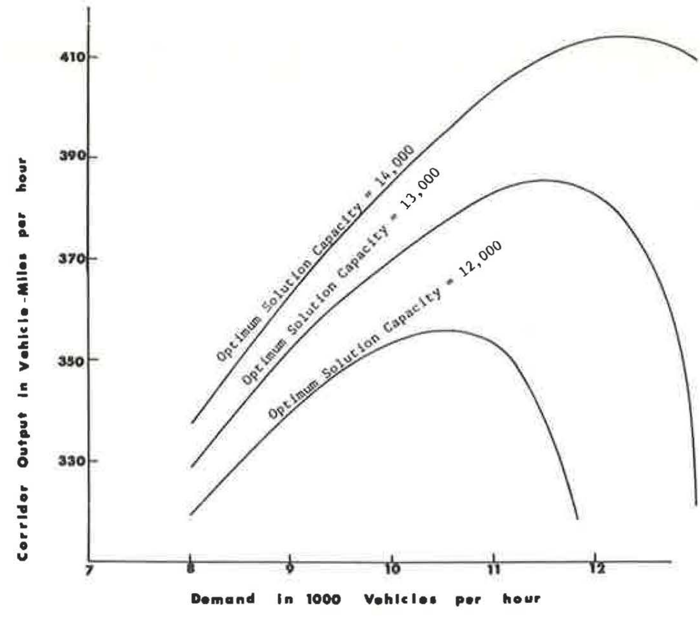


Figure 6. Changes in corridor output with increased capacity.

to 412,000 vehicle-miles per hour. This maximum value occurs at a volume of 10,500 vph for the lower volume, and at 12,500 for the higher volume.

Second, this example illustrates that it is probably better to optimize flow by ramp metering than to increase capacity for demands less than those associated with maximum output. The difference between the equal-time curve and the lower optimum curve is greater than the difference achieved by increasing the capacity. The cost of achieving each result would have to be determined before either alternative was chosen, but the effect on model output of optimization is greater than that of increasing capacity until congestion is reached.

Application of Model Results

Although the development of a technique for measuring differences in system output is the primary objective of this study, the application of these results in the decision-making process is the ultimate goal. The range of possible uses of any quantitative measure in decision-making is very broad and subject to the individual assuming the responsibility of making the decision.

The application of these results to determine the preferred operating policy for a specified decision paradigm are illustrated in the following examples. These examples are presented for three specified measures of the level of service: (a) a minimum acceptable average velocity; (b) a maximum acceptable difference between the demand input and the output; and (c) the maximum output per dollar expenditure. The same base corridor is used.

Assumed values are used for the rate of traffic growth and for the cost of improvements. Since the purpose of this example is to illustrate a technique as opposed to conducting a real corridor analysis, the use of assumed values is not felt to be critical. To be sure, the generalization of these results would depend on an accurate determination of these values.

The first criterion measure for an acceptable level of service is a minimum acceptable average velocity in the corridor. The value selected as acceptable is 36.0 mph. The figures in Table 4 would allow any velocity to be chosen without changing the approach:

The assumptions used in the development of Table 4 are (a) incremental increases in rush-hour volumes of 1000 vph; (b) capacity can be increased in increments of 500 vph on the surface streets; (c) volume-speed relationships given previously are valid; and (d) changing from two-way operation to one-way operation does not change the demand.

Table 4 illustrates alternative solutions to adding a new facility. The present system, operating without any changes, is adequate for the first three time units. At that time, some change is required. If the capacity is increased to 13,000 vph, the system will operate above the minimum speed for one additional time unit. If this alternative is selected, then another change will be required after the fourth time unit. Once again, the system capacity could be increased, this time to 14,000, and one additional unit of satisfactory operation obtained. At the end of the fifth unit, some other change in the operation would be required. The only possible change at this time would be to optimize the system to provide another unit of acceptable service. This alternative solution requires a change in operation each time unit from the third to the sixth to maintain an acceptable level of service. Another alternative is to change to a one-way operation in the third unit, and optimize this system for the fifth unit. This requires fewer changes, but it would require additional facilities one unit sooner than was required by the optimization of a two-way operation.

There are various combinations of improvements which would satisfy the criterion selected. For example, a new facility could be added at the first unit the present system was unacceptable. If this were done, then no further changes would be required. The cost of making these improvements and the ease with which they can be accomplished, must be considered in selecting the schedule for improvements.

The second criterion used to describe an acceptable level of service is a maximum acceptable difference between input and output volumes. This is a method of describing

TABLE 4

SCHEDULE OF IMPROVEMENTS TO MAINTAIN MINIMUM ACCEPTABLE VELOCITY

Time Period	Demand Volume	Average Velocity For:									
		Equal Time C = 12,000	Equal Time C = 13,000	Equal Time C = 14,000	Optimized C = 12,000	Optimized C = 13,000	Optimized C = 14,000	One-Way Equal Time C = 11,000	One-Way Equal Time C = 11,500	One-Way Optimum C = 11,000	One-Way Optimum C = 11,500
1	6,000	39	40	42	45	45	45	42	43	45	40
2	7,000	37	39	40	42	43	44	41	42	43	44
3	8,000	36	38	39	40	41	42	39	40	40	41
4	9,000	35	36	37	37	39	40	37	38	38	39
5	10,000	31	34	36	33	37	38	33	35	34	36
6	11,000	29	32	34	30	35	36	28	31	28	32
7	12,000	25	29	31	25	32	34	–	–	–	–
8	13,000	–	25	27	–	25	31	–	–	–	–
9	14,000	–	–	24	–	–	24	–	–	–	–

TABLE 5

SCHEDULE OF IMPROVEMENTS TO MAINTAIN MAXIMUM ACCEPTABLE DIFFERENCE BETWEEN INPUT AND OUTPUT

Time Period	Demand Volume	Input-Output in Vehicles Per Hour For:									
		Equal Time C = 12,000	Equal Time C = 13,000	Equal Time C = 14,000	Optimized C = 12,000	Optimized C = 13,000	Optimized C = 14,000	One-Way Equal Time C = 11,000	One-Way Equal Time C = 11,500	One-Way Optimum C = 11,000	One-Way Optimum C = 11,500
1	6,000	150	–	–	–	–	–	–	–	–	–
2	7,000	525	175	–	–	–	–	–	–	–	–
3	8,000	800	400	200	–	–	–	200	–	–	–
4	9,000	1125	900	675	675	225	–	675	450	450	225
5	10,000	2250	1500	1000	1750	750	500	1750	1250	1500	1000
6	11,000	3030	2200	1650	2750	1375	1100	3300	2480	3300	2200
7	12,000	4500	3300	2700	4500	2400	1800	–	–	–	–
8	13,000	–	4875	4225	–	4875	2920	–	–	–	–
9	14,000	–	–	5600	–	–	5600	–	–	–	–

the allowable storage in the system. The results of the analysis (Fig. 5) illustrate the need for considering this type of criterion. (See Table 5.)

The result is somewhat different from that with the previous criterion, due to the circumstance that a decrease in velocity at higher volume levels compounds itself when expressed in terms of stored vehicles. This is probably a more realistic criterion than the acceptable speed, since these stored vehicles also influence the operation of the cross streets in the system as well as the arterial streets. A decrease in speed is of less significance if it occurs when the demand is lower than it would be at a high speed. This difference is considered if the criterion measure is an allowable difference in the input rate and the output rate as determined in the model.

For example, an allowable difference of 2000 vph is selected. The present system is adequate for the first four time units of operation. At that time, the system operation must be changed by increasing capacity, optimizing the system, or by changing to one-way operation. If the decision is made to increase the capacity, the system will operate satisfactorily for an additional two units. Optimizing the system with the increased capacity provides one additional unit of operation.

The advantage of one-way operation is minimal for this criterion. The system will operate for only one extra unit, whereas the optimization of the system provides three units of acceptable service beyond the present one.

The third criterion measure is that of maximizing output per dollar and can be used as part of a total benefit-cost analysis. It provides the same type of information on which a yearly decision can be made, but has the additional benefit of providing a means of determining the best schedule for these improvements.

This analysis requires some cost figures, so the following values were assumed:

1. The cost of increasing the capacity of the arterial streets by 500 vph is 1.5 X dollars.
2. The cost of increasing the capacity of arterial streets by an additional 500 vph is X dollars.
3. The cost of ramp metering devices to optimize the system is X dollars.
4. The cost of converting to one-way operation is 0.75 X dollars.
5. The cost of reverting to two-way operation is 0.75 X dollars.

Table 6, which lists the increased output per dollar expended, provides a method of scheduling improvements. For this example, the minimum value, for which a change is considered, is $\frac{500}{X}$. The increased output is figured on the present system and the equal time model as a base.

Once again, there are various possible changes that could be made. However, the total benefits will be different for the different alternatives. Reading across the row of figures for any time unit, the highest value indicates the alternative which has the highest benefit-cost ratio. This alternative solution would then be maintained until the cumulative benefits of another alternative exceeds that of the one chosen.

The optimum schedule of improvements for this example would be: (a) operate the present system for the first unit; (b) change to an optimum system without an increase in capacity for units 2, 3, and 4; and (c) increase the capacity to 13,000 vph for units 5, 6, and 7. The total benefit for this program will be:

$$\frac{0}{X} + \frac{525}{X} + \frac{800}{X} + \frac{450}{X} + \frac{600}{X} + \frac{660}{X} + \frac{840}{X} \text{ or } \frac{3875}{X}$$

This program can be compared with the total benefit of a single solution, $\frac{3050}{X}$, which is realized by optimizing and increasing the capacity to 14,000 the first unit.

The results presented in Tables 4, 5, and 6 are calculated on assumed values for volume growth and improvement costs. It should be stressed that the results will vary considerably with changes in the assumed values. The values selected for these examples were merely to illustrate the application of the model results for various criterion measures.

TABLE 6

SCHEDULE OF IMPROVEMENTS TO MAINTAIN MAXIMUM OUTPUT PER DOLLAR EXPENDED

Time Period	Demand Volume	Increased Output Per Dollar Expended For:									
		Equal Time C = 12,000	Equal Time C = 13,000	Equal Time C = 14,000	Optimized C = 12,000	Optimized C = 13,000	Optimized C = 14,000	One-Way Equal Time C = 11,000	One-Way Equal Time C = 11,500	One-Way Optimum C = 11,000	One-Way Optimum C = 11,500
1	6,000	–	$\frac{100}{X}$	$\frac{60}{X}$	$\frac{150}{X}$	$\frac{60}{X}$	$\frac{45}{X}$	$\frac{200}{X}$	$\frac{65}{X}$	$\frac{85}{X}$	$\frac{45}{X}$
2	7,000	–	$\frac{330}{X}$	$\frac{270}{X}$	$\frac{675}{X}$	$\frac{270}{X}$	$\frac{195}{X}$	$\frac{900}{X}$	$\frac{300}{X}$	$\frac{385}{X}$	$\frac{210}{X}$
3	8,000	–	$\frac{600}{X}$	$\frac{510}{X}$	$\frac{1475}{X}$	$\frac{590}{X}$	$\frac{420}{X}$	$\frac{1700}{X}$	$\frac{655}{X}$	$\frac{845}{X}$	$\frac{455}{X}$
4	9,000	–	$\frac{750}{X}$	$\frac{690}{X}$	$\frac{1925}{X}$	$\frac{950}{X}$	$\frac{745}{X}$	$\frac{2300}{X}$	$\frac{955}{X}$	$\frac{1230}{X}$	$\frac{730}{X}$
5	10,000	–	$\frac{1250}{X}$	$\frac{1190}{X}$	$\frac{2425}{X}$	$\frac{1550}{X}$	$\frac{1245}{X}$	$\frac{2965}{X}$	$\frac{1400}{X}$	$\frac{1655}{X}$	$\frac{1115}{X}$
6	11,000	–	$\frac{1800}{X}$	$\frac{1740}{X}$	$\frac{2710}{X}$	$\frac{2210}{X}$	$\frac{1795}{X}$	$\frac{2965}{X}$	$\frac{1645}{X}$	$\frac{1655}{X}$	$\frac{1370}{X}$
7	12,000	–	$\frac{2600}{X}$	$\frac{2460}{X}$	$\frac{2710}{X}$	$\frac{3050}{X}$	$\frac{2565}{X}$	–	–	–	–
8	13,000	–	–	–	–	–	–	–	–	–	–
9	14,000	–	–	–	–	–	–	–	–	–	–

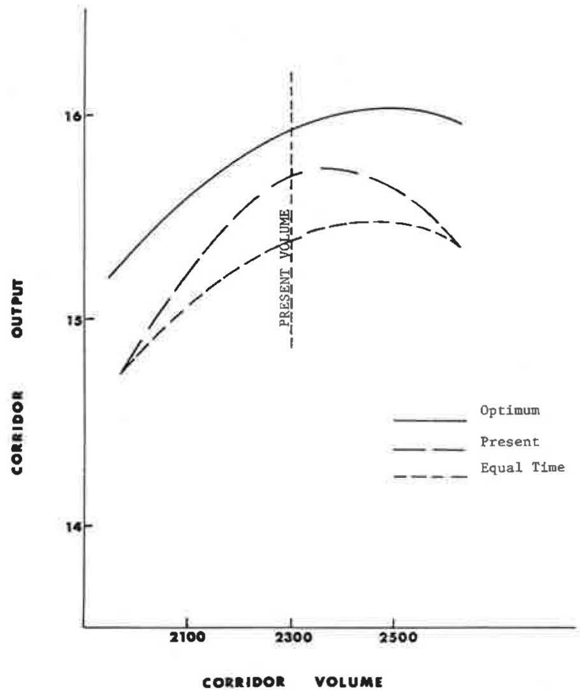


Figure 7. Corridor output for each of the three possible solutions.

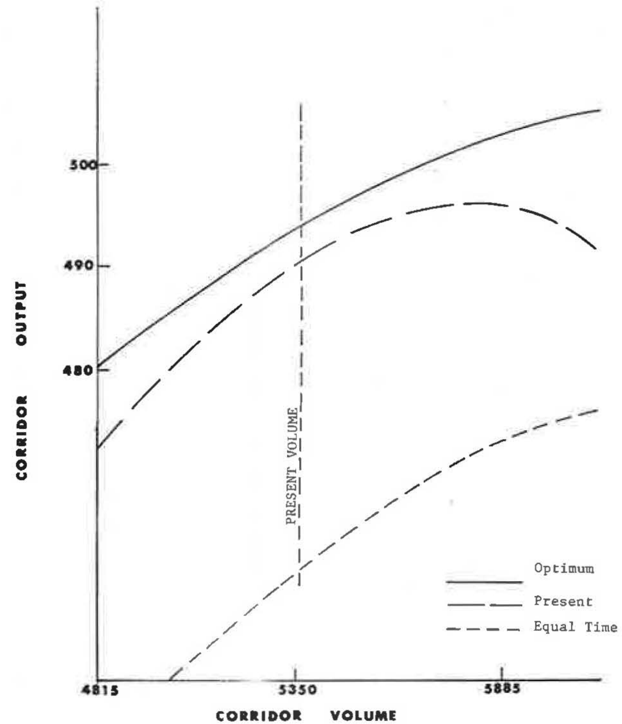


Figure 8. Corridor output for each of the three possible solutions.

Finally, the East Freeway corridor in Columbus, Ohio, is used to illustrate the complete process. The model is written to include the elements found in this corridor. The velocity-flow relationships are developed from aerial photographs, and the model inputs are taken from these same photographs and from traffic counts made by the city of Columbus. The model is solved to obtain the maximum output.

Results are compared with the output from an equal time model solution and with the output of the corridor operating on the driver's choice of routes. The data available for study permitted the solution of only two points. One was taken before the East Freeway was open, and the other was taken after the opening. Each point indicates that the driver's choice of routes is nearly the same as the optimization model requirement. The more commonly used equal time model was shown to be considerably inferior, when judged by the selected objective function. These results are shown in Figures 7 and 8.

The apparent shift in the driver's choice of routes through the corridor recorded between the before and after study indicates a further use for the model. The choice of routes selected by the driver is close to the optimum split for the present system demand in the East Freeway corridor. The fact that the optimum solution describes the present split more closely than the equal time solution indicates that the model should be used for traffic assignment. This may be due to the fact that inequalities brought about by low volumes assigned to certain elements by the equal time solution are eliminated.

Since this model does not require an iterative process, but relies on a single analysis solution, the processing time should be considerably less than the equal time solution. Thus, two of the major obstacles in the use of the equal time model are eliminated by the model type developed in this study.

SIGNIFICANCE

There are three major areas in which the use of this model type would add significantly to the field of traffic assignment and control.

1. It provides an objective measure by which alternate street systems and methods of control can be compared;
2. It provides a means of determining the effect of changes in one element of a system on the operation of the entire system; and
3. It provides a better measure of driver's route selection than the equal time solution.

SUMMARY

The objective of this study was to develop a technique for measuring, in quantitative terms, the relative merits of various possible alternative decisions on the operation of a street system. The system was confined to a corridor with a defined number of elements. The technique developed is applicable to other system configurations, however.

An objective function was selected, and a mathematical model developed to optimize the dependent variable. The model is an application of the technique of maximization under constraint, with the objective function being the number of vehicle-miles per hour produced within the corridor. The measure of effectiveness is the increase in the number of vehicle-miles per hour compared to the equal time assignment model.

The dependent variable, subject to the constraints of continuity, is the traffic volume on each element of the system. The independent variables include the system configuration, the system operation, the velocity-flow relationship for the traffic stream on each element, and the input demand function.

The model was designed to select the set of volume splits that maximizes the corridor output. The effect on model output of variations in the independent variables was also investigated. The results of these analyses indicated that the system operation is the independent variable which changes the model output most significantly.

REFERENCES

1. Wattleworth, Joseph A., and Berry, Donald S. Peak-Period Control of a Freeway System—Some Theoretical Investigations. Highway Research Record 89, p. 1-24, 1965.
2. Campbell, E. Wilson. A Mechanical Method for Assigning Traffic to Expressways. HRB Bull. 130, p. 27-41, 1956.
3. Saaty. Mathematical Methods of Operations Research. McGraw-Hill, 1959.
4. Carter, Arthur A., Jr. Increasing the Traffic-Carrying Capability of Urban Arterial Streets: The Wisconsin Avenue Study. U.S. Government Printing Office, May, 1962.
5. Highway Capacity Manual. HRB Special Report 87, 1965.

* * *

6. Brokke, Glen E. Program for Assigning Traffic to a Highway Network. HRB Bull. 224, p. 89-97, 1959.
7. Claffey, Paul J. Time and Fuel Consumption for Highway User Benefit Studies. HRB Bull. 276, p. 20-34, 1960.
8. Dean, Thomas B. The Use of Travel Time as a Factor in Rating Urban Streets. Traffic Engineering, p. 15-18 and 24, Jan. 1960.
9. Drew, Donald R., and Keese, Charles J. Freeway Level of Service as Influenced by Volume and Capacity Characteristics. Highway Research Record 99, p. 1-38, 1965.
10. Garrison, William L., and Marble, Duane F. Analysis of Highway Networks: A Linear Programming Formulation. HRB Proc., Vol. 37, p. 1-14, 1958.
11. Greensberg, H. An Analysis of Traffic Flow. Operations Research, Vol. 7, 1959.
12. Greenshields, Bruce D. Quality of Traffic Transmission. HRB Proc., Vol. 34, p. 508-522, 1955.
13. Hall, Edward M., and George, Stephen, Jr. Travel Time—An Effective Measure of Congestion and Level of Service. HRB Proc., Vol. 38, p. 511-529, 1959.
14. Hansen, Walter G., et al. Evaluation of Gravity Model Trip Distribution Procedures. HRB Bull. 347, p. 67-76, 1962.
15. Heathington, Kenneth W., and Jones, Andrew D. An Analysis of Peak Period Freeway Volume Characteristics. Traffic Engineering, p. 34-38, May 1966.
16. Higgins, Herbert C., and Dunn, Robert E. Urban Highway Route Evaluation. HRB Proc., Vol. 31, p. 425-429, 1952.
17. Hooper, Curtis J. The Need for Further Research on Traffic Assignment. HRB Bull. 61, p. 66-70, 1952.
18. Institute of Traffic Engineers. Traffic Engineering Handbook, p. 700-701, 1965.
19. Kennedy, Norman, et al. Fundamentals of Traffic Engineering. Univ. of California, 1962.
20. Miller, David W., and Starr, Martin K. Executive Decisions and Operations Research. Prentice-Hall, 1960.
21. Rothrock, C. A. Urban Congestion Index Principles. HRB Bull. 86, p. 26-34, 1954.
22. Schneider, Morton. A Direct Approach to Traffic Assignment. Highway Research Record 6, p. 71-75, 1963.
23. Voorhees, Alan M., and Morris, Robert. Estimating and Forecasting Travel for Baltimore by Use of a Mathematical Mode. HRB Bull. 224, p. 105-114, 1959.