TRANSPORTATION RESEARCH RECORD

509

Formerly issued as Highway Research Record

Traffic Flow: Theory and Practice

4 reports prepared for the 53rd Annual Meeting of the Highway Research Board

subject areas

53 traffic control and operations
54 traffic flow

TRANSPORTATION RESEARCH BOARD

NATIONAL RESEARCH COUNCIL

Washington, D. C., 1974

NOTICE

These papers report research work of the authors that was done at institutions named by the authors. The papers were offered to the Transportation Research Board of the National Research Council for publication and are published here in the interest of the dissemination of information from research, one of the major functions of the Transportation Research Board.

Before publication, each paper was reviewed by members of the TRB committee named as its sponsor and accepted as objective, useful, and suitable for publication by the National Research Council. The members of the review committee were chosen for recognized scholarly competence and with due consideration for the balance of disciplines appropriate to the subject concerned.

Responsibility for the publication of these reports rests with the sponsoring committee. However, the opinions and conclusions expressed in the reports are those of the individual authors and not necessarily those of the sponsoring committee, the Transportation Research Board, or the National Research Council.

Each report is reviewed and processed according to the procedures established and monitored by the Report Review Committee of the National Academy of Sciences. Distribution of the report is approved by the President of the Academy upon satisfactory completion of the review process.

The National Research Council is the principal operating agency of the National Academy of Sciences and the National Academy of Engineering, serving government and other organizations. The Transportation Research Board evolved from the 54-year-old Highway Research Board. The TRB incorporates all former HRB activities but also performs additional functions under a broader scope involving all modes of transportation and the interactions of transportation with society.

Transportation Research Record 509
International Standard Book Number 0-309-02352-1
Library of Congress Catalog Card Number 74-25504
Price: $2.40

Transportation Research Board publications may be ordered directly from the Board. They are also obtainable on a regular basis through organizational or individual supporting membership in the Board; members or library subscribers are eligible for substantial discounts. For further information write to the Transportation Research Board, National Academy of Sciences, 2101 Constitution Avenue N.W., Washington, D.C. 20418.

CONTENTS

FOREWORD

This RECORD contains two papers that deal with signalized street networks and two that are reviews or syntheses of traffic flow studies and freeway simulation models. Researchers, flow theorists, and practicing traffic engineers will find the material useful in their attempts to facilitate the orderly movement of traffic.

Traffic engineers today have the tools to permit the frequent review of signal system timing plan effectiveness and to change timing plans equally as frequently if desired. The transition in timing plans can be disruptive in itself, however, producing what some refer to as a hiccup in the system. Lieberman and Wicks present a new algorithm designed to minimize the length of the transition period while still providing acceptable service along all signal approaches in the system.

Davies, Grecco, and Heathington describe a simulation model that reproduces in the laboratory the traffic flow on any moderate-sized signalized street network. Tests and validation procedures lead them to conclude that the model yields accurate and realistic flow simulation.

Fifteen freeway traffic simulation models were compared against a baseline of eight features regarded by Hsu and Munjal as desirable and independent of specific simulation purpose. A comparison table for the models is presented, along with the authors' recommendations for the desirable features of a future general-purpose simulation model.

In the final paper, Weiner summarizes results of study by several researchers into distributions of speed and other traffic characteristics. Examples are given to illustrate for the practitioner the underlying methodology and its application in real-life situations.

A RAPID SIGNAL TRANSITION ALGORITHM

Edward B. Lieberman and David Wicks, KLD Associates, Inc., Huntington, New York

An increasing number of traffic signal systems are being brought under computer control. Although the designs of these control systems vary in detail, nearly all can be described as a sequence of fixed-time signal patterns. One of the primary advantages of computer control is the ability to change signal patterns in response to variations in traffic conditions. Associated with each change is a transition period wherein the signal settings transform from one fixed-time pattern to the next. Experience has shown that these transition periods can have disruptive effects on traffic operations. Hence, the more frequent these signal-pattern changes are, the greater is the need for careful design of the signal settings during the transition period. This paper presents a new algorithm that is easily implementable in real time by a digital computer and that is designed to minimize the duration of the transition period and yet service traffic demand along all approaches to every intersection. The results of an evaluation effort that compared this logarithm with two others are given. The results indicate that this method compares favorably with respect to the others considered.

•TRAFFIC signal systems controlled by computers are becoming increasingly common. Computer control permits the changing of signal patterns to respond to changing traffic conditions. For each change in pattern there is a transition period during which the signal settings convert from one fixed-time pattern to another. Because these transition periods can have disruptive effects on traffic, careful design of the signal settings during the transition period is necessary.

In general, transition policies may be classified as (a) a smooth, staged transition that restricts the change in offset per cycle or (b) a procedure designed to minimize delay. The first is "smooth" from the viewpoint of the control system; such a policy could, however, produce poor offsets over the duration of the transition period at several network nodes. The second is intuitively appealing and could well produce excellent results; such procedures, however, imply certain assumptions in the delay-control model that may not be realistic.

The algorithm presented here is designed to minimize the time to complete the signal transition, subject to the condition that all traffic demands can be serviced during this period. As such, there are no assumptions embedded in the methodology; the objective is explicitly satisfied. There is some question, of course, as to whether such an objective provides good service both during transition and subsequently, relative to other candidate procedures.

As part of the effort to extend and apply the UTCS-1 simulation model (1), an activity was undertaken to evaluate the performance of three signal transition algorithms:

1. "Immediate" transition;
2. "Second-generation" policy transition; and
3. The subject Rapid Signal Transition (RAST).

The first of these is normally implemented by standard multi-dial controllers. Essentially, the signal dwells in green for the main street until the new offset is attained. The "second-generation" transition (2-GT) algorithm is designed to minimize the sum of offset changes that must be experienced at all signals in the network; the transition parameter that satisfies this objective is computed in the process (2). The RAST algorithm is described below.

RAST ALGORITHM

Objectives

The RAST algorithm is designed to satisfy the following objectives:

1. Transform the signal control system from one offset pattern to the next immediately; i.e., minimize the number and duration of intervening nonoptimal signal intervals.
2. Keep the interval sequence unchanged.
3. Provide intervals of sufficient duration to service demand during transition.
4. Minimize the duration of transition at the "worst" node.

Parameter Definitions

The following terms are used in describing the algorithm:

X = reference time displacement in seconds.

$a_i^{(k)}$ = elapsed time to the offset of interval k, node i, from the reference time R for the existing pattern. Note that

$$-\frac{C_1}{2} < a_i^{(k)} \leq \frac{C_1}{2}$$

where R is the start of transition and C_1 is the cycle length of the signal pattern being terminated at time R.

$A_i^{(k)}$ = offset of interval k of signal i with respect to reference time $\hat{R} = R + X$ for new signal pattern of cycle length C_2.

$p_i^{(k)}$ = duration of signal interval k at i.

$\bar{p}_i^{(k)}$ = minimum duration of signal interval k at node i.

$V_{ji}^{(k)}$ = critical lane traffic volume serviced by signal interval k at node i for approach j, expressed in vehicles per hour.

$\hat{C}_i$ = minimum signal cycle time possible during transition period at node i,

$$= \sum_{k=1}^{K_i} \bar{p}_i^{(k)}$$

K_i = total number of intervals for signal at node i.

$\hat{k}$ = "key" interval at beginning of transition (may be modified by algorithm).

k_o = interval used as reference for defining the signal offset for the new pattern.

The parameters a_i, A_i, p_i, X, R, and $\hat{R}$ are shown in Figure 1.

Preliminary Determinations

The necessary preliminary steps are as follows:

1. For each interval $\bar{k}$, determine

$$\bar{p}_i^{(\bar{k})} = \max_j \left\{ (p_i^{(\bar{k})})_{min},\ S_j + H_j \left(\frac{V_{ji}^{(\bar{k})} \cdot C_1}{3,600} - 1 \right) \right\}$$

where

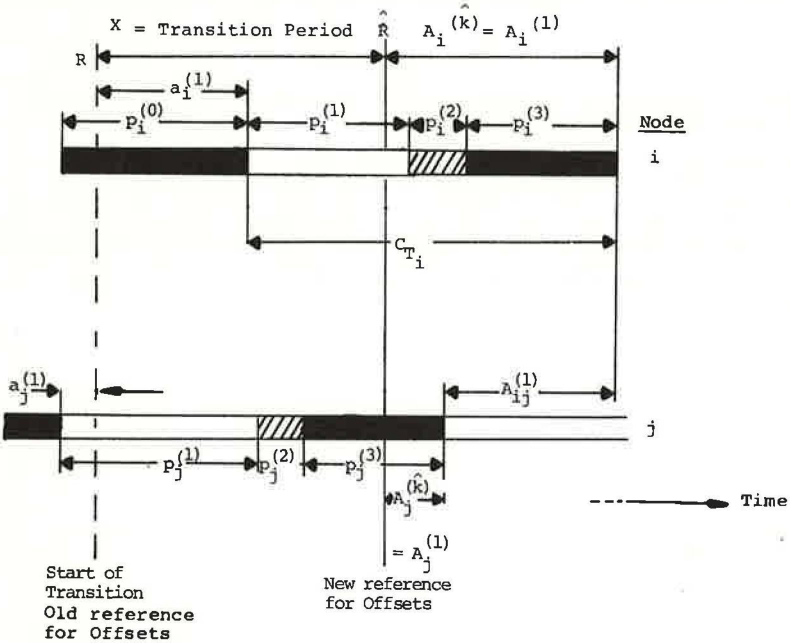
$(\bar{p}_i^{(\bar{k})})_{min}$ = minimum allowable green interval duration servicing the appropriate component of traffic volume, $V_{ji}^{(\bar{k})}$. This value is specified externally and dictated by practical considerations such as pedestrian crossing time.

S_j = start-up loss for queue on approach j.

H_j = mean queue discharge headway for approach j.

$\bar{k}$ = "major" intervals, i.e., that subset of the K_i intervals that are variable during the transition period. For example, amber, all-red intervals, and possibly intervals servicing turning traffic may be held fixed.

Figure 1. RAST parameter definition.



$A_{ij}^{(1)}$ new specified offset = $A_i^{(1)} - A_j^{(1)}$, for interval 1.

$\hat{k}$ =1 at both intersections

$\bar{p}_i^{(\bar{k})}$ = minimum duration of green interval $\bar{k}$ during transition period servicing traffic on the appropriate approaches.

2. At each node i, identify as the "key interval" that major interval $\bar{k}$ that is active at time R. If a minor (e.g., amber) interval is active, then the following major interval is identified as such. Denote this interval by the superscript $\hat{k}$. If $a_i^{(\hat{k})} < 0$ and $|a_i^{(\hat{k})}| > \bar{p}_i^{(\hat{k})}$, then select the following major interval as $\hat{k}$ instead, and recompute $a_i^{(\hat{k})}$. In Figure 1, $\hat{k} = 1$ at both intersections: $a_i^{(1)} > 0$, $a_j^{(1)} < 0$.

3. The offset $A_i^{(\hat{k})}$ of interval $\hat{k}$ for the new signal pattern at node i with respect to some common reference time $\hat{R}$ for the entire network is specified externally. This new reference time $\hat{R}$ is displaced from the current reference time R by a period X, which is a solution variable.

4. From Figure 1,

$$A_i^{(\hat{k})} = a_i^{(\hat{k})} + \sum_{\substack{k=\hat{k},\hat{k}+1,\ldots,K_i,\\ 1,2,\ldots,\hat{k}-1}} p_i^{(k)} - X$$

$$= a_i^{(\hat{k})} + C_{T_i} - X$$

where C_{T_i} is the (unknown) transition signal cycle at node i.

Necessary Conditions

The following conditions are to be met:

1. Satisfy all specified $A_i^{(\hat{k})}$ (new offsets).
2. Minimize transition time at the critical node (i.e., the node that takes the longest time to complete transition).
3. Determine all $p_i^{(k)} \geq \bar{p}_i^{(k)}$.

At each node i in the subnetwork the given parameters are $(p_i^{(k)})_{min}$; S_i; $V_{ji}^{(k)}$; H_i; $A_i^{(k_o)}$; $(p_i^{(k)})_{new\ pattern}$. For the subnetwork, C_1 = existing cycle length and C_2 = new cycle length.

Procedure

This algorithm, which is essentially an application of elementary game theory, is described in terms of a step-by-step procedure:

1. Determine $\bar{p}_i^{(k)}$ for all major signal intervals at each node i; then $\hat{C}_i$.
2. Identify interval $\hat{k}$ and its reference offset $a_i^{(\hat{k})}$ at each node i.
3. Calculate $\hat{X}_i = a_i^{(\hat{k})} + \hat{C}_i$ at each node i. This is the minimum possible value of X_i.
4. Determine the new reference offsets $A_i^{(\hat{k})}$ at each node with respect to the (unknown) reference time $\hat{R} = R + X$ consistent with the specified new offsets $A_i^{(k_o)}$ as follows: Initially, determine $A_i^{(\hat{k})}$,

$$A_i^{(\hat{k})} = (A_i^{(k_o)} + \emptyset_i) \bmod C_2$$

where

$$\emptyset_i = \begin{cases} 0 \text{ if } \hat{k} = k_o \\ \sum_{k=k_o}^{\hat{k}-1} (p_i^{(k)})_{new} \text{ if } \hat{k} \neq k_o \end{cases}$$

5. Scan all nodes, n = 1, 2, . . . , N, in sequence, where node n denotes an "anchor" node. For each n, compute

$$A_{in}^{(\hat{k})} = (A_i^{(\hat{k})} - A_n^{(\hat{k})}) \bmod C_2;\ i = 1, 2, \ldots, N$$

and then

$$\delta_{in} = A_{in}^{(\hat{k})} + \hat{X}_n \quad \text{for } i = 1, 2, \ldots, N$$

This value of δ_{in} is the time required for the signal at node i to attain the new synchronization if the transition period is set at $\hat{X}_n$, i.e., if the time to complete transition at the anchor node n is minimized. This value is subject to the constraint

$$\delta_{in} \geq \hat{X}_i$$

which, if satisfied, ensures that the constraints on minimum-interval durations for the signal at node i will be satisfied. If

$$\delta_{in} < \hat{X}_i$$

then the value of $\hat{X}_n$ must be revised (increased),

$$\hat{X}_n = (\hat{X}_n)_{old} + \hat{X}_i - \delta_{in},$$

and all previously calculated values of δ_{in} must be suitably revised prior to continuing the sweep over i.

6. Define $\Delta_n = \max_i (\delta_{in})$. This is the worst case (longest transition) in the network, if we minimize the time to complete transition at the anchor node n. The algorithm seeks the minimum value of Δ_n. Each node in sequence takes on the role of the anchor node, generating a new set of δ_{in} and another value of Δ_n. Then the critical anchor node $\hat{n}$ is located:

$$\Delta_{\hat{n}} = \min_n [\Delta_n] = \min_n \left\{ \max_i (\delta_{in}) \right\}$$

The transition period is $X = \hat{X}_{\hat{n}}$. It is seen that the minimum time of transition for the entire network is $\Delta_{\hat{n}}$ and the transition time at node $\hat{n}$ is the minimum possible, subject to constraints on minimum phase durations.

7. The excess time available at each signal i is

$$E_i = \delta_{i\hat{n}} - a_i^{(\hat{k})} - \hat{C}_i$$

This excess is allocated in proportion to requirements of the dominant approaches. Determine the excess of green time provided at each node to the dominant approach (i.e., the approach that services the higher volume in that direction) j serviced by interval $\bar{p}^{(\bar{k}_j)}$:

$$\Delta s_j = \bar{p}^{(\bar{k}_j)} - \left\{ S_j + H_j \left(\frac{V_{ji}^{(\bar{k}_j)} \cdot C_i}{3,600} - 1 \right) \right\}$$

As noted earlier, $\Delta s_j > 0$ only if $\bar{p}_i^{(\bar{k}_j)} = (p_i^{(\bar{k}_j)})_{min}$. Denote as $\bar{k}_1$ and $\bar{k}_2$ the two dominant (variable) intervals that service the demand in the respective directions, j_1 and j_2. (This analysis can be extended to consider additional phases.) The objective is to allocate the excess green time E_i between these intervals as follows:

$$\frac{\Delta s_{j_1} + e_{\bar{k}_1}}{V_{j_1}} = \frac{\Delta s_{j_2} + e_{\bar{k}_2}}{V_{j_2}}$$

$$e_{\bar{k}_1} + e_{\bar{k}_2} = E_i$$

Hence, the total excess green time above that required to satisfy demand is allocated so as to be proportional to the critical per-lane demands $V_{ji}^{(\bar{k}_j)}$ in the respective directions serviced by the signal intervals $\bar{k}_1$ and $\bar{k}_2$. Then,

$$e_{\bar{k}_1} = \frac{V_{j_1}}{V_{j_1} + V_{j_2}} \left\{ \Delta s_{j_2} + E_i - \frac{V_{j_2}}{V_{j_1}} \Delta s_{j_1} \right\}$$

(Note that $0 \leq e_{\bar{k}_1} \leq E_i$ must be asserted.) Then, $e_{\bar{k}_2} = E_i - e_{\bar{k}_1}$. Hence, $p^{(\bar{k}_1)} = \bar{p}^{(\bar{k}_1)} + e_{\bar{k}_1}$ and $p^{(\bar{k}_2)} = \bar{p}^{(\bar{k}_2)} + e_{\bar{k}_2}$.

8. Step 6 has yielded $X = \hat{X}_{\hat{n}}$; step 7, the interval durations during this transition period. Note that the determination of $\bar{p}_i^{(k)}$ and $\hat{C}_i$ in step 1 utilized C_i; $C_{T_i} = A_i^{(\hat{k})} + X - a_i^{(\hat{k})}$ should have been utilized, but X was unknown at that point. It may, therefore, be necessary to iterate for X. Note also that the node-specific cycle length during the transition period C_{T_i} varies from one node to another.

The offsets for the active intervals prior to the beginning of transition are "optimal" for the old pattern. Starting with interval $\hat{k}$, signal offsets depart from these optimal values and do not attain their new optimal values until the completion of transition at each node. Hence, minimizing the time to complete transition at the critical (worst) node should serve to restrict the duration of turbulence arising from these nonoptimal offsets. The interval durations during transition, however, do satisfy demand on a volume basis. An illustrative case showing all calculations is given in the Appendix.

EVALUATION OF TRANSITION ALGORITHMS

The logic representing all the transition algorithms was introduced into the UTCS-1 model and exercised to perform tests on two networks—the Washington, D.C., UTCS grid (Figure 2) and the Wisconsin Avenue arterial (Figure 3)—for two traffic conditions—morning peak to off-peak and off-peak to morning peak. The final results are shown in Figures 4 through 7. A summary of all results is given in Tables 1 and 2. Full details are provided elsewhere (2).

CONCLUSIONS

The impact of a signal transition policy on the effectiveness of a responsive control

Figure 2. Washington UTCS grid network.

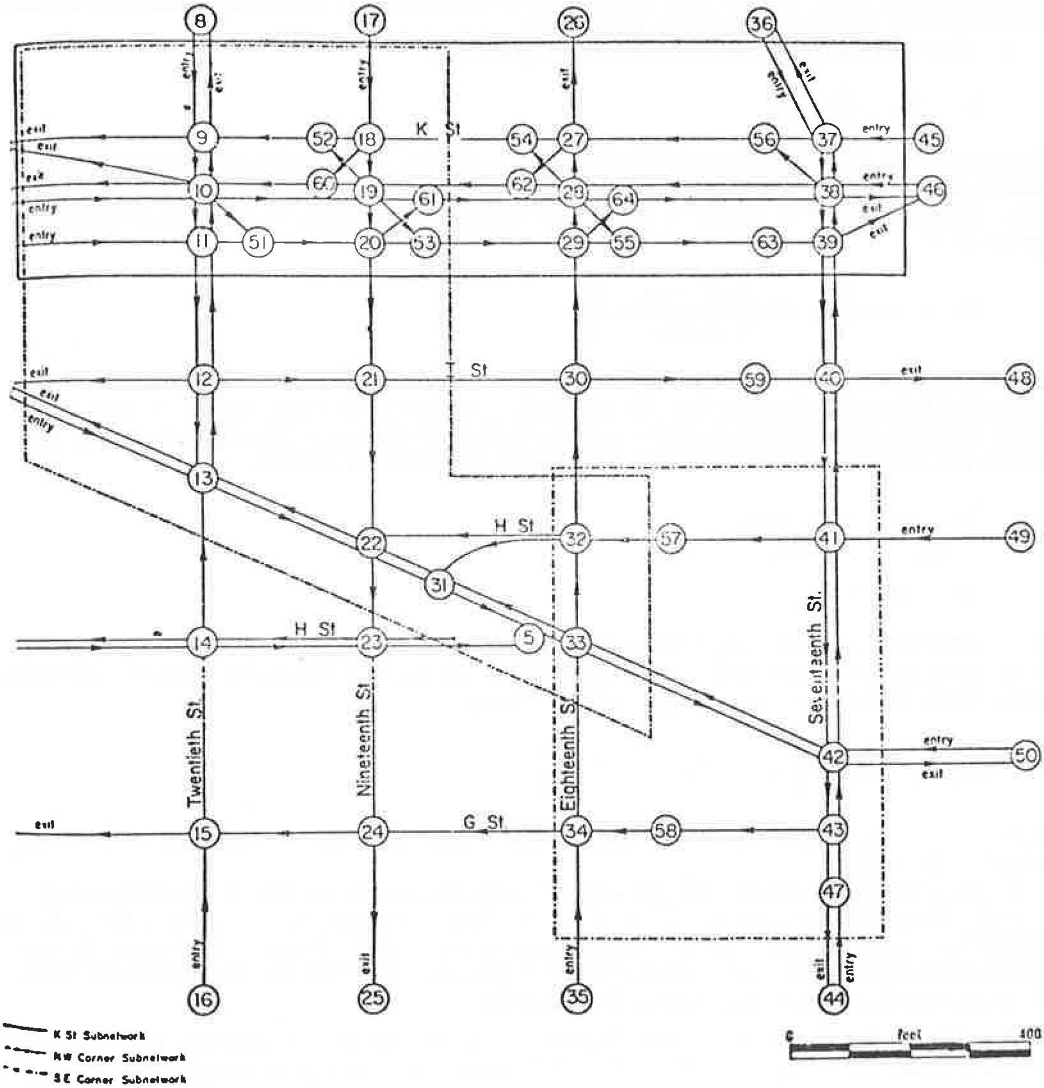


Table 1. Total delay, vehicle-minutes per hour.

Test	Network	Transition Algorithm: Immediate	2-GT	RAST
Peak to off-peak	Grid	12,733	11,427	11,127
Off-peak to peak	Grid	8,566	7,801	7,845
Peak to off-peak	Arterial	4,279	4,185	4,272
Off-peak to peak	Arterial	3,361	4,211	3,483
Total		28,939	27,624	26,727

Table 2. Percent reduction in delay relative to immediate transition.

Network	Transition Algorithm: 2-GT	RAST
Grid	+9.0	+10.9
Arterial	-9.9[a]	-1.5[a]
Overall	+4.5	+7.6

[a]Increase in delay.

Figure 3. Wisconsin Avenue transition test network.

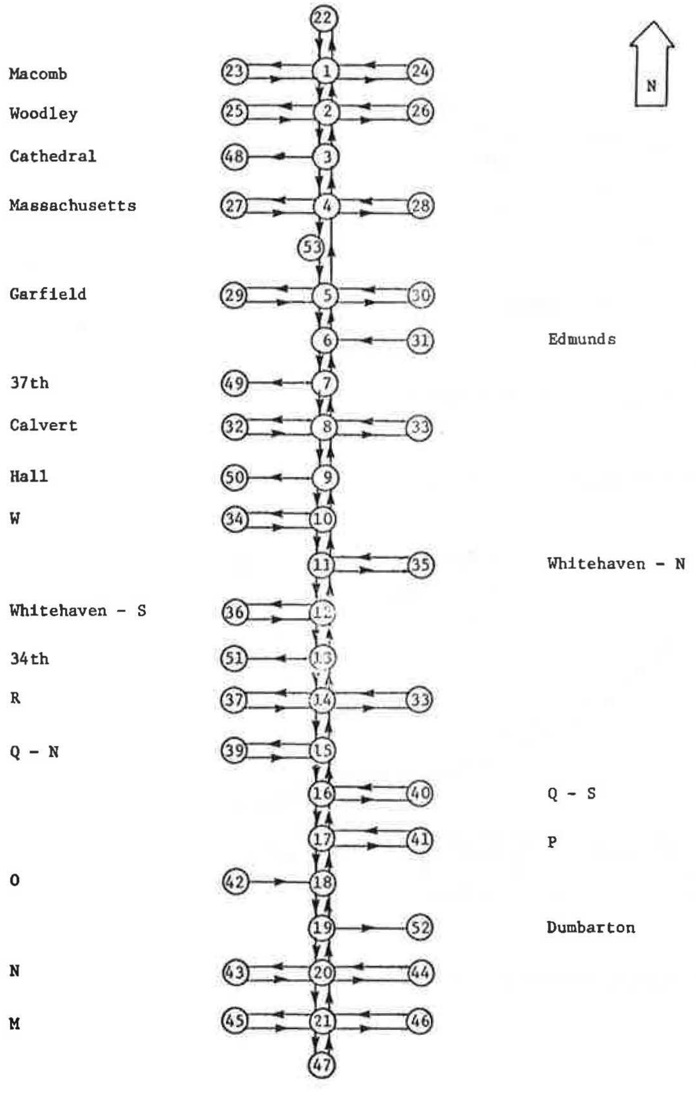


Figure 4. Delay per subinterval, UTCS-1 network, peak to off-peak.

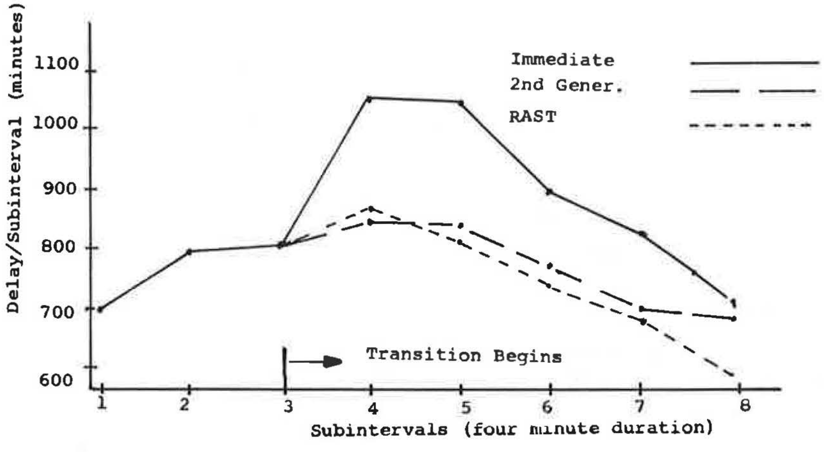


Figure 5. Delay per subinterval, UTCS-1 network, off-peak to peak.

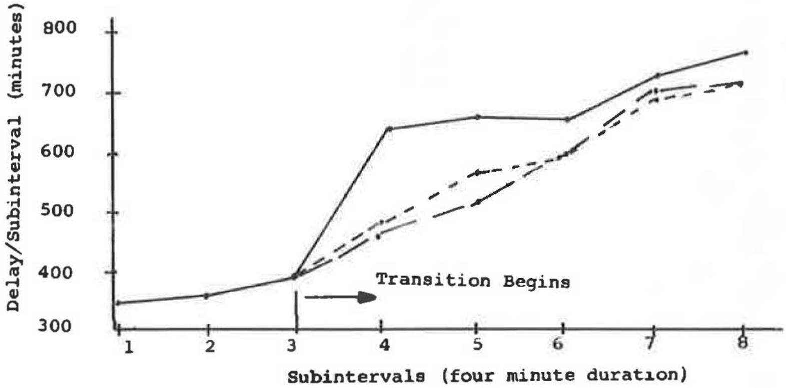


Figure 6. Delay per subinterval, Wisconsin Avenue, peak to off-peak.

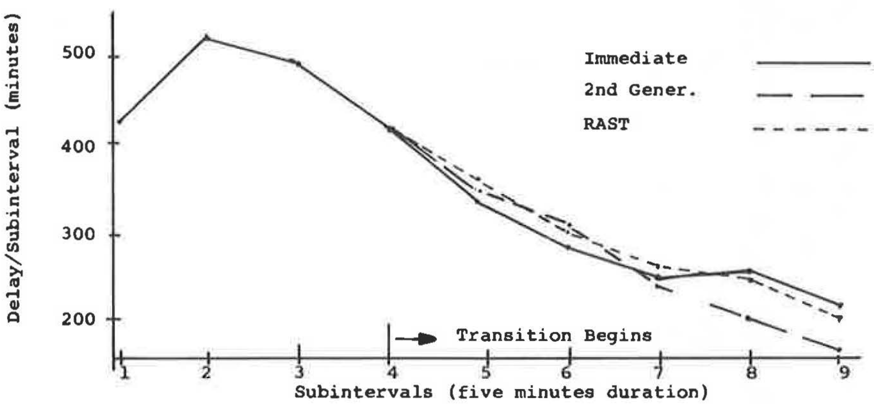
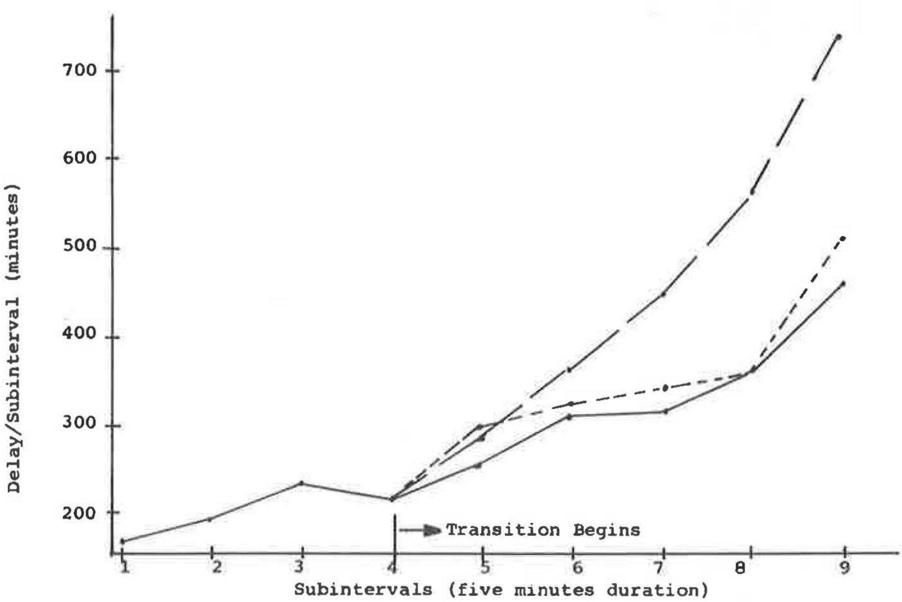


Figure 7. Delay per subinterval, Wisconsin Avenue, off-peak to peak.



system is comparable in importance to the signal optimization procedure for fixed-cycle control policies. In this study, both the 2-GT and RAST algorithms offer significant improvements in traffic operations relative to the "immediate" transition for grid networks (Tables 1 and 2). As should be expected for the Wisconsin Avenue study, the immediate transition policy, which dwells in green facing the arterial until the new offsets are realized, performed relatively well, while the RAST algorithm was strongly competitive. The sensitivity of traffic operations on grid networks to signal transition methodology, however, produced markedly different results, which emphasizes the need for careful treatment of this aspect of control, particularly for a system that changes signal patterns frequently. The RAST algorithm compares favorably with respect to the others considered. This study also demonstrates the utility of traffic simulation as a medium for conducting such evaluations within the framework of a controlled experiment.

ACKNOWLEDGMENTS

We wish to thank R. D. Worrall and R. Ferlis of Peat, Marwick, Mitchell and Co. for designing the evaluation tests and reducing the simulation results. Our thanks also go to G. Radelat of FHWA for his comments. This work was performed under subcontract to PMM under contract DOT-FH-11-7885.

REFERENCES

1. Worrall, R., Lieberman, E., and Bruggeman, J. Network Flow Simulation for Urban Traffic Control System. Report No. FH-11-7462-2, June 1971.
2. Worrall, R., Lieberman, E., and Ferlis, R. Network Flow Simulation for Urban Traffic Control System, Phase II. Report No. DOT-FH-11-7885-4, June 1973.

APPENDIX

ILLUSTRATIVE CASE

Given: $S_i = 4$; $\bar{H}_i = 2.4$ seconds/vehicle on all approaches at all nodes i; $(\bar{p}_i)^{(\bar{k})}_{min} = 15$ seconds for all i, $\bar{k}$; V_{ji} as follows:

	Approach No. j			
	Main	Minor	Main	Minor
i	1	2	3	4
1	333	300	467	250
2	367	300	433	200
3	400	300	433	325
4	367	300	400	250

Old interval lengths (seconds), with 4-second amber period prior to each red, are as follows:

$p_i^{(k)}$	Interval No. k				Main Street
	M.S.G.				Green Off-
Node No. i	1	2	3	4	sets (k_o = 1)
1	25	4	27	4	10
2	40	4	12	4	55
3	30	4	22	4	40
4	21	4	31	4	36

Figure 8. Old signal pattern facing main street.

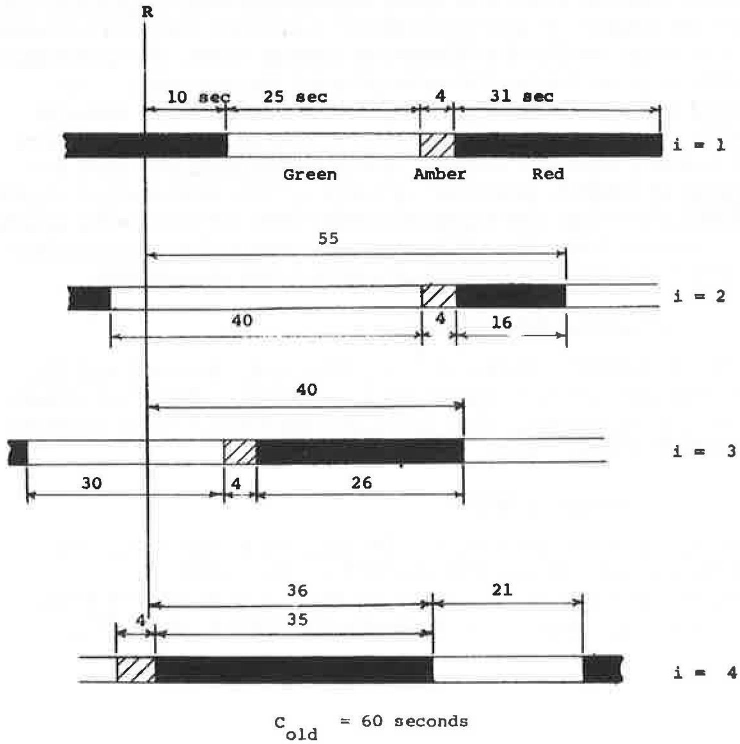
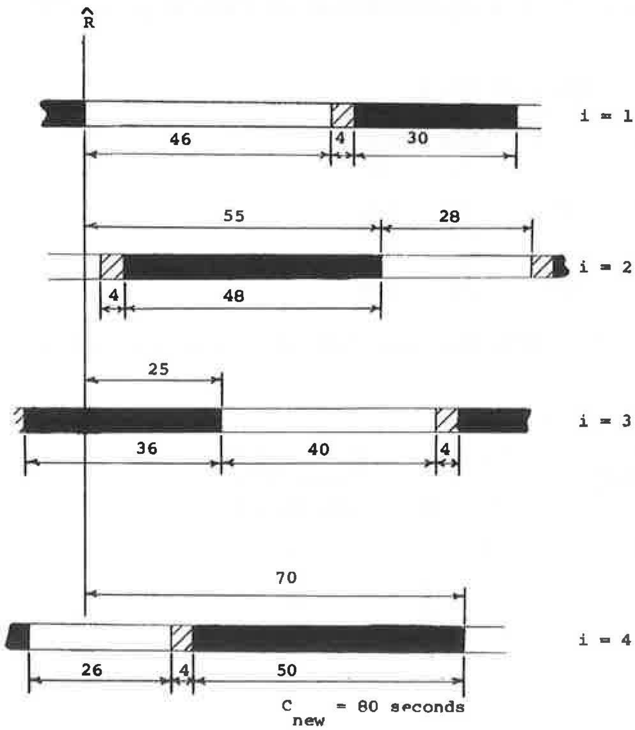


Figure 9. New signal pattern facing main street.



New interval lengths (seconds), with 4-second amber period prior to each red, are as follows:

$p_i^{(k)}$ Node No. i	Interval No. K M.S.G. 1	2	3	4	Main Street Green Offsets (k_o = 1)
1	46	4	26	4	0
2	28	4	44	4	55
3	40	4	32	4	25
4	26	4	46	4	70

Figure 8 is a schematic of the old signal pattern, and Figure 9 is a schematic of the new signal pattern. Note the disparate cycle lengths between the two patterns. Figure 10 shows an example transition diagram.

Procedure

1. $$\bar{p}_1^{(1)} = \max\left[15,\ 4 + 2.4\left(\frac{\overbrace{333}^{j=1} \cdot 60}{3,600} - 1\right),\ 4 + 2.4\left(\frac{\overbrace{467}^{j=3} \cdot 60}{3,600} - 1\right)\right]$$

$$= \max\,[15,\ 14.9,\ 20.3] = 20.3 \text{ seconds}$$

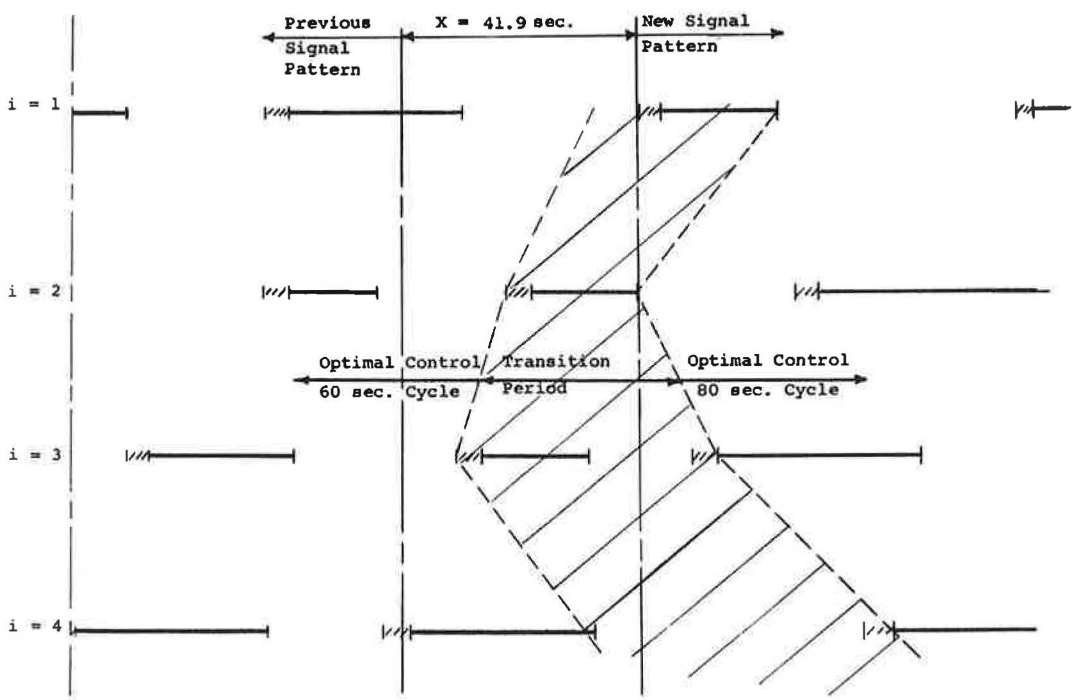
$\bar{p}_1^{(3)} = \max\,[15,\ 13.6] = 15.0$

$\bar{p}_2^{(1)} = 18.9;\ \bar{p}_2^{(3)} = 15$

$\bar{p}_3^{(1)} = 18.9;\ \bar{p}_3^{(3)} = 15$

$\bar{p}_4^{(1)} = 17.6;\ \bar{p}_4^{(3)} = 15$

Figure 10. Illustrative example transition diagram.



Then

$\hat{C}_1 = 20.3 + 4 + 15 + 4 = 43.3$ seconds
$\hat{C}_2 = 41.9$
$\hat{C}_3 = 41.9$
$\hat{C}_4 = 40.6$

2. At $i = 1, \hat{k} = 3; i = 2, \hat{k} = 1; i = 3, \hat{k} = 1; i = 4, \hat{k} = 3$ ($k = 1$: green; $k = 3$: red, as shown facing main street). Then $a_1^{(3)} = -21$; $a_2^{(1)} = -5$; $a_3^{(1)} = -20$; $a_4^{(3)} = 1$. Since $a_1^{(3)}$, $a_2^{(1)}$, and $a_3^{(1)}$ are all <0, we must test whether there is sufficient time remaining in these active intervals (3, 1, and 1 respectively) to permit shortening if necessary. Since $|a_1^{(3)}| > \bar{p}_1^{(3)}$, we must identify the other major interval as $\hat{k}$, i.e., define $\hat{k} = 1$ for node 1. Similarly, with $|a_3^{(1)}| > \bar{p}_3^{(1)}$, $\hat{k} = 3$ for node 3, while $\hat{k}$ for node 2 remains 1, since $|a_2^{(1)}| < \bar{p}_2^{(1)}$. Hence we have

i	$\hat{k}$	$A_i^{(1)}$	j_1	j_2	$a_i^{(\hat{k})}$
1	1	0	3	2	10
2	1	55	3	2	-5
3	3	25	3	4	14
4	3	70	3	2	1

where j_1 is the dominant main street approach and j_2 is the dominant minor street approach (values of j_1, j_2 assigned for this example). Note that

$$a_1^{(1)} = a_1^{(3)} + (p_1^{(1)} + p_1^{(2)})_{old} = -21 + 31 = 10;$$
$$a_3^{(3)} = -20 + 30 + 4 = 14$$

3. $\hat{X}_1 = a_1^{(\hat{k})} + \hat{C}_1 = 10 + 43.3 = 53.3$; $\hat{X}_2 = -5 + 41.9 = 36.9$

$\hat{X}_3 = 14 + 41.9 = 55.9$; $\hat{X}_4 = 1 + 40.6 = 41.6$.

4. $A_1^{(\hat{k})} = A_1^{(k_o)} = A_1^{(1)} = 0$

$A_2^{(\hat{k})} = A_2^{(k_o)} = A_2^{(1)} = 55$

$A_3^{(\hat{k})} = A_3^{(3)} = A_3^{(1)} + (p_3^{(1)} + p_3^{(2)})_{new} = 25 + 40 + 4 = 69$

$A_4^{(\hat{k})} = A_4^{(3)} = A_4^{(1)} + (p_4^{(1)} + p_4^{(2)})_{new} = 70 + 26 + 4$
$\mod (80) = 20$

5 and 6. At $n = 1$:

$A_{11}^{(1)} = A_1^{(1)} - A_1^{(1)} = 0$; $\delta_{11} = A_{11}^{(1)} + \hat{X}_1 = 53.3$

$A_{21}^{(1)} = A_2^{(1)} - A_1^{(1)} = 55 - 0 = 55$

$A_{21}^{(1)} = 55$; $\delta_{21} = 55 + 53.3 = 108.3 > 36.9$ ok

$A_{31}^{(3)} = 69$; $\delta_{31} = 69 + 53.3 = 122.3 > 55.9$ ok

$A_{41}^{(3)} = 20$; $\delta_{41} = 20 + 53.3 = 73.3 > 41.6$ ok

$\Delta_1 = \max [53.3, 108.3, 122{,}3, 73.3] = 122.3$

At $n = 2$:

$A_{12}^{(1)} = A_1^{(1)} - A_2^{(1)} = 0 - 55 = -55 \mod (80) = 25$

$\delta_{12} = A_{12}^{(1)} + X_2 = 25 + 36.9 = 61.9 > 53.3$ ok

$A_{22}^{(1)} = A_2^{(1)} - A_2^{(1)} = 0$; $\delta_{22} = 0 + 36.9 = 36.9 \geq 36.9$ ok

$A_{32}^{(3)} = A_3^{(3)} - A_2^{(1)} = 69 - 55 = 14$; $\delta_{32} = 14 + 36.9 =$
$50.9 < 55.9$ NG

Revised $\hat{X}_2 = \hat{X}_{2old} + \hat{X}_3 - \delta_{32old} = 36.9 + 55.9 - 50.9 = 41.9$. (This revision applied only to the analysis for n = 2.)

Return to beginning of calculation (n = 2) with this new value of $\hat{X}_2$ and repeat procedure:

$A_{12}^{(1)} = 25$; $\delta_{12} = 25 + 41.9 = 66.9 > 53.3$ ok

$A_{22}^{(1)} = 0$; $\delta_{22} = 0 + 41.9 = 41.9 \geq 41.9$ ok

$A_{32}^{(3)} = 14$; $\delta_{32} = 14 + 41.9 = 55.9 \geq 55.9$ ok

$A_{42}^{(3)} = 25 - 55 \bmod (80) = 45$;

$\delta_{42} = 45 + 41.9 = 86.9 > 41.6$ ok

$\Delta_2 = 86.9$

At n = 3:

$A_{13}^{(1)} = 11$; $\delta_{13} = 66.9 > 53.3$ ok

$A_{23}^{(1)} = 66$; $\delta_{23} = 121.9 > 36.9$ ok

$A_{33}^{(3)} = 0$; $\delta_{33} = 55.9 \geq 55.9$ ok

$A_{43}^{(3)} = 31$; $\delta_{43} = 86.9 > 41.6$ ok

$\Delta_3 = 121.9$

At n = 4:

$A_{14}^{(1)} = A_1^{(1)} - A_4^{(3)} = 0 - 20 \bmod (80) = 60$;

$\delta_{14} = 60 + 41.6 = 101.6 > 53.3$ ok

$A_{24}^{(1)} = 35$; $\delta_{24} = 76.6 > 36.9$ ok

$A_{34}^{(3)} = 49$; $\delta_{34} = 90.6 > 55.9$ ok

$A_{44}^{(3)} = 0$; $\delta_{44} = 41.6 \geq 41.6$ ok

$\Delta_4 = 101.6$

$\Delta_{\hat{n}} = \min [122.3, 86.9, 121.9, 101.6] = 86.9$; $\hat{n} = 2$

Then, with node 2 as the critical one, $X = \hat{X}_{\hat{n}} = 41.9$. Also, $A_1^{(1)} = A_{12}^{(1)} = 25$, $A_2^{(1)} = 0$, $A_3^{(3)} = 14$, $A_4^{(3)} = 45$.

7. $E_1 = 66.9 - 10 - 43.3 = 13.6$

$E_2 = 41.9 - (-5) - 41.9 = 5.0$

$E_3 = 55.9 - 14 - 41.9 = 0$

$E_4 = 86.9 - 1 - 40.6 = 45.3$

Node 1

$$\Delta s_3 = 20.3 - \left[4 + 2.4\left(\frac{467 \cdot 60}{3,600} - 1\right)\right] = 0$$

$$\Delta s_2 = 15 - \left[4 + 2.4\left(\frac{300 \cdot 60}{3,600} - 1\right)\right] = 1.4$$

Node 2

$\Delta s_3 = 0$

$$\Delta s_2 = 15 - \left[4 + 2.4\left(\frac{300 \cdot 60}{3{,}600} - 1\right)\right] = 1.4$$

Node 3

$$\Delta s_3 = 0$$

$$\Delta s_4 = 15 - \left[4 + 2.4\left(\frac{325 \cdot 60}{3{,}600} - 1\right)\right] = 0.4$$

Node 4

$$\Delta s_3 = 0$$

$$\Delta s_2 = 15 - \left[4 + 2.4\left(\frac{300 \cdot 60}{3{,}600} - 1\right)\right] = 1.4$$

The allocation of excess time to each of the major intervals is

Node 1

$$e_1 = \frac{467}{467 + 300} [1.4 - 0 + 13.6] = 9.1$$

$$e_3 = 13.6 - 9.1 = 4.5$$

$$p_1^{(1)} = \bar{p}_1^{(1)} + e_1 = 20.3 + 9.1 = 29.4$$

$$p_1^{(3)} = 15 + 4.5 = 19.5$$

Node 2

$$e_1 = \frac{433}{433 + 300} [1.4 - 0 + 5] = 3.8$$

Then $e_3 = 5 - 3.8 = 1.2$

and

$$p_2^{(1)} = \bar{p}_2^{(1)} + e_1 = 18.9 + 3.8 = 22.7 \text{ seconds}$$

$$p_2^{(3)} = \bar{p}_2^{(3)} + e_3 = 15 + 1.2 = 16.2 \text{ seconds}$$

Check: $X + A_{22}^{(1)} \stackrel{?}{=} a_2^{(1)} + \sum_{k=1}^{4} p_2^{(k)}$

$$41.9 + 0 \stackrel{?}{=} -5 + 22.7 + 4 + 16.2 + 4$$

$$41.9 = 41.9 \quad \text{ok}$$

Node 3

Since $E_3 = 0$, $p_3^{(1)} = \bar{p}_3^{(1)} = 18.9$

$$p_3^{(3)} = \bar{p}_3^{(3)} = 15$$

Node 4

$$e_1 = \frac{400}{400 + 300} [1.4 - 0 + 45.3] = 26.7$$

$$e_3 = 45.3 - 26.7 = 18.6$$

$$p_4^{(1)} = \bar{p}_4^{(1)} + e_1 = 17.6 + 26.7 = 44.3$$

$$p_4^{(3)} = \bar{p}_4^{(3)} + e_3 = 15 + 18.6 = 33.6$$

Transition cycle lengths

$$C_{T_1} = \sum_{k=1}^{4} p_1^{(k)} = 29.4 + 4 + 19.5 + 4 = 56.9 \text{ seconds}$$

$$C_{T_2} = \sum_{k=1}^{4} p_2^{(k)} = 22.7 + 4 + 16.2 + 4 = 46.9 \text{ seconds}$$

$$C_{T_3} = \sum_{k=1}^{4} p_3^{(k)} = 18.9 + 4 + 15 + 4 = 41.9 \text{ seconds}$$

$$C_{T_4} = \sum_{k=1}^{4} p_4^{(k)} = 44.3 + 4 + 33.6 + 4 = 85.9 \text{ seconds}$$

As indicated in step 8, we now obtain improved values of $\bar{p}_i^{(k)}$ based on C_{T_i} rather than C_{old}. Only node 4 need be considered, where $C_{T_4} > C_1$:

$$\bar{p}_4 = \max \left[15,\ 4 + 2.4\left(\frac{400 \cdot 85.9}{3{,}600} - 1\right)\right]$$

$$= 24.4 < 44.3 \quad \text{ok} \quad \text{No iteration is necessary.}$$

A GENERALIZED STREET NETWORK SIMULATION MODEL

G. W. Davies, Wilbur Smith and Associates, New York; and
W. L. Grecco and K. W. Heathington,
Department of Civil Engineering, University of Tennessee

This paper describes a microscopic simulation model that reproduces traffic flow on a signalized street network under laboratory conditions. The input format and structure of the program facilitate application to any moderate-sized network. The program is economical to use, achieving a 6.5-to-1 real-time to simulation-time ratio for an 85-link network. The model has undergone a testing and validation procedure in which simulated vehicular travel times have been compared with actual travel times recorded on the test network to evaluate overall model accuracy. An analysis of individual model segments has been conducted to test their sensitivity to changes in key parameters. Results of these tests indicate that the model accurately and realistically simulates traffic flow through the street network.

•COMPUTER simulation is a tool that has been used effectively in business and scientific fields to gain more understanding of complex processes and to facilitate decision-making. Its application to traffic studies is especially appropriate because it allows a stochastic process, traffic flow, to be studied under controlled laboratory conditions.

To be useful, traffic simulation must satisfy three basic considerations:

1. The results of the simulation must fit the facts. Observations obtained as a result of simulation must agree with similar results obtained from observations of actual traffic flow.
2. The time required to simulate a problem must be reasonable. The ratio of simulated time to real time must be such that computer simulation of a street network is economically feasible. Before embarking on a study, its objectives should be thoroughly reviewed and alternative techniques for meeting them compared. The technique that does the job effectively at the least cost should then be chosen. When viewed from this standpoint, the use of simulation as a study tool can become much more attractive economically.
3. The results of simulation must be accessible in a format that is meaningful to those using them. The actual simulation takes place within the computer and is, of course, unobservable to the user (in the absence of some type of on-line visual display device). Thus it is necessary to devise some means of displaying simulation results in a form convenient to the user.

This paper describes a computer simulation program that was developed for use in the analysis of a signalized street network. The model is general so that it can be applied to any moderate-sized network, and the inputs and outputs can be understood by a traffic engineer not oriented toward computers.

MATHEMATICS OF VEHICLE BEHAVIOR

There are various methods that may be used to represent the flow of traffic within the computer. Early traffic simulations employed a physical notation (1, 2). Binary "1's" were used to represent vehicles and "0's" were used to indicate the spaces between vehicles. Groups of memory cells were figuratively placed end to end to represent the roadway. Algebraic manipulations caused the "1's" to change position, thereby simulating the flow of traffic. With this mode of representation the vehicles could occupy only certain specified locations (bit positions) along the roadway and individual vehicles had no identity as such.

The memorandum notation utilizes an entire word to represent a vehicle. Various parts of the word are used for such individual characteristics as its time of entry into the system and its desired velocity. These parts may be extracted and interpreted as desired. This method is more versatile in that each vehicle's characteristics are identifiable as it moves through the network, making it possible to compute delays associated with the individual vehicle.

A third method of representation has been called a mathematical notation (1). This form of representation is similar to the memorandum notation except that, in addition to its other characteristics, each vehicle is associated with its own position indicator. Its position is therefore continuous within the accuracy of the computer. A vehicle's new position can at any time be computed as a function of its last position, its velocity, its acceleration, and the time increment. Spacings between vehicles are available from their respective coordinates and the vehicle length.

A fully mathematical notation requires more complicated program logic. Maneuvers such as turns, which must be accomplished at a specified location, are more difficult when the vehicle can occupy any position at the start of the maneuver. Furthermore, mathematical processing of vehicles is more complex, thereby increasing execution time required. On the other hand, elimination of limitations on the position increment allows some increase in the size of the time increment for the same model accuracy and provides increased versatility.

The SIGNET model (SIGNET is the name of the simulation model developed in this project) employs a fully mathematical notation. The advantages to be realized from a virtually continuous position vector outweigh the additional execution time required. Furthermore, the high execution speed of the CDC 6500 computer somewhat offsets this loss.

Each vehicle in the SIGNET model is completely represented by the information contained in four computer words. Current position, current velocity, and current acceleration are each individual words (POSN, VEL, and ACCEL respectively). The fourth word, ICAR, contains information on ten variables, as shown in Figure 1. Each variable is easily accessed via an unpacking function.

The philosophy on which the SIGNET model is based is relatively simple in principle but involves complex programming for its implementation. The basic premise is that all drivers have a target velocity at which they would prefer to travel if conditions meet certain minimum requirements. Acting to limit the driver in the pursuit of his target velocity are limitations generated by interactions with other vehicles and the physical environment, including leading vehicles moving at a slower speed, red signal indications, turning movements, obstructions to lane changes, and conflicts with vehicles from other links at intersections.

The mathematical relationships describing vehicle behavior can be divided into nine separate areas: vehicle generation, car-following, free behavior, vehicle updating, amber acceptance, stopping performance, queue discharge, lane-changing, and turning performance.

Vehicle Generation

Vehicles are generated at the zero coordinate of each input link on a per-lane basis using a translated negative-exponential distribution. Traditionally the unmodified negative-exponential distribution has been used to obtain intervehicle headways. It is of the form

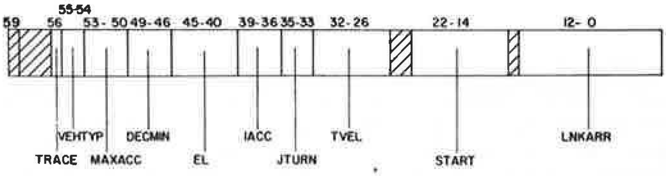
$$P\,(h \geq t) = \exp\,(-\beta_1 t)$$

where

$P\,(h \geq t)$ = probability of headway being greater than or equal to t; and
β_1 = vehicle flow rate in vehicles per second.

However, being distributed in $(0, \infty)$, the negative-exponential does not compensate for a minimum headway that of course exists for every vehicle. Therefore, as proposed by Gerlough (3), a better approximation is a negative-exponential with a translated axis:

Figure 1. Content of ICAR word.



LNKARR = link arrival time

START = initial coordinate at arrival on link

TVEL = target velocity

JTURN = turning movement

IACC = acceleration type

1 = stopped	8 = car following
2 = stopping	9 = going in delayed left turn
3 = stopped, waiting for turn	10 = going in normal turn
4 = stopping for turn	11 = slowing for turn (no conflict)
5 = stopping for lane change	12 = stopping for red signal
6 = stopped for lane change	
7 = free-behavior	

EL = effective length

DECMIN = minimum desired deceleration

MAXACC = maximum desired deceleration

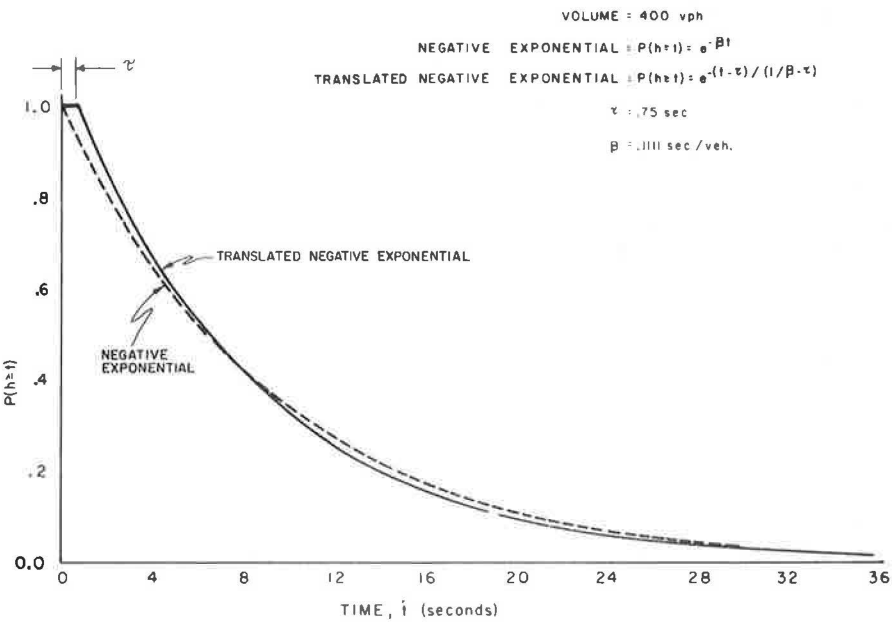
VEHTYP = vehicle type

- 1 = car
- 2 = truck

TRACE = vehicle being traced on travel-time route

- 0 = no
- 1 = yes

Figure 2. Comparison of negative-exponential and translated negative-exponential distributions.



$$P(h \geq t) = \exp\left(\frac{-(t - \tau)}{(1/\beta_1 - \tau)}\right)$$

where

τ = amount of translation; equivalently, in a physical sense, the minimum headway (Figure 2).

Work by Dawson and Chimini (4) in fitting the hyperlang probability distribution to intervehicular headways indicates a good value of τ (δ_1 in their study) to be 0.75 second under nonsignalized conditions. Input links to the SIGNET network are assumed not to be within the influence of upstream signals. Presence of a signal would necessitate reduction in this value of τ to reflect the lower headways of discharged vehicles.

In the SIGNET model the generation headway is independent of the simulation scan cycle. For each input lane in the network a variable is maintained to indicate the next arrival time of a vehicle in that lane. This tally is updated by the translated negative-exponential distribution at generation time to indicate the exact arrival time of the next vehicle in that lane.

Several descriptors of the vehicle's behavior are set at generation time, some stochastically and others deterministically. The vehicle's target velocity (TVEL), minimum desired deceleration (DECMIN), maximum desired acceleration (ACCMAX), and effective length (EL) are generated probabilistically according to their respective distributions. In addition, the vehicle type (IVEHTYP) is randomly determined from the link truck percentage input with the link traffic volume. The turning movement (JTURN) to be pursued at the link head is also determined randomly from the link turning probabilities. Finally, the link arrival time (LINKARR) is set to the current time, and the acceleration type (ACCTYP) is set to free behavior. All of the above are then packed into the ICAR word (Figure 1).

Car Following

At the heart of the simulation model are the free-behavior and car-following relationships. Much of the full spectrum of behavior at an intersection involves a tracking or following process, as seen in the case of queue discharge. Therefore the stimulus-response equations of car-following theory developed by Herman and associates (5, 6) are used to describe certain patterns of intersection performance.

Herman's works pertaining to car-following theory apply directly to the problem of processing vehicles in a digital simulation. With two basic exceptions (turning movements and stopping), one of the alternative vehicle behavior equations can be used to describe the behavior of individual vehicles within an intersection system.

Herman's equations have the general form:

response = sensitivity × stimulus

The best specific equation of this form Herman found was (in the notation of this study):

$$\text{ACCEL}(J+1, I+T) = a_0 \frac{\text{VEL}(J, I) - \text{VEL}(J+1, I)}{\text{POSN}(J, I) - \text{POSN}(J+1, I)}$$

where

ACCEL(J+1, I+T) = acceleration of car J+1, the follower, initiated at time I+T;
T = the car/driver lag;
VEL(J, I) and VEL(J+1, I) = the velocities of the leader and follower, initiated at time I;
POSN(J, I) and POSN(J+1, I) = the positions of the leader and follower initiated at time I; and
a_0 = the characteristic speed.

This equation is termed the reciprocal spacing model.

Free Behavior

Not all vehicles in a real system act as followers, however. An example of such a vehicle is the leader of a queue being discharged from a signal. In such a case behavior can be described by

$$ACCEL(J, I+T) = K\ (TVEL(J) - VEL(J, I))$$

where

ACCEL(J, I+T) = acceleration of car J initiated at time I+T;
K = proportionality coefficient;
TVEL(J) = target velocity of car J; and
VEL(J, I) = velocity of car J at time I.

This equation is termed the free-behavior model.

Stopping Performance

Two types of stops occur in the model: (a) stopping first in line at the intersection and (b) stopping behind another stopped vehicle. Empirical work at Ohio State University (7) indicated that use of a constant deceleration stopping model was realistic. It was found that, when given the choice, drivers tended to decelerate at an approximately constant rate throughout the duration of their stop.

In the SIGNET stopping model, the parameters of a minimum desired deceleration rate are supplied and a value of DECMIN(J) is randomly selected for each vehicle J at its generation time. During each scanning cycle the required stopping rate for the first vehicle on the approach is computed. When this rate is less than the minimum desired acceleration (implying a more severe stop), the vehicle begins stopping at the computed rate and continues to do so until zero velocity is reached. The stochastic nature of DECMIN thus accounts for different deceleration rates produced by each driver-vehicle combination.

A similar model is used for stopping behind another vehicle, with the principal difference being in the computation of the target stopped position. Vehicles stop at the position of the effective rear of the previously stopped vehicle, as determined by its effective length.

The effective length of a vehicle is equivalent to the average stopped spacing of vehicles stopped in queue, measured from the front bumper of the leading vehicle to the front bumper of the following vehicle, and therefore including the vehicle length and a clear space. Field studies have shown that it has a value of approximately 22 ft (6.7 m) for cars (8, 9). In SIGNET the vehicle's type (car or truck) is determined probabilistically at generation time. Then its effective length is determined stochastically from the parameters input for each of the vehicle types.

Turning Performance

Vehicles that desire to turn left or right at an intersection must at some point cease operating under the stimulus-response model and undertake an independent fixed turning schedule. The principal requirement is that vehicles must not exceed a given maximum speed during the turn. Maximum turning velocity is related to turning radius and side friction by the equation

$$VTURN = \sqrt{fgr}$$

where

VTURN = maximum turning velocity, feet per second;
f = coefficient of friction;
r = turning radius; and
g = acceleration of gravity.

The AASHO Policy on Geometric Design of Rural Highways (10) indicates that the 95-percentile turning speed is associated with side friction f = 0.3 for medium- to low-speed turns. Therefore,

$$VTURN = \sqrt{0.3\ (32.2)\ r} \ = \sqrt{9.66\ r}$$

where VTURN is in feet per second and r is in feet. Thus turning radii are supplied as input data and the maximum speeds associated with them are computed.

As a turning car approaches the intersection it is scanned at each simulation cycle. If the current velocity is greater than maximum turning velocity, the deceleration rate required to reach maximum turning velocity exactly at the start-turn point is computed. If this rate is less than the vehicle's minimum desired deceleration rate, the vehicle begins slowing to a maximum turning velocity, as in the stopping model. It continues to do so until reaching the start-turn position, unless affected by more stringent conditions. Maximum turning speed is maintained through the turn, whereupon the vehicle resumes behavior under one of the stimulus-response models, car-following or free-behavior, whichever is more stringent.

The foregoing does not imply that all vehicles will make their turns at maximum speed. Some will be affected by other vehicles in queue or vehicles on the receiving link so that their turns will be made at considerably lower speeds. However, none will exceed maximum turning speed.

Vehicle Updating

The various behavior relationships yield a negative or positive acceleration rate that begins after some reaction lag and continues for the rest of the scan cycle. In SIGNET both the reaction lag (REACT) and the scan cycle (CYCLE) are specified with the input data.

A theoretical analysis of driver reactions and highway events indicates the importance of driver reaction in safety and highway design. Therefore the inclusion of this parameter helps achieve realism within the model. In practice a wide range of values are used for reaction time; the Traffic Engineering Manual (11), however, recommends a time between 0.75 and 1.0 second for design purposes in urban traffic. To maintain continuity with the Carstens data (12) used in validating the queue-discharge model, a time of 0.75 second is recommended for use in SIGNET.

Movements of vehicles between scans are computed by adaptations of the equations of motion:

$$V2 = VEL(J, I) + ACCEL(J, I) \cdot REACT$$

$$\begin{aligned} POSN(J, I{+}CYCLE) = {} & POSN(J, I) + VEL(J, I) \cdot REACT + 0.5 \cdot ACCEL(J, I) \\ & \cdot REACT^2 + V2 \cdot (CYCLE - REACT) + 0.5 \\ & \cdot ACCEL(J, I{+}REACT) \cdot (CYCLE - REACT)^2 \end{aligned}$$

$$VEL(J, I{+}CYCLE) = V2 + ACCEL(J, I{+}REACT) \cdot (CYCLE - REACT)$$

where

V2 = the velocity of vehicle J after the reaction period;

REACT = the reaction time;

CYCLE = scan cycle;

POSN(J, I) and POSN(J, I+CYCLE) = the position of vehicle J at time I and I+CYCLE respectively;

VEL(J, I) and VEL(J, I+CYCLE) = the velocity of vehicle J at time I and I+CYCLE respectively; and

ACCEL(J, I) and ACCEL(J, I+REACT) = acceleration of vehicle J at time I and I+REACT respectively.

Geometric Configuration

The roadway is not represented physically as such, in the computer. However, it is necessary to specify certain roadway references in order to make meaningful the positions of vehicles contained in the POSN word. Thus there is a need for a coordinate system in which vehicles operate. The zero coordinate of each link is taken to be the point where all contributing turning movements are complete. In turn, the end of the link, or discharge boundary, is the point where all other contributing movements to the receiving link are complete. In reality there could be several contributing movements at the tail (i.e., zero coordinate) of the link, each ending at a different point. It is necessary to figuratively add a tangent section to each except the longest, thereby making all end at the same point.

Individually these computations are quite simple. When performed for an entire network, however, they become tedious and awkward. Therefore, the program PRESIG was developed as a companion to SIGNET. Its function is to compute a complete set of geometric data ready for input to SIGNET, based on easily acquired measurements from the network.

The logic of PRESIG may be divided into three areas as shown in Figure 3. Region I contains the primary input and initialization functions. Included in the input block is the verification of the input data, primarily achieved through checking card types and link numbers. Additional means for verification are obtained by printing out all input data.

Region II is concerned with the computation of a tentative set of discharge boundaries and begin-turn points for each link.

In Region III this tentative set is compared to revised discharge boundaries and begin-turn points, which are read from the input data. Any differences are changed to agree with the revised values. This option is provided to enable the user to easily correct any discrepancies that may arise between the PRESIG output and actual conditions. The final function of PRESIG is to output, both on the line printer and card punch, all the geometric data needed for input to SIGNET.

SIGNET PROGRAM STRUCTURE

The SIGNET program consists of a set of 33 nested, closed subroutines. The program is modular in construction, making it possible to insert new routines or modify existing ones. Twenty-eight routines occupy a fixed place in the program structure, while five routines serve auxiliary functions that are called on at various places in the program. All routines are written in FORTRAN IV for operation under the MACE operating system of the CDC 6500 computer.

Tasks performed by the main SIGNET program include

1. Input of all program data, program instructions, and link data, followed by formatting and writing these data.
2. Primary initializing tasks, which include initializing of all variables not connected with the statistical summary.
3. Activation of the traffic data input and initialization routines (INITRN and INIVOL) at the proper times as expressed on the program instruction cards.
4. Maintenance of the time loop, which includes control over the simulation time cycle and proper calling of the vehicle generation, vehicle update, and interlink transfer routines.
5. Program termination, which entails calling the final summary routine and writing final counts.

The main program chain is shown in Figure 4.

SIMULATION OUTPUT

The simulation program output is a detailed statistical tabulation of traffic characteristics in the network. The first section of output is a listing of input data supplied to the simulation program. Included are program parameters, link parameters, link geometry descriptors, traffic signal settings, turning probabilities, and traffic volumes.

Figure 3. PRESIG program chain.

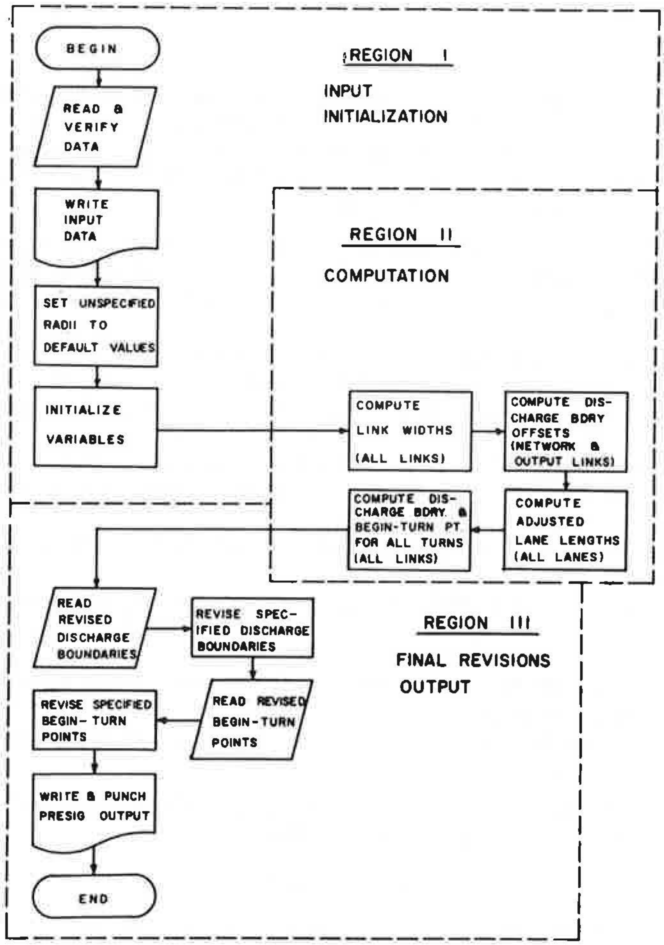
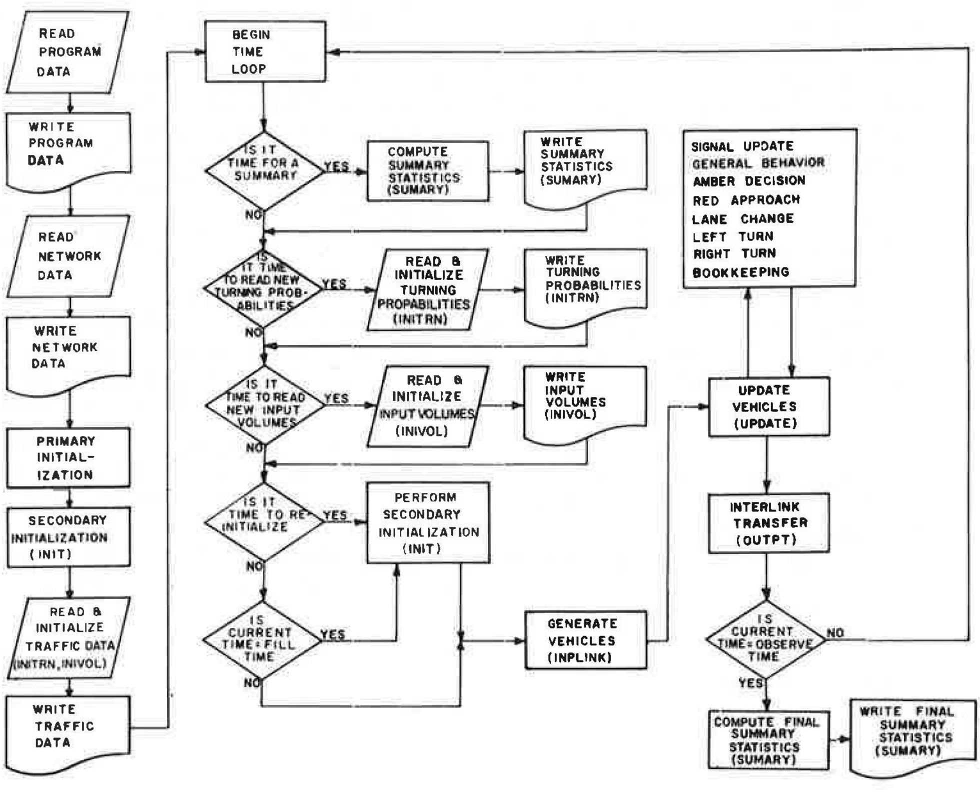


Figure 4. SIGNET main program chain.



The second section of output presents traffic operation data. Traffic magnitude data related to generated volumes, link exit volumes, and volumes traced along the specified route are reported, followed by statistics computed from the individual vehicles' performance. Values are presented for the overall system and for individual links. The following performance measurements are reported:

1. Total vehicle-miles—The total distance traveled on the link by all vehicles during the study period. Bookkeeping occurs when a vehicle leaves its link; the distance traveled by vehicles still on the link at the time of the summary, therefore, is not included in this tally.
2. Total delay (in seconds)—The sum of all vehicles' delays on a link. A vehicle's undelayed travel time is based on its target velocity and distance traveled. The delay encountered by the vehicle, then, is the difference between its actual and undelayed travel time.
3. Average delay (in seconds)—The total delay divided by the number of vehicles leaving the link. In this study average system delay is taken to be the primary measurement of system performance.
4. Delay standard deviation—The standard deviation associated with the above average delay.
5. Average delay (in seconds per vehicle-mile)—An extended form of the average delay. Its purpose is to facilitate the comparison of average delays among links.
6. Total travel time (in seconds)—The travel time for all vehicles.
7. Average travel time (in seconds per vehicle)—The average time required to traverse the link under consideration.
8. Average speed (in miles per hour)—The average speed achieved by all vehicles over the entire link or network.

The final sections of the output report contain frequency tables of queue lengths on the specified links and travel times on the route links.

TESTING AND VALIDATION

The SIGNET model was tested and validated to determine its accuracy in representing real-world conditions and its sensitivity to changes in input parameters. The first phase of the sensitivity analysis compared travel times produced by the simulation model with equivalent travel times obtained in the field.

Four simulation runs were made using different random-number generator seeds and network traffic volumes, turn percentages, and auto-truck ratios determined during the travel-time studies. Histograms were made for each link showing the number of vehicles having various simulated travel times. Such a histogram is shown in Figure 5 for a typical link during the test period. As indicated, the simulated travel time distribution was closely correlated to the actual travel times, with the mean simulated time being approximately equal to the actual travel times. This analysis indicated that the model working as a whole produced acceptable results.

The objective of the second phase of the testing was to determine the sensitivity of the model output to changes in various input parameters. An inherent characteristic of computer simulation, the ability to control all input parameters precisely, enables the investigator to change only one parameter and thereby to determine its influence on the output. This phase of the sensitivity analysis was useful for two reasons:

1. It provided additional data that could be used in confirming the reasonableness of the model. The direction of change in the output caused by modifying a parameter could be examined to see if it was logical, and the magnitude of the change could be checked for reasonableness.
2. It indicated whether particular portions of the model logic were indeed operating. If a parameter modification produced no change where one was expected, there was good evidence of logic flaws in the model.

Sensitivity tests were conducted by evaluating the effect of variation in seven different parameters:

Figure 5. Comparison of typical link travel times.

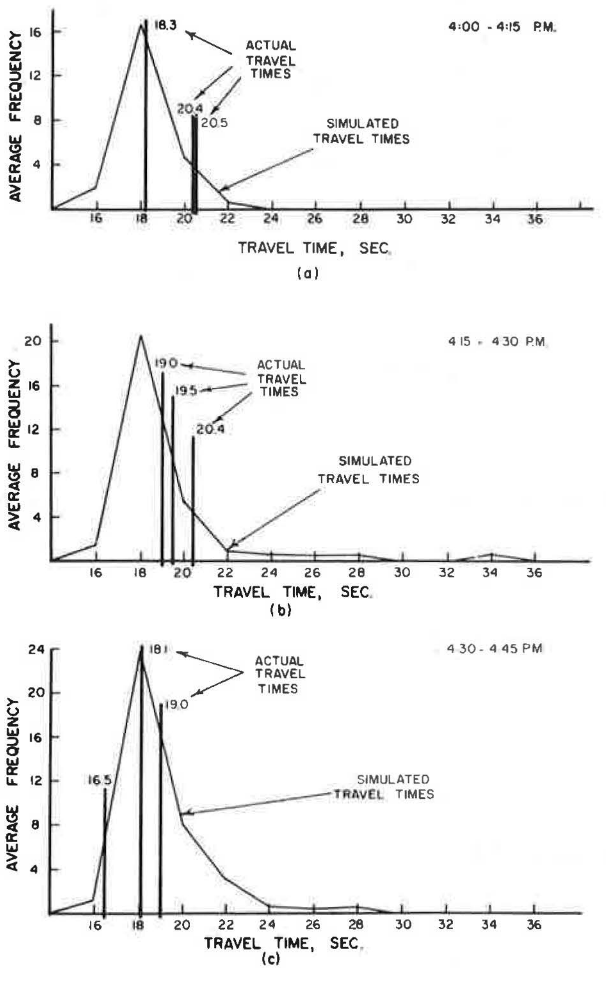


Figure 6. Effect of traffic volume on average vehicle delay.

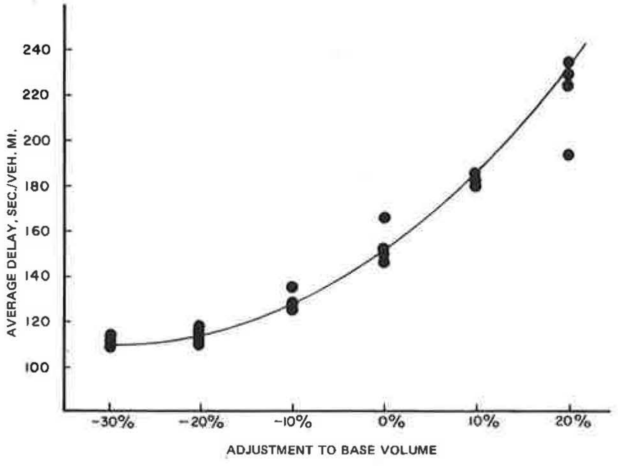


Figure 7. Effect of trucks on average vehicle speed.

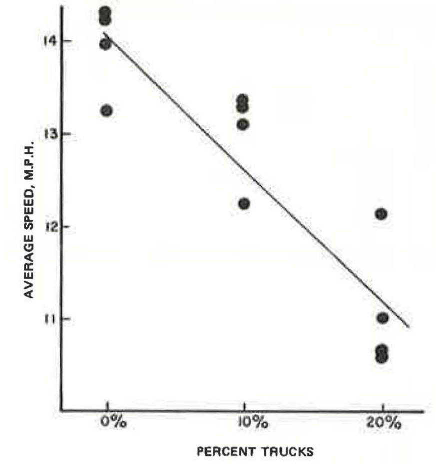


Table 1. Analysis of variance: effect of traffic volume on average vehicle delay.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
Among treatments:	43,631.71	3	14,543.90	79.0[a]
Linear effect	39,856.58	1	39,856.58	216.0[a]
Quadratic effect	3,747.26	1	3,747.26	20.4[a]
Cubic effect	27.86	1	27.86	0.16
Experimental error	3,683.79	20	184.19	
Total	47,315.50	23		

[a]Effect is significant at $\alpha = 0.10$. Critical region: $F > 2.97$.

Table 2. Analysis of variance: effect of trucks on average vehicle speed.

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Square	F Ratio
Among treatments:	16.815	2	8.408	24.5[a]
Linear effect	16.188	1	16.188	47.2[a]
Quadratic effect	0.627	1	0.627	1.83
Experimental error	3.084	9		
Total	19.899	11		

[a]Effect is significant at $\alpha = 0.10$. Critical region: $F > 3.36$.

1. Input traffic volume;
2. Target velocity, which affects the free-behavior model;
3. Effective length of vehicles, which affects queue discharge;
4. Truck percentage, which affects overall traffic throughput;
5. Maximum desired acceleration, which affects the acceleration rate and therefore queue discharge;
6. Characteristic speed (a_0), which affects the car-following model; and
7. Proportionality coefficient (K), which affects the free-behavior model.

Four simulation runs were made for each value of the input parameters, providing data for an analysis of variance for each effect.

It was desirable to know not only that the simulation output varied with changes in the input parameters but also what form the variation took (i.e., if the response curve was linear, quadratic, or cubic in nature). If the output variation leveled out at some point, this response could be compared with expectations, providing added evidence of the model's acceptability. For this reason the treatment sum of squares was broken down into sums of squares associated with linear, quadratic, and cubic effects using the orthogonal polynomials method.

The effects of variation in two typical variables, input traffic volume and truck percentage, are shown in Figures 6 and 7. As shown, decreased traffic volumes resulted in less delay. The effect was reduced, however, as volumes became lower and signal delays became critical. The analysis of variance in Table 1 confirms this quadratic effect with 90 percent confidence.

Increased percentages of trucks in the traffic stream caused a reduction in average speed. As indicated in the analysis of variance in Table 2, this effect was linear within the range of percentages studied.

Adherence of these sensitivity test results to anticipated patterns indicated that the various individual portions of the model were functioning as intended. When viewed in combination with favorable travel time combinations, it was concluded that the model provided a realistic and accurate simulation of actual conditions.

SUMMARY

SIGNET is a microscopic simulation model that reproduces traffic flow on a signalized street network under laboratory conditions. Its input format and structure enables it to be readily applied to any moderate-sized signal street network.

The program is economical to run, achieving a 6.5-to-1 real-time to simulation-time ratio for an 85-link network. This ratio is of course variable, depending on network size and configuration and on input traffic volumes.

The model underwent a testing and validation procedure that included comparison of travel times with actual field data and a comprehensive sensitivity analysis of individual model segments. These tests indicate that the model realistically and accurately simulates traffic flow through the street network.

REFERENCES

1. Gerlough, D. L. Simulation of Freeway Traffic by an Electronic Computer. HRB Proc., Vol. 35, 1956, pp. 543-547.
2. Goode, H. H., Pollmar, C. H., and Wright, J. B. The Use of a Digital Computer to Model a Signalized Intersection. HRB Proc., Vol. 35, 1956, pp. 548-557.
3. Gerlough, D. L. Simulation of Traffic Flow. HRB Spec. Rept. 79, 1964, pp. 97-118.
4. Dawson, R. F., and Chimini, L. A. The Hyperlang Probability Distribution: A Generalized Traffic Headway Model. Highway Research Record 230, 1968, pp. 1-14.
5. Herman, R., and Potts, R. B. Single Lane Traffic Theory and Experiment. Theory of Traffic Flow, Elsevier, 1961, pp. 120-146.
6. Herman, R., and Rothery, R. W. Car Following and Steady State Flow. Proc. International Symposium on Theory of Traffic Flow, London, 1963, pp. 1-11.
7. Study of Electronic Devices as Traffic Aids. Transportation Engineering Center, Ohio State University, Rept. 202-1, 1962.

8. Bellis, W. R. Capacity of Traffic Signals and Traffic Signal Timing. HRB Bull. 271, 1960, pp. 45-67.
9. Greenshields, B. D., Schapiro, D., and Erickson, E. L. Traffic Performance at Urban Street Intersections. Bureau of Highway Traffic, Yale University, Tech. Rept. 1, 1947.
10. A Policy on Geometric Design of Rural Highways. American Association of State Highway Officials, Washington, D.C., 1965.
11. Traffic Engineering Handbook. Institute of Traffic Engineers, Washington, D.C., 1965.
12. Carstens, R. L. Some Traffic Parameters at Signalized Intersections. Traffic Engineering, Vol. 41, Aug. 1971, pp. 33-36.

FREEWAY DIGITAL SIMULATION MODELS

Y. S. Hsu and P. K. Munjal, System Development Corporation, Santa Monica, California

This paper reviews 15 simulation models associated with various aspects of freeway vehicular traffic. They range from special-purpose programs directed toward studying the impact of trucks on the traffic flow to general-purpose programs that include most known variables of importance. The models are compared against a baseline of eight features that are regarded as both desirable and independent of specific simulation purpose. Most of these features typically represent characteristics that are of value to a potential user in making a choice as to which model would best serve his needs. Concurrently it is represented that these features could furnish a baseline to accomplish a certain amount of standardization. Each model is treated briefly in terms of these features and other special attributes. An overall comparison table is developed for easy reference as to the basic purpose and characteristics of each model.

•SIMULATION of vehicular traffic on digital computers has attracted considerable interest since the late 1950s. It began with the simulation of vehicles approaching and departing from isolated signal-controlled intersections. It was not until the late 1960s that digital simulation was applied to freeway traffic and related features.

A decade ago Gerlough (1) presented a detailed discussion on simulation techniques and what could be achieved toward improving traffic flow theory and practices by the application of digital computer simulation. Today, more than a dozen general-purpose or special-purpose freeway digital simulation models have been developed.

A careful examination of the existing models indicates that there was a lack of coordination in the development of models. There were no standards for the models and no application guidelines, which makes it difficult for the user to determine what model to select for his needs. Because of the lack of a universally accepted traffic flow theory and varying operational characteristics, each model was developed largely through intuition. Validation is a very expensive and time-consuming process, and no extensive validation covering a wide range of freeway geometrics and traffic patterns has been conducted on any model. Therefore, the realism and utility of the existing traffic simulation models are still doubtful.

These limitations do not imply that all of the model development effort was wasted. On the contrary, considerable fresh knowledge of traffic flow phenomena has been obtained through simulation. It is believed that further advancement in digital simulation can be achieved through a well-planned, coordinated effort.

The purpose of this paper, in addition to a review of existing simulation models, is to address the desirable characteristics and features of models. We hope to attract the attention of highway research personnel to the need for future standardization model development and documentation and provide summaries of existing models for those who want to select a model for their use but do not have the time to study each model's capabilities and limitations in detail.

The 15 models under consideration are

1. Arizona Transportation and Traffic Institute Traffic Simulation Model,
2. Midwest Research Institute Freeway Simulation Model,
3. Midwest Research Institute Mountainous Terrain Model,
4. Northwestern University Lane-Changing Model,
5. Sinha Freeway Simulation Model,

6. Connecticut Department of Transportation Expressway Simulation Model,
7. Texas Transportation Institute Freeway Merging Model,
8. System Development Corporation Diamond Interchange Model,
9. System Development Corporation Freeway Simulation Model,
10. Mikhalkin Freeway Simulation Model,
11. Georgia Model,
12. SCOT Corridor Model,
13. Priority Lane Model,
14. Aggregate Variable Models, and
15. Aerospace Corporation Freeway Simulation Model.

These simulation models varied in purpose and structure because of different user requirements. However, there are desirable general characteristics that each model should possess and other features that will add more application value to the model. They include

1. Realism for representing freeway flow phenomena,
2. Existing features built into models to handle anticipated applications,
3. Logic complexity,
4. Computer running efficiency,
5. Extent of model validation,
6. Flexibility and expandability,
7. Suitability for incident detection and ramp control, and
8. Completeness of program documentation.

The following sections give a detailed discussion of these characteristics and a critical review of each model with respect to them.

DESIRABLE CHARACTERISTICS

The eight characteristics and features mentioned certainly do not cover the complete spectrum of simulation models. They do, however, more or less reveal the value and capability of each simulation model.

Realism for Representing Freeway-Flow Phenomena

This characteristic reflects how closely a simulation model is able to describe traffic flow behavior, particularly when certain important freeway traffic components are neglected or some convenient assumptions are made to simplify the model.

Existing Features Built Into Models to Handle Anticipated Applications

Most of the 15 simulation models listed were developed with a single purpose in mind. The purpose for the individual models and existing features and capabilities are addressed later.

Logic Complexity

Because each model is developed for a different application, some have more features than others. In general, car-following and lane-changing are the two most important elements found in common. Because of space limitations, we shall not describe the logic in detail but rather overview the logic complexity of each model. A comprehensive summary of the car-following and lane-changing rules is given elsewhere (2).

Computer Running Efficiency

This refers to the ratio of computer time to simulated real-time freeway traffic. More accurately, it refers to the computer costs associated with the simulation per unit of real time. To compare strictly the efficiency of two different simulation models, simulation runs should be conducted under identical conditions of freeway geometry and traffic volumes. Since each model is developed for different purposes and has different freeway geometry, it is not possible to run them under identical conditions or to use reported results to compare the efficiency. However, by quoting the simulated condition and the running time plus the computer model and core size used, the reader will have some idea of the relative efficiency of each model.

Extent of Model Validation

Model validation is a time-consuming process that involves data collection, data reduction, and statistical analysis and testing. Validation of the models as a whole has been less than adequate, particularly in terms of statistical analysis and testing.

Flexibility and Expandability

The flexibility of the computer simulation model refers mainly to the structuring of the computer program in terms of facilitating the improvement or integration of the logic into a more general-purpose model. The expandability refers to the extension of a model to cover a more general freeway configuration or traffic flow conditions. Detailed discussion of flexibility and expandability of each model is difficult without a complete examination of program listing and documentation. These have not been obtained to date, and thus only general comments will be made regarding these aspects.

Suitability for Incident Detection and Ramp Control

A major use of traffic simulation models is to test the effectiveness of freeway surveillance and control strategies before they are implemented into an operational system. This avoids the expensive testing of ineffective strategies on-line. One important element in surveillance is the density and total number of detectors required to measure various traffic characteristics at specific locations so that the occurrence of freeway incidents can be identified. A counterpart in control is the provision of on-ramp signal-control capability so that the simulation model is able to test the effectiveness of different on-ramp control strategies.

Completeness of Program Documentation

All the simulation models except one were programmed in FORTRAN, but some models have subroutines written in lower-level languages. The exception is model 8, which is in JOVIAL. The extent of program documentation varies considerably from model to model. We shall rate the program documentation in three levels:

1. Availability of a user's manual in addition to comprehensive documentation. This will allow users unfamiliar with the program details to proceed step by step and complete a successful simulation run.
2. A comprehensive explanation of the program. This provides to users the detailed capability of the model and the structure of individual components of the model and thereby allows the user to visualize the possibility of program modification and expansion.
3. A brief explanation of the various routines of the model. This implies that program documentation is inadequate and therefore makes it difficult to evaluate the efficiency and usefulness of the model. It is to be noted that the completeness of program documentation is based on reports that are currently available.

Each of these points will be reflected in the following discussions of the individual models. Because program documentation is not prepared at the same level of sophistication, some characteristics of specific models are either unclear or completely unknown.

CHARACTERISTICS OF THE MODELS

Model 1—Arizona Transportation and Traffic Institute Traffic Simulation Model

This model was developed by Richard, Baker, and Sheldon (3) to simulate freeway traffic that may be used to establish freeway interchange design criteria. The freeway geometry is restricted to 3 through lanes, 1 ramp, and an acceleration lane or an auxiliary lane. However, with minor program changes, 1 and 2 through-lane systems may be simulated. Ramps are restricted to direct connections and loop connections. Freeway grade is handled in this study by changing both the operating speeds and the vehicle acceleration and deceleration rates. Simple logic for vehicle distribution among lanes, car-following, and lane-changing is provided.

In preparing simulation runs, freeway volume and ramp volume are specified. There are three alternatives in choosing vehicle processing time, starting from 1.5 seconds, with an increment of 1.5 seconds. Vehicles are generated from a binary decision rule within the review period according to the input volume. This gives essentially a negative exponential distribution. Desired speeds are generated from a normal distribution with modification for trucks and grades.

This model is probably flexible enough to allow some simple additional capabilities. Since the overall logic is very simple, it is doubtful that the model realistically represents traffic flow in any detail.

The model has not been validated and the simulations did not provide figures that related to running efficiency. The program is written in FORTRAN for use on an IBM 7072/1401 computer and requires 8K of core storage. A summary of the characteristics of this model as well as other models is given in Table 1.

Model 2—Midwest Research Institute Freeway Simulation Model

The purpose of this model (4, 5) is to assist in the design of interchanges by providing a method for assessing the efforts of design variables on traffic capacity, safety, and level of service. Special emphasis is therefore placed on traffic flow in the vicinity of entrance and exit ramps.

The freeway section can be up to 80,000 ft long, with 2 to 4 through lanes and up to 6 right-hand and 6 left-hand ramps (a maximum possible total of 12 ramps). The ramps can be any combination of on- and off-ramps located arbitrarily along the freeway section. Any or all of the on-ramps can be equipped with traffic signals, and therefore the model is capable of testing ramp control strategies.

The simulation vehicles are designated by driver type, vehicle type, desired speed, and cooperator, which indicates that the vehicle will oblige would-be lane changers by trying to provide a usable gap in front. The volumes are specified for each lane and ramp.

The car-following and lane-changing logic is very complex. The acceleration and deceleration capabilities are reflected by vehicle types. Vehicles are generated in a fashion similar to those in model 1 but with the application of different weights for the red and green periods to the ramps. Desired speeds are generated from truncated normal distributions. The review time period is 1 second, which is shorter than that of model 1. Under this processing time interval it would typically require 20 minutes of computer time on an IBM 360/50 per minute of simulated time to simulate 1 mile of freeway with 3 lanes in 1 direction with an average traffic density of 45 vehicles/mile/lane.

The program is written in FORTRAN IV, with a few subroutines written in assembler language. The model has not been validated. The program documentation is adequate.

Model 3—Midwest Research Institute Mountainous Terrain Model

This model (6) was developed to study traffic characteristics on 4-lane divided highways in mountainous terrain. The geometric configuration of the model allows simulation of a freeway section up to 131,000 ft long with the 2 lanes and an intermittent right climbing lane. There is no provision for on- or off-ramps. The grade and the front and rear sight distances are defined for the entire section. Different vehicle characteristics are also defined, and curve-limited or downgrade-limited maximum speeds may be specified within certain zones.

Most simulation dynamics are the same as those of model 2 except where the desired speeds and acceleration capabilities are functions of grades and horizontal curvature.

Validation was performed at two levels, the microscopic and the macroscopic. The former includes vehicle performance characteristics, car-following behavior, and gap

Table 1. Summary of the 15 models.

Model No.	Purpose	Freeway Geometry	No. of Ramps	Vehicle Generation	Desired Speed Distribution	Simulation Dynamics	Grade and Curvature Effect	Validation	Ramp Signal Capability
1	Ramp design criteria	Straight section, 3 lanes	1	Negative exponential	Normal for both freeway and ramp	Lane distribution logic, car-following and lane-changing	Grade effect included	None	None
2	General purpose	Straight section, up to 4 lanes	6 right, 6 left	Negative exponential	Truncated normal	Complex car-following and lane-changing rules	None	None	Yes
3	Mountainous road	Mountainous terrain	None	Negative exponential	Truncated normal	Complex car-following and lane-changing rules	Yes	Yes	None
4	Lane-changing	Straight section	None	Shifted exponential	Normal	Car-following and simple lane-change and gap-acceptance logic	None	Very little	None
5	General purpose	Straight section	4 on, 6 off	Freeway, shifted exponential; ramp, hyper-Erlang	Normal	Car-following and lane-changing, merging	None	Yes	None
6	Design tool	Straight section	10 on, 10 off	Negative exponential	Near normal from field data	Simple car-following and lane-changing rules	None	Yes	None
7	Merging	Straight section	2 off, 6 through and on	Poisson	Truncated normal	Simple car-following and lane-changing logic but extensive ramp merging logic	None	Very little	Yes
8	Diamond interchange design and operation	Diamond interchange	1 on, 1 off	Truncated exponential	None	Microscopic on arterial and macroscopic on freeway	None	Yes	Yes
9	General purpose	Arbitrary network	Unlimited	Negative exponential	None	Car-following, lane-changing, merging	None	None	None
10	Freeway surveillance	Straight section	None	Not available	Truncated normal	Car-following, lane-changing, sensor simulation	None	Yes	None
11	Truck behavior	Straight section	Not available	Shifted exponential	Normal	Car-following, lane-changing, ramp merging	None	Yes	None
12	Freeway corridor operation	Freeway corridor	Not available	Not available	Not available	Macroscopic on freeway and microscopic elsewhere	None	Yes	Yes
13	Priority lane	Straight section	50	Not available	Not available	Compressible fluid	None	Yes	None
14	Ramp control	Straight section	Not available	Not available	Not available	Continuum model	None	Yes	Yes
15	General simulation	Arbitrary network	Unlimited	Poisson	Normal	Car-following, lane-changing, ramp merging, collision	Yes	Yes, but no statistical test	None

Model No.	Detector Capability	Starting Mechanism	Warm-Up Time	Maximum No. of Vehicles Allowed	Computer	Programming Language	Core Requirement	Computer Time/Simulation Time Ratio	Documentation
1	None	Not available	Not available	Not available	IBM 7072/1401	FORTRAN plus Autocoder statements	8K	Not available	Poor
2	None	Empty	Not available	3,000	IBM 360/50	FORTRAN IV plus 2 subroutines in assembler	Not available	20:1	Good
3	None	Preloaded	Not available	Not available	CDC 6400	FORTRAN IV plus assembler for 1 subroutine	32K (60 bit)	20:1 to 10:1	Good
4	None	Empty	Not available	300/lane	CDC 6400	FORTRAN IV plus SPURT simulation language	Not available	1:4 to 1:20	Poor
5	None	Preloaded	Not available	800/lane	IBM 360/65	FORTRAN IV plus assembler	110K bytes	1:2 to 1:10	Fair
6	None	Empty	Not available	1,000 at any time	Univac III	FORTRAN IV	30K (24 bit)	3:1	Poor
7	None	Preloaded	1 minute	500 at any time	IBM 7094 Model I	FORTRAN IV	Not available	Not available	Poor
8	None	Preloaded	15 minutes	2,000	IBM 360/67	JOVIAL plus machine language	60K	1:75	Poor
9	Yes	Preloaded	15 seconds	2,500	IBM 360/67; Univac 1108	FORTRAN IV	65K (36 bits)	1:1 to 5:1	Poor
10	Yes	Preloaded	Not available	Not available	IBM 360/144; 360/67	FORTRAN IV	120K bytes	4:1	Fair
11	None	Preloaded	None	Not available	IBM 360	FORTRAN IV plus BAL	Not available	0.5:1 to 0.75:1	Fair
12	None	Not available	Not available	Not available	Not available	FORTRAN IV	213K bytes	Not available	Fair
13	None	Not available	Not available	Not available	CDC 6400	FORTRAN IV	Not available	Not available	Fair
14	None	Not available	Not available	Not available	IBM 360/44	FORTRAN IV	Not available	1:11	Poor
15	None	Not available	Not available	Unlimited	CDC 7600	FORTRAN IV and COMPASS machine language	100K (60 bit) est., 8,000 statements	1:7	Poor

acceptance in lane change. The latter considers gross flow characteristics such as flow to lanes, lane change frequencies, spot speed distributions, time headways, and overall travel speeds. Reported validation involves the comparison of real-world data from other studies to simulations using the model. Continuing efforts are being conducted as more extensive validations of platoon behavior and passing logic using newly collected photographic data taken near Pacheco Pass on California Route 152 and Topanga Canyon on California Route 23.

Program documentation is extensive and includes a comprehensive user's manual, but the program listing is not available. Because of the detailed logic and good validation results the model is believed to possess sufficient realism for its purpose. However, the model is not very efficient. Reported simulation runs have a ratio of computer time to simulated time in the order of 20:1 to 10:1 using a CDC 6400. The program is written in FORTRAN IV except for one short subroutine in assembly language. It requires 32K words of computer memory.

Model 4—Northwestern University Lane-Changing Model

A detailed examination of freeway lane-changing behavior was the motivation for developing this model by Worrall and Bullen (7). For this reason the freeway geometry is limited to a 4-lane straight section without ramps. The simulated freeway length can be up to a few miles.

The car-following logic is fairly complex. The lane-changing logic is based on a lane-changing desire flag for each vehicle and the available gap. Vehicles are generated from a shifted negative-exponential distribution with desired speeds chosen from a normal distribution. The model produces output showing lane frequencies, lane-change delays, vehicle redistribution, etc., but does not accept a mix of vehicle types. The computer running efficiency of the model is relatively high, the computer-time to simulated real-time ratios ranging from 1:4 to 1:20 for 2-, 3-, and 4-lane situations with volume ranging from 600 to 1,800 vehicles/lane/hour. The flexibility of the model is fairly high, although the relative ease with which the model can be recalibrated is not. Lane-changing frequency and speed-volume outputs of the model match favorably with field data collected at various Chicago freeways.

The model is programmed in FORTRAN IV for a CDC 6400 computer, with some subroutines written in SPURT simulation language developed at Northwestern University. Only a small-scale calibration has been made on the model. Program documentation consists of a brief description of the various routines and a program listing.

Model 5—Sinha Freeway Simulation Model

This is a general-purpose simulation model developed by Sinha (8, 9) for use as a tool in the analysis of freeway phenomena. The model has a capacity for the simulation of 5 lanes, 4 on-ramps, and 6 off-ramps. The ramps may be located either on the right-hand or left-hand side of the freeway. It can simulate up to 3½ miles in length using a 256K IBM 360/65 system.

The car-following and lane-changing logic is fairly complex. The gap-acceptance logic is similar to that of model 7. Only two types of vehicles are assumed. Freeway mainline traffic was generated from a shifted exponential distribution, while ramp vehicles were generated from a hyper-Erlang distribution. Desired speeds were generated from a normal distribution.

Because the model was developed as a general-purpose tool for analyzing traffic operating characteristics, the computer output provides detailed information at each of the several control points (points of interest):

1. Distribution of headways in each lane,
2. Distribution of speeds in each lane,
3. Distribution of traffic volumes in each lane,
4. Distribution of exiting, entering, and through vehicles, and
5. Distribution of exiting, entering, and through vehicle speeds in each lane.

The computer program is written in both FORTRAN IV and IBM 360 assembler language for an IBM 360/65 computer. The review period is 1 second. However, the core requirement is not given. A maximum of 800 vehicles can be processed for each lane. Reported simulation results show the ratios of computer time to simulated time vary from 1:2 to 1:10.

Data for model validation were from the Eisenhower Expressway in Chicago, the Long Island Expressway in New York, and the 1965 Highway Capacity Manual. Program documentation consists of a good description of the main program and subroutines plus program listing.

Model 6—Connection Department of Transportation Expressway Simulation Model

The Connecticut model (10) is similar to model 1 and was also developed for the purpose of investigating, evaluating, and solving freeway design problems. It allows a 5-mile, 7-lane section with 10 on-ramps and 10 off-ramps. On-ramps are restricted to direct connections. It can handle 1,000 vehicles in the system during any given second.

Driver characteristics include the assignment of acceptance gaps to individual vehicles, desired speeds from a near-normal distribution based on field data, the generating of vehicles similar to that of model 1, and the acceleration and deceleration capability as linear functions of speeds. The car-following and lane-changing logic is very simple as compared to most of the other models.

The reported computer running efficiency of this model is about 3 minutes of computer time for every minute of real time on a Univac III. The program is written in FORTRAN IV with a requirement of 30K (24 bits) words of core storage. Program documentation was not available at the time of preparing this paper.

The model was validated by using the chi-square and Kolmogorov-Smirnov test to compare the simulation results with data on speed and headway distributions from the 1965 Highway Capacity Manual.

The chi-square test yields a confidence level for a cell-by-cell comparison of two distributions while the Kolmogorov-Smirnov is a nonparametric test of the maximum difference between two accumulative distributions. For the three volumes investigated, i.e., 1,000, 1,500, and 1,800 vehicles per lane, the chi-square test showed a level of confidence of 85, 90, and 95 percent respectively, while the Kolmogorov-Smirnov test showed a confidence level of 95 percent for all three volumes.

Model 7—Texas Transportation Institute Freeway Merging Model

This model was developed by Buhr et al. (11) for the purpose of simulating traffic operations under different modes of on-ramp control. The number of off-ramps is limited to 2, while the number of entrance ramps plus freeway lanes is limited to 6, with a maximum freeway length of 6,000 ft.

During simulation road sections must be preloaded. New vehicles are generated from a Poisson distribution. Each vehicle is assigned a number of characteristics such as length, current speed, desired speed, and distance from the zero reference point (beginning of the simulation section). The desired speed is generated from a normal distribution, but, if the generated speed is higher than the designated maximum speed, it is then reduced to the maximum speed. The simulation program consists of one 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 various ramp control modes the model can handle include (a) no control, (b) fixed-time metering, (c) demand-capacity metering, and (d) gap acceptance control. The computer scan time is 1 second. Besides rather simple car-following and lane-changing logic, the model provides extensive ramp merging logic for the purpose of testing the various ramp control strategies.

A simple validation study was performed using the geometrics and data of the outbound Cullen on-ramp on the Gulf Freeway in Houston. However, no statistical tests were conducted to indicate the level of confidence.

The computer running efficiency was not reported. The program is written in FORTRAN IV for the IBM 7094 Model I computer. Complete program documentation is not available.

Model 8—System Development Corporation Diamond Interchange Model

This is one of the only two models among the 15 that have both the microscopic and macroscopic features built into the model. This model, developed by Nemeczky and Widdice (12) as a tool to aid in the design and operation of signalized diamond interchanges, contains both freeway and signalized arterial submodels, the freeway section being limited to 2 on- and 2 off-ramps.

The program logic is fairly complicated in terms of a macroscopic model. Vehicles are generated from a truncated exponential distribution to eliminate the possibility of unusually large time headways, and five driver types were allowed.

Since the model was developed to evaluate selected operational or design alternatives, the following outputs were provided:

1. Average travel time through the system,
2. Average travel time through each individual model region,
3. Average speed through the system,
4. Average speed through each individual model region,
5. Average delay through the system,
6. Average delay through each individual region,
7. Number of stops, and
8. Acceleration noise.

The model allows inspection of changes of a diamond interchange geometry (full, split, or partial diamond, changes in through and turning lanes and pockets, etc.) as well as the change in signal control parameters. Model validation was conducted by comparing the simulation outputs of the number of cars through the system and through each section of the model by origin-destination and the travel times by origin-destination with data collected at the Coldwater Canyon diamond interchange of the Ventura Freeway in Los Angeles. The Wilcoxon signed-rank tests indicated the model was valid at the 5 percent level of significance.

Model 9—System Development Corporation Freeway Simulation Model

A series of general-purpose freeway simulation models were developed in the order of increasing complexity. Here we shall discuss only the most recently developed one (13) because it is the improved and generalized version of all its predecessors. A unique feature of this model is that multilane highways were modeled by circular tracks. This creation generated many advantages in the simulation. The model is considered very general, so that any reasonable freeway configuration (including freeway interchanges) can be modeled. The network size is limited primarily by the number of cars that can be handled. For a 65K core (36 bits) computer the number is 2,500 cars.

The model provides extra capabilities such as the generating of position-time plots, and its structure allows direct simulation of sensors, controls, and control algorithms. The logic is not complicated in terms of the capabilities it provides.

The position-time plots could be viewed as a computer-generated movie, so that it is easy to bring out the turbulent aspects of the overall flow or so that one can focus on the behavior of individual vehicles to determine the realism of the simulation logic.

The model advances roughly 500 cars in 1 minute of computer time for 1 minute of real time on a Univac 1108. It is expected, with some modifications, that 15 minutes of Univac 1108 computer time will allow 5,000 cars to be advanced for 3 minutes of simulation time.

No validation has been done on the current version, but limited validation performed on an earlier version of the model included the comparison of flow-concentration data obtained from the simulation with data collected on a 2-lane expressway in Virginia. Program documentation has not been prepared at the present time.

Model 10—Mikhalkin Freeway Simulation Model

The model, developed by Mikhalkin (14), provides a means for systematic experimentation with the capability of controlling factors of the driver-vehicle-roadway system that usually cannot be controlled in real traffic flow. The simulated roadway is a straight, level freeway with no ramps, up to 4 lanes wide and 20,000 ft long, and with sensors. Simulation is on the IBM 360/65 system with 120K core storage. However, this restriction can be easily removed if larger core size is available.

Simulation can begin at a high concentration, low concentration, or even an empty system. The scan interval is 0.75 second, which is equivalent to the driver reaction time in the model.

The car-following logic is based on the nonlinear car-following rule of Gazis et al. (15), and the lane-changing logic is based on that of model 4. Detectors were simulated in the model to provide volume and occupancy measurements. The sensor sampling rate was 15 per second. Procedures for estimating the roadway local density and space mean speed from detector measurements were developed with a high degree of accuracy. Vehicles were generated randomly, but no specific distribution was mentioned. Vehicle desired speeds and vehicle lengths were obtained from truncated normal distributions whose parameters were input data.

Because of the various traffic parameter estimating algorithms implemented in the model that use detector-measured data, this model is extremely useful for freeway surveillance and incident-detection purposes.

The model was written in FORTRAN IV and is modular in form so that each routine can be easily modified. The computer running efficiency is low—of the order of 4 units of computer time to 1 unit of simulated real time for a 4.6-mile section of a 4-lane freeway.

The model validation was based in part on the similarity in form and magnitude of the flow-concentration relationships obtained from the simulation and published data. Although good agreements were reported, no statistical tests were conducted to justify the observations. However, extensive statistical work was performed on sensor simulation.

Model 11—Georgia Model

The Georgia model developed by Wildermuth (16) was primarily concerned with the assessment of truck effects on freeway flow characteristics. Model development was based on an extensive evaluation of models 2, 4, and 7. Successful components from the earlier models were adapted and modified so that trucks could be properly introduced as a distinct element into the traffic flow simulation.

The basic structural elements of the Georgia model closely resemble those of model 4. Each vehicle is associated with a vector containing 12 specific characteristics such as the desired speed, current speed, and vehicle type. Vehicles are generated from a shifted exponential distribution. The desired speed is generated from a normal distribution, with the mean and standard deviation as separate input variables for each lane.

Simulation starts with a preloaded condition without requiring a warm-up time to achieve a stable flow.

Model validation was done in terms of comparing the generation of different vehicle types, headway distributions, lane volume, speed distribution, and lane-changing frequencies from the simulation runs to those of the real data, and good results were shown.

For a 1-mile freeway section with 3 lanes, simulation times on an IBM 360/30 computer ranged between 2.4 and 3.0 times the simulated real time depending on the traffic volume. It is expected the ratio will be 0.5 to 0.75 on the IBM 360/50.

The model was written in FORTRAN IV, with several minor routines in assembly language. A complete program documentation is not available, but instructions on running the model are given.

Model 12—SCOT Corridor Model

The SCOT (simulation of corridor traffic) model (17) was originally conceived as a concatenation of two existing models—the UTCS-1 simulation model of urban traffic (18) and the DAFT simulation model of freeway traffic (19). The SCOT model is also a dual microscopic (the UTCS-1) and macroscopic (the DAFT) model like model 8.

The SCOT model treats vehicles microscopically on the arterial street system (including ramps) and macroscopically (as platoons) on the freeway. Any arbitrary freeway and surface street network containing up to 200 intersections may be represented, and traffic flow is described by specifying either origin-destination volumes along the peripheral entry links or turning movements at each node.

The objective of developing this model was to use it as a medium for assisting in defining the surveillance and control requirement for both existing and planned freeway corridors.

The crux of the freeway component of the SCOT model is the speed-density formulation resulting from a general form of the non-integer car-following rule by Gazis et al. (15). The unknown parameters of the speed-density equation have to be determined experimentally before the simulation run.

Complete program documentation, including a user's manual, is available. However, the description of the system is poor as compared to other model documents, and therefore much key information such as detailed system capability and computer running efficiency does not appear in the system's description.

Model validation consists of a comparison of simulation on a 0.4-square-mile network in Dallas containing two short freeway sections to the aerial photographic field data; good agreement is shown.

Model 13—Priority Lane Model

This model (20, 21) was directed toward evaluating traffic operations on freeways with priority lanes such as those allocated for buses or vehicles containing a required minimum number of passengers. The model geometry allows a maximum of 50 freeway subsections, with not more than 1 ramp in each subsection.

The model logic is sophisticated and efficient. It is essentially macroscopic, with input data provided for each 15 minutes. The computer program has a modular structure, and additional capability can easily be obtained by modifying or including the appropriate subroutines. Future changes can be made with minimum effort to match model results with empirical data. Running instructions and input data formats are provided for using the model.

The model idealizes physical queues, and this may obscure some of the effects being studied. Furthermore, subsection capacities and demand are assumed to remain constant over 15-minute time slices. The study section is limited to 10 miles, and no off-ramp queuing calculations are attempted if off-ramp demand exceeds ramp capacity. Otherwise the model affords sufficient realism for representing traffic flow on freeways with any kind of priority lanes or reversible lanes and for ramp control schemes for priority vehicles. Validation of earlier versions of the model was made with data collected on the San Francisco-Oakland Bay Bridge, where vehicles with at least 3 passengers (to encourage the formation of car pools) were allowed the use of a faster-moving priority lane.

The computer program is written in FORTRAN IV for the CDC 6400 system.

Model 14—Aggregate Variable Models

In the aggregate variable models developed by Payne (22), a freeway is partitioned into sections and the freeway traffic is described by a set of dynamic equations in terms of the aggregate variables of flow rate, section density, and section speed. The purpose of the models is to study the problem of developing ramp control strategies with a high simulated real-time to computer-time ratio. The model does not distinguish flow by lanes, and the traffic flow is described as an extension of the simple continuum models. However, it seems that the model produces good results for ramp control purposes.

Except for a listing, no program documentation is available. The program is written in FORTRAN IV for IBM 360 computer systems. The computer running efficiency is high because of the macroscopic nature of the model—80 seconds of computer time on an IBM 360/44 for simulating 3 hours of real time for a 4.6-mile stretch of 4-lane freeway. The logic employed is relatively simple, and program flexibility appears to be poor. Model validation consists of a crude comparison between simulated results and data collected on a 4.6-mile section of the Hollywood Freeway in Los Angeles with 9 on-ramps, 7 off-ramps, and fixed-time (time-of-day) ramp metering.

Model 15—Aerospace Corporation Freeway Simulation Model

This very general model developed by Harju et al. (23) is capable of simulating any freeway network. A special feature of the model, similar to model 9 of System Development Corporation, is the capability to produce computer-generated traffic flow movies of any specific subarea of the network under simulation. The movies appear as stationary overhead aerial shots and can be used for detailed flow analysis, for program debugging, and as an aid during the validation process. The simulation is microscopic, with random assignment of individual driver attributes. Other features include introducing any type of traffic control system, simulation of individual collision situations, and grade and curve effects.

The model can accept any freeway road configuration under any possible traffic condition for up to 50 miles of a 4-lane freeway. The car-following, lane-changing, and off-ramp exiting logic used is relatively simple as compared to other microscopic types of models. The on-ramp gap-acceptance merge algorithms are much more complex; on-ramp configurations allow merging with or without an acceleration lane or with an auxiliary weaving lane connecting adjacent on- and off-ramps. The model also allows lane restrictions to be specified (trucks to remain in right lane; bus expressway lanes). The computer program is written in FORTRAN IV with the exception of a group of small machine-language (COMPASS) routines for fast data packing and unpacking. The computer running efficiencies for the freeway sections simulated vary from a computer-time to simulated-time ratio of 1:7 for a 2-mile, 4-lane section with 1 on-ramp and 1 off-ramp, a flow of 6,000 vehicles/hour, and a 1-second scan interval to 1:0.75 for a 6.5-mile section of the Los Angeles Hollywood Freeway with 8 on-ramps and 7 off-ramps, a maximum flow of 8,400 vehicles/hour, and a ½-second scan interval, on the CDC 7600. Model validation includes the following:

1. Time-headway study based on data collected on the Eisenhower Expressway in Chicago;
2. Passenger-car velocity distributions at different lane volumes, using results published in the 1965 Highway Capacity Manual;
3. Off-ramp exiting behavior based on data collected by the Institute of Transportation and Traffic Engineering, UCLA, at the White Oak off-ramp on the Ventura Freeway in Los Angeles; and
4. Merging and weaving studies to test the capability of the merge algorithm based on freeway and ramp volume data collected on a 4.1-mile section of the Hollywood Freeway in Los Angeles.

Although this model possesses many desirable features, there is insufficient program documentation for more detailed evaluation.

A summary of the 15 models is given in Table 1 for a clear-cut comparison of each model's capability and structure.

CONCLUSIONS AND RECOMMENDATIONS

As we have seen from the foregoing, each model has its own merit and may be uniquely qualified for a particular application. One prime consideration in simulating large sections of freeway is the computer costs involved. To this end, separate special-purpose models may have an edge on general-purpose models. For example, a model developed for the purpose of incident detection could be much simpler and more economical than a general freeway simulation model. Such a simple model would be one that includes only the characteristics relevant to the occurrence of an incident. However, it is desirable for individual special-purpose models to be modular and compatible to each other so that they can be put together to simulate a variety of operational practices. This is the place where, without standardization, the adaptation can hardly be achieved.

A general-purpose simulation model should have the following features:

1. Unrestricted freeway geometry, or a collection of geometrics so that the right one can be selected for each simulation.
2. Simple car-following roles. Different rules may be required for free-flow and constraint-flow regions. Lane densities under 20 vehicles/mile would constitute the free-flow region where traffic is sparse and vehicles behave essentially independently of one another. A mean free-flow speed would be specified in this region.
3. Simple lane-changing logic that needs only to be statistically valid. A good example is the logic implemented in model 4. The simplicity of the car-following and lane-changing logic requirement is mainly for the gain of computer running efficiency.
4. Ramp control capability, so that different control strategies can be tested and evaluated.
5. Merging algorithms, cooperation with merging vehicles, and driver accommodation to temporarily accept low headways. Separate subroutines are provided for each of these features in model 2.
6. Varied vehicle characteristics, to distinguish between passenger and commercial vehicles. Many existing models provide such variety.
7. Varied driver characteristics.
8. Lane restrictions. This would include priority lanes as in model 13 as well as other restrictions as discussed in connection with model 15.
9. Incident generation procedures. This feature would allow the model to simulate and detect an incident.
10. Simulation of vehicle sensors at various freeway locations, as discussed in model 10. This feature is of particular interest in freeway incident detection.
11. Grade and curvature effect and weather and environment effect. This allows the model to adjust its parameters due to changes in these factors.

There may exist some other important features that need to be included. A complete list of requirements can be reasoned only after a careful investigation of the needs and an in-depth discussion with highway personnel. If each component can be built modularly and tested separately, then it is a simple matter to add more components to the general-purpose simulation model as the need arises.

Therefore, we can use any general-purpose simulation model such as model 9 or model 15 as a framework to modularize its individual components and include all the required capabilities. The resulting model is thus flexible, expandable, and economical to use. Furthermore, the computer program should be written in a high-level language such as FORTRAN that is suitable for execution in any general-purpose computer, and it should be well-documented so that it can be widely used with minimum effort.

REFERENCES

1. Gerlough, D. L. Simulation of Traffic Flow. HRB Spec. Rept. 79, 1964, pp. 97-188.
2. Goodwin, D. N. Examination of Nine Freeway Simulation Models. Final Report, System Development Corp., contract FH-11-7628 for FHWA, July 1972.
3. Richard, R. M., Baker, R. L., and Sheldon, W. P. Simulation of Traffic Flow as a Basis for Interchange Design. Final Report, Arizona Transportation and Traffic Institute, for Arizona Highway Dept. and FHWA, June 1965.
4. Kobett, D. R. A Digital Simulation Model of Freeway Traffic, Volume I: Model Description. Final Report, Midwest Research Institute, contract CPR-11-3661 for FHWA, Dec. 1968.
5. Kobett, D. R., Levy, S. L. Study of Expressway Traffic Flow Through Digital Simulation. Final Report, Midwest Research Institute, contract CPR-11-0963 for FHWA, May 1965.

6. St. John, A. D., Kobett, D. R., Sommerville, D., and Colanz, W. D. Traffic Simulation for the Design of Uniform Service Roads in Mountainous Terrain, Volumes I-IV. Final Report, Midwest Research Institute, contract CPR-11-5093 for FHWA, May 1970.
7. Worral, R. D., and Bullen, A. G. R. Lane-Changing on Multi-Lane Highways. Final Report, Northwestern Univ., contract CPR-11-5228 for FHWA, Aug. 1969.
8. Sinha, K. C. The Development of a Digital Simulator for the Analysis of Freeway Traffic Phenomena. Univ. of Connecticut, PhD dissertation, 1969.
9. Sinha, K. C., and Dawson, R. F. The Development and Validation of a Freeway Traffic Simulator. Highway Research Record 308, 1970, pp. 34-47.
10. Leland, S. D. A General Traffic Flow Simulation Model for Freeway Operation. Final Report, Connecticut Dept. of Transportation, May 1970.
11. Buhr, J. H., Meserole, T. C., and Drew, D. R. A Digital Simulation Program of a Section of Freeway With Entrance and Exit Ramps. Highway Research Record 230, 1968, pp. 15-31.
12. Nemeczky, J. A., and Widdice, B. D. Development and Validation of a Digital Simulation Model for Research in Geometric Design and Control of Diamond Interchanges. Proc. Summer Simulation Conference, Vol. II, sponsored by ACM/SHARE/SCI, 1970.
13. Warnshuis, Paul. Simulation of Traffic Flows Through Interchange Systems of Arbitrary Configuration. System Development Corp., contract FH-11-7628 for FHWA, March 1972.
14. Mikhalkin, B. Estimation of Roadway Behavior Using Occupancy Detectors. Univ. of Southern California, PhD dissertation, Feb. 1971.
15. Gazis, D., Herman, R., and Rothery, R. Non-Linear Follow-the-Leader Models of Traffic Flow. Operations Research, Vol. 9, No. 4, 1961, pp. 545-567.
16. Wildermuth, Bruno R. Effect of Lane Placement of Truck Traffic Freeway Flow Characteristics. Final Report, Wilbur Smith and Associates for Georgia Highway Department, Project 6909, April 1971.
17. Wicks, David. Traffic Flow Simulation Study: The SCOT Model, Volumes I-IV. Final Report, General Applied Science Laboratories, Inc., contract DOT-TSC-161, Feb. 1972.
18. Network Flow Simulation for Urban Traffic Control System. Final Technical Report prepared by Peat, Marwick, Mitchell and Co. and General Applied Science Laboratories, Inc., FHWA Report FH-11-7462-2, 1971.
19. Lieberman, E. Dynamic Analysis of Freeway Corridor Traffic (DAFT). General Applied Science Laboratories, Inc., TR-744, May 1970.
20. Blankenhorn, R. C., and May, A. D. FREQ2: A Revision of the "FREQ" Freeway Model. Special Report, Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, June 1972.
21. May, A. D., Minister, R. D., Lew, L. P., and Ovaici, K. A Computer Simulation Model for Evaluating Priority Operations on Freeways. Special Report, Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, contract FH-11-7704 for FHWA, June 1972.
22. Payne, H. J. Aggregate Variable Models of Freeway Traffic. Dept. of Electrical Engineering, Univ. of Southern California, NSF grant GK-24520, June 1972.
23. Harju, G. W., Bush, L. R., Kramer, R. F., and Porges, H. S. An Advanced Computer Concept for Freeway Traffic Flow Modeling. Presented at Summer Simulation Conference, June 1972.

A REVIEW OF THE TRAFFIC FLOW PROCESS

Sidney Weiner, Federal Highway Administration,
U.S. Department of Transportation

This paper summarizes the results of Wardrop, Edie, Haight, Breiman, and others in the study of space and time distributions of speed and other traffic characteristics. In order to help the traffic engineer properly assess mean speeds and other traffic characteristics, the underlying methodology is clarified and its application illustrated by examples of real-life situations. Such applications involve data from manual counts, several types of detectors, and aerial photography.

•THIS expository paper reviews the work of various investigators on the definition and measurement of traffic parameters. An attempt is made to unify these results to make them more understandable to traffic engineers. The use of the theory is illustrated by examples taken from realistic situations. These examples help the traffic engineer to apply the theory properly in the measurement of mean speeds and other characteristics when various detecting methods are used to record traffic data.

Traffic flow is a rather complex process when one considers some of the characteristic variables that can be associated with an individual vehicle: speed at an instant of time, speed at a particular location, location at an instant of time, number of passengers, distance and time separation of vehicles, and quite a few more. To study this field it is necessary to focus attention on several important variables and consider how they behave under uniform conditions of roadway and environment and under conditions of light to medium flow. As shown by Wardrop's results (17), it is convenient initially to make a simplifying assumption about one of the variables, speed: namely, that any vehicle is considered to have one speed associated with it in order to reflect uniform conditions. Another assumption by Wardrop has to do with the arrival process of a vehicle having a given speed. This paper will dwell mainly on the theory relating to stationary flow.

In the basic works of Wardrop (17) and Lighthill and Whitham (11), the important quantities such as flow, concentration, time-mean speed, and space-mean speed were defined and the relations between them explored. The results of these examinations are remarkable in view of the limited data base available to these and other researchers at that time. Wardrop looked at uniform traffic that was fairly homogeneous in space (over a stretch of roadway) at any instant of time and in time (period of observation) at any location on the road. He then developed several important relationships by means of an ingenious intuitive argument. Lighthill and Whitham required only time homogeneity and developed a local relationship for the uniform condition between flow and concentration by considering road traffic by analogy as a stream and by building a fluid continuum model involving three characteristics of streams: flow (quantity per unit time), concentration (quantity per unit space), and speed (space per unit time). Mathematical relations were studied as to how they varied over space and time so that the situation of traffic on long, crowded roads could be formally modeled under nonuniform conditions. For these time-inhomogeneous conditions, Lighthill and Whitham used their results to study congested flow and bottlenecks.

The variety of definitions applied to measurements in the traffic stream was reviewed by Edie (7). Relations between apparently different definitions of the same characteristics were clarified. First it was brought out that relations between flow, concentration, and speed are meaningful only when their averages are considered. Next it was advocated that the correct type of average be employed, space-mean or time-mean, in forming such relations. This would depend on the type of measurement that was employed: one type made at a point (or short distance) in space taken over a long interval of time and the other type taken at an instant of time (or short interval) taken over a long roadway. Arithmetic means computed from the first type are referred to as time means (averages over time) while those computed from the latter type are referred to as space means (average over distance). In another paper, Edie et al. (8) examine a large sample of speed, concentration, and flow data gathered through the use of electronic instrumentation in the Holland Tunnel in order to study time- and space-inhomogeneous situations.

Recently, three fundamental studies on traffic data and models were made by Breiman (2, 3, 4). The first paper reviews the data base, models, and statistical results for one-way homogeneous multilane traffic flow. The second paper, employing the methodology of stochastic processes, first derives the following relation developed by Lighthill and Whitham under locally homogeneous flow:

$$q = k\,\bar{v}_s \tag{1}$$

where q and k are average flow and concentration and $\bar{v}_s$ is the space mean speed defined by Wardrop. The paper then establishes the relation between the space and time distribution of speeds. In the third paper Breiman provides a further clarification as to interpretation of reduced aerial data and derives the fundamental theorem that relates the space distribution of speeds and headways to obtainable synchronous data involving these variables.

In the following sections a heuristic development is made that reflects the results contained in the papers of Wardrop, Edie, and Breiman. It is important, however, that full recognition be given to the many contributions and studies by other researchers that preceded or were contemporaneous to these. Some, such as those by Weiss and Herman (19), Brieman (1), Thedéen (16), and Renyi (15), consider the statistical properties of traffic under low density, while others, such as Miller (12), Buckley (5), Gafarian et al. (6), and Munjal and Hsu (13), explore the behavior of traffic by empirical investigations and application of the theory.

HOMOGENEOUS DISCRETE TRAFFIC STREAM MODELS OF WARDROP

To develop Wardrop's relations, it is necessary to make some formal assumptions as to the possible behavior underlying traffic characteristics in order to study its measurement. We will consider three basic quantities that need to be measured. These are flow, concentration, and speed. As a start, consider a simple model in which the overall process of vehicle speeds $\{V\}$ may be considered as a stream that consists of C (finite) superimposed substreams $\{S\}$. In this model the following assumptions are made to describe it:

1. Any vehicle has associated with it only one speed, v_i.
2. Any vehicle belongs to the ith substream, S_i, only if its speed is exactly equal to v_i.
3. Vehicles are considered as moving points determined from the corresponding locations on the vehicle (i.e., front bumper) and as traveling without interfering with one another.
4. Vehicles proceed on the right lane of a 2-lane section of a 4-lane divided highway, and whenever a point overtakes and passes any point it does so by using the left lane prior to overtaking and immediately merging to the right lane upon passing.
5. For each substream, the vehicles enter one end of a very long roadway at completely random instants of time, constituting a Poisson process of events.

Some of these assumptions could be modified, but in any case there is eventually achieved a homogeneity of traffic after some amount of time has elapsed from the initial entering if one looks at a large section of the road downstream. After such time has elapsed, traffic is called time-homogeneous, which means that any and all time averages converge to a limiting average for long time intervals.

Similarly, traffic is called space-homogeneous if space averages converge to a limiting average for a long enough distance. It has been shown (3) that the limiting averages for both space- and time-homogeneous traffic are equivalent.

It is interesting to note that Wardrop defines a random series of events in time as a series of events in which (a) each event is completely independent of any other event and (b) equal intervals of time are equally likely to contain a given number of events. But this implies that assumption 5, involving Poisson events, would hold. However, Breiman (1) has shown that one can start with an arbitrary homogeneous speed distribution in space and obtain a limiting Poisson spatial distribution under the assumption that cars can pass freely. Similarly, Thedéen (16) concludes that both time and space counts eventually tend to form a Poisson process.

Relationship Involving Space Mean Speed

With these preliminaries we can now present Wardrop's relations on a statistical basis or in a frequency interpretation setting. First look at the process of vehicles in an individual substream, S_i. Since the vehicles in S_i are identified with their own arrival process, which is Poisson or completely random, the quantity q_i (cars per hour) is associated with the arrival rate or traffic intensity parameter, while the time interval T_i between the instants of arrivals of such vehicles obeys the exponential distribution whose density function is given by

$$f_{T_i}(t) = q_i \exp(-q_i t) \qquad 0 \le t < \infty$$

From the expectation of T_i, the average time interval between vehicles passing an observer stationed at a fixed point adjacent to the road is then $1/q_i$. During this time interval, the vehicle is going at fixed speed v_i so that the average distance traveled in this average time is v_i/q_i. This means that, on the average, each vehicle in the ith substream is separately located somewhere along a distance of road that is v_i/q_i units long at any instant of time. It then follows that the average number of substream vehicles per unit length of road (concentration) is given by the reciprocal of this distance or

$$k_i = (v_i/q_i)^{-1} = q_i/v_i \qquad (i = 1, 2, \ldots, C) \tag{2}$$

If $k = \sum_1^C k_i$ denotes the concentration of the entire traffic stream, the discrete frequency distribution in space of vehicles whose speed is v_i is then defined by the multinomial probability

$$p_s(i) = \text{Prob}\,(V_s = v_i) = k_i/k \qquad (i = 1, 2, \ldots, C) \tag{3}$$

Thus $p_s(i)$ is the assigned probability space measure to the body of vehicles in the ith substream. Taking expectations, the space mean speed is obtained as follows:

$$\bar{v}_s = E(V_s) = \sum_1^C v_i p_s(i) = \sum_1^C v_i k_i/k$$

$$\bar{v}_s = \sum_1^C q_i/k \qquad \text{(applying Eq. 2)}$$

$$\bar{v}_s = q/k \tag{4}$$

where $q = \sum_1^C q_i$ is the composite flow of all the substreams or simply the sum of the arrival rates. Equation 4 was first developed by Wardrop and is identical to Eq. 1 here. This is the only valid relation that connects average flow, average concentration, and average speed.

Equation 4 can be used directly in obtaining the space mean speed if one has an unbiased estimate of each k_i/k, which necessitates the observation of vehicle separations on a long roadway at an instant of time. Even reduced aerial data do not provide a long enough distance, as pointed out by Breiman (4). Therefore, we will next consider the alternative method of estimating $\bar{v}_s$ by examining time measurements at a point on the road.

Relationship Involving Harmonic Mean of Time Speeds

Let us consider measurements of speeds as vehicles in the composite stream pass a given point on the road over a long interval of time. We shall designate this time-speed process by $\{V_T\}$ to distinguish it from the process of speeds over space $\{V_s\}$ previously examined. By applying the frequency interpretation for the probability that any vehicle passing the point will have speed $V_t = v_i$, we obtain

$$\text{Prob}\{V_t = v_i\} = p_t(i) \tag{5}$$

where $p_t(i)$ is approximated by n_i/n, n_i being the number of vehicles having speed v_i and $n = \sum_1^C n_i$. Hence we can approximate $p_t(i)$ by

$$\hat{p}_t(i) = \frac{n_i}{n} = \frac{n_i/T}{n/T} = \frac{\hat{q}_i}{\hat{q}} \tag{6}$$

where T is the period of observation and $\hat{q}_i$ is the observed arrival rate per unit of time. For large T, we can assume that $\hat{q}_i$ and $\hat{q}$ are equivalent to the underlying corresponding traffic intensities q_i and q, so that we have

$$p_t(i) = q_i/q \qquad (i = 1, 2, \ldots, C) \tag{7}$$

upon applying the frequency interpretation for probabilities.

Now consider the expected value of V_t^{-1} given by

$$E\{1/V_t\} = \sum_1^C \frac{1}{v_i} p_t(i)$$

Upon substitution of Eqs. 7 and 2 in turn we get

$$E\{1/V_t\} = \sum_1^C \frac{1}{v_i}\frac{q_i}{q} = \frac{1}{q}\sum_1^C k_i = \frac{k}{q} \tag{8}$$

From Eq. 8 we learn that the reciprocal of an individual time speed is an unbiased estimate of k/q or

$$q = k\left[E(1/V_t)\right]^{-1} \tag{9}$$

in contrast to the relation involving q, k, and space mean speed $\bar{v}_s$.

As a practical consideration, no one would use a single observation on V to estimate the expected value of the population in this case. Any individual speed could only relate to one of C denumerable substreams. One then considers a random sample of n successive speeds passing a point denoted by $V_1, V_2, \ldots, V_n$. Employing stationarity, these speeds can be considered to be identically distributed in the multinomial population defined by Eq. 7 and in addition may be dependent on each other. These variables obey the law of large numbers under certain conditions that imply that the covariance between any two sample speeds V_i and V_{i+m} tends to zero as the lag m increases (14, chapter 10). The law of large numbers informs us that the sample mean approaches the population mean so that, for large n,

$$E\left\{\frac{1}{n}\sum_{1}^{n}\frac{1}{V_i}\right\} \approx \frac{k}{q} \qquad (10)$$

But, from our sample,

$$(1/n)\sum_{1}^{N}(1/v_i) = \bar{v}_h^{-1}$$

where $\bar{v}_h$ is the harmonic mean of the time speeds. Hence the harmonic mean $\bar{v}_h$ of the time speeds can be employed as an asymptotic estimate in the homogeneous traffic relation, as follows:

$$q \approx k\,\bar{v}_h \qquad (11)$$

in contrast to Eq. 1. By employing the harmonic mean of speeds obtained from the time process to estimate the mean speed in the space process, one can thus correctly formulate the fundamental relation in Eq. 1.

In the foregoing treatment, the symbols q and k were used to indicate population parameters, where q represented an underlying flow and k represented an underlying concentration. This was done in order to be consistent with their historical treatment in the literature. It is unfortunate that this same treatment has confused these symbols with their observed measurements. Thus, on presenting the following section on examples involving the harmonic mean, the quantitites q and k will be perceived to represent measured quantities in order to be consistent with another body of the literature on traffic measurements. It would have been preferable to use the symbols λ for the underlying flow (replacing q) and ψ for the underlying concentration (replacing k). It is hoped that this dual use of the symbols q and k will not prove to be confusing to the reader.

Examples Using Harmonic Mean

Example 1: Manual Volume Counts—Manual traffic counts are used to obtain flow in traffic surveys where perhaps it is desired to know only the volume of traffic that affects an intersection in order to establish a warrant for signalized control or redesign. This type of method is also employed when other mechanical equipment cannot be readily installed. It is customary to start the count at the start of an hour or the start of a 15-minute period. This is called asynchronous counting, relating to the fact that a vehicle may not be at the location at the start of the count and similarly the count does not end specifically at the instant of arrival of the last (uncounted) vehicle. Synchronous counting refers to initiating the time period at the arrival of a vehicle and terminating the count at the arrival of an uncounted vehicle (10). The asynchronous count data are typically easier to acquire and for large counts would closely approximate the synchronous method.

If N are the number of vehicle counts in a time period T, then the flow (vehicles per unit time) is computed as $q = N/T$. Under light flow, the observed flow per unit time can be considered to have a Poisson distribution with arrival rate λ. Hence for time period T, the number of vehicles N has a Poisson distribution with mean λT. The observed ratio, q, has a mean equal to λ, since

$$E(q) = E(N/T) = \frac{1}{T}E(N) = \frac{\lambda T}{T} = \lambda$$

However, the value T may itself be considered to represent approximately the sum of N vehicle headways (times between front bumpers) so that $T = \sum_{i=1}^{N} h_i$. We may therefore write

$$q = \frac{N}{T} = \frac{N}{\sum_{1}^{N} h_i} = \left[\frac{1}{N}\sum_{1}^{N} h_i\right]^{-1} = \frac{1}{\bar{h}}$$

where $\bar{h}$ is the mean headway. This is only an approximation of the actual situation because, if N vehicles were actually counted, then the time interval T would represent the sum of N-1 headways and 2 partial headways.

Now consider the total observed flow q to consist of the sum of N individual flows, or let $q = \sum_{1}^{N} q_i$, where $q_i = 1/h_i$ in which q_i and h_i are respectively the instantaneous flow and headway associated with each vehicle.

Then we may write

$$q = 1/\bar{h} = \frac{1}{\frac{1}{N}\sum h_i} = \left(\frac{1}{N}\sum\frac{1}{q_i}\right)^{-1}$$

Thus the average flow, when computed from individual flows associated with each vehicle, is the harmonic mean of the individual flows.

Example 2: Detector Measurements—There are various methods of reducing traffic data from measurements taken from a detector or pair of detectors at a location.

Method 1: Pneumatic Tubes (BPR Traffic Analyzer)—A pair of pneumatic tubes are stretched across a given lane on the roadway; these tubes are usually separated by a distance d, 8 to 10 ft apart. When a vehicle's front tires cross over the first tube a signal is sent to a counter to register its arrival time and when the front tires cross the next tube another counter records another arrival time. The difference of these arrival times, t_i, represents the traversal time—the time it took the vehicle to traverse the known distance d. Let us assume that T is the entire period of observation while N is the total count (9). Then the flow, speed, and concentration may be obtained by means of the following formulas.

Individual speeds: $v_i = d/t_i \quad (i = 1, 2, \ldots, N)$

Space mean speed estimate:

$$\bar{v}_s \approx \frac{Nd}{\sum t_i} = \frac{N}{\sum t_i/d} = \left[\frac{1}{N}\sum\frac{1}{v_i}\right]^{-1}$$

= harmonic mean of the spot speeds

$= \bar{v}_h$

Flow: $q \approx N/T$

Concentration: $k = q/\bar{v}_s \approx \frac{N}{T}\cdot\frac{\sum t_i/d}{N} = \frac{1}{d}\frac{\sum t_i}{T}$

Method 2: Tape Switch or Occupancy Detector—Another type of method to directly measure flow q and speed $\bar{v}_s$ is from a simple occupancy detector or tape switch over a particular lane during a data sampling period of duration T. Let

N_T = number of vehicles that traversed the detector during interval T;
O_T = estimated portion of the time T that the axles of the vehicle were sensed by the detector (occupancy);
t_i = traversal time of the ith vehicle length sensed by the detector; and
L = average length of vehicles.

Then we can form the following computations.

Occupancy: $O_T = \sum t_i$

Individual speeds: $v_i = L/t_i$
Space mean speed estimate:

$$\bar{v}_s \approx N_T \frac{L}{O_T} = N_T \frac{L}{\sum t_i} = N_T \left(\sum t_i/L\right)^{-1}$$

$$= \left(\frac{1}{N_T}\sum_{1}^{N_T}\frac{1}{v_i}\right)^{-1} = \bar{v}_h$$

= harmonic mean of the spot speeds

Flow: $q \approx N_T/T$

Concentration: $k \approx q/\bar{v}_h = \frac{N_T}{T}\frac{\sum t_i/L}{N_T}$

$$= \frac{\sum t_i}{TL} = O_T/(TL)$$

Method 3: Detectors Involving Two Classes of Vehicles—Consider two classes of vehicles such as passenger and commercial vehicles that can be distinguished by height sensors installed under overpasses. Denote the measurements on occupancy, number of vehicles, and average vehicle length by O_i, n_i, and L_i, where $i = 1, 2$ designates passenger and commercial vehicles respectively.

In order to use Eq. 11, it is necessary to obtain an estimate for

$$\bar{v}_h = \left[\frac{1}{\sum_{1}^{2} n_i}\sum_{i=1}^{2}\sum_{j=1}^{n_i} v_{ij}^{-1}\right]^{-1}$$

where v_{ij} is the jth measurement for the speed of a vehicle in the ith class. However, the quantity $\sum_{1}^{n_i} v_{ij}^{-1}$ is estimated by $\sum_{1}^{n_i} t_{ij}/L_i = O_i/L_i$, where t_{ij} is the time measurement for the jth vehicle in the ith class.

Hence the approximation for $\bar{v}_s$ is given by

$$\bar{v}_h = \sum_{1}^{2} n_i \left[\sum_{1}^{2} O_i/L_i\right]^{-1}$$

If we now write

$$O_i/L_i = \sum_{1}^{n_i} t_{ij}/L_i = \sum_{1}^{n_i} v_{ij}^{-1} = n_i\ (\bar{v}_{h,i})^{-1}$$

where $\bar{v}_{h,i}$ is the harmonic mean speed in category i, we have for the combined harmonic mean

$$\bar{v}_h = \sum_{1}^{2} n_i \left[\sum_{1}^{2} n_i/\bar{v}_{h,i}\right]^{-1}$$

as the estimate of the space mean speed. This relation is easily extended to apply to several classes of vehicle lengths instead of only two. Reference is made to the corresponding formula found in Weinberg et al. (18), which differs from the above.

The calculations for flow and concentration are $q = \sum_{1}^{2} n_i \Big/ T$ and $k = q/\bar{v}_h$ respectively, where T is the total observation time.

Relationships in Wardrop's Discrete Model

One can find basic relationships that relate time and space distributional properties. The first to consider is that

$$p_s(i) = \frac{\bar{v}_s}{v_i} p_t(i) \qquad i = 1, 2, \ldots, C \tag{12}$$

where $p_s(i)$ and $p_t(i)$ are the corresponding space and time probabilities for Wardrop's substream or discrete model. This is the discrete analogue for the corresponding relation found in the continuous case by Haight (10) and by Breiman (3). Equation 12 is directly obtained by using Eqs. 2 and 4 in the definition of $p_s(i)$; i.e.,

$$p_s(i) = k_i/k = \frac{q_i/v_i}{q/\bar{v}_s} = \frac{\bar{v}_s}{v_i} p_t(i)$$

Another important relationship is that found by Wardrop:

$$\sigma_s^2 = \bar{v}_s(\bar{v}_t - \bar{v}_s) \tag{13}$$

where σ_s^2, $\bar{v}_s$, and $\bar{v}_t$ are respectively the space speed variance, space mean speed, and time mean speed. This is proved by employing the definitions of the variance and mean, as follows:

$$\sigma_s^2 = \sum_{i=1}^{C} (v_i - \bar{v}_s)^2 p_s(i) = \sum_{i=1}^{C} v_i^2 p_s(i) - \bar{v}_s^2 \tag{14}$$

where $\bar{v}_s = \sum_{1}^{C} v_i p_s(i)$ is the expectation of V_s.

By using Eq. 12, the summation term on the right reduces to

$$\sum_{i=1}^{C} v_i^2 \left(\frac{\bar{v}_s}{v_i} p_t(i)\right) = \bar{v}_s \sum_{1}^{C} v_i \frac{q_i}{q} = \bar{v}_s \bar{v}_t$$

which when substituted in Eq. 14 yields Eq. 13. It may be seen that more general relations involving the moments of the space and time speed distributions can be derived from Eq. 12. Thus, if $\mu_s'^{(r)}$ and $\mu_t'^{(r)}$ designate the corresponding r th moments of the space and time distributions about the origin, we have

$$\mu_s'^{(r+1)} = \bar{v}_s \mu_t'^{(r)} \tag{15}$$

For the corresponding moments about the mean (central moments) between the time ($\mu_t^{(r)}$) and space ($\mu_s^{(r)}$) distributions, the following formula can be established:

$$\mu_t^{(r)} = \sum_{j=0}^{r} \binom{r}{j} \left(\frac{\mu_s^{(j+1)}}{\bar{v}_s} + \mu_s^{(j)}\right) (\bar{v}_s - \bar{v}_t)^{r-j} \tag{16}$$

CONTINUOUS SPACE AND TIME SPEED DISTRIBUTIONS

The speed of a vehicle is generally considered to obey some unknown continuous distribution such as a Gaussian or gamma distribution instead of the discrete distribution considered by Wardrop in his substream model. In Haight (10) there is introduced a basic relation that connects space and time distribution of speeds for which an intuitive argument was provided. If one lets the space and time distribution of speeds be represented by the corresponding probability density functions $f_s(v)$ and $f_t(v)$, then, analogous to Eq. 12, the following is obtained:

$$f_s(v) = \frac{\overline{V}_s}{v} f_t(v) \qquad 0 \le v < \infty \tag{17}$$

where $f_s(v)$ and $f_t(v)$ are identically zero for $v < 0$.

Breiman (3) provides a rigorous proof for Eq. 17 that involves an analysis of the time-space process of speeds, and in fact his result is applicable to a more general distribution function that may involve discontinuities. Wardrop's substream model, involving a completely discrete or discontinuous set of probabilities, is in fact a special case of Breiman's result.

A heuristic proof of Eq. 17 may be developed from the discrete relation of Eq. 12. This development follows.

Let the range of V be finite, with minimum and maximum values of 0 and M respectively. Now partition the closed intervals (0, M) into n subintervals defined by (v_{i-1}, v_i) for $i = 1, 2, \ldots, n$, where $v_0 = 0$, $v_{i-1} < v_i$ and $v_n = M$. We can designate this partition by $I_n = \{v_{i-1}, v_i\}$.

If the randon variable V belongs to the ith interval, we can arbitrarily assign the value v_i to V. Thus for the partition of n intervals we can associate the probability that V assumes the value v_i by means of the probability $p(v_i)$. If the random variable is in a space process, we designate the probability by $p_s(v_i)$, and if it is in the time process, we designate the probability by $p_t(v_i)$.

Thus for the partition I_n we know from Eq. 12 that $p_s(v_i) = \overline{v}_s p_t(v_i)/v_i$.

Let the number of subdivisions be increased, with each interval being made sufficiently small so that with good approximation we have

$$p_s(v_i) \approx f_s(v)\,\Delta v \text{ and } p_t(v_i) \approx f_t(v)\,\Delta v$$

where

$$\Delta v = v_i - v_{i-1} \text{ and } v_i \approx v.$$

For any such fine partition we then have, applying Eq. 12,

$$f_s(v)\,\Delta v \approx \frac{\overline{V}_s}{V} f_t(v)\,\Delta v$$

which, upon division of both sides by Δv, completes our proof.

BREIMAN'S FUNDAMENTAL THEOREM

Previous results have provided us with an essentially unbiased expression for the mean space speed $\overline{v}_s$. Thus the harmonic mean $\overline{v}_h$ of the synchronous time speeds at a given point on the road is used to estimate $\overline{v}_s$. This may be put in the form (see Eq. 9)

$$\overline{v}_s = E_s(V) = \left[E_t(1/V)\right]^{-1} = \left[E_{sy}(1/V_o)\right]^{-1} \tag{18}$$

wherein the subscripts t and sy on the right side have almost identical meanings. Although the subscript t was previously used to indicate the synchronous speed of the vehicle or "sy" as it passed a ground detector, it could have been applied to the asynchronous time case discussed previously under Example 1: Manual Volume Counts. Similarly, the quantity V_o represents the observed speed of a car, C_o, when it reaches a designated location L_o. It would be equivalent to the quantity V in the above expression under E_t.

The estimate of the right side of Eq. 18 is the harmonic mean $\overline{v}_h$ of the synchronous speeds at L_o, or

$$\overline{v}_s \approx \left\{\frac{1}{n} \sum_{i=1}^{n} \frac{1}{v_{oi}}\right\}^{-1} = \overline{v}_h \tag{19}$$

where $v_{o1}, v_{o2}, \ldots, v_{on}$ represent the set of observed speeds of each successive car, C_{oi}, ascertained when it reaches L_o.

Breiman's theorem (4) allows us to form the unbiased estimate of any function of speed and headway in the space process in terms of a similar function in the synchronous time process. It therefore allows us to develop valid analyses of traffic data reduced from aerial photographs. From aerial data, not only joint speed characteristics but also headway measurements of successive vehicles are obtained.

As mentioned earlier, one wants to compile synchronous time data. On each frame a particular location on the roadway is referenced, say L_o. If one is interested in a particular lane, then we first see what the traffic looks like at the instant of time when the front part of the vehicle passes directly over that location. This vehicle is labeled C_o, and its downstream predecessors are $C_1, C_2, \ldots$, and its followers are $C_{-1}, C_{-2}, \ldots$.

At the instant of time when C_o reaches location L_o, say, the following joint set of synchronous time measurements is simultaneously obtained (provided of course that they appear on the same frame):

Symbol Identification	←Upstream	At L_o	Downstream→
Vehicle:	$\ldots, C_{-2}, C_{-1},$	$C_o,$	$C_1, C_2, C_3, \ldots$
Location:	$\ldots, L_{-2}, L_{-1},$	$L_o,$	$L_1, L_2, L_3, \ldots$
Space headway (in ft) or gap:	$\ldots, X_{-1}, X_{-1},$	$X_o,$	$X_1, X_2, X_3, \ldots$
Speed (ft/sec):	$\ldots, V_{-2}, V_{-1},$	$V_o,$	$V_1, V_2, V_3, \ldots$

In practice, the front bumper (location) of C_o may not be at L_o for any frame. In general, its location is ascertained by the linearly interpolated distance between two successive frames. This interpolation is similarly performed for the other vehicles. The speeds can be obtained by simply dividing the distance moved for each vehicle from one frame to the next by the frame lapsed time. Each time that vehicle C_{-1} reaches L_o, then C_{-1} is redesignated C_o, and all other vehicles are similarly relabeled.

By applying Breiman's powerful fundamental theorem on synchronous data (4), any function of the space headway and speed process, say $\phi(X, V) = \phi(X_1, \ldots, X_n; V_1, \ldots, V_n)$, may be estimated by

$$E_s\, \phi(X, V) = \overline{v}_s\, E_{sy}\left[\phi(X, V)/V_o\right] \tag{20}$$

where the left side represents the average of any arbitrary function ϕ of the space headway and speed process while the expectation E_{sy} on the right represents the average of the same function of the synchronous (time) headway and speed process, each divided by the synchronous speed V_o at L_o. Some examples for the use of Eq. 20 follow.

Example 1: Equation 18

Let ϕ be identically equal to 1 in Eq. 20. This is allowed because ϕ is arbitrary. Then, since on the left the expected value of a constant equals that constant, we have

$$1 = \overline{v}_s\, E_{sy}\,(1/V_o)$$

which is the well-known result that $\overline{v}_s$ is the harmonic mean of the speeds at a fixed spot and is estimated by Eq. 19 in terms of the sample harmonic mean $\overline{v}_h$.

Example 2: Equation 17

In Eq. 20, let $\phi(X, V) = \phi(V)$, which is the outcome of successive values of V_o or speeds when cars pass the origin, L_o. The expectation on the right side of Eq. 20 is the expectation of $\phi(V)/V$ under the time distribution of speeds while the left side is the expectation of $\phi(V)$ under the space distribution of speeds. Thus we write Eq. 20 as

$$\int \phi(v) f_s(v)\, dv = \bar{v}_s \int \frac{\phi(v)}{v} f(v)\, dv$$

Since this holds for all functions $\phi(v)$, it certainly holds for

$$\phi(v) = \begin{cases} 1 & \text{where v is included in the interval } (v', v' + \Delta v') \\ 0 & \text{where v is not included in the interval } (v', v' + \Delta v') \end{cases}$$

From this we obtain for any v

$$f_s(v') = \bar{v}_s \frac{1}{v'} f_t(v')$$

which gives an alternate proof for relation 17.

Example 3: Variance of the Space Speed (σ_s^2)

By letting $\phi(V) = V^2$ in Eq. 20, we have

$$E_s V^2 = \bar{v}_s E_t(V) = \bar{v}_s \bar{v}_t$$

from which we obtain

$$\sigma_s^2 = E_s V^2 - \bar{v}_s^2 = \bar{v}_s (\bar{v}_t - \bar{v}_s)$$

which is another well-known result.

Example 4: Expectation of Headway Distances in Space

Let $\phi = X_o$ and apply the fact that $E_s X_o = 1/k$. Equation 20 then becomes

$$\frac{1}{k} = \bar{v}_s E_{sy}(X_o V_o)$$

or

$$\frac{1}{q} = E_{sy}(X_o/V_o) \approx \frac{1}{n} \sum_1^n X_{oi}/V_{oi}$$

which is similar to one of Edie's formulas (7, Table 1) for measurements at a point. That is, let X_{oi} approximate $1/k_i$ while we let $\frac{1}{V_{oi}}$ approximate k_i/q_i (using Wardrop's Eq. 2.2). Then we obtain the corresponding relation,

$$\frac{1}{q} = \frac{1}{n} \sum_1^n (1/q_i)$$

Example 5: Expression as an Arithmetic "Mean"

In general, for any function ϕ, Eq. 20 provides an operational method of estimating the space expectation of any function of speeds and headways. For large N, we can write the theorem as

$$E_s\,\phi(X, V) \approx \overline{v}_h \cdot \frac{1}{N} \sum_{i=1}^{N} \phi\left\{\left(X^{(i)}, V^{(i)}\right)/V_o^{(i)}\right\}$$

where $\overline{v}_h = \left\{\frac{1}{N} \sum_{1}^{N} \frac{1}{V_o^{(i)}}\right\}^{-1}$

Thus Breiman's recipe for estimation is

1. Look at those time instants at which C_o passes L_o.
2. At those instants, calculate speeds and headways counting upstream and downstream.
3. Find the value of the function ϕ for these measurements; i.e., if $\phi = (X_o + X_1)\,V_o$, then the succession of measurements is represented by the vector

$$\{\phi_i\} = \left[\frac{X_{o1} + X_{11}}{V_{o1}}, \frac{X_{o2} + X_{12}}{V_{o2}}, \ldots, \frac{X_{oN} + X_{1N}}{V_{oN}}\right]$$

where, at the first time instant, X_{o1} and X_{11} are headways of C_o and C_1 while V_{o1} is C_o's speed, and so on.

4. Divide each ϕ by the corresponding value V_{oi}, the speed of the car at the origin, L_o.
5. Take the arithmetic mean of ϕ_i/V_{oi} and multiply by the harmonic mean of the V_{oi}; i.e., the space distribution estimate is

$$E_s\left\{\frac{X_o + X_1}{V_o}\right\} \approx \left[\frac{1}{N} \sum_{1}^{N} \frac{1}{V_{oi}}\right]^{-1} \left[\frac{1}{N} \sum_{1}^{N} \frac{X_{oi} + X_{1i}}{V_{oi}^2}\right]$$

This example is given only for purposes of illustrating Eq. 20.

Many more illustrative examples can be formulated for the application of Breiman's theorem. One can obtain useful formulas for the dependence of successive headways or speeds of various order lags. It should be noted that the headways in this section were expressed in units of distance. Breiman calls these space headways as distinguished from time headways, which are expressed in units of time. Conventional nomenclature by highway engineers refers to Breiman's space and time headways as gaps and headways respectively. It could cause some confusion to discuss the space and synchronous distributions of "space headways", so it would perhaps be preferable to refer to Breiman's result as "the relation between the space and synchronous time distributions of gaps and speeds".

However, this relation also holds between the space and synchronous time distributions of headways and speeds, since a gap can always be expressed as a headway by simply dividing it by the speed, e.g.,

$$\phi = \frac{X_o^2}{V_o} = \frac{X_o^2}{V_o^2} \cdot V_o = H_o^2 \cdot V_o$$

where H_o is the corresponding headway for car C_o.

It should be stressed that the function $\phi(\,)$ does not have to involve V. In fact, it may be deduced from the derivation of Eq. 20 that any function of traffic involved in the carrier space process of speeds could have been substituted for gap (X) or headway (H). Thus, any of the characteristic traffic variables mentioned in the introduction could be substituted for X in order to obtain an unbiased estimate of its space mean. For example, one could obtain an unbiased estimate of the average number of vehicle occupants by using Eq. 20. This has been examined for several extreme cases as well as for an intermediate joint set of speeds and number of car occupants where the true space mean speed and mean number of occupants were known. It was ascertained that the usual method of the arithmetic mean number of occupants (ignoring speeds) may produce a slight bias for the intermediate case but could present a large bias in the extreme cases. In every instance, however, it was shown that the synchronous method of Breiman produced an unbiased estimate. This indicates the utility of traffic flow theory in allowing one to examine the validity of alternative methods as well as to provide an unbiased method of estimating traffic characteristics. This type of analysis is applicable to other measurement variables such as energy, age or make of vehicle, and proportion of heavy vehicles.

REFERENCES

1. Breiman, Leo. The Poisson Tendency in Traffic Distributions. Annals of Mathematical Statistics, Vol. 34, 1963, pp. 308-311.
2. Breiman, Leo. Data and Models in Homogeneous One-Way Traffic Flow. Transportation Research, Vol. 3, No. 2, July 1969, pp. 235-249.
3. Breiman, Leo. Space-Time Relationships in One-Way Traffic Flow. Transportation Research, Vol. 3, No. 3, Sept. 1969.
4. Breiman, Leo. A Relationship Between Space and Time Distributions in Homogeneous Traffic Flow and Its Applications. Transportation Research, Vol. 8, No. 1, Feb. 1974, pp. 39-44.
5. Buckley, D. J. A Semi-Poisson Model of Traffic Flow. Transportation Science, Vol. 2, No. 2, 1968, pp. 107-134.
6. Gafarian, A. V., Lawrence, R. L., Munjal, P. K., and Pahl, J. An Experimental Validation of Various Methods for Obtaining Relationships Between Traffic Flow, Concentration, and Speed on Multilane Highways. Highway Research Record 349, 1971, pp. 13-30.
7. Edie, L. C. Discussion of Traffic Stream Measurements and Definitions. Proc. Second Internat. Symp. on Theory of Traffic Flow, London, 1965.
8. Edie, L. C., Foote, R. S., Herman, R., and Rothery, R. Analysis of Single Lane Traffic Flow. Traffic Engineering, Vol. 33, No. 4, Jan. 1963, pp. 21-27.
9. Everall, Paul. Urban Freeway Surveillance and Control: The State of the Art. Federal Highway Administration, 1972.
10. Haight, Frank. Mathematical Theory of Traffic Flow. Academic Press, 1963.
11. Lighthill, M. J., and Whitham, G. B. On Kinematic Waves, II: A Theory of Traffic on Long Crowded Roads. Proc. Royal Society, London, Series A, Vol. 229, No. 1178, pp. 317-345.
12. Miller, A. J. An Empirical Model for Multi-Lane Road Traffic. Inst. of Highway and Traffic Res., Univ. of New South Wales, 1968.
13. Munjal, P. K., and Hsu, Y. S. Characteristics of Freeway Traffic. System Development Corp., TM-4638/017, 1972.
14. Parzen, E. Modern Probability Theory and Its Applications. John Wiley and Sons, 1960.
15. Renyi, A. On Two Mathematical Models of the Traffic on a Divided Highway. Journal of Applied Prob., Vol. 1, 1964, pp. 311-320.
16. Thedéen, T. A Note on the Poisson Tendency in Traffic Distributions. Annals of Mathematical Statistics, Vol. 35, 1964, pp. 1823-1824.
17. Wardrop, J. G. Some Theoretical Aspects of Road Traffic Research. Proc. Inst. Civil Eng., Part II, Vol. 1, No. 2, 1952, pp. 325-362.
18. Weinberg, Morton I., et al. Surveillance Methods and Ways and Means of Communicating With Drivers. NCHRP Report 28, Highway Research Board, 1966.
19. Weiss, G., and Herman, R. Statistical Properties of Low Density Traffic. Quarterly of Applied Math., Vol. 20, No. 2, 1962, pp. 121-130.

SPONSORSHIP OF THIS RECORD

GROUP 3—OPERATION AND MAINTENANCE OF TRANSPORTATION FACILITIES
Lloyd G. Byrd, Byrd, Tallamy, MacDonald, and Lewis, chairman

Committee on Traffic Flow Theory and Characteristics
Donald G. Capelle, Alan M. Voorhees and Associates, Inc., chairman
Patrick J. Athol, John L. Barker, Martin J. Beckmann, Martin J. Bouman, Kenneth A. Brewer, Donald E. Cleveland, Kenneth W. Crowley, Lucien Duckstein, Leslie C. Edie, H. M. Edwards, A. V. Gafarian, Denos C. Gazis, Daniel L. Gerlough, John J. Haynes, Edmund A. Hodgkins, James H. Kell, John B. Kreer, Leonard Newman, O. J. Reichelderfer, Richard Rothery, August J. Saccoccio, A. D. St. John, William C. Taylor, Joseph Treiterer, William P. Walker, Sidney Weiner, W. W. Wolman

K. B. Johns, Transportation Research Board staff

The organizational units and the chairmen and members are as of December 31, 1973.