TRANSPORTATION RESEARCH RECORD 869

Highway Capacity and Traffic Characteristics

TRANSPORTATION RESEARCH BOARD

NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1982

TRANSPORTATION RESEARCH BOARD
National Research Council
ERRATA 1982-1985

Special Report 201

page 17, column 2, second paragraph, should read
"Tools also need to change as the nature of options changes significantly. Emerging policy options are not largely focused on network-expansion investments, whereas traditional models were developed long ago to deal with such options."

Special Report 200

page 3, column 1
Change the caption for the bottom figure to
"A new AM General trolley bus starts down the 18 percent grade on Queen Anne Avenue North in Seattle in October 1979 (photograph by J. P. Aurelius)".

Transportation Research Record 1040

page ii
Under "Library of Congress Cataloging-in-Publication Data," delete "Meeting (64th: 1985: Washington, D.C.)" and "ISBN 0361-1981"

Transportation Research Record 1020

page 7, Figure 1
The histogram should reflect that the rail mode is represented by the black bar and that the highway mode is represented by the white bar.

Transportation Research Record 1017

page 19, column 1, 7 lines above Table 1
Change "ranged from 1 in.^{2} to nearly 30 in.^{2} of runoff" to "ranged from 1 area inch to nearly 30 area inches of runoff"

page 22, column 1, last line
Change "1 to nearly 30 in.^{2}" to "1 to nearly 30 area inches"

page 22, column 2, first line
Change "13 in.^{2}" to "13 area inches"

Transportation Research Record 1011

page 12, Figure 4
Figure does not show right-of-way structure for O-Bahn. See discussion on page 11, column 1, paragraph 3.

Transportation Research Record 996

page 49
Insert the following note to Figure 2:
"The contour lines connect points of equal candlepower."

page 49
Insert the following note to Figure 3:

"The candlepower contours are superimposed on a 'headlight's-eye-view' of a road scene. The candlepower directed at any point in the scene is given by the particular candlepower contour light that overlays that point.

For example, 1400 candlepower is directed at points on the pedestrian's upper torso. For points between contour lines, it is necessary to interpolate."

page 50
Insert the following note to Figure 3:

"Where

ρ = the azimuth angle from the driver's eye to a point P on the pavement;
θ = the elevation angle from the driver's eye to a point P on the pavement;
EZ = the driver's eye height above the pavement; and
DX, DY, DZ = the longitudinal, horizontal, and vertical distance between the headlamp and eye point.

Then

$$EX = EZ/Tan\ \theta \qquad HZ = EX-DZ$$
$$H1^2 = EZ^2 + EX^2 \qquad HX = EX-DX$$
$$EY = H1\ Tan\ \rho \qquad HY = EY-DY$$
$$H2^2 = H1^2 + EY^2 \qquad H3^2 = HX^2 + HZ^2$$
$$\alpha = Tan^{-1}(HZ/HX),\ \beta = Tan^{-1}(HY/H3),\ H4^2 = H3^2 + HY^2"$$

Transportation Research Record 972

page 30, column 2, 22 lines up from bottom
Reference number (5) should be deleted

page 31, column 2, 5 lines up from bottom
Reference number should be 5, not 4

page 34, column 2, 8 lines above References
Reference number (5) should be deleted

Transportation Research Record 971

page 31, reference 3
Change to read as follows:

Merkblatt fürLichtsignalanlagen an Landstrassen, Ausgabe 1972. Forschungsgesellschaft für das Strassenwes en, Köln, Federal Republic of Germany, 1972.

Transportation Research Record 965

page 34, column 1, Equation 1
Change equation to
$$r_u = \gamma_w \cdot h/\gamma \cdot z$$

where

γ_w = unit weight of water,
γ = moist unit weight of soil,
h = piezometric head, and
z = vertical thickness of slide.

Transportation Research Record 905

page 60, column 1, 9 lines up from bottom
Change "by Payne (6)" to " by us"

Table 1. Summary of interactions between signal-timing parameters and MOEs.

Timing Method	Parameter	Total Delay	Stops	Fuel Consumption	Emissions		
					HC	CO	NO$_x$
Manual	Cycle length	⊕	⊕	⊕	⊕	⊕	⊕
	Speed of progression	+	⊕	+	+	+	+
	Priority policy	+	+	+	+	+	+
	Split method	+					
TRANSYT	Cycle length	⊕	⊕	⊕	⊕	⊕	⊕
	K-factor	+	⊕				
	Priority policy				+		

Note: + = main effect detected from TRANSYT output, and ◯ = main effect detected from NETSIM output.

Contents

Authors of the Papers in This Record

Ballard, John L., Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588

Chang, Edmund Chin-Ping, Texas Transportation Institute, Texas A&M University System, College Station, TX 77843

Crabtree, Joseph D., Kentucky Transportation Research Program, University of Kentucky, Lexington, KY 40506

Cunagin, Wiley D., Texas Transportation Institute, Texas A&M University System, College Station, TX 77843

Deacon, John A., Department of Civil Engineering, University of Kentucky, Lexington, KY 40506

Dudek, Conrad L., Texas Transportation Institute, Texas A&M University System, College Station, TX 77843

Easa, Said M., School of Engineering, Lakehead University, Thunder Bay, Ontario, P7B 5E1, Canada; formerly with Institute of Transportation Studies, University of California, Berkeley

Eck, Ronald W., Department of Civil Engineering, West Virginia University, Morgantown, WV 26506

Huber, Matthew J., Department of Civil and Mineral Engineering, University of Minnesota, Minneapolis, MN 55455

Lieberman, Edward B., KLD Associates, Inc., 300 Broadway, Huntington Station, NY 11746

Machemehl, Randy, Department of Civil Engineering, Ernest Cockrell, Jr. Hall 6.900, University of Texas at Austin, Austin, TX 78712

McCoy, Patrick T., Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588

McNees, Roger W., Texas Transportation Institute, Texas A&M University System, College Station, TX 77843

Nemeth, Zoltan A., Department of Civil Engineering, Ohio State University, 2070 Neil Avenue, Columbus, OH 43210

Polus, Abishai, Department of Civil Engineering, Technion-Israel Institute of Technology, Haifa 3200, Israel; formerly with Department of Civil Engineering, West Virginia University, Morgantown, WV 26506

Richards, Stephen H., Texas Transportation Institute, Texas A&M University System, College Station, TX 77843

Ross, Paul, Office of Research, HRS-31, Federal Highway Administration, 400 7th Street, S.W., Washington, DC 20590

Rouphail, Nagui M., Department of Materials Engineering, University of Illinois at Chicago Circle, P.O. Box 4348, Chicago, IL 60680

St. John, A.D., Midwest Research Institute, 425 Volker Boulevard, Kansas City, MO 64110

Tobin, John R., Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588

Tsou, Kai-Chu, Transportation Planning Division, Taipei City Government, Taipei, Taiwan, R.O.C.; formerly with Department of Civil Engineering, West Virginia University, Morgantown, WV 26506

Wijaya, Yahya H., Department of Civil Engineering, University of Nebraska-Lincoln, Lincoln, NE 68588

1

Use of Additional Through Lanes at Signalized Intersections

PATRICK T. McCOY AND JOHN R. TOBIN

One method sometimes used to increase the capacity of signalized intersections is to widen the roadways through the intersections to provide additional through lanes. The degree to which these additional through lanes actually improve the efficiency of traffic operations at the intersections depends on the extent to which they are used by through vehicles. The objectives of this research were to (a) observe the use of additional through lanes, (b) evaluate the effect of the length of these lanes on their use by through vehicles, and (c) incorporate the findings into the critical-movement-analysis technique in Transportation Research Board (TRB) Circular 212. Lane-use studies were conducted during peak periods on five signalized intersection approaches that had additional through lanes that ranged from 800 to 1200 ft in length. The lane-use data were collected for more than 700 signal cycles that were fully utilized by through vehicles. Analyses of these data determined that use of these lanes by through vehicles fit a Poisson distribution with a mean that was a linear function of lane length and green time. It was concluded that (a) the lane-use factors of the critical-movement-analysis procedure in TRB Circular 212 generally overestimate the use of additional through lanes by through vehicles and (b) length requirements for additional through lanes based on vehicle storage considerations are too short to achieve an average use of additional through lanes of more than 1.5 through vehicles per cycle.

To increase capacity, roadways are often widened at signalized intersections to provide additional lanes. Usually these lanes are provided for turning movements, but, in some cases, additional through lanes are also required. A common situation on two-lane, two-way roadways is the addition of both a left-turn lane and a through lane, as illustrated in Figure 1. In such cases, the additional through lane is a curb lane, which is used by both through and right-turning vehicles.

The degree to which this additional through lane improves the efficiency of traffic operations at the intersection depends on the extent to which it is used by through vehicles. According to the assumptions of previously developed lane-distribution models ($\underline{1},\underline{2}$), use of the additional through lane by through vehicles depends on drivers' perceptions of the travel-time savings to be realized by using it. If a driver perceives that use of the lane will minimize his or her travel time, it will be used; otherwise, it will not. Therefore, to use the additional through lane, the driver must perceive the delay to be experienced by entering it in advance of the intersection and merging from it beyond the intersection to be less than the additional delay to be experienced on the intersection approach by not using it. Obviously, the delay of entering and merging from the additional through lane relative to the additional delay of not using it is reduced as the length of the additional through lane is increased. Thus, use of the additional through lane by through vehicles would be expected to be greater as its length is increased.

PROBLEM

The critical-movement-analysis procedure in Transportation Research Board (TRB) Circular 212 ($\underline{3}$) does not include lane-use factors specifically for through lanes of limited length. In fact, few capacity-analysis procedures do account for the effects of limited through-lane length on lane use and/or capacity at signalized intersections.

Capacity-analysis techniques developed by Leisch ($\underline{4}$) do include a theoretical method for determining

Figure 1. Additional through-lane geometry.
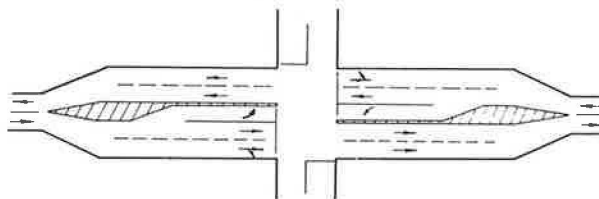
minimum-length requirements of additional through lanes. This method is based on vehicle-storage requirements, both in advance of and beyond the intersection, which are intended to prevent (a) the blocking of access to the additional lane by vehicles waiting on the intersection approach and (b) the blocking of the intersection itself by vehicles waiting to merge from the additional lane on the exit side of the intersection. However, these techniques are not applicable to capacity analysis of signalized intersections with additional through lanes that do not satisfy these minimum-length requirements nor has the adequacy of these length requirements been verified in the field ($\underline{5}$).

The Australian Road Capacity Guide ($\underline{6}$) does address the question of use of through lanes of limited length. With this guide, an additional through lane would be treated in the same way as a through curb lane blocked by parked vehicles on the approach and exit sides of the intersection. For approaches with three or more lanes and no vehicles parked within 100 ft back from the stop line, an average of 1.5 through vehicles per cycle would be assumed to use the blocked lane. This average use by through vehicles was determined from field studies conducted in Sydney ($\underline{7}$). It was also concluded from these studies that one parked vehicle 500 ft downstream from the intersection has as much effect on lane use by through vehicles as one parked only 200 ft downstream. This conclusion implies that to achieve an average through-vehicle use of an additional through lane of more than 1.5 through vehicles per cycle, the length of this lane beyond the intersection must be considerably more than 500 ft.

The Australian data suggest that the lane-use factors in TRB Circular 212 ($\underline{3}$) would be inappropriate for determining the level of service of a signalized intersection with additional through lanes. In addition, these data indicate that the minimum-length requirements of additional through lanes developed by Leisch ($\underline{4}$) may be too short to achieve the level of through-vehicle use that would be implied by the use of these lane-use factors. Thus, there is a need to determine the appropriate lane-use factor to be used in the critical-movement analysis of signalized intersections with additional through lanes.

OBJECTIVES

The objectives of the research reported in this paper were to (a) observe the use of additional through lanes at signalized intersections, (b)

quantify the effect of the length of additional through lanes on their use by through vehicles, and (c) incorporate the findings into the critical-movement-analysis procedure in TRB Circular 212 (3). This paper presents the procedure, findings, and conclusions of this research. Also, the application of the research results to the critical-movement analysis of signalized intersections is presented.

PROCEDURE

Lane-use studies were conducted during peak periods on five signalized-intersection approaches in Lincoln, Nebraska, in the spring of 1980. The approaches were on two-lane, two-way streets that had been widened at the intersections several years earlier. The widenings were done to add a left-turn lane and another through lane on the approaches. Although the approaches studied all had lane geometries similar to that shown in Figure 1, the total length of the additional through lanes ranged from just more than 800 ft to nearly 1200 ft. All the studies were conducted during fair weather and under dry pavement conditions.

The lane-use data collected during these studies were analyzed to determine the use of the additional through lanes by through vehicles. Comparisons were made among the approaches to determine the effects of additional-through-lane length on its use by through vehicles. The results of this analysis were then formulated for use in the critical-movement-analysis procedure in TRB Circular 212 (3), in order to make it more directly applicable to signalized intersections with additional through lanes.

Lane-Use Study

A lane-use study consisted of counting the number of vehicles that entered the intersection per cycle from each of the two through lanes on a study approach. Thus, during each cycle, a count was made on the approach of (a) the number of through vehicles discharging from the additional through (curb) lane, (b) the number of right-turning vehicles discharging from the additional through (curb) lane, and (c) the number of through vehicles discharging from the other (inside) through lane. The number of vehicles discharging from the left-turn lane was not recorded, because in the critical-movement-analysis procedure the left-turn volume would be assigned to the left-turn lane and adjusted separately from the through and right-turn volumes in the other two approach lanes. However, the left-turn operations were noted during the conduct of the lane-use studies to ensure that data were not collected if these operations interfered with those in the through lanes.

In an attempt to observe the most frequent use of the additional through lanes by through vehicles, the lane-use studies were conducted only during peak periods when the green times on the study approaches were fully utilized by the through vehicles in the inside through lane. Also, only data collected during these fully utilized cycles were used in the subsequent analysis. A cycle was considered to be fully utilized when the following conditions applied: (a) there are vehicles stopped waiting in the inside through lane on the approach when the signal turns green, (b) vehicles in the inside through lane continue to be available to enter the intersection during the entire phase and there is no unused time or exceedingly long spacings between the vehicles at any time due to lack of traffic, and (c) at least one vehicle in the inside through lane is stopped at the end of the phase when the signal turns red. These conditions are similar to those used in the 1965 Highway Capacity Manual (8) to define a loaded cycle.

Study Sites

All five of the study sites were level and were located at four-legged, right-angle intersections. There was no parking at any time on the streets on which the study sites were located. The posted speed limit on these streets was 35 mph.

As mentioned previously, the total lengths of the additional through lanes at these sites ranged from around 800 ft to nearly 1200 ft. The taper and lane lengths of these lanes, both in advance of and beyond the intersections, are given in Figure 2. Also given are the widths of the through lanes at each site, which were all within the 11- to 13-ft range.

The minimum-length requirements recommended by Leisch (4) for a 40-mph design speed are presented in Table 1. Comparison of these requirements with the lengths given in Figure 2 indicates that except for the taper lengths in advance of the intersections, all the lengths at the study sites exceeded these requirements.

During the periods when the lane-use studies were conducted, the signals at the study sites operated in an isolated, pretimed mode. The cycle lengths were 60 s, and the green times for the through and right-turn movements are shown in Figure 2. Although only one of the study sites (site 3) had a separate left-turn phase, all the approaches had sufficient left-turn capacity to accommodate their left-turn volumes without their interfering with the through and right-turn movements.

Traffic volume data obtained from the City of Lincoln prior to the conduct of the lane-use studies indicated that there would be less than 1 percent trucks and little pedestrian activity at the study sites in the peak hours during which the studies were to be conducted. Also, there were no scheduled local bus stops at any of the study sites. A critical-movement analysis of these volume data determined that according to TRB Circular 212 (3), an intersection level-of-service C existed at the study sites during the peak hours when the lane-use studies were conducted.

FINDINGS

Lane-use data were collected for more than 700 fully utilized cycles at the study sites. During these cycles, there was no pedestrian activity at the intersections, and, as expected, very few of these cycles had any trucks and/or buses included in the lane-use counts. Therefore, those few that did were excluded from the subsequent analysis. Thus, this analysis included only passenger-car data.

A summary of the sample means and standard deviations of the lane-use counts made on each approach is given in Table 2. Of course, with respect to the objectives of this research, the sample statistics for the number of through vehicles discharging from the additional through lane per cycle (STR) were of primary interest. It was noted that the mean values of STR were in general agreement with the Australian findings (7). Also, none of them exceeded the 1.5 vehicles per cycle, which is the value assumed in the Australian Road Capacity Guide (6) for the use of through lanes of limited length by through vehicles.

The relative frequency distributions of STR on the approaches are shown in Figure 3. Chi-square tests ($\alpha = 0.10$) of these distributions determined that each of them fit a Poisson distribution that had a mean equal to the sample mean value of STR shown in Table 2.

Figure 2. Study sites.
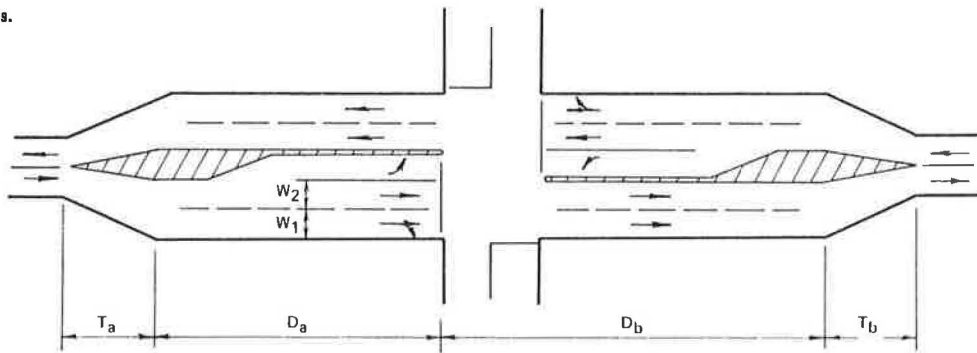
Study Site	T_a (ft)	D_a (ft)	D_b (ft)	T_b (ft)	$D_a + D_b$ (ft)	W_1 (ft)	W_2 (ft)	Green Time (sec)
1	155	365	465	260	830	11	11	23
2	145	475	495	205	970	12	11	21
3	65	580	460	240	1,040	12	11	20
4	155	600	480	580	1,080	13	12	30
5	130	515	680	285	1,195	13	12	27

Table 1. Minimum-length requirements.

Study Site	$T_a{}^a$ (ft)	$D_a{}^b$ (ft)	$D_b{}^c$ (ft)	$T_b{}^d$ (ft)
1	175	300	275	200
2	175	275	250	200
3	175	250	240	200
4	175	375	360	200
5	175	400	325	200

[a] Length of taper in advance of intersection.
[b] Length of additional through lane in advance of stop line.
[c] Length of additional through lane beyond stop line.
[d] Length of taper beyond intersection.

Table 2. Summary of lane-use data statistics (passenger cars).

Study Site	THRU[a] Mean	THRU[a] SD	STR[b] Mean	STR[b] SD	RT[c] Mean	RT[c] SD	No. of Cycles
1	9.19	1.32	1.23	1.00	1.02	0.98	160
2	8.35	1.42	1.37	1.09	1.39	1.09	175
3	7.62	1.76	1.41	1.24	0.78	0.85	189
4	13.04	2.18	1.23	1.04	0.54	0.69	99
5	11.41	0.95	1.36	1.10	2.83	1.70	103

[a] Number of through vehicles discharging from other (inside) through lane per cycle.
[b] Number of through vehicles discharging from additional through (curb) lane per cycle.
[c] Number of right-turning vehicles discharging from additional through (curb) lane per cycle.

It was anticipated that the mean number of through vehicles discharging from an additional through lane per cycle ($\overline{\text{STR}}$) might be affected by traffic, geometric, and signal-timing factors. Therefore, a stepwise multiple-linear-regression analysis was performed that used $\overline{\text{STR}}$ as the dependent variable and various traffic, geometric, and signal-timing factors as the independent variables. The specific geometric and signal-timing factors used in this analysis were the taper and lane lengths and the green times given in Figure 2. The traffic factors used were the mean number of through vehicles discharging from the other through lane per

Figure 3. Relative frequency distributions of $\overline{\text{STR}}$.
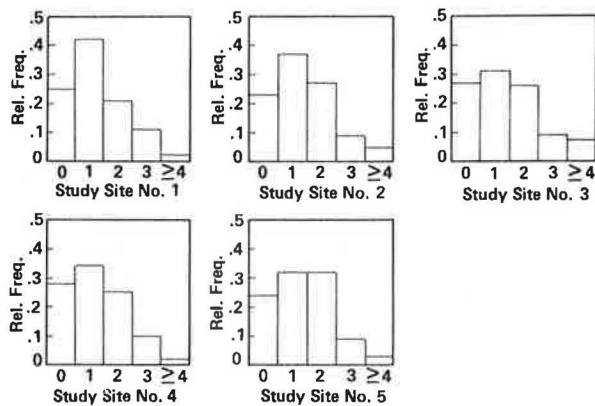
cycle ($\overline{\text{THRU}}$) and the mean number of right-turning vehicles discharging from the additional through lane per cycle ($\overline{\text{RT}}$). The values used for $\overline{\text{STR}}$, $\overline{\text{THRU}}$, and $\overline{\text{RT}}$ are shown in Table 2.

As a result of the regression analysis, the following relationship was found to be statistically significant ($\alpha = 0.01$):

$$\overline{\text{STR}} = 1.24 + 0.000\,58\,(D_a + D_b) - 0.021G \tag{1}$$

where

$\overline{\text{STR}}$ = mean number of through vehicles (passenger cars) discharging from additional through lane per cycle,
D_a = length of additional through lane in advance of stop line (ft),
D_b = length of additional through lane beyond stop line (ft), and
G = green time for through and right-turn movement on approach (s).

This relationship explained 99 percent of the variation in the observed values of $\overline{\text{STR}}$.

The relationship in Equation 1 is consistent with the assumption discussed earlier that the use of an additional through lane by through vehicles is directly proportional to the travel-time savings that drivers perceive would result from its use. The longer the additional-through-lane length ($D_a + D_b$), the greater the likelihood that travel-time savings would result from its use, and, as in Equation 1, the value of $\overline{STR}$ would be higher. Conversely, the longer the green time for the through and right-turn movement on an approach, the greater the number of through vehicles that can be accommodated per cycle in the other through lane and the lower the probability that travel-time savings would result by using the additional through lane. Also, as in Equation 1, the value of $\overline{STR}$ would be lower.

It is interesting to note that the results of the regression analysis, as well as those of the chi-square tests, indicated that the number of right-turning vehicles discharging from the additional through lane did not significantly affect its use by through vehicles. However, it must be remembered that the lane-use data were collected at intersections that had little pedestrian activity and where, as indicated in Table 2, the right-turn volumes were less than 25 percent of their respective through volumes. Therefore, the effect of right-turning traffic might be significant at intersections that had more pedestrians and higher right-turn percentages.

Also, it should be noted that meaningful application of Equation 1 is limited to signalized intersections that have a 60-s cycle and where the lengths of the additional through lanes and the green times for the through and right-turn movement are within the ranges of those of the study sites [800 ft < ($D_a + D_b$) < 1200 ft and 20 s ≤ G ≤ 30 s]. Within these ranges, as shown in Figure 4, $\overline{STR}$ varies from 1.1 to 1.5 passenger cars per cycle, which for a 60-s cycle amounts to a range in hourly flow rate from 66 to 90 passenger cars per hour. However, in any case where the length of the additional through lane is less than 1200 ft, the mean number of through vehicles discharging from the additional through lane should not be assumed to be greater than 1.5 passenger cars per cycle.

APPLICATION TO CRITICAL-MOVEMENT ANALYSIS

In the critical-movement-analysis procedure for operations and design presented in TRB Circular 212 (3), lane-use factors are applied in Step 8 to passenger-car volumes that have been adjusted in preceding steps for the effects of trucks, local buses, peaking, and turning movements. For approaches with lane geometries like that shown in Figure 1, which were the subject of this research, a lane-use factor of 1.05 would be applied to the through and right-turn movement volume. However, in view of the findings of this research, this lane-use factor probably overestimates the use of the additional through lane by through vehicles. Therefore, the derivation of a lane-use-factor equation in terms of the findings of this research follows.

For cases like those studied in this research, where the traffic volume in the additional through lane is less than or equal to that in the other through lane, the lane-use factor is computed as follows:

$$U = (2 \cdot \overline{THRU})/(\overline{THRU} + \overline{STR} + \overline{RT}) \qquad (2)$$

where

U = lane-use factor (1.00 ≤ U ≤ 2.00),
$\overline{THRU}$ = mean number of through vehicles (passenger

Figure 4. $\overline{STR}$ over study site ranges.
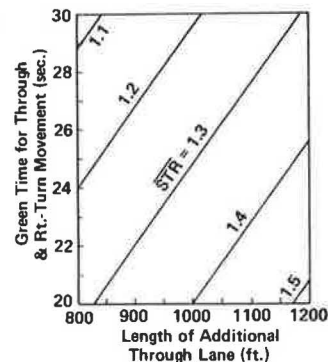
Figure 5. Lane-use factors for $\overline{STR}$ = 1.5 and C = 60 s.
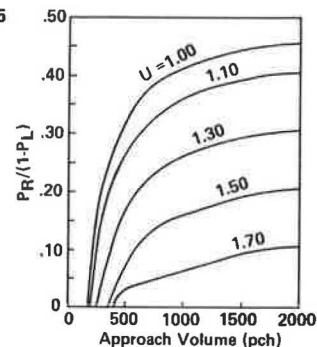
cars) discharging from other through lane per cycle ($\overline{THRU} > \overline{STR} + \overline{RT}$),
$\overline{STR}$ = mean number of through vehicles (passenger cars) discharging from additional through lane per cycle, and
$\overline{RT}$ = mean number of right-turning vehicles (passenger cars) discharging from additional through lane per cycle.

At intersections where there is little pedestrian activity (0-99 pedestrians per hour) (3), as was the case at the study sites, $\overline{RT}$ is expressed as follows:

$$\overline{RT} = [P_R/(1 - P_L)] \, (C/3600) \, V \qquad (3)$$

where

P_R = proportion of approach volume turning right,
P_L = proportion of approach volume turning left,
C = cycle length (s), and
V = approach volume (passenger cars per hour).

Also, V is expressed as follows:

$$V = (3600/C)(\overline{THRU} + \overline{STR} + \overline{RT}) \qquad (4)$$

from which $\overline{THRU}$ is determined:

$$\overline{THRU} = (C/3600) \, V - (\overline{STR} + \overline{RT}) \qquad (5)$$

If we substitute Equations 3 and 4 into Equation 2 and simplify, the equation for the lane-use factor becomes the following:

$$U = 2 \cdot \left(\left\{ 1 - [P_R/(1 - P_L)] \right\} - (3600/C)(\overline{STR}/V) \right) \qquad (6)$$

STR can be determined from Equation 1, or if Equation 1 is not applicable, a value of not more than 1.5 passenger cars per cycle can be used for addi-

tional through lanes of less than 1200 ft.

A plot of Equation 6 is shown in Figure 5 for a 60-s cycle and an $\overline{STR}$ of 1.5 passenger cars per cycle. This illustrates that the lane-use factor of 1.05, which is given in TRB Circular 212 (3), is appropriate for approaches with additional through lanes only when the percentage of right turns is relatively high. When the percentage is not high, Equation 6 should be used to avoid overestimating the use of additional through lanes by through vehicles.

CONCLUSIONS

Based on the findings of this research, the following conclusions were reached concerning the use of additional through lanes by through vehicles on signalized intersection approaches with lane geometries similar to that shown in Figure 1:

1. The lane-use factors of the critical-movement-analysis procedure in TRB Circular 212 (3) generally overestimate the use of additional through lanes by through vehicles. Therefore, in using this capacity-analysis procedure to evaluate the operations and design of signalized intersections with additional through lanes similar to those studied in this research, lane-use factors computed by the method presented in this paper should be used instead.

2. Length requirements for additional through lanes based on vehicle-storage considerations, such as those developed by Leisch (4), are too short to achieve an average use of the additional through lane by through vehicles of more than 1.5 passenger cars per cycle.

3. Use of additional through lanes by through vehicles is a function of the total length of the lane and the green time provided for the through and right-turn movement on the approach. It is positively correlated with length and negatively correlated with green time.

4. Use of additional through lanes by through vehicles is independent of the right-turn volume on signalized intersection approaches where there is little pedestrian activity and on which the right-turn volume is less than 25 percent of the through volume

5. The numbers of through vehicles per fully utilized cycle that use an additional through lane fit a Poisson distribution

Although the findings of this research were conclusive and consistent with those of Australian studies

(6,7), they are applicable to a limited range of traffic, geometric, and signal-timing conditions. Therefore, there is a need for further research to study the use of additional through lanes over wider ranges of lane lengths, signal timings, and traffic volumes than was possible within the limited resources of this research. Also, further studies should be conducted in other urban areas, where drivers' attitudes toward the confrontations associated with the use of an additional through lane and their perceptions of delay might be different from those of the driver population in Lincoln, Nebraska.

REFERENCES

1. E.B. Lieberman. Determining the Lateral Deployment of Traffic on an Approach to an Intersection. TRB, Transportation Research Record 772, 1980, pp. 1-5.
2. K.-L. Bang. Swedish Capacity Manual--Part 3: Capacity of Signalized Intersections. TRB, Transportation Research Record 667, 1978, pp. 11-28.
3. Interim Materials on Highway Capacity. TRB, Transportation Research Circular 212, Jan. 1980.
4. J.E. Leisch. Capacity Analysis Techniques for Design of Signalized Intersections. Public Roads, Vol. 34, No. 9, Aug. 1967.
5. Research Problem Statements. TRB, Transportation Research Circular 227, April 1981.
6. Australian Road Capacity Guide. Australian Road Research Board, Nunawading, Australia, Bull. 4, June 1968.
7. A.J. Miller. The Capacity of Signalized Intersections in Australia. Australian Road Research Board, Nunawading, Australia, Bull. 3, March 1968.
8. Highway Capacity Manual. HRB, Special Rept. 87, 1965.

Highway Sizing

JOSEPH D. CRABTREE AND JOHN A. DEACON

A critical examination is made of the conventional method for highway sizing, that is, the determination of lane requirements. Ranked hourly traffic-volume distributions, obtained from 1977 Kentucky volume stations, are examined to test certain assumptions common to the conventional approach. Assumptions regarding the existence and location of "knees" within these distributions, a common requirement of current procedures, are found to be of questionable validity. However, the fundamental fallacy of the conventional procedure rests with its focus on a single design hour and its orientation toward conditions experienced by the highway rather than by the user. This can readily be overcome by basing size decisions on an alternative criterion such as the percentage of vehicles that suffer congestion during the design life. An example demonstrating this concept is presented. At the same time, more significant improvement in sizing methodology can be achieved by directly computing the economic efficiency of investment in additional lanes. An example is presented to demonstrate current capabilities for such computations. The example also demonstrates that current procedures do not always yield the most economical designs and that the most economical highway size is affected by the specific form of the traffic-volume distribution. Use of economic efficiency analysis as a standard tool in evaluating important sizing decisions is highly recommended.

In many highway construction or reconstruction projects, one important decision is the number of lanes to be provided. Procedures used to determine lane requirements (highway sizing) are normally based on identification of a single design hour within which the anticipated demand volume [commonly the 30th highest hourly volume (HHV) in the design year] is balanced against supply volumes (capacities or service volumes) for the alternative highway sizes under consideration.

During the past three decades, conventional highway-sizing procedures have remained virtually unchanged. During this same period, other highway decisionmaking processes have changed markedly as emphasis has highlighted broad social concerns and environmental impacts and as competition for the public dollar has intensified. In view of this situation, it is appropriate to reexamine conventional sizing procedures. The project reported here was initiated to determine the soundness of these procedures and to identify, if necessary, possible techniques for improvement.

CURRENT METHODOLOGY

Development of the current sizing methodology is credited to Peabody and Normann. In 1941, by using the single design-hour volume versus capacity, they recommended use of a design-hour volume within the range of the 30th to the 50th HHV (1). Endorsements for use of the 30th HHV soon came from the American Association of State Highway Officials (AASHO) and the Committee on Highway Capacity of the Highway Research Board. AASHO, in 1945, adopted the 30th HHV for a year 20 years from the date of construction as the design-hour volume for the national system of Interstate highways, an adoption that, with only slight modifications, has remained in subsequent design standards (2). In 1950, the Committee on Highway Capacity recommended use of the 30th HHV as the normal design-hour volume (3). However, the Committee cautioned, as had Peabody and Normann, that the 30th HHV was not necessarily applicable in every instance and that it would "not always result in the best engineering practice" (3).

To understand the rationale for these recommendations, it is necessary to visualize the characteristic shape of the plot of a ranked hourly volume dis-

Figure 1. Typical ranked hourly volume distribution (station 16).
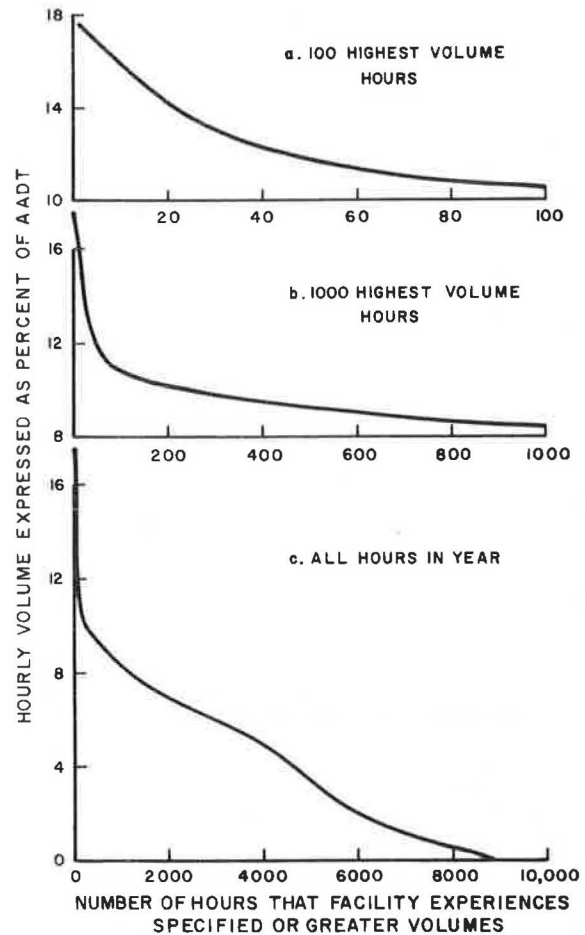
tribution. Figure 1a, constructed from hourly volume data obtained from one automatic traffic recorder (ATR) in Kentucky during 1977, is one such plot. The resulting curve seems to show a "knee" (a small region with a rapid change in slope) at or about the 30th HHV. After observing the regularity with which such a knee occurred in the region between the 30th and 50th HHV for a large number of highway locations, Peabody and Normann concluded that it was impractical to design for volumes greater than the 30th HHV and further that designs for volumes less than the 50th HHV would likely result in only small savings in construction cost but great loss to the expedition of traffic movement (1). Through the years, use of the 30th HHV seems to have been based to a large degree on the assertion that it yielded the most economical design, or, as stated by the Committee on Highway Capacity, it is at this point that the "ratio of benefit to expenditure is near the maximum" (3). Matson, Smith, and Hurd more subjectively argued (4): "The most equitable ratio between the service provided by the road and its costs will be achieved when the design volume is selected near the knee of the curve."

Although endorsement of the 30th-HHV design concept by these respected authorities contributed

greatly to its rapid and widespread adoption, at least one other factor was also of importance. The Committee on Highway Capacity had concluded that the 30th HHV, when expressed as a percentage of the annual average daily traffic (AADT) volume, changed very little from year to year (3). A future-year AADT prediction could be easily and accurately converted to the design-hour volume through application of what has come to be called the K-factor (the frequently measurable ratio of the 30th HHV) to the AADT. Confidence of the designer in the design-hour volume prediction was thus greatly enhanced.

The most authoritative current recommendations for highway sizing are those of the American Association of State Highway and Transportation Officials (AASHTO). AASHTO recommends use of an hourly volume representative of flows at the end of the design life, that is, 10-20 years following completion of construction. For rural highways with normal flow variations, the 30th HHV should be used. For rural highways with unusual or highly seasonal traffic fluctuation, the design hourly volume should be as follows (5):

about 50 percent of the volumes expected to occur during a very few maximum hours of the design year....A check should be made to insure the expected maximum hourly traffic does not exceed possible capacity.

For urban streets and highways, the design hourly volume should be the average of the 52 highest afternoon peak-hour volumes for each of the weeks in the design year. After observing that this average is not significantly different from the 30th HHV, AASHTO concluded (6):

Therefore, for use in urban design the 30th highest hourly volume can be accepted since it is a reasonable representation of daily peak hours during the year. Exception may be necessary in those areas or locations where concentrated recreational or other travel during some seasons of the year results in a distribution of traffic volume of such nature that a sufficient number of the hourly volumes are so much greater than the 30 HV that they cannot be tolerated and a higher value must be considered in design.

CRITIQUE

To evaluate the soundness of sizing procedures, one would prefer to examine a large number of past sizing decisions and determine, in retrospect, the fraction that was successful. Unfortunately, such an evaluation is very difficult, if not impossible, both because of the difficulty of acquiring the necessary data and because of the absence of an accepted criterion for defining success. The approach taken in this critique is therefore to focus on the identification of procedural difficulties and on an assessment of the validity of assumptions that undergird the decisionmaking process.

In applying the conventional procedure, the designer is continually challenged to determine when the 30th or 50th HHV should be used (for normal flows) or when other more appropriate measures should be sought (for unusual flows). This choice is one of increasing difficulty: There is simply a continuum of traffic-flow patterns reflecting the wide variety of travel desires served by individual facilities and their varying degrees of operational adequacy. Not only does this difficulty raise questions about procedural technique, but also an analysis of flow patterns suggests possible fallacies in underlying assumptions.

The conventional highway-sizing procedure draws its strength in part from the following four basic assumptions: (a) the ranked hourly volume distribution exhibits a discernible knee; (b) this knee occurs at or near the 30th HHV; (c) the knee defines the point of most economical sizing; and (d) the 30th HHV, expressed as a percentage of the AADT, remains constant over time.

To examine the first two of these assumptions, traffic-volume data collected in 1977 from 45 Kentucky ATR stations were analyzed. Three ranked hourly volume-distribution graphs for each station, similar to those of Figure 1, were constructed for use in the visual component of the analysis. Although most prior analyses had examined in detail only the 200 or so highest volume hours, the three different data sets were used here to identify any possible bias in the more conventional but also more limited examination.

The first portion of the analysis was a subjective one. Four observers were asked to independently examine each ranked hourly volume-distribution graph and to determine whether a knee could be discerned. They were told only that a knee was a small region on either side of which the slopes of the curve were markedly different. Figure 1 is typical of the situation in which there was general agreement among the observers that knees did exist. In Figure 1, the four observers located knees on the 100-h, 1000-h, and 8760-h graphs within the following ranges in ranks, respectively: 23rd-25th HHV, 70th-84th HHV, and 100th-200th HHV. Figure 2 is representative of graphs for which the observers had more difficulty locating knees. Three of the four observers were unable to locate knees on the 100-h and 1000-h graphs, and two did not find a knee on the 8760-h graph. The difficulty with the graphs in Figure 2 was that the curves, although well behaved, exhibited slopes that changed quite gradually with increases in rank. Any knee was therefore very difficult to identify.

The first part of Table 1, which summarizes this portion of the analysis, shows that there was a discernible knee in most instances and that the likelihood of finding a knee increased as the size of the data set increased. However, in a substantial percentage of cases (approximately 16 percent for the 100-h graphs), no knee could be found: These cases cannot be dismissed as mere exceptions. Also noted, although not shown by Table 1, is the fact that there were many cases in which individual observers disagreed over the existence of knees. Assuming that the observers were reasonably competent, this type of disagreement effectively demonstrates the subjective and somewhat vague nature of the knee-of-curve concept.

Observers were also asked to determine, where possible, the location of each knee. This subjective analysis was augmented by a more objective one employing a nonlinear regression program of the Statistical Analysis System (SAS). SAS was used to fit a segmented model to each set of volume data. This involved the optimal separation of each set of data into two subsets and the fitting of independent curves to each of the two subsets. Figure 3 typifies the results. The knee was assumed to occur at the intersection of the two fitted curves, the location labeled "boundary" in Figure 3. The remarkable similarity between the observer-reported knee locations and those determined by SAS gave much credibility to the SAS analysis. Although both linear and quadratic models were tested, they were found to yield similar boundary locations, and only results from the quadratic models are reported here.

The results of the analysis of knee-of-curve location are also summarized in Table 1. The first

striking observation is that the location of the knee is influenced drastically by the extent of the data set. This fact became readily apparent early in the research when graphs for individual stations were compared (see, for example, Figure 1): It was confirmed by both the visual and SAS analyses when the average ranks of Table 1 were determined. The sensitivity of the location of the knee to the amount of data is sufficient to cast serious doubt on the efficacy of knee-of-curve procedures. A knee

the location of which varies for a given data set with the method for graphically portraying that data would seem to be of questionable reliability.

Originally there was great interest in whether the knee occurred at or near the 30th HHV: Interest waned when it was conclusively established that the knee location was influenced by the number of hours within the data subset. A quick glance at the average ranks in Table 1 suggests that by selection of some subset of data between the 100 and 1000 highest volume hours, the location of the knee would average at or near the 30th HHV. At the same time, Table 1 shows that most of the knees were located outside the accepted range of the 30th-50th HHV for the data groupings employed here.

There was also much variability from station to station in the location of the knee. Results of the visual observations of the 1000-h graphs are shown in the following tabulation:

Location of Knee	Percentage of Stations
None	14
Between	
1st and 20th HHV	16
21st and 40th HHV	20
41st and 60th HHV	15
61st and 120th HHV	20
121st and 300th HHV	10
300th HHV and above	5

Certainly, those using knee-of-curve sizing proce-

Figure 2. Ranked hourly volume distribution showing indistinct knee (station 46).

Figure 3. Typical ranked hourly volume distribution showing segmented quadratic model of best fit (station 7-SB).
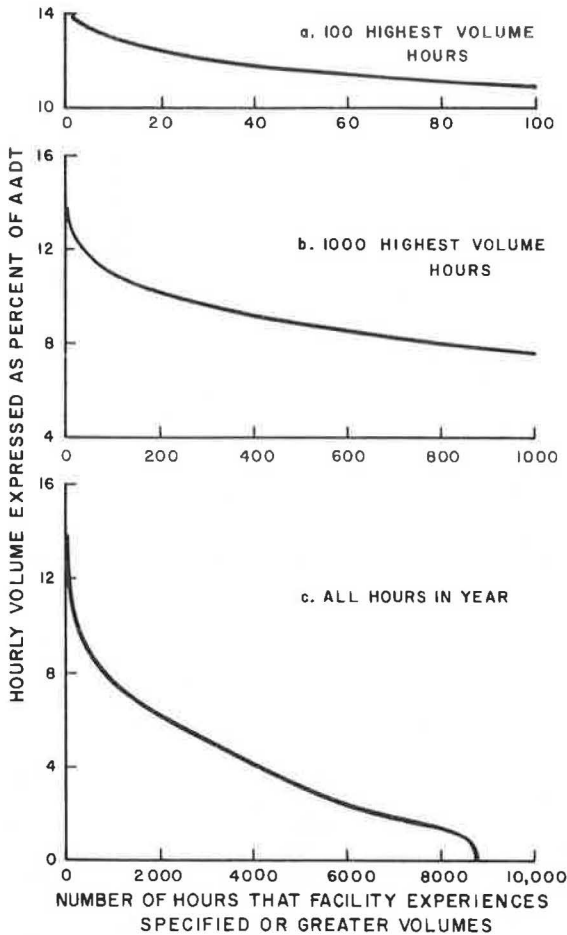
Table 1. Existence and location of knee for ranked hourly volume distributions.

Item	Graph of 100 Highest Volume Hours	Graph of 1000 Highest Volume Hours	Graph of All Hours in Year
Percentage of graphs with discernible knee (total for four observers)	83.8	86.2	91.2
Average rank of hour at knee location			
Range for four observers	6.6-9.9	47-82	310-620
Segmented model	19	110	360
Percentage of knee locations within 30th-50th HHV interval			
Average for four observers	0.6	33.3	0.0
Segmented model	11.1	0.0	0.0

dures would be well advised to determine the location of the knee of curve for each situation rather that to assume that it lies within the 30th-50th HHV range. This recommendation supports earlier work of Werner and Willis (7), who showed that the knee was not necessarily located at the 30th HHV and that it tended to lie within the 200th-600th HHV range for the larger AADTs.

A third assumption implicit in the conventional sizing procedure is that the knee defines the point of most economical sizing. Unfortunately, it has been impossible to conclusively prove or disprove this assumption. There is certainly an intuitive appeal to the argument that as one considers volumes to the left of the knee, construction costs would increase greatly while only a very few more hours or users would be accommodated. As one considers volumes to the right of the knee, very little is likely to be saved in construction cost but much would be sacrificed by the user as many additional hours would become congested. At the same time, it seems obvious that such a conclusion might be seriously distorted by focusing, as has been common in the past, on the few heaviest volume hours (perhaps 200) in some year 10-25 years in the future. In effect, the design to accommodate the future-year 30th HHV is very similar to the design to accommodate the maximum hourly volume in the design life, a design that most designers would consider to be inappropriate and uneconomical. Further to the point of economy in highway sizing, no study has been discovered in which any tests have been made or other objective evidence presented that supports the assumption that the knee defines the point of most economical sizing. At the same time, it is possible to demonstrate, as is done later in this paper, specific examples for which the knee does not define the most economical size.

The fourth assumption that has been important to widespread adoption of the conventional sizing procedure is that the 30th HHV, expressed as a percentage of the AADT, remains constant over time. Following such an assertion by the Committee on Highway Capacity in 1950 (3), a number of important studies have shown that the K-factor is not invariant and typically decreases with the increasing volumes that often accompany the passing of time. Among these studies are those of Walker (8), Bellis and Jones (9), Reilly and Radics (10), Chu (11), and Cameron (12). With this rather conclusive analysis, it was not imperative to examine the matter fully during this investigation. A superficial examination was made, however, of data from Kentucky ATR stations for 1973 and 1977. Between 1973 and 1977, the K-factor decreased for 28 of the 40 common ATR stations, increased for 8, and remained the same for 4. The average K-factor for the 40 stations decreased during this period from 11.5 to 11.2 percent. It is obvious, therefore, that the K-factor for a specific highway location is a time-variant quantity.

Conventional sizing procedures have been used with much success for many years, they are viewed quite favorably by design agencies, and their widespread use is likely to continue for many years. Those continuing to use these procedures, however, should consider implementation of changes suggested by the above analysis. The design-hour volume should be selected at the knee of the ranked hourly volume-distribution graph rather than at some arbitrarily chosen point such as the 30th HHV. In addition, the graph should contain all hourly volume data collected throughout the year rather than some arbitrarily chosen subset such as the 200 highest volume hours. Finally, since the pattern of traffic flow is likely to be different from location to lo-

cation, each site must be individually analyzed to ascertain what volume corresponds to the knee and how the K-factor is likely to vary through time. Other improvements, as identified and addressed in the following section, should also be considered for adoption.

EXTENSIONS

In examining the highway-sizing literature, two promising extensions to the conventional procedure were discovered. Because of their relative ease of implementation and because they overcome certain valid objections to the conventional procedure, they are described in this paper and their use is illustrated by means of examples. Hourly traffic-volume distributions used in these and subsequent examples are shown in Figure 4 and other traffic characteristics are described as part of the list of assumptions for the economic analysis given in the next section. The standard traffic distribution of Figure 4 is representative of the 1977 median for Kentucky ATR stations, whereas the alternate represents 1977 data for one particular station chosen because the hourly flows were less variable than those for the standard. Both distributions have K-factors of 11.2 percent, the 1977 median for Kentucky ATR stations.

The first extension, attributed to Glauz and St. John (13) and reported by the Institute of Traffic Engineers (ITE) Technical Council Committee 6F-2 (14), suggests a user orientation to design instead of the traditional facility orientation. The focus here becomes the percentage of time that the typical user experiences high-volume conditions rather than the percentage of time that the facility experiences such conditions. In the traditional approach, the highway is sized so that it will be congested no more than 30 h during the year or about 0.34 percent of the time. In the user-oriented approach, the highway would be sized so that the user would experience congestion no more than some acceptable percentage of the time. The difference between these approaches derives from the fact that a proportionally greater number of users travel during high-volume hours as compared with low-volume hours.

Figure 5 show the first 200 h of the traffic volume data of Figure 4 replotted to convert from number of hours to percentage of time and extended to show the difference between the user and facility orientations. To modify the conventional sizing procedure to the user approach requires use of ranked volume distributions for users rather than for facilities. The ITE report (14) describes the procedure in some detail. An individual plot, similar to that of Figure 5a, could be used to select a knee to support a specific design decision or a large number of such plots could be examined to locate the characteristic position of a knee or to otherwise derive an acceptable decision criterion.

The user approach is conceptually superior to the traditional one in that it more nearly recognizes the primary purpose of many highway developments--to provide an improved level of service to the road user. Practically, as suggested by Glauz and St. John (13), it offers a superior way to recognize and emphasize peculiar characteristics of recreational and other routes that have peaked-flow characteristics.

A second useful extension to the conventional sizing procedure derives from work of DeVries (15), also reported by ITE (14). To demonstrate the significance of DeVries' contribution, it is first necessary to emphasize that the conventional procedure is based on the concept of a single design hour. Lane requirements are determined by comparing the

Figure 4. Ranked hourly volume
distributions for examples.

NUMBER OF HOURS THAT FACILITY EXPERIENCES SPECIFIED OR GREATER VOLUMES
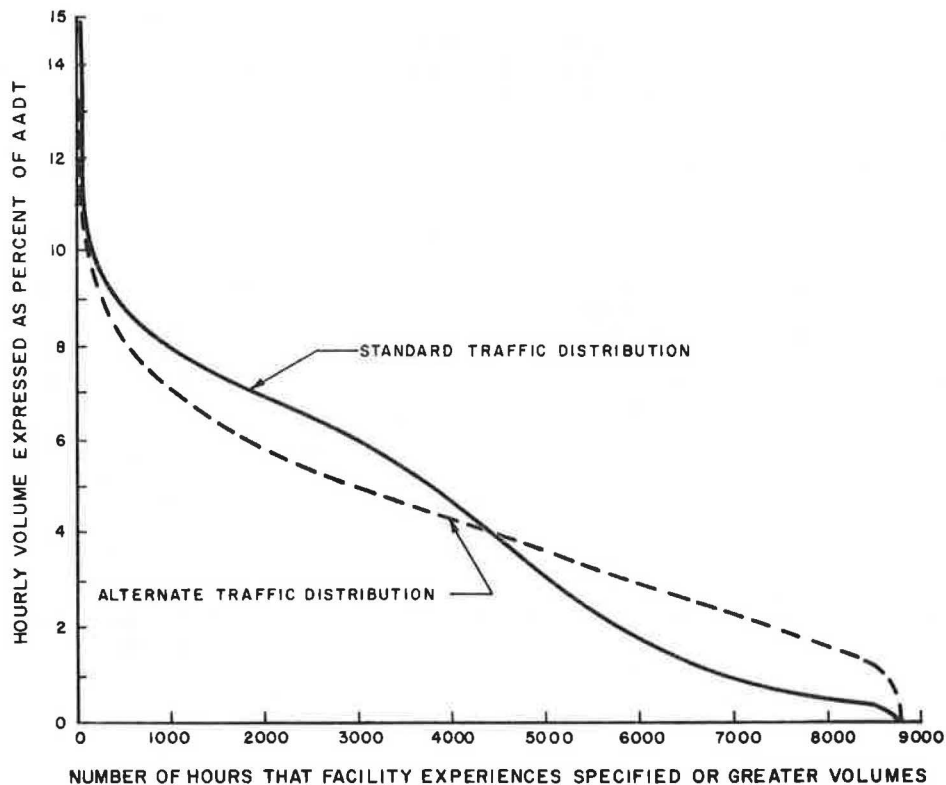
Figure 5. Ranked hourly volume distributions for both users and highway.

PERCENTAGE OF TIME THAT USERS OR HIGHWAY
EXPERIENCE SPECIFIED OR GREATER VOLUMES
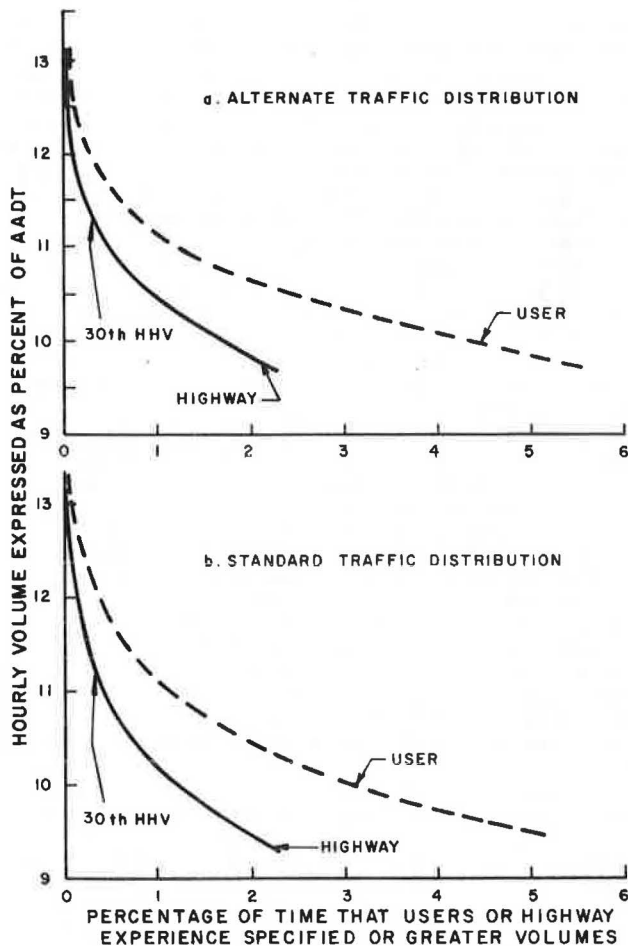
demand volume (design-hour volume) with the supply volume (service volume or capacity) for one particular hour during the entire design life of the highway. Is it not presumptuous to ignore conditions occurring during that overwhelming portion of the design life in which flow is more or less congested than in the design hour? Is it not also presumptuous to base such a design on demand and supply volumes that have been rather arbitrarily selected on the basis of the designer's intuition as to what conditions are acceptable to the traveler and what conditions result in the most economical design? Questions such as these lend credence to attempts to expand the focus from a single hour to a range of hours within the design life.

DeVries suggested that more prudent investment decisions for independent project analysis might result from investigations of the range of top hours (perhaps the highest volume 500) that were encompassed within the desired level of service. This concept might be implemented in any of several ways, which include specification of a minimum number of the top 500 h that must be included within the desired level of service.

As a variation of the DeVries proposal, which includes the Glauz-St. John user emphasis, sizing decisions might be based on the percentage of vehicles during the design life that suffer congestion. A simple but reasonable way to define congestion is in terms of operating conditions representative of D-, E-, or F-levels of service. The objective would be to base size decisions on a congestion level acceptable to the design agency. Figure 6 illustrates the output of such an analysis.

Figure 6 shows the traffic volume that would be subject to congestion on two-lane roadways for a range of future-year AADTs and the two different traffic distributions described earlier. Similar analysis showed that no congestion would be anticipated on four-lane facilities with volumes no greater than a future-year AADT of 14 000. The spe-

Figure 6. Influence of traffic volume on congestion of two-lane sample highway.
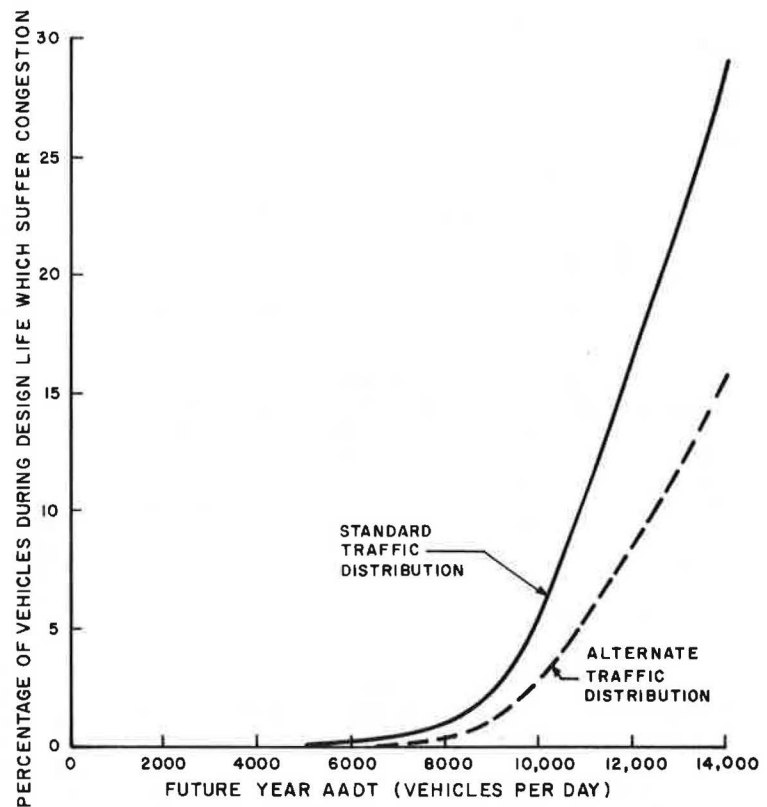
cific criterion for highway sizing in this example would have to be selected by the designer. Alternatives might be no congestion, some fixed level of congestion such as 2 percent, or even the knee of the curve. The knee is reasonably well defined in this example, and if that should prove to be true in other circumstances as well, the knee might furnish an acceptable heuristic decision point.

In summary, design to accommodate a single hour in the design life of a facility masks the reality of variable operational-flow conditions through time. This difficulty can and should be overcome by broadening the analysis to include a much larger time frame. Use as the decision criterion of the percentage of vehicles during the design life that suffer congestion accomplishes this objective as well as that of properly focusing on the user rather than the facility. Further testing and use of such a criterion seems warranted.

RECOMMENDED PROCEDURE

Highway-sizing decisions rank among the more important decisions confronting the designer or planner. Differential construction costs are measured in the hundreds of thousands of dollars per kilometer, and the cost of an additional pair of lanes will, in some circumstances, almost double construction outlays. Because of their importance, sizing decisions merit very critical analysis and should not be based on hunch and intuition. Although the conventional procedure can certainly be improved as indicated above, to accomplish what is really necessary requires a completely different perspective on the sizing task.

We contend that sizing decisions should be reached in the same manner as other major investment decisions. In whatever way has been found to be acceptable to each responsible agency, the gamut of both favorable and unfavorable consequences of the

sizing decision needs to be identified and evaluated. One such consequence that is often evaluated in public decisions involving the allocation of scarce resources is the economic efficiency of the investment. Economic analysis seems tailor-made to the sizing decision, since the primary impacts are often limited to savings to the road user and costs to the highway agency.

The technical literature abounds in information regarding economic analysis and its application to highway investment decisions. Maring (16) and Hutchinson (17) were among those who specifically advocated use of economic analysis in highway-sizing decisions. Although both presented useful examples to demonstrate their recommendations, effectiveness of these examples was limited by the data that were readily available when their work was performed. Publication of the authoritative manual on user benefit analysis by AASHTO (18) has helped to eliminate many of the earlier constraints to effective analysis. At the same time, it must be emphasized that economic analysis still involves a number of very important assumptions, any one of which can possibly affect the decision. Sensitivity analysis is a recommended technique for assessing the potential significance of the critical assumptions.

To demonstrate application of economic analysis, a hypothetical situation was defined in which a sizing decision was required on a new, 16.1-km highway. Future-year AADT was varied and two ranked hourly volume distributions, as shown in Figure 4, were independently investigated. Details of the analysis are identified below:

Traffic:

1. Growth of 3 percent compounded annually;
2. Composition of 85 percent cars, 10 percent single-unit trucks, and 5 percent combination trucks;

3. Directional split of 55 percent in direction of greatest flow; and
4. Ranked hourly volume distributions as shown in Figure 4.

Roadway:

1. Uninterrupted flow in rural area;
2. Design speed of 96.6 km/h and speed limit of 88.5 km/h;
3. Length of 16.1 km with 3.66-m lanes and 3.05-m shoulders;
4. No access control but four-lane highway has median;
5. Paved surface;
6. Rolling terrain with 11.3 km level, 3.2 km on a 1 percent grade, and 1.6 km on a 2 percent grade;
7. Tangent sections for 11.3 km and horizontal curvature of 1 and 2 degrees on lengths of 3.2 and 1.6 km, respectively; and
8. 100 percent of two-lane highway with passing sight distance in excess of 460 m.

Analysis:

1. 25-year period of analysis;
2. All costs expressed in constant (1975) dollars;
3. Discount rate of 5 percent;
4. Hourly time costs of $3.00 for cars, $7.00 for single-unit trucks, and $8.00 for combination trucks;
5. Construction costs of $615 000/km and $957 000/km for two-lane and four-lane highways, respectively;
6. Maintenance costs of $2660/(km·year) and $4320/(km·year) for two-lane and four-lane highways, respectively;
7. Residual value of $349 000/km and $560 000/km for two-lane and four-lane highways, respectively; and
8. Accident costs of $10.03/1000 vehicle-km and $8.78/1000 vehicle-km for two-lane and four-lane highways, respectively.

Insofar as practical, recommendations and data given by AASHTO (18) were used. Construction and maintenance costs were estimated on the basis of the Ken-

tucky experience, and accident costs as reported by AASHTO (18) were used.

The criterion chosen to represent economic efficiency was the net present worth (NPW) of four-lane as compared with NPW of two-lane construction. Benefits of the four-lane construction included savings in travel time and accident costs and an increase in the residual value of the investment. Greater costs for the four-lane facility were attributed to those of construction and maintenance as well as to increased vehicle operating costs occasioned primarily by increased speed.

Figure 7 summarizes the analysis. For the standard traffic distribution, two-lane construction is seen to be preferable for future-year AADTs less than about 9300 vehicles per day. This break-even volume increased to 9800 vehicles per day for the alternative traffic distribution. The fact that two different traffic distributions, although they have identical K-values and design hourly volumes, had different break-even volumes suggests that factors other than the location of the knee of the ranked hourly volume-distribution curve also influence the most economical design.

A comparison was also made between the break-even volumes of Figure 7 and those determined by the conventional sizing procedure. In the latter case, the break-even volume depends on which level of service is selected to represent acceptable congestion in the design hour. The future-year, break-even AADTs for the conventional analysis were determined to be approximately 4500, 7400, and 9300 vehicles per day for B, C, and D service volumes, respectively. Results from the conventional analysis and the economic analysis thus become comparable only for a level of service (D) that has normally been thought to be intolerable for all but exceptional design purposes. The conclusion, therefore, is that, for this sample problem and a rather wide range in future-year AADTs, the conventional sizing analysis would lead to a design decision different from that of an economic analysis. Of course, specific numbers reported here are unique to the given conditions, and generalizations based thereon are to be avoided.

The example of this section has demonstrated application of the techniques of engineering enonomy to the highway-sizing decision. It has also identi-

Figure 7. Economic efficiency of four-lane versus two-lane construction in sample highway.
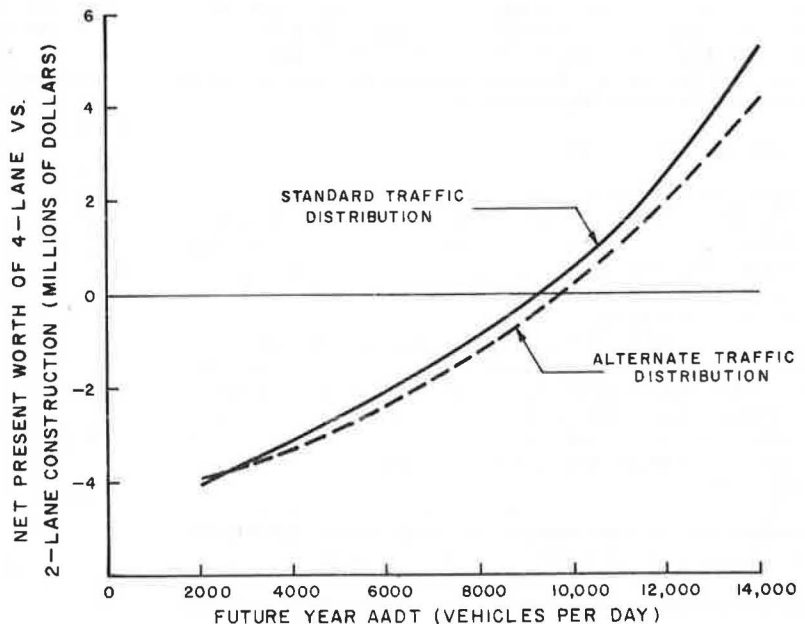
fied at least one situation in which the conventional sizing procedure yields a decision different from one based on the criterion of economic efficiency. We are convinced that techniques and data for performing competent economic analyses are readily available and are becoming more and more sophisticated. Further, we are convinced that the economic efficiency of additional-lane investments is one impact that should never be neglected in the sizing decision. At the same time, we are aware that other impacts are sometimes of paramount importance. Who cannot describe a situation in which a nearby cemetery, a row of stately shade trees, a bordering park, or any of a number of other situations has served to constrain the size of a highway improvement? The point is simply that economic efficiency, albeit important, is only one of the many impacts of the sizing decision that must be evaluated if prudent decisions are to be reached.

SUMMARY AND CONCLUSIONS

A critical examination has been made of the conventional method for highway sizing, that is, determination of lane requirements. Although this method has served admirably in the past, improvements can readily be made that will lead not only to more informed but also to more easily defensible decision-making.

The fallacy of the conventional method, which determines lane requirements by balancing a design-hour volume (demand) against service volumes for the alternative highway sizes (supply), rests with its focus on a single design hour as well as with its orientation to the facility rather than the user. It does not explicitly consider, therefore, the normal reason for increasing highway size, namely, the benefits that accrue through time to the users.

Further, some of the basic premises on which the conventional sizing methodology is based have been found to be invalid. Many ranked hourly volume distributions (nth-highest-hour plots) do not exhibit discernible knees, or small regions within which their slopes change markedly. Of those that seem to exhibit knees, knee locations vary among observers and are unquestionably and most inappropriately influenced by the number of hours of volume data being examined. Further, knees usually lie outside the normally accepted 30th-50th HHV interval. Traffic-volume data reported in this paper offer support to the prior conclusion of others that, at a given location, the K-factor (ratio of 30th HHV to the AADT) cannot be expected to remain constant through time and for underutilized facilities typically decreases as traffic volume increases. Finally, the conventional sizing methodology, although it has minimal data requirements and is quite simple to apply, cannot be expected to necessarily yield the most economical highway size decision.

Similar care and attention should be given to decisions regarding highway size as to other major highway investment decisions. The entire gamut of differential impacts, including such factors as the degradation of parks and historic places, aesthetics, and noise and air pollution, should, if possible, be evaluated. Of particular importance to this evaluation is the economic efficiency of the highway investment.

The capacity for using conventional highway economic analysis to aid highway-sizing decisions is well developed and readily available for immediate implementation. Its use is strongly recommended as a rational and defensible basis for supporting sizing decisions. However, for those who find this recommendation unacceptable, other improvements to the conventional methodology are suggested. The first involves focusing on the user instead of the facility by appropriately changing the abscissas of the ranked hourly volume-distribution plots and selecting the design-hour volume at the position of the relocated knee. The second would be more significant but would require a conceptual transition from a single-hour to a range-of-hours approach. A suitable decision criterion in this situation appears to be the percentage of vehicles during the entire design life that suffer congestion for the alternative highway sizes. A decision to increase highway size would be justifiable when the percentage of vehicles suffering congestion on the smaller facility was considered unacceptably large by the design agency.

REFERENCES

1. L.E. Peabody and O.K. Normann. Applications of Automatic Recorder Data in Highway Planning. Public Roads, Vol. 21, No. 11, Jan. 1941, pp. 203-222.
2. F.W. Cron. Highway Design for Motor Vehicles--A Historical Review: Part 2, The Beginnings of Traffic Research. Public Roads, Vol. 38, No. 4, March 1975, pp. 163-174.
3. Highway Capacity Manual: Practical Applications of Research. Bureau of Public Roads, U.S. Department of Commerce, 1950.
4. T.M. Matson, W.S. Smith, and F.W. Hurd. Traffic Engineering. McGraw-Hill, New York, 1955.
5. A Policy on Geometric Design of Rural Highways. AASHO, Washington, DC, 1966.
6. A Policy on Design of Urban Highways and Arterial Streets--1973. AASHTO, Washington, DC, 1973.
7. A. Werner and T. Willis. Cost-Effective Level of Service and Design Criteria. TRB, Transportation Research Record 699, 1979, pp. 1-7.
8. W.P. Walker. Trends in the 30th-Hour Factor. HRB, Bull. 167, 1957, pp. 75-83.
9. W.R. Bellis and J.E. Jones. 30th Peak Hour Factor Trend. HRB, Highway Research Record 27, 1963, pp. 1-13.
10. E.F. Reilly and R.D. Radics. 30th Peak Hour Factor Trend. HRB, Highway Research Record 199, 1967, p. 78.
11. B.P. Chu. Michigan's Statewide Traffic Forecasting Model--Volume III: Design Hour Volume Model Development. Michigan Department of State Highways and Transportation, Lansing, Nov. 1972.
12. N. Cameron. Determination of Design Hour Volumes. Department of Civil Engineering, Univ. of Calgary, Calgary, Alberta, May 1975.
13. W.D. Glauz and A.D. St. John. Recreational Travel Impacts--Procedures for Dealing with the Impacts. Midwest Research Institute, Kansas City; U.S. Department of Transportation, Interim Rept., Dec. 1977.
14. ITE Technical Council Committee 6F-2. Reexamination of Design Hour Volume Concepts. ITE Journal, Vol. 48, No. 9, Sept. 1979, pp. 45-49.
15. N.R. DeVries. The Role of Design Hour Volumes in Highway Planning and Design. Wisconsin Department of Transportation, Madison, Working Paper, Nov. 1975.

16. G.E. Maring. Weekend Recreational Travel Pat-
     terns. FHWA, Tech. Rept. 18, Feb. 1971.
17. B.G. Hutchinson. Economic Criterion for High-
     way Capacity Determination. Transportation En-
     gineering Journal, Vol. 98, No. TE3, Aug. 1972,
     pp. 465-475.

18. Manual on User Benefit Analysis of Highway and
     Bus-Transit Improvements, 1977. AASHTO, Wash-
     ington, DC, 1978.

Traffic Capacity Through Urban Freeway Work Zones in Texas

CONRAD L. DUDEK AND STEPHEN H. RICHARDS

Findings of capacity studies conducted at urban freeway maintenance and
construction work zones in Houston and Dallas are summarized. Studies
were conducted on five-, four-, and three-lane freeway sections. The results
indicate that the per-lane capacities are affected by the number of lanes open
during the roadwork. For example, the average capacity on a three-lane
section with two lanes open was 1500 vehicles per hour per lane (vphpl),
whereas the average capacity with one lane open was only 1130 vphpl.
Also illustrated is how the data can be used to estimate the effects of the
lane closure. The results of the study can be used in scheduling work that
involves lane closures on freeways.

Findings of capacity studies conducted at 28 main-
tenance and construction work zones on freeways in
Houston and Dallas are summarized. All these stud-
ies were made at sites where one or more traffic
lanes were closed. A total of 37 studies were con-
ducted at work zones while the work crew was at the
site; 4 studies were conducted while the work crew
was either not at the site or not occupying a closed
lane directly adjacent to one of the open lanes.

FREEWAY WORK-ZONE CAPACITY

Capacity with Work Crew at Site

Figure 1 illustrates the range of volumes measured
at several work sites while the work crew was at the
site. All volumes were measured while queues were
formed upstream from the lane closures and thus es-
sentially represent either the capacities of the
bottlenecks created by the lane closures or the ef-
fects of drivers staring because of the work crew
and machinery. Each point in Figure 1 represents
the volume observed during one study; therefore, it
is easy to view how the data cluster for each lane-
closure situation.

The formula (A,B) is used in this paper to iden-
tify the various lane-closure situations evaluated:
A represents the number of lanes in one direction
during normal operations; B is the number of lanes
open in one direction through the work zone.

The average capacity for each closure situation
studied is shown in the table below. The data show
that the average lane capacity for the (3,2) and
(4,2) combinations was approximately 1500 vehicles
per hour per lane (vphpl).

| No. of Lanes | | No. of | Avg Capacity | |
Normal	Open	Studies	Vehicles per Hour	Vehicles per Hour per Lane
3	1	5	1130	1130
2	1	8	1340	1340
5	2	8	2740	1370

| No. of Lanes | | No. of | Avg Capacity | |
Normal	Open	Studies	Vehicles per Hour	Vehicles per Hour per Lane
4	2	4	2960	1480
3	2	8	3000	1500
4	3	4	4560	1520

The studies conducted at work sites with (5,2)
and (2,1) closure situations indicate significant
reductions in capacity (compared with 1500 vphpl).
The average capacity for these two situations was
approximately 1350 vphpl.

Studies at (3,1) sites revealed an even greater
reduction in capacity. The average capacity was
found to be only 1130 vphpl.

Figure 2 shows the cumulative distribution of the
observed work-zone capacities. The function of Fig-
ure 2 is to assist the users in identifying risks in
using certain capacity values for a given lane-
closure situation to estimate the effects of the
lane closures (e.g., queue lengths).

For example, the 85th percentile for the (3,1)
situation is 1020 vphpl. This means that 85 percent
of the studies conducted on three-lane freeway sec-
tions with one lane open through the work zone re-
sulted in capacity flows equal to or greater than
1020 vphpl. The capacity flow was equal to or
greater than 1330 vphpl in only 20 percent of the
cases studied. Thus, to assume a capacity of 1500
vphpl for (3,1) work zones would tend to underesti-
mate the length of queues caused by the lane reduc-
tion at the vast majority of these work zones.

Because of the limited amount of data, no attempt
was made to statistically correlate capacity to the
type of road work. There are characteristics at
each work site that affect the flow through the work
zone. Presence of on ramps and off ramps, grades,
alignment, percentage of trucks, etc., also affect
the flow. These factors were not evaluated in the
studies performed as part of this research.

It is also interesting to note that, even at the
same site, there were variations in maximum flow
rate. Work activities (e.g., personnel adjacent to
an open traffic lane and trucks moving into and out
of the closed lanes) caused these variations.

Table 1 is an attempt to summarize typical capa-
cities observed in California by Kermode and Myra
(1) and those observed in Texas by the Texas Trans-
portation Institute. The California data represent
expanded hourly flow rates, whereas most of the
Texas data are full-hour counts. The reader is cau-
tioned that the typical capacities by type of work
zone shown in Table 1 for Texas freeways are based

Figure 1. Range of observed work-zone capacities for each lane-closure situation studied (work crew at site).
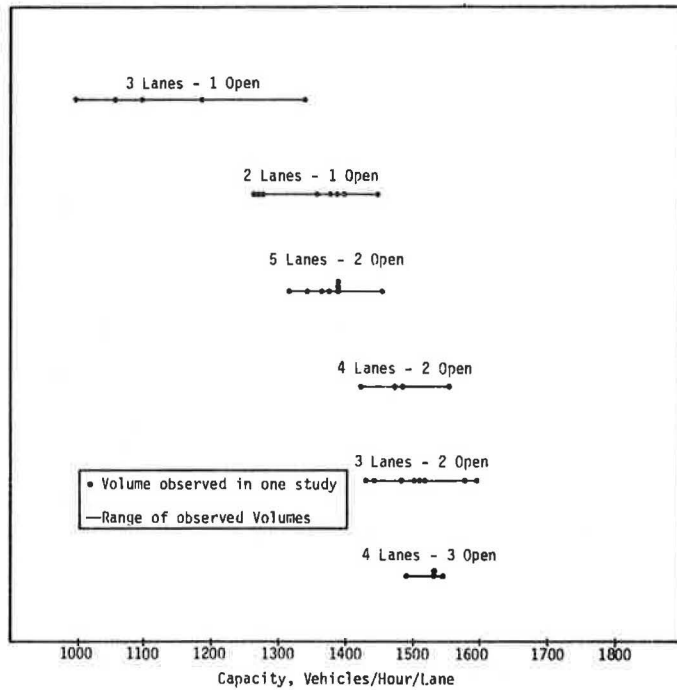
Figure 2. Cumulative distribution of observed work-zone capacities.
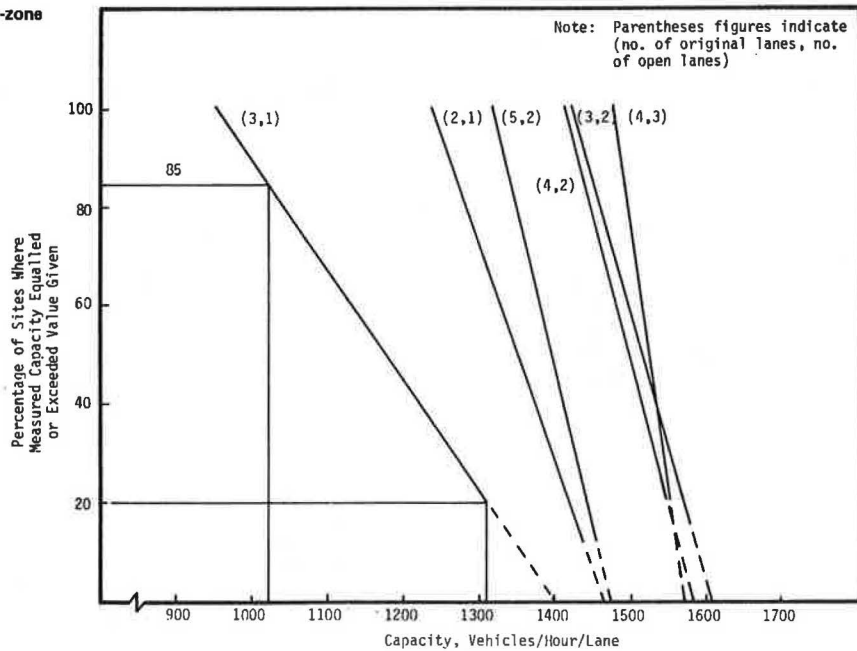
on limited data. The amount of data used to develop capacity rates for California was not indicated (1).

Capacity with No Work Activity at Site

Three studies were conducted at construction sites during the peak period while the work crew was not at the site. These studies were conducted in Houston on a three-lane section of southbound I-45. Two lanes were open during the studies. The average capacity for this (3,2) lane-closure situation was 1800 vphpl.

One study was conducted on the north I-610 loop in Houston. The right two lanes of a four-lane section were closed. There was no work activity in the closed lane immediately adjacent to traffic. A work crew and their machinery did occupy the shoulder lane, however, which was one lane removed from moving traffic. The volumes measured on the two open lanes over a period of 30 min were as follows: 926 vehicles in the lane adjacent to the closure and 730 vehicles in the median lane. These 30-min volumes are equivalent to flow rates of 1850 and 1475 vph. It was apparent from field observations that the demand volumes were lower than the capacity of the two open lanes. Queues did not form upstream from the work activity or the cone taper. There was available capacity in the median lane. The work crew (one lane away from an open traffic lane) did not affect flow through the work zone. It is estimated

Table 1. Summary of capacity for some typical operations.

| Type of Work | Capacity (vph) | | | | |
| | Lane-Closure Situation | | | | |
	3,1	2,1	5,2	3 or 4,2	4,3
Median barrier or guard-rail repair or installation	N/A	1500[a]	N/A	3200[a] 2940	4800[a] 4570
Pavement repair	1050	1400[a]	N/A	3000[a] 2900	4500[a]
Resurfacing, asphalt removal	1050	1200[a] 1300	2750	2600[a] 2900	4000[a]
Striping, slide removal	N/A	1200[a]	N/A	2600[a]	4000[a]
Pavement markers	N/A	1100[a]	N/A	2400[a]	3600[a]
Bridge repair	1350	1350	N/A	2200[a]	3400[a]

Note: N/A = not available.

[a]These volumes represent capacity rates observed in California (1). Other volumes are average capacities observed in Texas.

that the capacity of the two open lanes under the above-cited conditions was about 1800 vphpl. This volume could probably be sustained as long as queues did not form.

Shoulder Use and Traffic Splitting on Three-Lane Section

Generally, when maintenance work is required on the middle lane of a three-lane section, both the middle lane and one of the exterior lanes are closed. Table 1 indicates that the average capacity on the open lane may be between 1050 and 1350 vph depending on the type of road work. Results summarized in earlier research (2) indicate that the capacity could be increased to 3000 vph by using a traffic-control approach called "shifting" whereby drivers are encouraged to use the shoulder as an additional travel lane. In effect, two lanes are open to traffic.

The research also indicates that the capacity could be increased to approximately 3000 vph by using a traffic-splitting approach. In this approach the middle lane is closed and traffic is allowed to travel on both sides of the work activity. It is important, however, that the lane-closure technique recommended by Richards and Dudek (2) be used to implement the splitting approach. Otherwise, considerable driver confusion could take place. The technique involves closing the left lane far upstream from the work area so that only two lanes of traffic enter the split area. Traffic is then funneled and split by using cones--one lane to the left and the other to the right.

APPLICATION TO WORK SCHEDULING AND TRAFFIC CONTROL

Maintenance work on urban freeways. even if performed during off-peak periods, can result in serious congestion and motorist delay. Because of increasing pressures from the motoring public to maintain acceptable levels of service on urban freeways, it is important to analyze the potential impacts of a lane closure in order to schedule the work during periods when the congestion would be minimized and/or to select the most effective alternative traffic-control techniques.

This portion of the paper illustrates how the capacity-study findings can be applied to assist the users in making decisions about scheduling freeway maintenance. It discusses the requirements and procedures for making estimates of traffic volumes and capacities.

Estimating Traffic Volumes

Work-zone volumes are usually estimated from data routinely supplied by automatic traffic counters installed at permanent locations. It is important that current hourly volumes be used to estimate the potential impacts of a lane closure. Volume maps showing average daily traffic are not adequate for this purpose. Hourly traffic volumes recorded by the automatic counters during the previous two weeks on the same day of the week as the scheduled work will provide reasonable estimates of traffic demands.

Anticipated demand volumes at a work zone can also be estimated with good accuracy by making an on-site traffic count (manned or machine) one or two days prior to the work activity. The cost and time involved in conducting this type of special count, however, restrict the use of this approach to special cases.

Hourly traffic-volume data from permanent counters are readily available to most users; however, there are some limitations in using the data. One limitation is that the permanent count data may not provide an accurate estimate of work-zone traffic volumes. Many freeway maintenance sites are a considerable distance from a permanent counter. The volumes recorded at the count stations can differ greatly from those at the work site, especially when there are several ramps between the count station and the work zone. Traffic volumes on a radial freeway, for example, may be much higher near the central business district (CBD) compared with those on the outskirts of the city. If the permanent counter is located near the city limits, then the traffic volumes at a work zone near the CBD may be underestimated. In this case, the congestion may be somewhat more severe than estimated.

It should be apparent from this discussion that there may be significant problems and inaccuracies in using existing permanent counter data to estimate work-zone volumes. However, until new urban freeway counting programs are developed and implemented, permanent counter data are probably the most practical.

The problem of estimating traffic demands at work zones is compounded by the phenomenon of natural diversion. When encountering unusual congestion on an urban freeway during the off-peak periods, many drivers will leave the freeway and travel on the frontage road to bypass the congestion or seek alternative routes to their destinations (3,4). The extent of this natural diversion is difficult to predict.

Estimating Capacity

Previously, 1500 vphpl was a common value used by many traffic-control planning analysts to estimate the flow through work zones. The capacity data presented earlier, however, provide better insight into typical capacities at work zones on Texas freeways. For example, a review of Figure 2 suggests that using a work-zone capacity of 1500 vphpl for (4,3), (4,2), and (3,2) lane-closure situations may not be too critical. However, this value seems too high for estimating the impacts of the (3,1), (2,1), and (5,2) closure situations.

As previously discussed, the cumulative distributions of observed work-zone capacities shown in Figure 2 can be used to identify risks associated with using certain capacity values for a given lane-closure situation to estimate the effects of the lane closures (e.g., queue lengths).

Estimating Queue Length and Delay

The delays associated with stop-and-go driving that occur at work zones where there is a lane closure are the result of a lack of capacity. These work zones, which have insufficient capacity to handle demand, are analogous to an hourglass. The neck of the hourglass can handle only so much sand, and there is nothing the excess sand on top can do but wait. When traffic demand at a work zone exceeds the capacity of the work zone, vehicles begin to stack up at the lane-closure taper to wait their turn to pass through the work area.

Figure 3 is a simple graphical procedure that can be used to obtain a rough estimate of queue length and delays at work zones. These estimates are obtained by plotting the cumulative demand volumes and the cumulative service volumes (capacity) versus time. As illustrated, the number of vehicles stored (or queued) and individual vehicle delay at any given time can be estimated.

The length of traffic backup or queue length can be roughly estimated by using the following relationship:

$$L_t = Q_t \ell / N \tag{1}$$

where

- L_t = estimated length of backup (queue length, ft) at time t,
- Q_t = estimated number of vehicles in queue at time t,
- N = number of open lanes upstream from lane closure, and
- ℓ = average space occupied by vehicle in queue (use ℓ = 40 ft).

Sample Problem

Figures 2 and 3 and the tables in this paper present information to assist the user in making decisions related to scheduling maintenance. The following
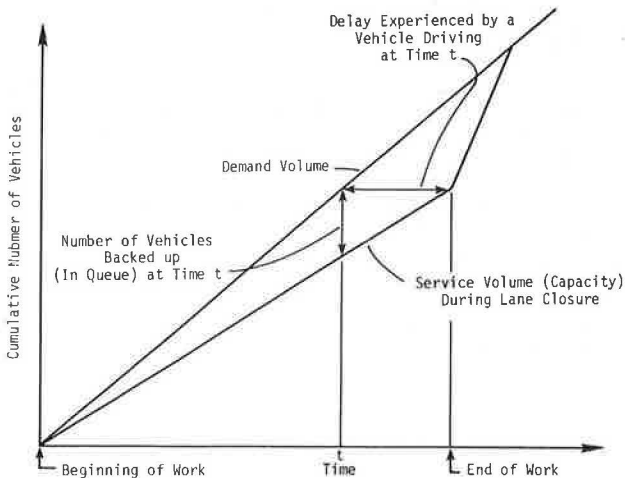
example demonstrates how the information may be used.

Assume that maintenance work must occupy a three-lane freeway section. The work will require that the median lane be closed as shown in Figure 4. The work will require approximately 4 h to complete. This includes the time required to install and remove traffic-control devices. Data obtained from a nearby permanent counter during the previous two weeks were used to estimate the following demand volumes:

Time	Volume Anticipated (vph)
9:00–10:00 a.m.	2920
10:00–11:00 a.m.	3120
11:00–12:00 a.m.	3200
12:00–1:00 p.m.	3500
1:00–2:00 p.m.	3830
2:00–3:00 p.m.	3940
3:00–4:00 p.m.	4620
4:00–5:00 p.m.	5520

It should be noted at this point that any estimates of the queue length and vehicle delays by using the procedure shown in Figure 3 will be influenced by the accuracy of the demand-volume data. The estimates are also greatly influenced by assumed work-zone capacity. The consequences of using different capacity estimates are explored in this sample problem.

Referring back to the table in the first section and Figure 2, it is seen that the average capacity for the (3,2) lane-closure situations studied was 1500 vphpl or 3000 vph. The 85th percentile was 1450 vphpl or 2900 vph, and the 100th percentile was 1420 vphpl or 2840 vph. If these capacities (3000, 2900, and 2840 vph) are assumed, the graphical technique discussed earlier has been used to estimate the resulting queue lengths and delays (see Figure 5).

In Figure 5, the work is assumed to begin at 9:00 a.m. The estimated queue length at 1:00 p.m., after 4 h of maintenance work and assuming a capacity of 3000 vph, is 2.1 miles. The estimate by using 2900 vph is 2.9 miles, almost 1 mile longer; and the estimate by using 2840 vph is 3.5 miles, about 1.5 miles longer. Therefore, the capacity value is a very sensitive parameter when queue length is estimated.

Figure 2 shows that the average capacity value of 3000 vph (1500 vphpl) is at the 60th percentile. This means that based on the data collected to date, there is a 40 percent chance that the actual capacity may be lower than 3000 vph and thus that the queue length will be longer than 2.1 miles. Likewise, there is only a 15 percent chance that the traffic will back up farther than 2.9 miles, if we assume that the maintenance work took 4 h to complete. These estimates should be helpful in deciding where to place the advance signs for the work zone.

It should be apparent that stop-and-go traffic extending for 2.9 miles would be very undesirable. Thus, other options should be explored, for example, the following:

1. Perform the work on a Saturday or Sunday when the volumes are lower,
2. Perform the work at night,
3. Reduce the work time or split the work into two shifts, or
4. Implement additional traffic-control strategies.

Curves similar to those shown in Figure 5 can be developed for weekend or night work. It is not the intent of this paper to discuss the merits or prob-

Figure 3. Graphical procedure for estimating queue length and delays at work zones.

Figure 4. Sample lane-closure problem.

Figure 5. Sample problem solution.

a.m. Closing ramps when available capacity still exists on the freeway promotes driver discontent and may create unnecessary operational problems on other facilities (e.g., frontage roads and arterial streets). Ramp-closure techniques are discussed in a report by Richards and Dudek (5). Provisions should be made to achieve improved signal coordination on the frontage road whenever ramps are closed.

Allowing traffic to use the shoulder is another way to increase work-zone capacity. Up to 1500 vph additional vehicles can be accommodated by using the shoulder. Traffic-control details for shoulder use have been presented by Richards and Dudek (2).

lems of performing road work during these times. However, it suffices to say that the lower volumes associated with these time periods will result in reduced congestion.

A review of Figure 5 indicates that if the work could be completed within 3 h or less, the amount of congestion would be greatly reduced. If a capacity of 3000 vph is assumed, the queue would extend an estimated 0.8 mile upstream from the lane closure, and with a capacity of 2900 vph (85th percentile), the queue would not extend more than 1.4 miles. If the work could be divided into two 2-h periods from 9:00 to 11:00 a.m. on two separate days, the expected queue length would be greatly reduced to approximately 0.5 mile (if comparable volumes are assumed for both days).

Another option would be to implement additional traffic-control strategies. These might include entrance-ramp closure and shoulder use. Each of these strategies should be evaluated for its merits before implementation.

Closing entrance ramps at and upstream from a work zone may possibly reduce the traffic demands and greatly reduce queues so that work could be performed for four continuous hours. Decisions concerning entrance-ramp closures, including the time of closures, should be based on the anticipated freeway and entrance-ramp traffic demands and the available capacity on the alternative route (e.g., frontage roads and arterial streets). Ramps should be closed when the combination of the freeway and the ramp volumes exceeds the work-zone capacity and there is available capacity on the alternative route. The ramps should remain open when the traffic demands are less than the work-zone capacity. In the sample problem, for example, the entrance ramps should not be closed until approximately 10:00 a.m. even though the maintenance begins at 9:00

REFERENCES

1. R.H. Kermode and W.A. Myra. Freeway Lane Closures. Traffic Engineering, Feb. 1970.
2. S.H. Richards and C.L. Dudek. Field Evaluation of Traffic Management Strategies for Maintenance Operations in Freeway Middle Lanes. TRB, Transportation Research Record 703, 1979, pp. 31-36.
3. J.M. Turner, C.L. Dudek, and J.D. Carvell. Real-Time Diversion of Freeway Traffic During Maintenance Operations. TRB, Transportation Research Record 683, 1978, pp. 8-10.
4. C.L. Dudek, W.R. Stockton, and D.R. Hatcher. San Antonio Motorist Information and Diversion System. Texas Transportation Institute, College Station, Rept. FHWA-RD-81/018, Dec. 1980.
5. S.H. Richards and C.L. Dudek. Special Traffic Management Requirements for Maintenance Work Zones on Urban Freeways. Texas Transportation Institute, College Station, Rept. FHWA/TX-82/1+228-8, Jan. 1982.

Lane Closures at Freeway Work Zones: Simulation Study

ZOLTAN A. NEMETH AND NAGUI M. ROUPHAIL

A study of freeway lane closures at work zones is described. It involved the development of a microscopic computer-simulation model. Vehicles in platoons are controlled by a car-following rule. The merging behavior is controlled by the information provided by the traffic-control devices, by personal preference for early or delayed merge, and by the availability of gaps in the open lane. The prescription of personal preference was based on a driver survey. The model also checks for the possible obscuring of signs by large vehicles. Field tests produced varied results, but average speeds and throughput (vehicle miles per hour squared) generated by the model fit between the classical Greenshield's model and those calculated by the 1980 revision of the Highway Capacity Manual. A factorial simulation study was conducted to investigate traffic behavior under a variety of conditions, represented by different volume levels, traffic compositions, merging preferences, speed control and compliance, and advance-warning distances. Delay and standard deviation of speed at the taper were generated for each factor-level combination. The results generally confirmed what was expected. Noteworthy is the indication that full compliance with a reduced speed limit of 45 mph would increase delayed merges within the taper area in the volume range simulated.

The problems associated with the safe and efficient conduct of traffic at work zones have received considerable attention in recent years. The Federal Highway Administration (FHWA) initiated a coordinated research program in 1975: Project 1Y--Traffic Management of Construction and Maintenance Zones. The purpose was to generate the basis for the development of new concepts, methods, and approaches to traffic management in construction, maintenance, and utility work zones. It resulted in the undertaking of numerous studies with wide scope and ranges of objectives.

Work activities that require lane closures and force traffic to merge into the open lane(s) represent a frequently encountered and a potentially hazardous situation. A study of road-under-repair accidents in Virginia found, for example, that of 426 accidents (for which the information on traffic-control characteristics was available), 47.9 percent occurred at lane closures (1). The same study found that close to 80.0 percent of the work-zone accidents can be attributed to driver error. Drivers approaching a work zone in the closed lane must receive and understand the information that they need to change lanes and merge into the open-lane traffic. Although this in itself does not appear to be an unusually demanding driving task, problems seem to develop that result in rear-end collisions, sideswipes, and single-vehicle/fixed-object accidents (2).

The objective of the research project described here was to study the operation of lane closures at construction sites on rural freeways. Two issues were addressed in particular:

1. Merging patterns from the closed lane into the open-lane traffic and
2. Speed reduction at work zones.

The approach taken by the study team was to build a simulation model supported by field studies and driver surveys.

CHARACTERISTICS OF SIMULATION MODEL

The microscopic digital simulation model Freeway Construction (FREECON) of lane closures at freeway construction sites is written in FORTRAN IV language and it uses the GASP IV simulation package of Pritsker (3). The model is based on the realistic description of the movement of a vehicle approaching a lane-closure site. The two major rules applied in previous freeway simulation models are the car-following and gap-acceptance rules. In this model, several additional features were needed relevant to traffic control at lane closures.

Driver Reaction to Merging Stimuli

Signs, arrowboards, and finally the delineation of the taper itself provide the information or stimulus to merge from the closed lane into the open lane. A driver survey was conducted to identify what proportion of drivers reacts to each stimulus. Each unit of driver and vehicle is randomly assigned by the model into one of the groups in an attempt to represent realistic merging behavior (i.e., a certain proportion of the drivers will begin searching for an acceptable gap at the first sign, whereas others might wait until the taper or the construction site becomes visible).

Driver Reaction to Speed Control

As an option, the model can also specify drivers' reaction to speed control (e.g., comply with advisory speed limit sign).

Traffic-Control Device (TCD) Design Constraint

The location can be specified for each TCD within the simulated freeway segment, and a recognition distance is also assigned to represent a particular design (e.g., size).

TCD Visibility Constraint

In many instances, a driver is unable to see a sign because his or her line of sight is blocked by another vehicle. In the model, vehicles are represented by their physical dimensions, and one of the subroutines checks for potential blockage of TCDs by large vehicles.

TCD Information Acquisition Constraint

Each TCD is assigned a minimum required information-processing time. In free-flow traffic, drivers should have ample time to look at signs long enough to understand the message. In high-density flow, however, more time is spent on fixating on other vehicles and less is available for sign recognition (4). An algorithm has been developed that relates the maximum duration of fixation on the various TCDs to the time headway between two vehicles in the car-following mode.

MODEL STRUCTURE

FREECON consists of a main program and 18 supporting subprograms and functions. The model is microscopic in nature; that is, each driver-vehicle unit is identified as a separate entity. Periodic updating of each vehicle's status is performed at 1-s intervals.

Figure 1 illustrates the general simulation logic as it applies to the microscopic driver-vehicle entities in the system. Also shown are the subprograms related to each step in the model.

Figure 1. Flowchart representation of model logic.
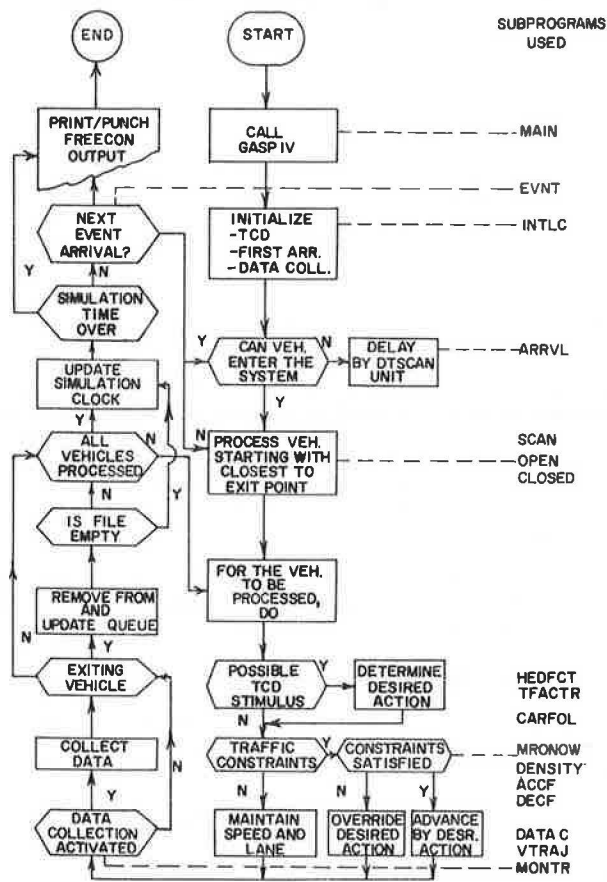
Eight basic and interacting components constitute the core of the model.

Driver-Vehicle Component

On entering the work zone, each individual driver or entity is assigned a set of 20 attributes, some of which are periodically updated during a simulation run. In this version of the model, all attributes were assigned independently. This was later revised to account for driver groups exhibiting 16 similar attributes such as speed and gap acceptance. A brief description of each attribute follows:

DENTR: time of entry (or entries), randomly assigned from a distribution of time headways in the corresponding lane or lanes.

DDSPD: desired speed (ft/s), randomly assigned from the distribution of vehicle speeds upstream of the work zone.

DLANE: lane of travel (1 = open lane, 2 = closed lane), randomly assigned based on the distribution of traffic among the approach lanes.

DSPL, DSPN: vehicle speed at end of last and current updating intervals, respectively (ft/s).

VTYPE: vehicle type (1 = passenger car, 2 = truck), randomly assigned based on the traffic composition upstream of the work zone.

DPSOL, DPSON: vehicle position at end of last and current updating intervals (ft); DPSON is the ranking attribute in the driver-vehicle file with a high-value-first (HVF) queue discipline.

TMRGS, SPST: selected merge and speed simulation codes, respectively. [Each driver entering the work zone is assigned a set of merge and speed stimuli.

This assignment is based on the survey of 229 drivers conducted at several freeway construction lane-closure sites. Among the results were the following: 45.4 percent of drivers merge at the earliest opportunity, 13.1 percent of drivers merge after having passed a few cars, 9.3 percent of drivers merge after having seen other drivers merge, 20.5 percent of drivers merge after having seen construction activity, and 11.7 percent gave no answer. From these results, as a first approach, a representative probabilistic distribution function of drivers' response to merge stimuli was formulated. More work is being done now on this aspect of the model. Similar treatment is applied in the development of a typical speed strategy.]

SFIXM, SFIXS: cumulative time fixations on a merge and a speed stimulus, respectively; merges and speed changes in the model are initiated as soon as either value exceeds the minimum information-processing time on the corresponding stimuli.

SLCH: cumulative time spent in the lane-change maneuver(s); the model assumes a 4-s lag between initiation and completion of a lane change.

MCODE: merging attempt code (1 = attempting, 0 = not attempting); introduced to ensure that once a lane-change attempt has aborted, consecutive attempts will be made until a successful maneuver has been completed.

VINDX: vehicle index register; used to trace vehicles' paths throughout the work zone.

TVLST: last time the vehicle's position was updated; introduced to prevent multiple processing of the same vehicle in the same interval.

SPEDG: cumulative speed gradient component; traces speed fluctuations throughout the work zone; final values for each vehicle are determined at the point of exit.

CGAP: driver critical gap (ft); randomly assigned from a gap-acceptance function derived in a related study (5).

DHEAD: desired headway(s); introduced to test whether a speed-control strategy based on reducing headway variance, instead of average speed, could improve the quality of traffic flow; DHEAD may be totally bypassed in the model logic.

TREACT: driver's brake reaction time; randomly assigned from a distribution of brake reaction times developed by Johannsson and Rumar (6).

TCD Component

Ten attributes describing each TCD are introduced in the initialization phase of the model. These are as follows:

SCODE: TCD code, unique to each device (e.g., arrowboard, signs, cones); SCODE is matched with TMRGS, SPST codes in the driver-vehicle component.

SL: TCD placement code; a code of 1 is given for TCD placed on the open-lane side of the road, 2 for those placed on the closed-lane side, 12 for both sides.

PS: location of TCD, measured from vehicle entry point along the longitudinal axis of the road (ft).

WS: lateral TCD placement, measured outward from lane edge (ft).

SLD: recognition distance, as measured in the field (ft).

SLP: upstream recognition point (= SL - SLD).

SDR: minimum information-processing time; SDR is compared with SFIXM or SFIXS in the driver-vehicle component in order to schedule lane and/or speed-change attempts.

SH: message height, measured from pavement level (ft).

ST: type of stimulus; TCDs are categorized as

either merge or speed stimuli (example: reduced-speed-limit sign).

SSPD: posted speed limit if different from freeway speed limit (ft/s).

Roadway and Data-Collection System Component

The location of data-collection points can be varied by the user; up to 20 data-collection points may be simulated, 10 for each approach lane. Vehicles crossing any of the simulated detectors activate the corresponding speed and headway registers. Although there is no physical limitation on the simulated length of the zone, the model logic makes it necessary that the termination point be in the single-lane zone of traffic.

In its present form, the model assumes a straight, level road alignment. However, horizontal and vertical curvature effects on TCD recognition distance can be readily manipulated in the TCD component.

Vehicle-Generation Component

Vehicle arrivals into the work zone are scheduled from a probability function of time headways. Nine such functions are available in the model, each of which is given a unique code provided by the user as input ($\underline{7}$). Desired speeds are generated in the same fashion. Tests are internally conducted in the model to ensure that the car-following rules are satisfied at the entry point. Modeling shifted distributions is readily available in the GASP IV input format.

Car-Following Component

The car-following model selected in the study closely follows the noncollision constraints developed in the INTRAS simulation model ($\underline{8}$) with some modifications. Three car-following rules are defined, as follows:

$$X_t - Y_t \geqslant L + CV_t \qquad 0 < V_t \leqslant U_t \qquad (1)$$

$$X_t - Y_t \geqslant L + CV_t + (V_t^2 - U_t^2)/2E \qquad V_t > U_t > 0 \qquad (2)$$

$$X_t - Y_t \geqslant L \qquad V_t = U_t = 0 \qquad (3)$$

where

X_t = position of lead vehicle at time t (ft),
Y_t = position of following vehicle at time t (ft),
U_t = speed of lead vehicle at time t (ft/s),
V_t = speed of following vehicle at time t (ft/s),
L = overall length of lead vehicle (ft),
C = brake-reaction time of following driver, and
E = maximum acceptable deceleration rate (ft/s²).

Since vehicles' positions are updated every second, it follows that the lead-vehicle position and speed are first determined at time t + 1; from Equations 1, 2, or 3, a maximum permissible acceleration rate (a_{max}) is determined. The actual acceleration rate (a) is computed as follows:

$$a = \min(a_d, a_{max}, a_v) \qquad (4)$$

where a_d is the desired acceleration rate based on current (V_t) and desired (DDSPD) speeds and a_v is the limiting acceleration rate based on current speed and vehicle type. The following-vehicle speed and position are then updated as follows:

$$V_{t+1} = V_t + a \qquad (5)$$

$$X_{t+1} = X_t + V_t + \tfrac{1}{2}a \qquad (6)$$

Lane-Switching Component

This component handles all merging maneuvers. When a lane change is attempted, tests are conducted to ensure that the car-following rules in the destination lane are satisfied. Additional tests are made to determine whether safe merging gaps are acceptable to the driver, based on the gap-acceptance function.

Another feature of this component is the automatic initiation of lane-change attempts for vehicles within the stopping-sight distance of the construction taper. It was assumed that only the car-following rules need to be satisfied to perform a successful merge in the region. Empirical evidence for this assumption can be found in Pahl's study of freeway exit ramps ($\underline{9}$).

TCD Information Acquisition Components

This component registers and updates the cumulative time fixations a driver makes on a TCD. In order to initiate a speed (lane-change) response, the cumulative time fixations on the speed (merge) stimulus should exceed the minimum information-processing time for the stimulus. Whether this condition is met before the driver passes the TCD location depends on two factors: (a) presence of obstruction to the driver-TCD line of sight and (b) current headway.

The impact of vehicle headway (h) on TCD information acquisition was modeled by using the following functions:

$$DT_h = 0 \qquad h \leqslant 0.5 \text{ s} \qquad (7)$$

$$DT_h = (2H - 1)/7 \qquad 0.5 < h < 4 \text{ s} \qquad (8)$$

$$DT_h = 1 \qquad h \geqslant 4 \text{ s} \qquad (9)$$

where DT_h is the fraction of time a driver spends fixating on TCD while traveling at headway h.

Determination of function parameters was based on preliminary results of driver test studies conducted at freeway lane closures ($\underline{10}$). Further refinement of the function may be necessary as additional data become available.

The final form of the TCD information acquisition constraint is stated as follows:

$$SFIXM_i (t + 1) = SFIXM_i (t) + B \times DT_h \geqslant SDR_i \qquad (10)$$

where

$SFIXM_i(t + 1)$ = cumulative time fixations on TCD_i after t + 1 s,
$SFIXM_i(t)$ = cumulative time fixations on TCD_i after t s,
B = binary variable that assumes a value of 1 (zero) if legibility rules are (are not) met, and
SDR_i = minimum information-processing time for TCD_i.

Output Component

The simulation model output component produces the following standard output:

1. Listing of user input data;
2. Descriptive statistics and histograms of speeds and time headways at each data-collection point;

3. Descriptive statistics and histograms of five performance measures: vehicle merging points measured from entry point (ft), vehicle delay, exit volume during simulation period, throughput (defined as the product of exit volume and exit speed and cumulative for all vehicles), and speed gradient (a measure of speed fluctuations for vehicles traveling in the work zone); and

4. Listing of vehicle trajectories at any point during simulation run; trajectory data can be routed to a plotting routine that produces a visual representation of the individual vehicles' paths in the approach, transition, and single-lane areas.

MODEL VALIDATION

The model was first tested by comparing outputs with generally accepted models of the speed-volume-density relationship. Average speeds and throughput (vehicle miles per hour squared) were calculated at volumes ranging from 1000 to 2000 passenger cars per hour per lane in the open lane. It was found that outputs fit between the classical Greenshield's model and those calculated by the 1980 revision of the Highway Capacity Manual.

Field studies were conducted at two construction sites (denoted A,B) with the purpose of testing the model logic. Specifically, the following traffic descriptions were targeted for comparison:

1. Means and distributions of vehicle time headways at each data-collection point,
2. Means and distributions of vehicle speeds at each data-collection point, and
3. Distribution of merging distances, defined as vehicle position (in feet) measured from the first construction sign at which a lane change is initiated into the through traffic lane.

Data Collection and Reduction

An instrumented data-acquisition system, developed by the Systems Research Group of the study team under the direction of T.H. Rockwell, was specifically designed for the purpose of collecting the above-mentioned traffic descriptions.

The system consists of eight 10-ft tapeswitches arranged in pairs. The tapeswitches were laid in the open lane of traffic and covered a distance of 1500 ft. Cable connectors were used to transmit vehicle actuations into a video cassette recorder via a 12-channel video box. A unique code for each tapeswitch (zero to seven) was assigned, which was displayed on a TV monitor for the duration of the actuation. Other elements in the system included a continuous five-digit clock and a video camera. Power was supplied to the various components by means of a portable 4-hp/1900-W, gasoline-powered A/C generator.

Supplementing the system was a number of manual observers who collected pertinent traffic data outside the system's 1500-ft range. Finally, a complete inventory of TCD design and performance characteristics was made prior to data collection.

The recorded vehicle arrival times at each tapeswitch were subsequently reduced and fed into a computer program for the determination of mean values and distributions of speeds, headways, and merging distances.

Results at Site A

Site A involved a left lane closure during a bridgedeck rehabilitation project on the southbound lanes of I-71. Statistical tests were conducted on speed and headway distributions for two independent ob-

servation periods (10 min each). Student's t-tests on mean values showed no statistically significant differences at α = 5 percent.

Distributions of speed and headways were then tested by means of the Kolmogorov-Smirnov two-sample test. Results indicated no significant differences except for speed distributions at the two downstream (i.e., last) tapeswitch pairs. The simulated speeds were slightly higher there than the observed speeds. The difference was always less than 2 mph.

Also, the statistical tests showed no significant differences in the cumulative distribution functions of merging distances, which suggests that the typical merging strategy indicated by the driver survey appears to be a valid indicator of drivers' preference in lane-closure situations under the conditions present at site A.

Results at Site B

A substantial difference was found between the field and simulated merging patterns at site B. Drivers were observed to merge much later at site B than at site A. The effective warning distance was only slightly shorter at site B and could not possibly account for the large difference. Geometrics were in general quite similar, but at site A the left lane was closed, whereas at site B, the right lane was closed. A closer look at the lane distribution of the volumes led to a plausible explanation for the difference in merging. Looking at the equivalent hourly approach volumes and approach speeds, we have found the following: approach volumes at site A were 330 vehicles per hour (vph) at 61 mph in the merging lane and 714 vph at 53 mph in the open lane; approach volumes at site B were 451 vph at 48 mph in the merging lane and 190 vph at 55 mph in the open lane.

Although traffic is not distributed uniformly over the roadway, it is worthwhile to express the above-described traffic flows in terms of average spacings in feet. While drivers are obviously not very sensitive to hourly volumes, they can observe and be influenced by the spacings of vehicles around them. At site A, average spacing is 978 ft in the closed lane and 390 ft in the open lane. At site B, the average spacing is 562 ft in the closed lane and 1510 ft in the open lane.

The open lane at site B must have looked empty to the drivers in the closed lane and thus there was no incentive to merge early.

The open lane at site A, however, looked fairly well traveled; thus it provided an incentive to merge early to at least some of the drivers; i.e., as expected, drivers use judgment regarding the urgency of lane changes.

It was concluded, therefore, that the application of the model should be limited to higher approach volumes, perhaps in the range of 1000 vph, combined on the two lanes, provided that the larger proportion of traffic occupies the open lane typical at left-lane closures. Research is under way to develop a merging model more general in scope.

EXPERIMENTS WITH MODEL

Several experiments were conducted with the model with the purpose of testing the sensitivity of some traffic-stream descriptors, merging and speed-control strategies, and TCD performance characteristics on a number of the system's performance measures, by using the site configuration shown in Figure 2.

A complete mixed factorial design was developed for the analysis of five independent variables. These were categorized into traffic-stream factors,

Figure 2. Simulated work zone for simulation study.
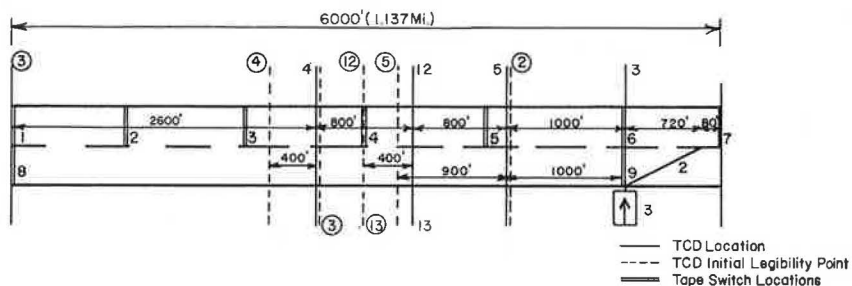
TCD	MESSAGE	LOCATION	LEGIBILITY	HEIGHT
2	Taper, Type II Barric.	5200'	1000'	3'
3(I)*	4'x8' Arrow Board	5200'	5200'	8'
(II)	4'x8' Arrow Board	5200'	2600'	8'
4	Right Lane Closed	2600'	400'	7'
5	Symbolic Sign N	4200'	900'	7'
12(I)	Allow Space for Merging Vehicles	3400'	400'	7'
13(II)	Speed Limit 45MPH	3400'	400'	7'

* I, II, Mutually Exclusive Conditions

driver-behavior factors, and TCD factors:

1. Traffic-stream factors:
   (a) Two-lane approach volume upstream of the work zone at levels of 100, 1250, and 1500 vph
   (b) Proportion of trucks at levels of 0 and 25 percent of approach volume
2. Driver-behavior factors:
   (a) Merging strategy at the following levels: early merging strategy (all drivers in the closed lane attempt to merge at first opportunity), typical merging strategy (as obtained from the driver's survey and validated at site A), and late merging strategy (all drivers in the closed lane attempt to merge only on recognizing the construction activity)
   (b) Speed-control strategy at the following levels: none (no special provisions for speed reduction), 45 mph (all drivers in the open lane comply with a posted 45-mph reduced-speed limit), and special sign [all drivers in platoons (headways $\leq$4 s) increase their headway in compliance with the experimental sign ALLOW SPACE FOR MERGING VEHICLES]
3. TCD factor: Effective warning distance at levels of 1 and 0.5 mile.

Variations in the effective warning distances are modeled by assigning two different legibility distances for the arrowboard.

Dependent variables included mean vehicular delay (DELAY), standard deviation of speeds at the start of lane taper (SSD), and proportion of merges prior to 400 ft from taper (MERG400). The latter variable reflects the relative frequency of occurrence of free versus forced merges. The 400-ft distance was computed as the stopping-sight distance for a vehicle traveling 55 mph and a brake reaction time of 1 s. Thus, all lane changes occurring within the last 400 ft were considered forced merges.

The analysis of variance (ANOVA) technique was used to formulate statistical models for the three performance measures. A level of significance of α = 5 percent was used throughout the analysis.

Interpretation of Results

Results of the ANOVA models are displayed in Figures 3, 4, and 5. All three-factor level interactions were found to be statistically insignificant.

Figure 3. Impact on delay.
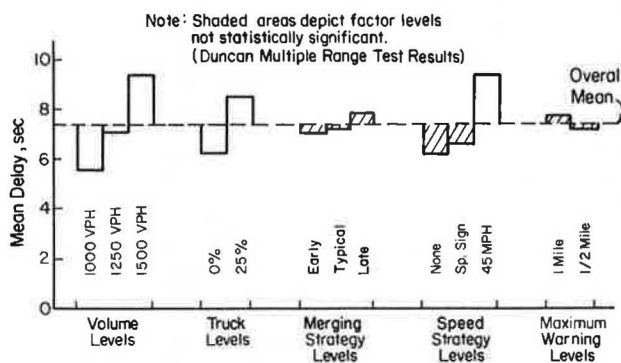
Figure 4. Impact on speed distribution.
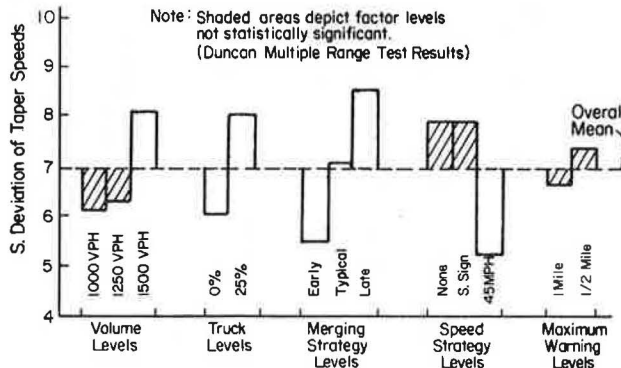
Delays (see Figure 3) were found to increase with volume and truck proportion, as expected. The main merging-strategy effect was not statistically significant, although a closer look at the details indicated that delays were highly sensitive to volume levels in the late merging strategy. At the other extreme, delays were all but unaffected by volumes under the early merging strategy. The 45-mph speed limit resulted in a mean delay increase of 50 percent, whereas no statistically significant increase in delays was noted with the experimental sign. The predicted increase in delay is obvious, since delay

Figure 5. Impact on forced merges.

Note: Shaded areas depict factor levels not statistically significant.
(Duncan Multiple Range Test Results)
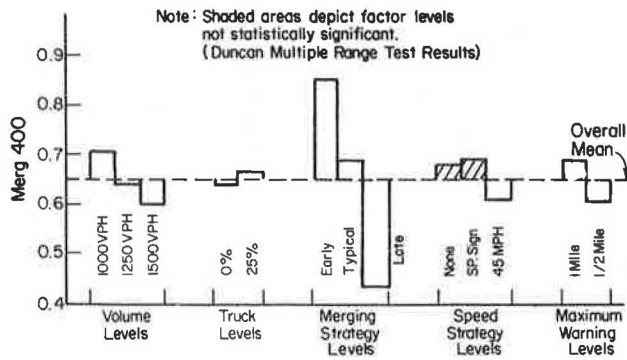
is defined as the difference between travel time at approach speed and travel time at actual speed (in this case, 45 mph or less).

ANOVA models for standard deviation of speeds at the taper (see Figure 4) indicated that speed variations increased significantly with both volume and truck levels. The early merging strategy resulted in the lowest observed speed variation, as did sites with a 45-mph speed limit. The latter was to be expected, since 100 percent compliance was assumed. These results tend to support the general concerns about the safety hazards associated with late lane changes (e.g., rear-end collisions). The problem of speed variation becomes even more acute at sites with short warning distances or with a large truck population.

The proportion of merges occurring prior to 400 ft (see Figure 5) from taper was found to decrease in a linear fashion with volume. Unexpectedly, however, the presence of trucks resulted in fewer late lane changes. It is suggested that the advantage truck drivers have in recognizing (hence responding to) the TCD outweighs the fact that trucks constitute a potential obstruction to following passenger cars. Of course, this interpretation considers similar responses from truck and passenger car drivers, a fact that could not be disproved from the driver survey results.

Merging strategy had a drastic impact on the frequency of free (as opposed to forced) merges (85 percent for early merging strategy versus 42 percent for late merging strategy). Furthermore, when an early merging strategy was coupled with a 1-mile effective warning distance, the proportion of free merges did not drop below 90 percent, even at volumes approaching the single-lane capacity.

Finally, the impact of speed-control strategies provided some revealing findings on the impact of reduced speed limits on merging in construction zones. When the 45-mph limit was in effect, the frequency of free merges was actually reduced by 15 percent compared with the no-speed-control strategy. It is suggested that since drivers are primarily concerned with maintaining safe headways in the open lane, a drastic speed reduction would in effect increase the traffic density near the transition zone. Consequently, the probability of finding acceptable gaps is reduced, hence the increase in the frequency of forced merges.

DISCUSSION OF RESULTS

This study was aimed at the investigation of merging and speed controls at freeway construction-lane closures through the use of computer-simulation techniques. A traffic model incorporating individual drivers' preference, traffic-stream descriptors,

and characteristics of TCDs was developed. The model was field tested with varied results at two sites. The work is continuing on the refinement of the model.

The findings of this study may be summarized as follows:

1. The field-study results indicate that the effectiveness of the advance-warning devices at freeway construction-lane closures is not determined solely by the design features of the individual devices but also, and perhaps more importantly, by the risk perceived by approaching drivers. Under low-volume conditions, drivers' merging patterns and travel speeds are virtually unaffected by the advance-warning devices at the site. Speeds and/or lane changes are initiated only when the construction activity is actually in sight. At higher volumes, however, many drivers merge early.

2. The simulation-study results indicate that at sites experiencing approach volumes in excess of 1000 vph, it is desirable that early merging be encouraged. Traffic-engineering measures that deter travel in the closed lane (i.e., lane to be closed ahead) should be contemplated. A recent study (11) indicated that changeable message signs were quite successful in that respect.

The implementation of the 45-mph maximum speed control and assumed 100 percent compliance resulted in higher percentages of forced merges in the taper area in the model. The assumption of 100 percent compliance was not meant to be a realistic assumption, but it is still interesting to note that from the point of view of smooth merging, the speed reduction may not even be desirable.

ACKNOWLEDGMENT

REFERENCES

1. B.T. Hargrove and M.R. Martin. Vehicle Accidents in Highway Work Zones. FHWA, Rept. FHWA/RD-80/063, Dec. 1980.

2. Z.A. Nemeth and D.J. Migletz. Accident Characteristics Before, During, and After Safety Upgrading Projects on Ohio's Rural Interstate System. TRB, Transportation Research Record 672, 1978, pp. 19-24.

3. A.A. Pritsker. The GASP IV Simulation Language. Wiley, New York, 1974.

4. T.H. Rockwell and V.D. Bhise. Development of a Methodology for Evaluating Road Signs. Department of Industrial and Systems Engineering, Ohio State Univ., Columbus, Project EES 315, 1973.

5. B. Miller. Gap Acceptance and Gap Estimation in a Freeway Driving Situation. Ohio State Univ., Columbus, M.S. thesis, 1980.

6. G. Johannsson and K. Rumar. Drivers' Brake Reaction Times. Human Factors, Vol. 13, No. 1, 1971.

7. N.M. Rouphail. A Model of Traffic Flow at Freeway Construction Lane Closures. Ohio State Univ., Columbus, Ph.D. dissertation, 1981.

8. A.G. Bullen and P. Athol. Development and Testing of INTRAS, a Microscopic Freeway Simulation Model, Volume 2. KLD Associates, Inc., Huntington Station, NY, Feb. 1976.

9. J. Pahl. Gap-Acceptance Characteristics in Freeway Traffic Flow. HRB, Highway Research Record 409, 1972, pp. 57-63.

10. T. Rockwell and Z. Nemeth. Development of a Driver-Based Method for Evaluating Traffic Control Systems at Construction and Maintenance Zones. Ohio State Univ., Columbus, Final Rept., EES 581, Oct. 1981.

11. F.R. Hanscom. Effectiveness of Changeable Message Displays in Advance of High-Speed Freeway Lane Closures. NCHRP, Rept. 235, Sept. 1981.

Selecting Two-Regime Traffic-Flow Models

SAID M. EASA

A procedure for selecting two-regime macroscopic models for a given set of traffic-flow data is presented. The procedure is based principally on the theoretical characteristics among the various regions of macroscopic models, which includes the limiting case and the convexity and concavity properties. The input to the procedure is represented by the basic traffic-flow criteria (free-flow speed, optimum speed, jam density, and so on) as well as auxiliary criteria to account for the variability of the traffic-flow relations in the intermediate ranges of flow. With these criteria, which are established from the data, the procedure can directly output model parameters, through simplified graphical tools, for the non-congested- and congested-flow regimes. Application of the procedure by using actual data was made to illustrate its use and to discuss some issues related to establishing the traffic-flow criteria from the data. This application also illustrates the flexibility of the procedure and the ease with which the specified criteria can be adjusted to further improve the data fitting. The procedure presented in this paper significantly reduces the need for using computer facilities in estimating traffic-flow relations and as such should prove useful in many transportation applications.

Macroscopic traffic-flow models have been widely used in the field of transportation, including freeway operations, highway levels of service, environmental studies, and transportation planning. Generally, these models can be used to describe the traffic-flow relations in two ways: single-regime and two-regime representations. In the former, the entire range of operation is represented by a single model, whereas in the latter, two models are used—one for the non-congested-flow regime and the other for the congested-flow regime. The idea of the two-regime representation was first proposed by Edie ($\underline{1}$). The general macroscopic models, their estimation approaches, and the scope of this paper are discussed first.

GENERAL MICROSCOPIC AND MACROSCOPIC MODELS

The general car-following (microscopic) equation developed by Gazis and others ($\underline{2},\underline{3}$) is given as follows:

$$\ddot{X}_{n+1}(t+T) = \alpha \left\{ \dot{X}_{n+1}^m(t+T)/[X_n(t) - X_{n+1}(t)]^\ell \right\} [\dot{X}_n(t) - \dot{X}_{n+1}(t)] \qquad (1)$$

where

$\dot{X}_n$, $\dot{X}_{n+1}$ = speed of leading and following vehicles, respectively;

$\ddot{X}_{n+1}$ = acceleration (or deceleration) rate of following vehicle;

T = time lag of response to stimulus;

α = constant of proportionality (referred to throughout as a model parameter); and

ℓ, m = model parameters.

By integrating Equation 1, the general form of macroscopic models has been developed by Gazis and others ($\underline{3}$). By using this general form, a matrix of macroscopic models has been established for different combinations of ℓ and m parameters by May and Keller ($\underline{4}$). This matrix has undergone some adjustments by Ceder ($\underline{5}$) and by Easa and May ($\underline{6}$). The final version of the matrix is shown in Figure 1, along with illustrations of its use for the two-regime representation.

Figure 1 shows the speed-density relations and consists of five regions. In regions 1 and 2, models have no intercept with the speed axis, $u_f \to \infty$. In regions 4 and 5, models have no intercept with the density axis, $k_j \to \infty$. Models in region 3 have intercepts with both the speed and the density axes. Single-regime representation is usually accomplished by using models from region 3. The two-regime representation can be made, as illustrated in Figure 1, by using models from regions 1, 2, or 3 for the congested-flow regime and from regions 3, 4, or 5 for the non-congested-flow regime.

ESTIMATION APPROACHES

Estimation of macroscopic models is an essential task. For a given set of traffic-flow data, one often needs to estimate model parameters that best represent these data. In this regard, an approach employing computer techniques has been developed by May and Keller ($\underline{4}$). This approach uses regression analysis to estimate model parameters for specified values of traffic-flow and statistical criteria. These criteria include free-flow speed u_f, optimum speed u_0, jam density k_j, optimum density k_0, maximum flow q_m, and a mean-deviation criterion.

In an attempt to significantly reduce the need for using computer systems, another theoretical-graphical approach has been recently proposed and applied to the estimation of single-regime models ($\underline{7}$). This approach is based principally on the theoretical relations among the first five criteria mentioned above and model parameters ℓ, m, and α. A simplified graphical tool was used to represent those relations and could directly provide model parameters that satisfy specified traffic-flow criteria. This approach was applied later to the estimation of a special case of two-regime models by Easa and May ($\underline{6}$). The procedure that has been developed for estimating the single-regime models corresponds to region 3 and that developed for the two-regime models corresponds to regions 2 and 4 and in a preliminary fashion to region 3. In both pro-

Figure 1. Macroscopic models and two-regime representation.
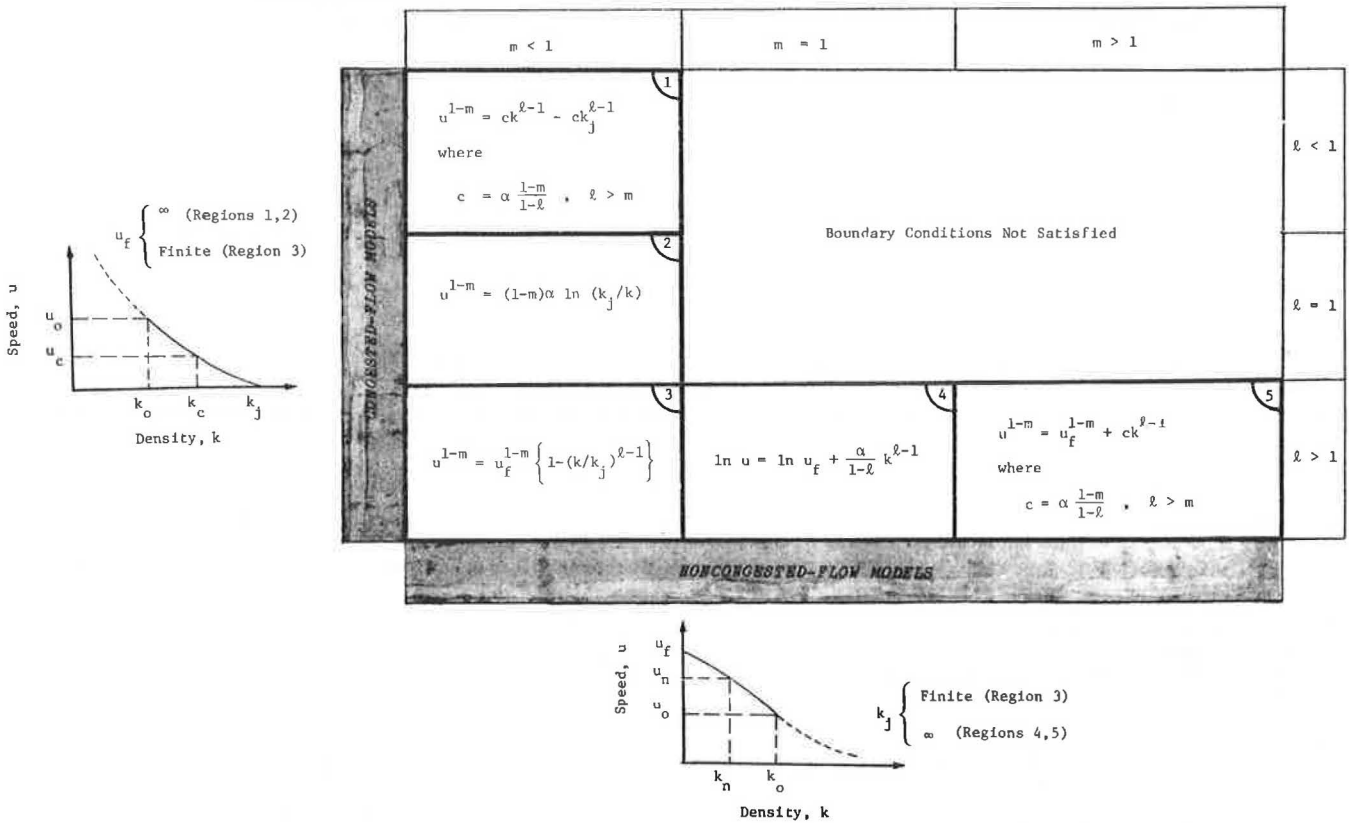
cedures, the traffic-flow criteria may be specified as single values or as ranges.

SCOPE OF PAPER

The primary purpose of this paper is to establish the theoretical relations of the remaining regions (1 and 5) and to develop a procedure for selecting the particular region--and the model within that region--that satisfies specified traffic-flow criteria for both non-congested-flow and congested-flow regimes. Specifically, the selection procedure will determine one of the three regions 1, 2, or 3 for the congested-flow regime and one of the three regions 3, 4, or 5 for the non-congested-flow regime and will provide the respective model parameters for each regime. The selection procedure is based principally on the theoretical properties of models among the five regions of Figure 1. In addition, two auxiliary criteria (u_n, k_n) and (u_c, k_c) were used to account for the variability of the traffic-flow patterns in the intermediate ranges of operations.

Throughout this paper, the traffic-flow criteria u_f, u_0, k_j, and k_0 (or a combination of them) will be referred to as the basic criteria, as distinguished from the auxiliary criteria. Furthermore, this paper is primarily concerned with selecting the two-regime models where the traffic-flow criteria--basic and auxiliary--are specified as single values. Thus, the maximum-flow criterion q_m is not needed since it is implicitly determined by the two criteria k_0 and u_0 ($q_m = k_0 u_0$).

The following section describes the theoretical properties of the models and regions of the non-congested-flow regime. The next section presents these properties for the congested-flow regime. Then, based on these properties, a description of

the procedure of region (and model) selection for each regime is given. After a section in which the procedure is applied to an actual set of traffic-flow data, there are some concluding remarks concerning practical aspects of the procedure.

NON-CONGESTED-FLOW REGIME

This section describes the theoretical properties of the non-congested-flow models of regions 3, 4, and 5. Those properties are presented in three parts: a summary of the previously reported results of the special case of region 4 and extended results of region 3, the theoretical relations of models in region 5, and between-region properties.

Previous Results

The speed-density equation of models in region 4 ($\ell > 1$, $m = 1$) is given as follows:

$$\ln u = \ln u_f + [\alpha/(1 - \ell)]k^{\ell-1} \tag{2}$$

The relations between the traffic-flow criteria u_f, u_0, and k_0 and model parameters ℓ and α are given by (6) the following:

$$\alpha = 1/k_0^{\ell-1} \tag{3}$$

$$u_0/u_f = \exp\left\{-[1/(\ell - 1)]\right\} \tag{4}$$

Equations 3 and 4 have been used to estimate model parameters based on specified traffic-flow criteria.

For models in region 3 ($\ell > 1$, $m < 1$), the speed-density equation is given as follows:

$$u^{1-m} = u_f^{1-m}[1 - (k/k_j)^{\ell-1}] \tag{5}$$

Based on Equation 5, the relations among the traffic-flow criteria u_f, u_0, k_j, and k_0 and model parameters ℓ and m are given by the following:

$$(k_0/k_j)^{\ell-1} = (1-m)/(\ell-m) \tag{6}$$

$$(u_0/u_f)^{1-m} = (\ell-1)/(\ell-m) \tag{7}$$

Equations 6 and 7 have been used for single-regime models to estimate model parameters based on specified traffic-flow criteria ($\underline{6,7}$).

In order to use region-3 models for the non-congested-flow regime, a particular modification should be made. In selecting a model for the non-congested-flow regime, the intercept with the density axis (k_j) is obviously not important. Therefore, it would seem reasonable to employ another (auxiliary) criterion within the non-congested-flow regime instead of the criterion k_j. This auxiliary criterion corresponds to a selected point from the data with speed and density denoted by u_n and k_n. The relation between this auxiliary criterion and the k_j criterion is given by the following:

$$k_j^{\ell-1} = k_n^{\ell-1}/[1-(u_n/u_f)^{1-m}] \tag{8}$$

Properties of Models in Region 5

Models in region 5 ($\ell > 1$, $m > 1$) are more general than the special-case models of region 4. For specific values of the basic criteria u_f, u_0, and k_0, region 5 provides a range of models, while region 4 provides only one model.

The speed-density relation of models in region 5, shown previously in Figure 1, is given by the following:

$$u^{1-m} = u_f^{1-m} + ck^{\ell-1} \qquad \ell > m \tag{9}$$

where

$$c = \alpha\,[(1-m)/(1-\ell)] \tag{10}$$

Equation 9 represents a model that has no intercept with the density axis, $k_j \to \infty$. The model contains three parameters--ℓ, m, and α. The relations between these parameters and the basic criteria can be established. From Equation 9, the relations between q and k and q and u can be obtained as follows:

$$q = k\,(u_f^{1-m} + ck^{\ell-1}) \tag{11}$$

$$q^{\ell-1} = (u^{\ell-1}/c)\,(u^{1-m} - u_f^{1-m}) \tag{12}$$

At maximum flow, $q'(k)_k = 0$ and $q'(u)_u = 0$. Therefore, by differentiating Equations 11 and 12 with respect to k and u, respectively, and equating the derivatives to zero, one obtains the following:

$$\alpha = [(\ell-1)/(\ell-m)]\,(u_f^{1-m}/k_0^{\ell-1}) \qquad \ell > m \tag{13}$$

$$(u_0/u_f)^{1-m} = (\ell-1)/(\ell-m) \qquad \ell > m \tag{14}$$

(Equation 14 is the same as Equation 7 of region 3.)

Equations 13 and 14 contain three parameters and in order to estimate these parameters, it is necessary to have a third condition. This condition may be represented by the auxiliary criteria (u_n,k_n). Thus, substituting these criteria into Equation 9, we have the following:

$$u_n^{1-m} = u_f^{1-m} + ck_n^{\ell-1} \tag{15}$$

Now, Equations 13, 14, and 15 can be solved for specified values of u_f, u_0, k_0, u_n, and k_n

to determine model parameters ℓ, m, and α.

It is important to note that models in region 5 are considered valid for only $\ell > m$. For $\ell \leq m$, neither the $q - k$ nor the $q - u$ relations will have a maximum point; $q'(k)_k > 0$ for all k and $q'(u)_u < 0$ for all u. Although such relations may be used to represent the non-congested-flow regime, they clearly do not describe properly the behavior of the traffic flow near capacity. Consequently, models corresponding to $\ell \leq m$ are considered undesirable for the non-congested-flow regime (they are certainly not meaningful for the single-regime representation).

Between-Region Properties

The theoretical properties of models among regions 3, 4, and 5 will now be presented. A fundamental property is that models in region 4 are a limiting case of both models in region 3 and models in region 5. This property is proved below for only models in regions 3 and 4; the proof for regions 5 and 4 will be similar.

Substituting for k_j from Equation 6 into Equation 5 (region 3) gives the following:

$$u^{1-m} = u_f^{1-m}\left\{1 - (k/k_0)^{\ell-1}\,[(1-m)/(\ell-m)]\right\}$$

To prove that the above model approaches the one corresponding to region 4, Equation 2, as $m \to 1$, let us first express k as a function of u:

$$k^{\ell-1} = k_0^{\ell-1}\,[(\ell-m)/(1-m)]\,[1-(u/u_f)^{1-m}] \tag{16}$$

As $m \to 1$, the limit of Equation 16 is equal to zero divided by zero. Therefore, one may apply L'Hospital's rule.

Let

$$f(m) = (\ell-m)\,k_0^{\ell-1}\,[1-(u/u_f)^{1-m}] \tag{17}$$

$$g(m) = 1 - m \tag{18}$$

Then

$$\lim_{m\to 1}\,[f(m)/g(m)] = \lim_{m\to 1}\,[f'(m)/g'(m)] \qquad g'(m) \neq 0 \tag{19}$$

where the prime represents the first derivative with respect to m.

The derivatives in Equation 19 can be obtained from Equations 17 and 18 as follows:

$$f'(m) = k_0^{\ell-1}\,[-(\ell-m)\,u^{1-m}\ln(u/u_f) - (u/u_f)^{1-m} + 1]$$

$$g'(m) = -1$$

Substituting into Equation 19 gives

$$\lim_{m\to 1}\,[f(m)/g(m)] = (\ell-1)\,k_0^{\ell-1}\ln(u/u_f)$$

Thus, as $m \to 1$, Equation 16 becomes

$$k^{\ell-1} = (\ell-1)\,k_0^{\ell-1}\ln(u/u_f)$$

or

$$\ln u = \ln u_f + [1/(\ell-1)]\,(k/k_0)^{\ell-1}$$

which is the same as Equation 2 of region 4; note that $\alpha = 1/k_0^{\ell-1}$ from Equation 3.

It follows from this limiting-case property that the characteristics of models in regions 3 and 5 should, as $m \to 1$, approach those of region 4. For example, it can be easily shown that by taking the limit of α and the limit of u_0/u_f (Equations 13 and 14) of region 5, as $m \to 1$, the correspond-

ing formulas of region 4 are obtained (Equations 3 and 4). The continuity of the u_0/u_f ratio in regions 3, 4, and 5 is illustrated in Figure 2.

Another important aspect that will be useful in model selection is the continuity of the location of the inflection point and the associated convexity and concavity characteristics of the speed-density curves. For $\ell \leq 2$, models in the three regions are entirely concave, as illustrated in Figure 2. That can be expressed mathematically as follows:

$$u''(k)_k > 0 \quad 0 \leqslant k \leqslant k_j \tag{20}$$

for the $(\ell \leq 2, m \geq 0)$ space except at $\ell = 2$, $m = 0$, where $u''(k)_k = 0$ (the double prime represents the second derivative with respect to k).

For $\ell > 2$, models in the three regions are mixed; they consist of two portions (convex and concave). The concave portion diminishes at $m = 0$; that is, models become entirely convex. When the inflection point coincides with k_0, the non-congested-flow regime becomes totally convex. If the density corresponding to the inflection point is denoted by k_t, the above characteristics can be expressed mathematically as follows:

$$u''(k)_k < 0 \quad 0 \leqslant k \leqslant k_j$$

for $\ell > 2$, $m = 0$, and

$$u''(k)_k < 0 \quad 0 \leqslant k < k_t$$
$$u''(k)_k > 0 \quad k_t < k \leqslant k_j \tag{21}$$

for $\ell > 2$, $m > 0$, and $0 < k_t < k_j$. The location of models where $k_t = k_0$ is given by $m = \ell - 2$. This linear function separates the models where the non-congested-flow regime is totally convex from those where it is mixed, as shown in Figure 2. Also shown in Figure 2 is the location of models with an inflection point lying at half the optimum density, $k_t = 0.5k_0$.

As noted in Figure 2, all convex, concave, and mixed models converge to a linear function at $\ell = 2$, $m = 0$ [Greenshields' model (8)]. Also, in region 4, the two models by Drew (9), $\ell = 1.5$, and by Underwood (10), $\ell = 2$, are totally concave. The model by Drake, Schoefer, and May, $\ell = 3$, is convex in the non-congested-flow regime and concave in the congested-flow regime; $k_t = k_0$.

A final property of the non-congested-flow models is that, for specific values of the basic criteria u_f, u_0, and k_0, the following relation can be shown for the entire range of the traffic density, except at $k = 0$ and $k = k_0$ where the values of u coincide:

$$u_{(3)} < u_{(4)} < u_{(5)} \tag{22}$$

The subscripts in Equation 22 refer to the region number, and the values of u correspond to Equations 5, 2, and 9, respectively. Both $u_{(3)}$ and $u_{(5)}$ approach $u_{(4)}$ as $m \to 1$, as proved earlier. This property is illustrated in Figure 3, which shows three models from regions 3, 4, and 5 drawn for the specific criteria $u_f = 50$ mph, $u_0 = 25$ mph, and

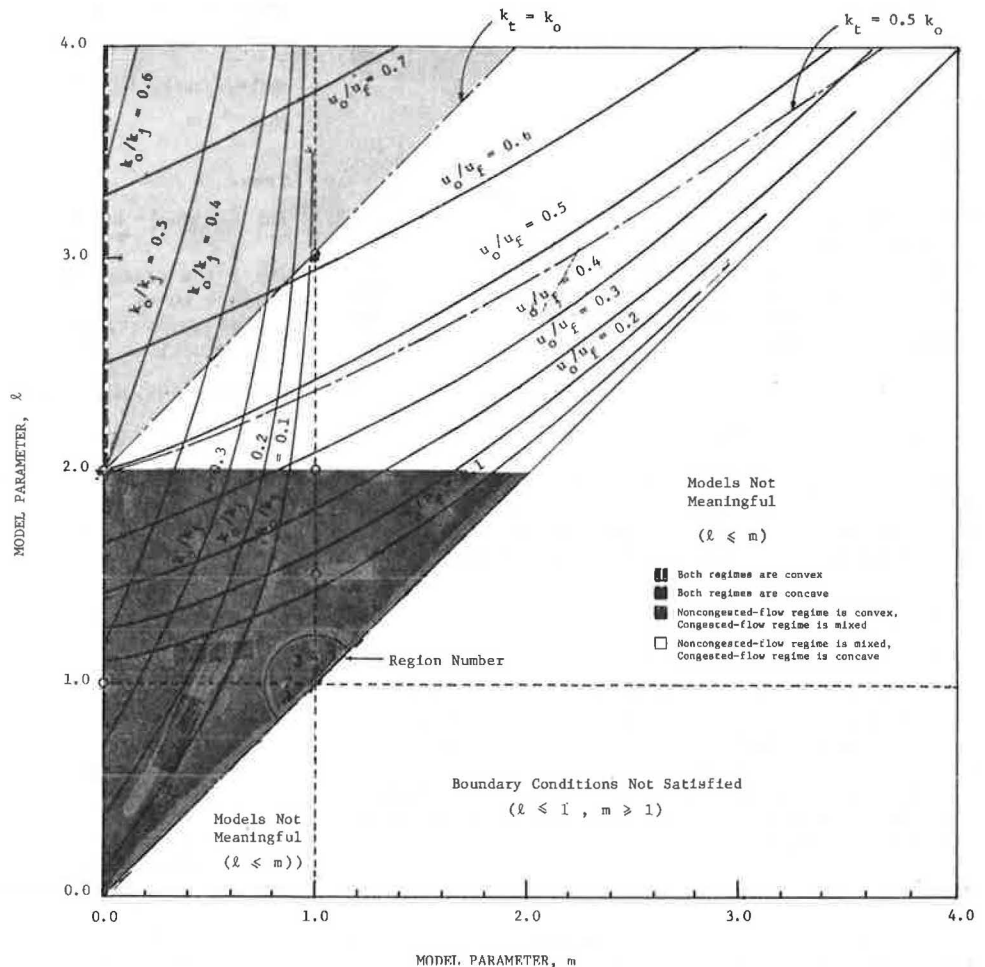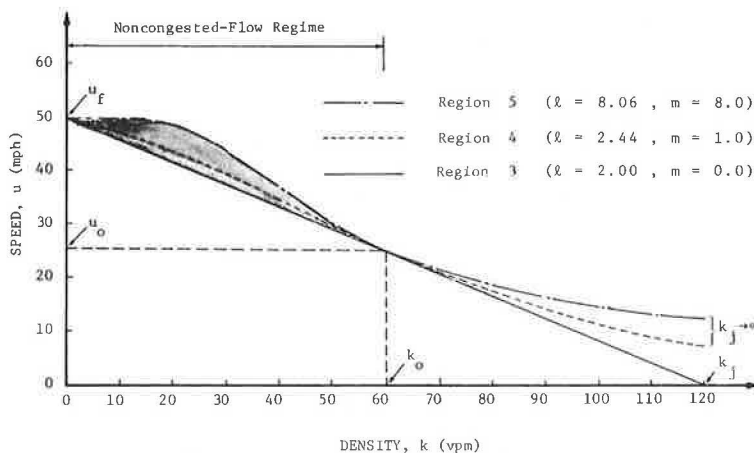Figure 2. Continuity of model characteristics, convexity, and concavity.

Figure 3. Continuity of non-congested-flow models.

$k_0 = 60$ vehicles per mile (vpm). There is only one model from region 4 (dashed curve) that satisfies the above criteria. The range below the dashed curve corresponds to models from region 3, while the range above it corresponds to models from region 5. For these specified criteria, the model from region 3 represents a lower bound of the models from that region since it corresponds to $m = 0$. On the other hand, there exist other models from region 5 that lie above the one shown in Figure 3.

Selecting a particular model from the range of models shown in Figure 3 can be accomplished by the auxiliary criteria. Such a model will satisfy the five criteria of that regime. In addition, an even better fit to the data between these criteria can be obtained by the knowledge of the convexity and concavity of the models. These aspects represent the basic elements of the procedure of model selection to be described later.

CONGESTED-FLOW REGIME

The congested-flow regime corresponds, as stated previously, to the models of regions 1, 2, and 3. The purpose of this section is to describe the theoretical properties of models and regions of this regime. The description is given in three parts: a summary of the previously reported results of region 2 and extended results of region 3, the theoretical relations of models in region 1, and between-region properties.

Previous Results

The speed-density relation of models in region 2 ($\ell = 1$, $m < 1$), shown previously in Figure 1, is given by the following:

$$u^{1-m} = \alpha (1 - m) \ln (k_j/k) \tag{23}$$

The relations between the traffic-flow criteria k_j, u_0, and k_0 and model parameters m and α are given by ([6]) the following:

$$\alpha = u_0^{1-m} \tag{24}$$

$$k_0/k_j = \exp\left\{ - [1/(1-m)] \right\} \tag{25}$$

For models in region 3, the speed-density relation and the relations between the basic traffic-flow criteria and the model parameters are as given previously in Equations 5, 6, and 7. In order to use these models for the congested-flow regime, a similar modification to the one employed previously should be made. Specifically, for the congested-

flow regime, the intercept with the speed axis, u_f, is not of major importance and is replaced here by an auxiliary criterion (u_c, k_c), which lies somewhere within the congested-flow regime. The relation between this auxiliary criterion and the u_f criterion is given by the following:

$$u_f^{1-m} = u_c^{1-m}/[1 - (k_c/k_j)^{\ell-1}] \tag{26}$$

Properties of Models in Region 1

Models in region 1 ($\ell < 1$, $m < 1$) are more general than the special-case models of region 2. For specific values of the basic criteria k_j, u_0, k_0, region 1 provides a range of models, while region 2 provides only one model.

The speed-density relation of models in region 1, shown previously in Figure 1, is given by

$$u^{1-m} = c (k^{\ell-1} - k_j^{\ell-1}) \qquad \ell > m \tag{27}$$

where $c = \alpha [(1 - m)/(1 - \ell)]$.

Equation 27 represents a model that has no intercept with the speed axis, $u_f \to \infty$, and contains three parameters. The relations between these parameters and the basic criteria can be established. From Equation 27 one can obtain q as a function of k and as a function of u as follows:

$$q = k [c (k^{\ell-1} - k_j^{\ell-1})]^{1/1-m} \tag{28}$$

$$q^{\ell-1} = u^{\ell-1} [(u^{1-m}/c) + k_j^{\ell-1}] \tag{29}$$

Following a similar procedure to that described for region 5, α and k_0/k_j can be obtained as follows:

$$\alpha = [(\ell-m)/(1-m)] (u_0^{1-m}/k_j^{\ell-1}) \qquad \ell > m \tag{30}$$

$$(k_0/k_j)^{\ell-1} = (1-m)/(\ell-m) \qquad \ell > m \tag{31}$$

(Equation 31 is the same as Equation 6 of region 3.)

The auxiliary criteria (u_c, k_c) can now be used to provide a third relation so that the three parameters may be estimated. Substituting for these criteria into Equation 27, one obtains the following:

$$u_c^{1-m} = c (k_c^{\ell-1} - k_j^{\ell-1}) \tag{32}$$

Equations 30, 31, and 32 can be solved to determine model parameters ℓ, m, and α for specified values of k_j, u_0, k_0, u_c, and k_c.

It is noted that models in region 1 are considered valid for only $\ell > m$. For $\ell \leq m$, the $q - k$ and $q - u$ relations will not have a maximum

point; $q'(k)_k < 0$ for all k and $q'(u)_u > 0$ for all u. For example, for $\ell = m = 0$, Equation 28 will be a straight line having an intercept α with the flow axis, and Equation 29 will be an increasing function; q approaches α as u approaches infinity. The characteristics of these models near capacity are considered here undesirable for representing the congested-flow regime. Moreover, these models are certainly not meaningful for the single-regime representation.

Between-Region Properties

The congested-flow models of regions 1, 2, and 3 exhibit similar properties to those of the non-congested-flow models described in the previous section. Specifically, as $\ell \to 1$, models in region 2 are a limiting case of both models in region 1 and models in region 3. As a result, there is also a continuity of model characteristics and for specific values of k_j, k_0, and u_0, models in the three regions exhibit a particular pattern.

The limiting-case property of only models in regions 1 and 2 (both regions have no intercept with the speed axis, $u_f \to \infty$) will now be proved. This will perhaps complement the presentation in the previous section where the limiting-case property of regions 3 and 4 (one region with finite k_j and the other with $k_j \to \infty$) was proved.

Substituting for c into Equation 27 (region 1) gives the following:

$$u^{1-m} = \alpha \left[(1-m)/(1-\ell)\right] (k^{\ell-1} - k_j^{\ell-1}) \tag{33}$$

Since the limit of the product is the product of the limits, the limit of Equation 33 is equal to the limit of α multiplied by the limit of the remainder, as $\ell \to 1$. The limit of α, Equation 30, is equal to u_0^{1-m} and the limit of the remaining part can be obtained by applying L'Hospital's rule, given previously in Equation 19.

Let

$$f(\ell) = (1-m)(k^{\ell-1} - k_j^{\ell-1}) \tag{34}$$

$$g(\ell) = 1 - \ell \tag{35}$$

Applying Equation 19 gives the following:

$$\lim_{\ell \to 1} \left[f'(\ell)/g'(\ell)\right] = (1-m)\ln(k_j/k) \tag{36}$$

Thus, as $\ell \to 1$, Equation 30 becomes the following:

$$u^{1-m} = u_0^{1-m}(1-m)\ln(k_j/k)$$

which is the same as Equation 23 of region 2; note that $\alpha = u_0^{1-m}$ from Equation 24.

It follows from this limiting-case property that the characteristics of models in regions 1 and 3, α and (k_0/k_j), should approach those of region 2 as $\ell \to 1$. The continuity of the (k_0/k_j) ratio is illustrated in Figure 2. It is noted that region 3 exhibits a wide range of this ratio, while regions 1 and 2 show a narrow range with a maximum value of $(k_0/k_j) = 0.37$, which corresponds to Greenberg's model (12).

The congested-flow models of regions 1, 2, and 3 may be convex, concave, or mixed. These characteristics were described previously for region 3. Models in regions 1 and 2 are entirely concave as illustrated in Figure 2. Thus,

$$u''(k)_k > 0 \qquad 0 \leq k < k_j \tag{37}$$

for $\ell \leq 1$ and $m \geq 0$.

Finally, for specific values for the basic criteria k_j, u_0, and k_0, it can be shown that the following relation holds:

$$u_{(3)} < u_{(2)} < u_{(1)} \tag{38}$$

for the entire range of the traffic density, except at $k = k_0$ and $k = k_j$, where the values of u coincide. Both $u_{(3)}$ and $u_{(1)}$ approach $u_{(2)}$ as $\ell \to 1$, as proved previously.

An illustration of this property is shown in Figure 4 for the specific criteria $k_j = 200$ vpm, $u_0 = 25$ mph, and $k_0 = 60$ vpm. The dashed curve represents the (only) model from region 2 that satisfies these criteria. The model from region 1 represents an upper bound of models from that region. On the other hand, the model from region 3 is a specific selection and other models below the solid curve do exist. Thus, the range above the dashed curve corresponds to models from region 1 and the range below it corresponds to models from region 3. Selecting a model from the range of models shown in Figure 4 can be accomplished by means of the auxiliary criteria. The selection procedure for the non-congested-flow and congested-flow regimes is described in the following section.

MODEL SELECTION

The continuity property of the non-congested-flow and congested-flow models described in the previous sections represents the basis for the selection procedure presented in this section. The procedure is described first for the non-congested-flow regime and then for the congested-flow regime.

Non-Congested-Flow Regime

As mentioned previously, the non-congested-flow regime can be represented by a model from regions 3, 4, or 5. For specified traffic-flow criteria u_f, u_0, k_0, u_n, and k_n, it is necessary to determine first the region that contains the model satisfying these criteria and then to determine that model. It is important to note that since models in the three regions exhibit the continuity property, these models do not intersect, and therefore there is only one specific model that satisfies the above five criteria. The relations between the traffic-flow criteria and model parameters for regions 3, 4, and 5 will now be obtained.

For region 5, the formula that considers the auxiliary criterion (Equation 15) can be written as follows:

$$u_n^{1-m} = u_f^{1-m} + c(\beta_n k_0)^{\ell-1} \tag{39}$$

where β_n is a standardized variable given by the following:

$$\beta_n = k_n/k_0 \tag{40}$$

Substituting for c from Equation 10 (and for α from Equation 13) and dividing both sides of Equation 39 by u_f^{1-m} gives the following:
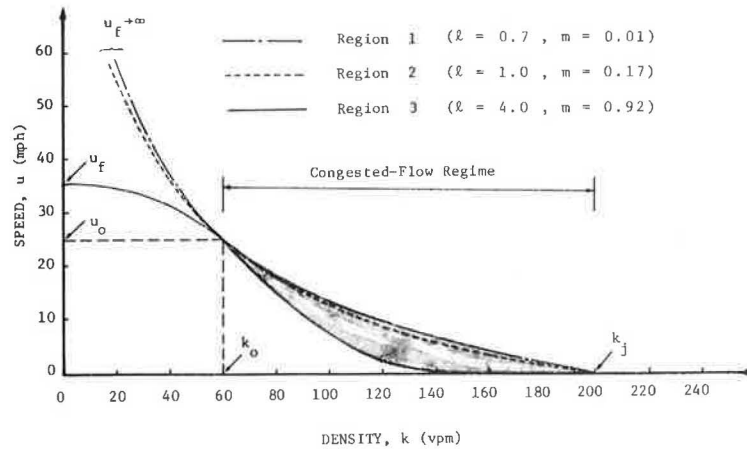
$$(u_n/u_f)^{1-m} = 1 - \left[(1-m)/(\ell-m)\right]\beta_n^{\ell-1} \tag{41}$$

From Equation 14, we have

$$\ell = \left[m(u_0/u_f)^{1-m} - 1\right]/\left[(u_0/u_f)^{1-m} - 1\right] \tag{42}$$

Equations 41 and 42 relate model parameters ℓ and m to the traffic-flow criteria u_f, u_0, u_n,

Figure 4. Continuity of congested-flow models.

and β_n (which incorporates k_n and k_0). For region 3, identical formulas to the above can be obtained from Equations 6, 7, and 8.

For region 4, the relations between ℓ and the traffic-flow criteria can be obtained, based on Equations 2, 3, and 4, as follows:

$$\ln(u_n/u_f) = [1/(1-\ell)] \beta_n^{\ell-1} \tag{43}$$

$$\ell = 1 - [1/\ln(u_0/u_f)] \tag{44}$$

The above formulas are shown in Figure 5 for $\beta_n = 0.5$ and for various ranges of the other traffic-flow criteria (the value $\beta_n = 0.5$ indicates that the auxiliary criterion u_n, k_n lies at a point where k_n is half the optimum density k_0). The thick solid curve in Figure 5 corresponds to region 4. This curve corresponds to Equations 43 and 44 (or equivalently to the limits of Equations 41 and 42 as $m \to 1$). The area above that curve corresponds to region 5 and the one below it corresponds to region 3. Models below the curve $m = 0$ (region 3) correspond to negative values of m and are considered undesirable; a negative m has the effect of shifting the speed term in Equation 1 from the numerator to the denominator. Figure 5 also shows the curvature properties of models.

Now, the selection procedure is as follows:

1. Determine from the data the criteria u_f, u_0, k_0, u_n, and k_n where k_n is to be chosen as $0.5 k_0$.
2. Follow the sample arrows shown in Figure 5 and determine the intersection point. Read the corresponding values of ℓ and m. If the intersection point lies above the thick solid curve, the model satisfying the specified criteria lies in region 5. Estimate the value of α from Equation 13 and obtain the traffic-flow relations corresponding to this model from Equations 9 through 12.
3. If the intersection point lies below the thick solid curve, the model we are after is located in region 3. Obtain the traffic-flow relations corresponding to the model from Equation 5 (the q-u and q-k relations can be easily obtained from this equation).
4. If the intersection point lies on the thick solid curve, the model lies in region 4. Estimate the value of α from Equation 3 and obtain traffic-flow relations corresponding to this model from Equation 2.

In order to help the user determine the values of ℓ and m accurately, the numerical values on which the nomograph of Figure 5 is based are provided in

Table 1. Table 1 will be particularly useful for the upper portion of the nomograph where ℓ becomes very close to m and the difference cannot be easily detected from Figure 5.

It is important to note that, based on the exhibited pattern of the non-congested-flow regime data, some adjustments of the traffic-flow criteria may be deemed necessary in order to select the model that is compatible with such a pattern. For example, if the data exhibit a convex shape, the traffic-flow criteria may be properly adjusted so that the intersection point in Figure 5 lies in the area delineated in the upper portion of this figure. This adjustment will ensure an even better fit to the data between the traffic-flow criteria.

Congested-Flow Regime

The objective here is to select one of the three regions--1, 2, or 3--and the model within the selected region that satisfies the specified criteria k_j, u_0, k_0, u_c, and k_c. Because of the continuity property of the congested-flow models, there is only one specific model that satisfies these criteria. Let us first estimate model parameters for regions 1 and 2 and then outline the procedure of model selection.

Similar to the procedure adopted for region 5 of the non-congested-flow regime, the formulas relating model parameters to the traffic-flow criteria in region 1 can be obtained, based on Equations 30, 31, and 32, as follows:

$$(u_c/u_0)^{1-m} = [(\ell-m)/(1-\ell)] [(\beta_c k_0/k_j)^{\ell-1} - 1] \tag{45}$$

$$m = [1 - \ell(k_0/k_j)^{\ell-1}]/[1 - (k_0/k_j)^{\ell-1}] \tag{46}$$

where β_c is a standardized variable given by the following:

$$\beta_c = k_c/k_0 \tag{47}$$

For region 3, identical formulas to the above can be obtained from Equations 6, 7, and 26.

For region 2, the corresponding relations, based on Equations 23, 24, and 25, are given:

$$(u_c/u_0)^{1-m} = (1-m) \ln(k_j/k_0\beta_c) \tag{48}$$

$$m = 1 + [1/\ln(k_0/k_j)] \tag{49}$$

The above formulas are shown in Figure 6 for $\beta_c = 1.5$ and 2.0. The thick solid curve corresponds to region 2. The area above that curve corresponds to region 1 and the one below it corre-

sponds to region 3. The curvature properties of models are also shown in Figure 6.

Now, the selection procedure is as follows:

1. Determine from the data the criteria k_j, u_0, k_0, and k_c, where k_c is to be chosen as either $1.5k_0$ or $2k_0$.

2. Follow the sample arrows shown in Figure 6 and determine the intersection point. Read the corresponding values of ℓ and m. If the intersection point lies above the thick solid curve for the corresponding β_c, the model lies in region 1. Es-

timate the value of α from Equation 30 and obtain the traffic-flow relations corresponding to this model from Equations 27 through 29.

3. If the intersection point lies below the thick solid curve, the model lies in region 3 and the corresponding traffic-flow criteria can be obtained from Equation 5.

4. If the intersection point lies on the thick solid curve, the model lies in region 2. Estimate the value of α from Equation 24 and obtain the corresponding traffic-flow relations from Equation 23.

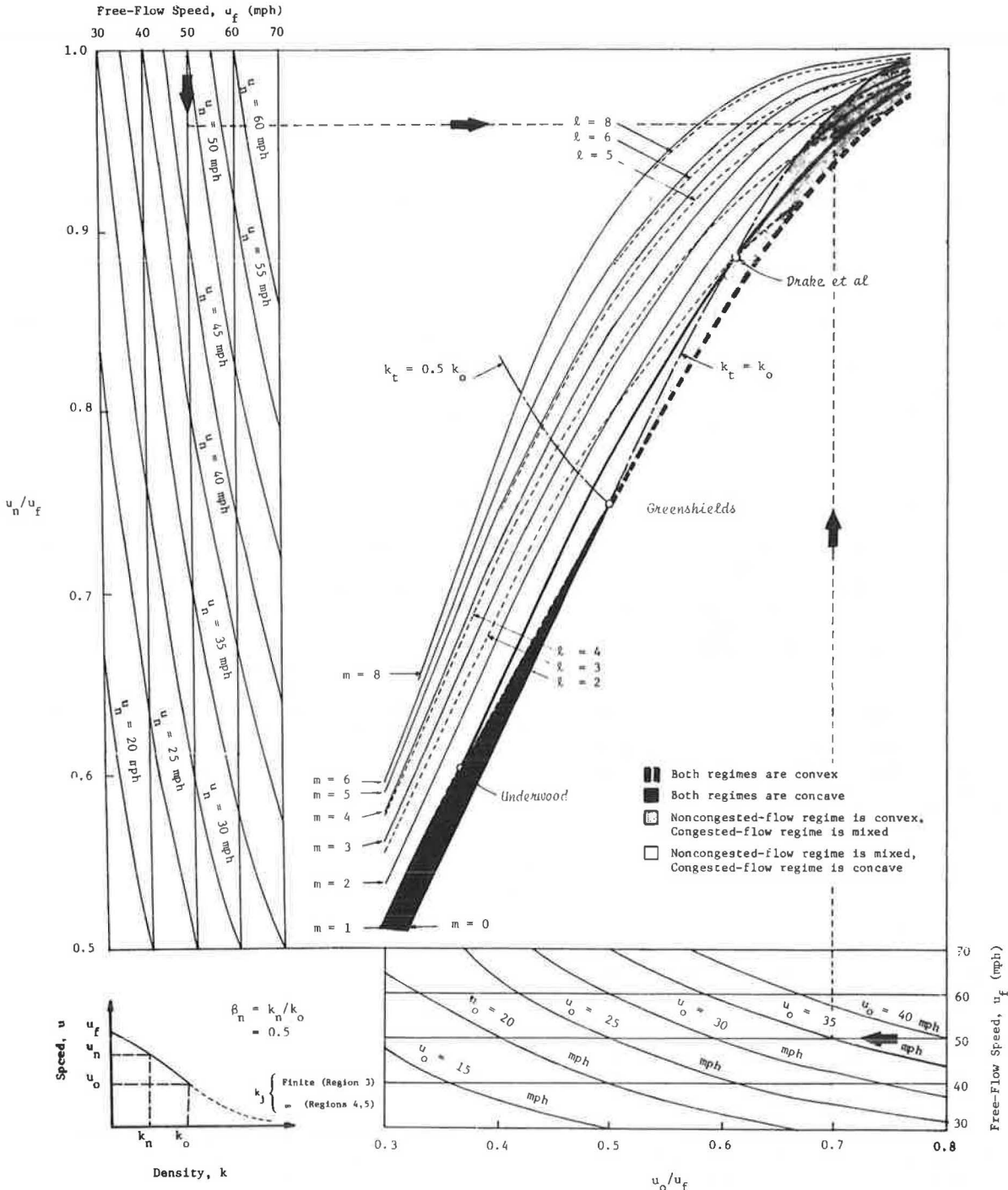Figure 5. Nomograph for non-congested-flow regime (regions 3, 4, and 5).

33

Table 1. Model parameters for non-congested-flow regime (regions 3, 4, and 5).

(u_0/u_f)	ℓ	m	(u_n/u_f)	(u_0/u_f)	ℓ	m	(u_n/u_f)
0.3	1.429	0.000	0.480	0.6	2.500	0.000	0.859
	1.494	0.200	0.485		2.585	0.200	0.862
	1.566	0.400	0.491		2.673	0.400	0.866
	1.647	0.600	0.497		2.764	0.600	0.870
	1.735	0.800	0.502		2.859	0.800	0.873
	1.831	1.000	0.508		2.958	1.000	0.877
	2.429	2.000	0.536		3.500	2.000	0.895
	3.198	3.000	0.559		4.125	3.000	0.911
	4.083	4.000	0.575		4.827	4.000	0.927
	5.033	5.000	0.586		5.596	5.000	0.941
	6.012	6.000	0.592		6.422	6.000	0.953
	7.004	7.000	0.596		7.294	7.000	0.962
	8.002	8.000	0.598		8.202	8.000	0.970
	9.001	9.000	0.599		9.137	9.000	0.977
0.4	1.667	0.000	0.662	0.7	3.333	0.000	0.940
	1.740	0.200	0.628		3.423	0.200	0.942
	1.819	0.400	0.633		3.514	0.400	0.944
	1.904	0.600	0.639		3.608	0.600	0.946
	1.994	0.800	0.645		3.705	0.800	0.948
	2.091	1.000	0.650		3.804	1.000	0.950
	2.667	2.000	0.679		4.333	2.000	0.959
	3.381	3.000	0.706		4.922	3.000	0.967
	4.205	4.000	0.729		5.566	4.000	0.974
	5.105	5.000	0.747		6.264	5.000	0.980
	6.052	6.000	0.761		7.010	6.000	0.985
	7.025	7.000	0.772		7.800	7.000	0.989
	8.011	8.000	0.779		8.628	8.000	0.992
	9.005	9.000	0.785		9.489	9.000	0.994
0.5	2.000	0.000	0.750	0.75	4.00	0.000	0.969
	2.079	0.200	0.755		4.091	0.200	0.970
	2.163	0.300	0.760		4.185	0.400	0.971
	2.252	0.600	0.765		4.280	0.600	0.972
	2.345	0.800	0.770		4.377	0.800	0.973
	2.443	1.000	0.775		4.476	1.000	0.974
	3.000	2.000	0.800		5.000	2.000	0.980
	3.667	3.000	0.824		5.571	3.000	0.984
	4.429	4.000	0.846		6.189	4.000	0.988
	5.267	5.000	0.866		6.851	5.000	0.991
	6.161	6.000	0.883		7.556	6.000	0.993
	7.095	7.000	0.897		8.299	7.000	0.995
	8.055	8.000	0.909		9.078	8.000	0.997
	9.031	9.000	0.918		9.890	9.000	0.998

Criterion	Regime	
	Non-Congested Flow	Congested Flow
u_f (mph)	50	NA
u_n (mph)	48	-
k_n (vpm)	25	-
u_0 (mph)	31	35
k_0 (vpm)	50	50
k_j (vpm)	-	200
u_c (mph)	-	12
k_c (vpm)	-	100

It is useful at this point to comment on establishing the traffic-flow criteria. As noted in the table above and Figure 7, the optimum speed is 35 mph for the non-congested-flow regime and 31 mph for the congested-flow regime. These values were determined by establishing first the maximum flow values in Figure 7 (bottom) corresponding to k_0 = 50 vpm, which was arbitrarily chosen. Then these values were used to determine the optimum speed. The auxiliary criteria were chosen at the mid-range of the non-congested-flow regime (β_n = 0.5) and at a point twice the optimum density for the congested-flow regime (β_c = 2.0). Establishing the remaining criteria u_f and k_j will generally require a knowledge of the characteristics of the highway facility. In particular, data points near the jam density are not usually obtainable and one would rely on previous data on the respective freeway or other similar freeways. As a general guideline, the jam density ranges from 180 to 250 vpm, which correspond to low-design and high-design facilities, respectively. Since the Eisenhower Freeway is an old-design facility of the 1950s, the jam density was arbitrarily chosen as 200 vpm. Now we will determine the non-congested-flow and congested-flow models that satisfy the specified traffic-flow criteria by following the procedure described in the previous section.

For the non-congested-flow regime, we first use Figure 5 and draw lines corresponding to the specified criteria shown in the table above. Since the intersection point lies above the thick solid curve, the model we are seeking lies in region 5. The values of model parameters can be read from the nomograph as ℓ = 4.3 and m = 2.0. Since the model lies in the upper-delineated portion of Figure 5, the result will be that the speed-density relation is convex in the entire non-congested-flow regime. Because the speed-density data also appear to exhibit a convex shape, the model will provide a good fit to the data between the traffic-flow criteria. In cases where the convexity and concavity characteristics are not the same for the data and the determined model, one may adjust the traffic-flow criteria to achieve this similarity. This adjustment can be made with the aid of the curvature properties shown in Figure 5.

With the above values of ℓ and m and the specified criteria, the value of α can be calculated from Equation 13. Thus, the three parameters in the traffic-flow relations of the non-congested-flow regime are known. The speed-density relation (Equation 9), for example, becomes the following:

$$u = 1/(0.02 + 2.15 \times 10^{-8} k^{3.3}) \qquad 0 < k < 50 \qquad (50)$$

The speed-density relation of Equation 50 is shown in Figure 7 (top) and is convex, as expected. The flow-density relation is also shown in Figure 7 (bottom). As seen, the selected model provides an excellent fit to the data, which is largely achieved by the proper selection of the traffic-flow criteria.

For the congested-flow regime, the intersection point corresponding to the specified criteria was determined as shown in Figure 6. As seen, this

It should be noted that if the intersection point lies above the curve m = 0 (models not desirable), the traffic-flow criteria should be adjusted so that the intersection point lies on or below that curve.

APPLICATION

The selection procedure described in the previous section was applied to traffic-flow data that have been collected on the Eisenhower Freeway at Harlem. The speed-density and flow-density data are shown in Figure 7 (top and bottom, respectively). The data exhibit some discontinuity, which can be seen more clearly from the flow-density graph. An excellent discussion of the possible reasons for the existence of discontinuity in some traffic-flow data is given by Gazis, Herman, and Rothery (3).

The traffic-flow criteria can now be established for both non-congested-flow and congested-flow regimes. For the non-congested-flow regime, the criteria include u_f, u_0, k_0, u_n, and k_n. Those for the congested-flow regime include k_j, u_0, k_0, u_c, and k_c. Because of the discontinuity, however, the optimum density u_0 need not be the same for both regimes. These criteria can be established from the data in such a way that the model corresponding to the respective criteria for each regime provides a good visual fit to the data of the particular regime. The specified criteria are given in the table below and are indicated by circles in Figure 7 (NA = not applicable):

point lies above the thick curve corresponding to $\beta_c = 2.0$. Therefore, the model we are seeking is located in region 1 and the speed-density relation will be concave in the entire congested-flow regime. The values of model parameters can be read from Figure 6 as $\ell = 0.5$ and $m = 0$.

With the above values of ℓ and m and the traffic-flow criteria, the value of α can be calculated from Equation 30. The three parameters for the congested-flow regime are now known and so are the traffic-flow relations. The speed-density rela-

tion (Equation 27), for example, becomes the following:

$$u = (438.4/k^{0.5}) - 31.0 \qquad 50 \leqslant k \leqslant 200 \qquad (51)$$

The speed-density relation of Equation 51 is shown in Figure 7 (top). The flow-density relation is also shown in Figure 7 (bottom). Again, the selected model provides excellent fit to the data.

CONCLUDING REMARKS

This paper has presented a theoretical procedure for

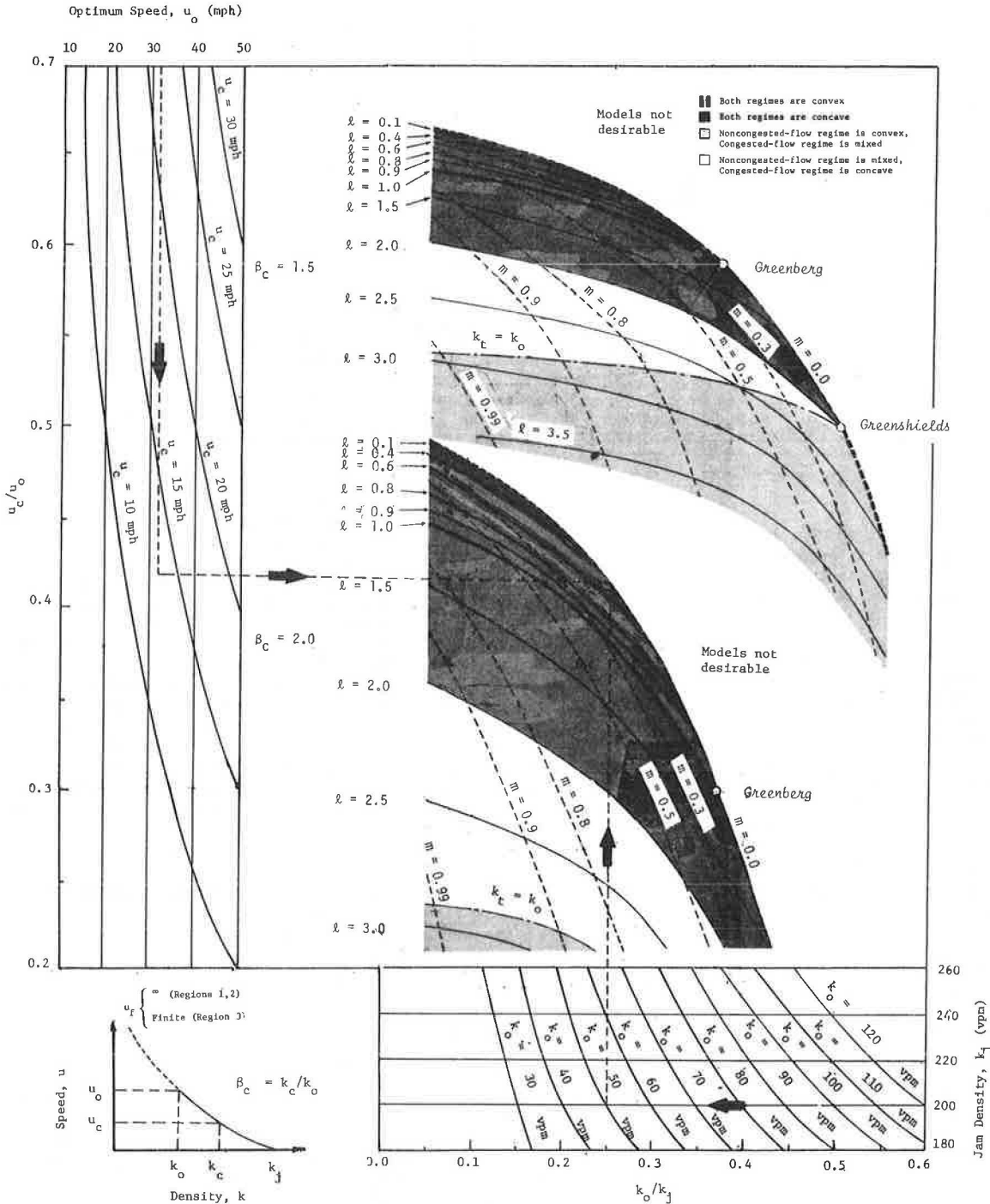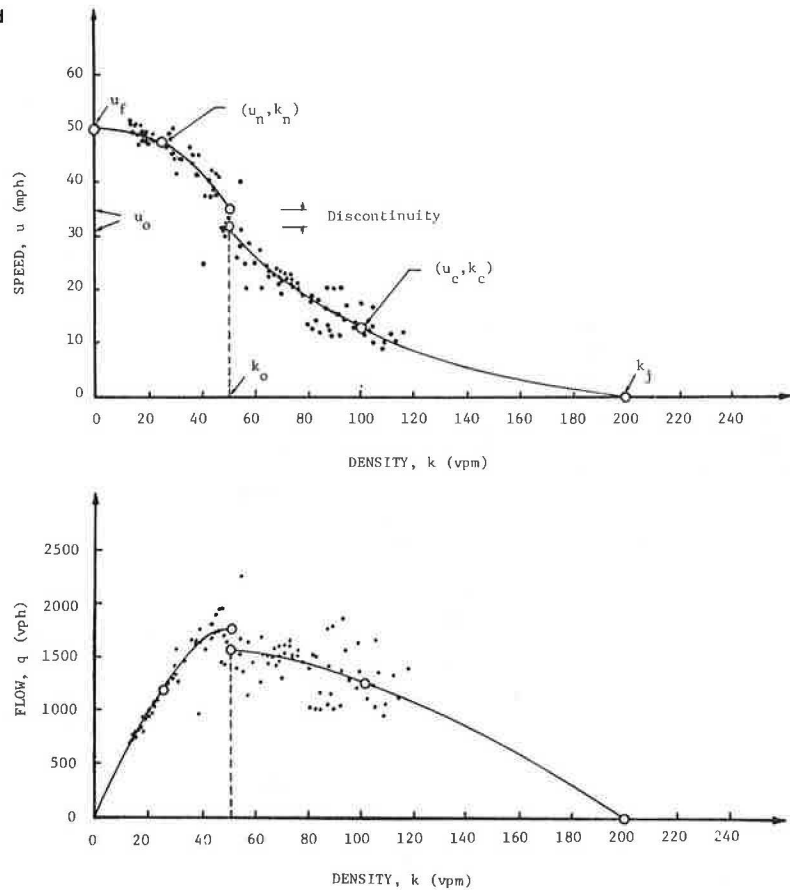Figure 6. Nomograph for congested-flow regime (regions 1, 2, and 3).

Figure 7. Speed-density and flow-density data and selected models.

DENSITY, k (vpm)

DENSITY, k (vpm)

selecting two-regime traffic-flow models and has demonstrated its use by applying it to actual data. The procedure is based principally on the between-region properties, including the continuity of models and the convexity and concavity character-istics. These properties have allowed a systematic procedure for selecting the region and the model within the region that satisfies the specified traffic-flow criteria and further maintains a good fit to the data in the intermediate ranges of flow. A few observations concerning the selection pro-cedure and some of its practical applications are worthy of note:

1. The two-regime representation will generally provide a better fit to the traffic-flow data than the single-regime representation, especially if there is a wide range of flow disturbance (discon-tinuity) near capacity. This is mainly due to the fact that two-regime models account for the vari-ability of the data in the intermediate ranges of operations, by means of the auxiliary criteria, while single-regime models consider only the basic criteria. This superiority of the two-regime over the single-regime representation has also been demonstrated by using the regression-analysis pro-cedure (13). It is therefore desirable that the two-regime representation be employed in practice whenever possible. In this regard, it is important to point out that the two-regime representation can be used whether there is discontinuity or not in the traffic-flow data.

2. As illustrated in the application described in this paper, the proposed procedure provides a level of data fit that appears to be reasonable for most practical applications. In situations in which

a particularly high degree of accuracy is deemed necessary, the regression-analysis procedure (4) may be used. In this case, however, the procedure pre-sented can be first used to determine the specific region of Figure 1 (or perhaps a portion of it) that will most likely contain the best-fit model. The regression-analysis procedure can then be used to search for that model within the identified region rather than within all three regions of each re-gime. This will obviously reduce the computer time and the data-analysis effort required for the appli-cation of this computer-based procedure.

3. The procedure presented requires that the traffic-flow criteria be specified as single val-ues. With a modest effort, however, the procedure can be extended to deal with ranges of the traffic-flow criteria, in which case the procedure would provide a feasible region (rather than specific values) of model parameters. Previous work on single-regime models and the special case of two-regime models (6,7) consider, in addition to single values, ranges of the traffic-flow criteria. Simi-lar logic to the one used in these references may be used here to determine the feasible region of model parameters.

4. The general family of models located in re-gion 5 exhibits certain features that would be of practical use. For large values of m, the speed-flow relation of this region is almost flat for a portion of the low-density end of the curve and then decreases rapidly as density increases. These char-acteristics have been observed on highways with a rigidly enforced speed limit that is lower than the average highway speed (14). The flattening of the curve occurs since normal average speed cannot be attained due to the speed limit; at some point, as

density increases, the speed limit would no longer govern. Thus, models in region 5 are particularly useful for representing such situations.

5. Another practical phenomenon that can be accounted for by the non-congested-flow models presented is that traffic-flow relations on highways with the same average highway speed exhibit different shapes in the range between the free flow and capacity as a result of a different number of lanes ($\underline{14}$). For such highways, the traffic-flow criteria u_f, u_0, and k_0 are generally the same; only the shape of the traffic-flow relations in the intermediate operation differs. Specifically, the Highway Capacity Manual shows that the speed-flow curve of an eight-lane highway is higher than that of a six-lane highway, which is higher than that of a four-lane highway. The proposed procedure can capture this variability in the traffic-flow relations by modifying the auxiliary criteria while retaining the basic criteria fixed.

6. The procedure of selecting two-regime models presented in this paper and the one reported previously for estimating single-regime models ($\underline{6},\underline{7}$) significantly reduce the need for using computer facilities in estimating traffic-flow relations. As such, the procedure may be useful in such transportation areas as highway capacity and level of service, freeway operations, transportation planning, and environmental studies.

REFERENCES

1. L.C. Edie. Car-Following and Steady-State Theory for Noncongested Traffic. Operations Research, Vol. 9, No. 1, 1961, pp. 66-76.
2. D.C. Gazis, R. Herman, and R. Potts. Car-Following Theory of Steady-State Traffic Flow. Operations Research, Vol. 7, 1959, pp. 499-505.
3. D.C. Gazis, R. Herman, and R.W. Rothery. Nonlinear Follow-the-Leader Models of Traffic Flow. Operations Research, Vol. 9, No. 4, 1961, pp. 545-567.
4. A.D. May and H.E.M. Keller. Non-Integer Car-Following Models. HRB, Highway Research Record 199, 1967, pp. 19-32.
5. A. Ceder. Investigation of Two-Regime Traffic-Flow Models at the Micro- and Macroscopic Levels. Univ. of California, Berkeley, Ph.D. dissertation, 1975.
6. S.M. Easa and A.D. May. Generalized Procedure for Estimating Single- and Two-Regime Traffic-Flow Models. TRB, Transportation Research Record 772, 1980, pp. 24-37.
7. S.M. Easa. Generalized Procedure for Estimation of Single-Regime Traffic-Flow Models. Institute of Transportation Studies, Univ. of California, Berkeley, 1979.
8. B.D. Greenshields. A Study in Highway Capacity. HRB Proc., Vol. 14, 1934, pp. 448-477.
9. D.R. Drew. Deterministic Aspects of Freeway Operations and Control. HRB, Highway Research Record 99, 1965, pp. 48-58.
10. R.T. Underwood. Speed, Volume and Density Relationships: Quality and Theory of Traffic Flow. Yale Bureau of Highway Traffic, Yale Univ., New Haven, CT, 1961.
11. J.S. Drake, J.J. Schoefer, and A.D. May. Statistical Analysis of Speed-Density Hypothesis. Proc., Third International Symposium on the Theory of Traffic Flow. Elsevier, New York, 1967.
12. H. Greenburg. An Analysis of Traffic-Flow. Operations Research, Vol. 7, No. 4, 1959, pp. 79-85.
13. A. Ceder and A.D. May. Further Evaluation of Single- and Two-Regime Traffic-Flow Models. TRB, Transportation Research Record 567, 1976, pp. 1-15.
14. Highway Capacity Manual. HRB, Special Rept. 87, 1965.

In Situ Study Determining Lane-Maneuvering Distance for Three- and Four-Lane Freeways for Various Traffic-Volume Conditions

ROGER W. McNEES

The objective of this research was to determine on the basis of driver performance the distance it takes a driver to maneuver across several lanes in light, medium, and heavy traffic. The distance was expected to vary with a number of situational variables, several of which were investigated in this research. To obtain actual freeway distances associated with components of the model, an instrumented vehicle study was performed. Twenty drivers from Houston, Texas, drove sections of two freeways near downtown Houston. Interstate 45 was used for the lane-maneuvering study. All drivers were required to drive a three- and a four-lane section of the freeway and maneuver from the extreme left lane to the extreme right lane in light, medium, and heavy traffic. To determine an estimate of maneuvering distance, each driver was required (by instructions) to perform in succession three lane-change maneuvers on both the three- and the four-lane sections in each of the three traffic volumes. The distances were determined indirectly by recording the time required for a particular test and the speed of the test vehicle during each particular test. The major contribution of this research was a set of empirically determined maneuvering distances based on actual driving performance on a three- and a four-lane freeway under various traffic-volume conditions. Rather than a single value, the research findings offer several distances appropriate under various assumptions regarding the number of lanes, traffic volumes and speed, visibility, driver familiarity, and the percentage of drivers to be accommodated by the distance. The results indicate that traffic volumes and the number of lanes have a significant effect on maneuvering distance. Another finding was that when a driver is traveling at low speed in heavy traffic, the distance required to maneuver is significantly less than that when the speed of the vehicle is higher.

One portion of a much larger project funded by the Federal Highway Administration (FHWA) entitled "Human Factors Requirements for Real-Time Motorist Information Displays" is presented. The objective of this research was to determine and evaluate current standards on the placement of the advanced-warning (exit) signs based on actual performance data relating to sign reading, lane maneuvering, and deceleration distances.

Due to the length of the research effort, only the lane-maneuvering portion of this study will be presented here. Although the intent is not to slight the sign-reading or deceleration portions, the lane-maneuvering portion has a greater impact on sign placement and many more applications in other areas unrelated to sign placement than do the other portions of this study.

RESEARCH OBJECTIVES

The objective of this research was to develop initial placement locations of advanced-warning (exit) signs relating to diversions from a freeway or transition from one freeway to another freeway as incurred during route guidance. These placement locations were derived from actual driving performance and are appropriate under various assumptions regarding the number of lanes, traffic volumes, visibility, driver familiarity, and percentage of drivers to be accommodated by the distance.

The locations developed in this research considered factors such as the distance required to make several lane changes, the distance required for the driver to read the sign and start the initial lane change, and the distance required to decelerate the vehicle to the exit-ramp speed so it would not impede traffic along the freeway.

The specific objectives of this research included the following:

1. To obtain human performance criteria related to distances required for sign reading, lane maneuvering, and decision;
2. To develop a model that will allow computation of the longitudinal distance from the gore point to the sign location based on a set of measured distances; and
3. To determine recommended sign-placement distances for exit direction signs upstream of the gore point based on the number of lanes on the freeway, ambient light conditions, and traffic volumes on the freeway.

LITERATURE REVIEW

The basic principles and standards that govern the design and use of all traffic control devices are set forth in the Manual on Uniform Traffic Control Devices for Streets and Highways (MUTCD) (1). The objective of this research was to determine sign-placement criteria. Therefore, it is directly relevant to the MUTCD. Traffic control devices include all signs, signals, markings, and devices placed on or adjacent to a public roadway by an agency or official that has jurisdiction to regulate, warn, or guide traffic. The principles and standards set forth in the manual apply on a national level, and each state develops its own manual, which must be in compliance with the national standards.

There are five basic considerations employed to ensure compliance with these standards--design, placement, operation, maintenance, and uniformity. Current standards for sign placement are directly relevant to the present research. The MUTCD states (1, p. 1A-2):

> Placement of the device should assure that it is within the cone of vision of the viewer so that it will command attention; that it is positioned with respect to the point, object, or situation to which it applies to aid in conveying the proper meaning; and that its location, combined with suitable legibility, is such that a driver traveling at normal speed has adequate time to make proper response.

Basically, there are three types of signs used in any exiting maneuver. These signs are the advance-guide sign, interchange-sequence signs, and the exit-direction sign. This research is directly related to the placement location of the exit-direction sign. The advance-guide sign is the sign that warns the driver well in advance of the upcoming exit. The interchange-sequence signs are the series of signs that warn the driver of the remaining distance prior to the exit. The exit-direction sign is the last sign prior to the exit at which the motorist will be able to make the appropriate lane changes and decelerate to a safe exit speed. The

gore sign is not considered in this research due to its placement position.

The present research was conducted to determine the actual distance required by drivers to perform several lane changes as a part of their exit maneuver.

The distance at which an exit sign should be placed is dependent on the distance required by a driver to maneuver across a freeway and enter the exit lane. One approach to the evaluation of sign-placement distances reported in the literature (2,3) has been determining whether drivers have sufficient distance to maneuver into the exit lane for existing sign-placement locations. To evaluate whether a particular sign location provides sufficient distance, several techniques have been used. One of these techniques was to determine whether the location affects the driver's behavior while he or she is preparing to exit the freeway. The driver's behavior in this respect relates to activities such as steering reversals, brake application and reversals, lane positioning, and passing. The second technique for evaluating the location is by analytic methods of estimating distances associated with task times presumably required to perform part of the task. The third technique is to evaluate the sign location by determining, in vehicles, the actual distance required by drivers to change lanes and exit a freeway.

Although these three techniques appear to be similar, each method provides a different approach to determine placement distance. The first technique investigates driver-related factors and the effect sign placement has on them. These factors may be studied either in a laboratory (simulator) study or in a field (instrumented-vehicle) study. This method evaluates existing placement locations by determining what effect a particular location has on the driver's behavior. Those placement locations that have no effect on the driver are assumed to provide sufficient distance to exit the freeway. The second and third techniques determine placement locations based on distances required to exit the freeway. The only difference between these two techniques is the manner in which the distances are calculated. In the second technique, distance is determined by using an analytic approach in the form of a task analysis. This technique does not involve actual driving to determine the distance. The third technique uses an instrumented vehicle in actual freeway traffic to determine the distance required to exit the freeway. Several studies employing these techniques will be reviewed.

A study using the first technique to evaluate the distance from an exit that the advanced-information sign should be placed was conducted by Mace, Hostetter and Seguin (2). To determine the effect on the driver's behavior, three methods of analyses were used. The first method was to use a conceptual model to determine the nature of the interface linking the individual driver to other components of the traffic system. The components relating to information lead distance were first studied by using a driving simulator, which was the second method employed. Lead distances were approximated by varying the location of the sign on a filmstrip, which was run at a speed corresponding to the speed of the vehicle. In addition to providing information on the effects of lead distance, these simulation studies were used to provide inputs to determine the amount of task loading required of the driver during the in situ phase.

To provide a more direct test of the hypothesis concerning lead distance, a third method was applied in situ by using an instrumented vehicle under actual traffic conditions. Variables associated with the driver, the signs, and the environment in which the signs exist were investigated. Directional information signs were not used: Rather, commands when to exit and the direction of the exit were displayed on a screen mounted on the dashboard of the vehicle. The 18 subjects were given the command to exit coinciding to lead distances of 0.25-mile, 0.50-mile, and 1-mile intervals.

In the instrumented-vehicle study, the following variables were recorded: (a) speed, (b) steering reversals, (c) brake applications, (d) turn signal use, (e) lane position, (f) passing, and (g) significant unpredictable occurrences. The variables recorded associated with the traffic and the environment were (a) experimental vehicle in the right lane, in the center lane, and in the left lane; (b) passing vehicle, (c) vehicle passed, (d) display activation, and (e) unpredictable events. To record these variables, an Esterline-Augus chart recorder with two discrete and three analog channels was used.

The test site was a section of Interstate 495, the Washington Beltway, between Exits 27 (College Park, Maryland) and 37 (Indian Head Highway). This test was performed during heavy traffic volumes.

The general conclusions derived from the study indicated that the effects of the information lead distances on driving behavior were negligible, except for the 0.25-mile lead distance. The 0.25-mile lead distance frequently resulted in late entries into the extreme right lane. It was also concluded that if a number of performance variables are considered, the 0.25-mile information lead distance is less desirable than the 0.50-mile information lead distance.

The 0.25-mile lead distance would result in either mainstream turbulences or missing an exit under moderate to heavy traffic conditions. One driver error, reduction of speed in the mainstream, was prevalent for all lead distances. This study provided a method of evaluating specific distances at which exit signs may be located. In general, the study indicated that the 0.25-mile lead distance is inadequate in providing drivers sufficient distance to exit the freeway and that both the 0.50-mile and the 1-mile lead distances had negligible effects on driver behavior.

Levin (3) used the third technique, directed at determining the accuracy of sign placements based on lane changing in traffic. Levin evaluated signs placed at the gore 0.25 mile in advance, 0.50 mile in advance, and 0.75 mile in advance for service levels B and C. Level-of-service B is associated with a speed of between 55 and 60 mph and a freeway volume of 2800-3200 vehicles/h.

Levin used two methods to determine the lane-changing distance. His first method was a mathematical model describing the lane-changing process from one lane to the next adjacent lane to either the left or right. His mathematical model used a gap-acceptance and/or gap-rejection concept. This model allows computation of the required distance to complete the maneuver if the probabilities of occurrences associated with each of the three forms of the process, the traffic volume, and the speed of the vehicle are given.

Levin attempted to validate his model by using one subject driving a test vehicle on a freeway in Houston, Texas. He made 1000 lane changes from one lane to an adjacent lane in service-level-B and service-level-C traffic conditions. He was interested in determining the distance required for the lane change in each of those traffic conditions. Distance was determined by using magnets on the brake drum and the chassis of the test vehicle. By using a constant tire pressure and knowing the revolutions per minute at a set speed, distance can be determined.

The results obtained from Levin's model indicate that in most situations, drivers can perform their lane change for an exit with a 0.50-mile or greater advanced warning. This model, however, assumes only one lane change to perform the exiting maneuver. In situations where drivers must perform several lane changes, the effectiveness of the sign at the 0.50-mile location and the 0.75-mile location may be reduced to an unacceptable level. Levin determined that for four-lane freeways (two lanes in each direction) the effectiveness of the 0.25-mile sign was low and that the effectiveness of the 1- and 2-mile signs was high in level-B and level-C traffic volumes.

One question raised by Levin was concerned with the necessity of having both a 1-mile and a 2-mile sign. It may be possible to have a 1-mile sign alone or the sign could be moved closer to the gore point since the effectiveness of the 0.50- and the 0.75-mile signs was the same.

Other significant results from Levin's research were as follows:

1. The behavior of the lane-changing vehicle within the accepted gap in delayed lane changes may satisfactorily be described by a model based on the concepts of the car-following model.
2. The higher the traffic volume on the freeway, the higher the sensitivity of the model to changes in lane mean speeds.
3. The delay and distance involved in the lane-changing process depend on traffic volume and increase as volume increases.
4. As the driver's critical gap increases, the delay and distance experienced with the lane-changing process also increase.

The results obtained by Levin, even though they support Mace's results, were based on an incomplete study of the exiting process. Levin used only one subject, who drove an instrumented vehicle on a Houston freeway. This subject was required to make one lane change for each test run. The subject performed 500 test runs in traffic level B and 500 in traffic level C. These traffic levels are associated with speeds that range from 45 to 60 mph and a traffic volume between 1800 and 3200 vehicles/h. This is equivalent to making one lane change in light and in medium traffic conditions in this research effort. Two or three lane changes in heavy traffic might well require a significantly greater distance than that determined by Levin.

Eberhard (4), by using the second technique, performed an analytic study of sign-placement distance. In this research effort, Eberhard established information lead distances based on estimates, rather than an actual driving test, of the times required by drivers with a wide range of capabilities to perform tasks involved in negotiating an intersection. The hypothetical situation posed in this study involved a driver who was required to change one lane from right to left and then turn left at the next intersection.

During the task analysis, two elements emerged as the most relevant for the determination of the information lead distance. These elements were changing lanes to prepare for a maneuver and changing speed to perform the maneuver.

Eberhard estimated the lane-changing distance and the speed-change distance for the worst-case situation. This worst-case situation assumed a truck merging left; traffic density of 1000 vehicles/h; approach speed of 40 mph; an aged driver with long perception, decision, and maneuvering times; and conditions of poor visibility on a wet surface. His results indicate that under the worst possible case

the driver requires 2459 ft from the point at which the sign is noticed until changing lanes has been accomplished.

In general, the distance reported by Eberhard seems extremely large for a one-lane change. He estimated from the results obtained from his questionnaire that it required approximately 50 s from the time the signal was presented to the driver until lanes had been changed. Eberhard estimated that it took approximately 18 s for a driver to detect the driving situation present and that there would then be a wait for a gap before changing lanes. Another 25.5 s are spent in waiting for the gap to occur. These worst-case response times appear to be out of line in relation to what is required by the 85th-percentile driver. It is very difficult to estimate times accurately by analytic methods. More realistic estimates should be obtainable by timing the drivers' responses in traffic.

These three studies (2-4) provide a basis for evaluating sign-placement locations; however, the authors did not provide an in-depth study of the exiting process. In his research, Mace performed the exiting in heavy traffic during the day. The major criticism of the study was that he did not determine the actual distance traveled by the vehicle to determine whether there was a correlation between the actual distance traveled and the drivers' behavior. He did not study the light or medium traffic conditions.

Levin, on the other hand, studied the light and medium traffic during the day but did not study the heavy traffic condition. He used only one subject to obtain his data. This subject will eventually incur a learning effect and distances will become progressively shorter as the number of tests increases. The subject was also required to make only one lane change. On most freeways, at least two lane changes will be required. The results obtained by Levin apply to an isolated situation in which the driver needs to change one lane.

Eberhard's task analysis assumed several conditions that are not consistent with freeway operations. This study was not designed to study a freeway exiting maneuver; however, the tasks required to negotiate a turning maneuver at an intersection are similar to those tasks required to exit a freeway. Eberhard's study estimated distances based on a task analysis and not on actual field test data. The study was based on a hypothetical situation in which a driver must detect guidance information, change one lane from the right to the left, and wait for an acceptable gap. The distance estimated by Eberhard could be shorter than those determined during actual driving tests, because the vehicle is traveling at a lower initial speed, the driver makes only one lane change, and the response times may be much shorter.

METHOD

The approach used to conduct this study was an in situ instrumented-vehicle study. The maneuvering test was a 2x3x20 repeated-measures design. There were two levels of the number of lanes on the freeway (three and four lanes); three levels of traffic volumes (light, medium, and heavy); 20 drivers; and three replications per driver per condition. This design would provide 360 data points from which maneuvering distance was determined.

Maneuvering distance was the major emphasis of this research. Maneuvering distance was dependent on factors such as the type of vehicle, maximum traffic volumes, number of lanes, and differences in driving behavior between drivers. All these factors were studied in this project except the type of vehicle.

To determine distance, the test vehicle recorded the speed of the vehicle and the time required for each event. It should be noted that it was thought that total distance would deviate significantly from longitudinal distance (distance along the freeway from the gore point to the sign-placement location) due to the width of the freeway. However, in practice the width of the freeway is so small in relation to the total maneuvering distance that the width of the lanes does not make a significant difference. Therefore, total linear distance was defined as equal to longitudinal distance. The sign-placement distance computed from the human performance measurements was then compared with the sign-placement distances as set forth in the MUTCD (1).

Subjects

The subjects for this research were 20 drivers from the Houston area. The drivers were obtained from two sources. Eight are employed by the Texas Transportation Institute (TTI) and 12 are employed by the Texas State Department of Highways and Public Transportation (TSDHPT) at the Urban Office in Houston.

The selection criteria in obtaining drivers were sex, age, and a valid driver's license. The distribution of these drivers was determined by using national statistics of the driving population based on age and sex (U.S. Statistical Abstract, 1971). The driver's sample consisted of 13 males and 7 females. The males constituted 65 percent and the females constituted 35 percent. Ten drivers, or 50 percent, were in the 18-34 age group. Based on a statistical distribution provided the employers, drivers were selected to participate on a completely voluntary basis in this research project. It is for this reason that the male-female ratio does not coincide exactly with the national norm of 55 percent males and 45 percent females (U.S. Statistical Abstract, 1971).

All drivers held current driver's licenses. They had an average of 22 years of driving experience.

Instrumented Vehicle

The instrumented vehicle was a 1969 Plymouth Fury I four-door sedan equipped with a V-8 engine, automatic transmission, power steering, power brakes, and air conditioning. The front and rear seats were removed and bucket seats were installed to replace the front bench seat. The rear compartment was left open to accommodate the instrumentation.

The instrumentation package consisted of a power inverter, power supply, master control panel, Rustrak four-channel event recorder, and Rustrak analog recorder.

This research required driving a test section along one freeway located near downtown Houston, Texas. The selection of this location was dictated by the requirements of the research. The maneuvering test required the use of a three-lane and a four-lane section of freeway, which were monitored for traffic volumes and traffic speed. These volumes and speeds were independent variables in this study. The Gulf Freeway (I-45 North) has three- and four-lane sections that were monitored by the Texas Transportation Institute-Freeway Surveillance Center (TTI-FSC).

The Gulf Freeway test strip was a 6.5-mile section beginning at Park Place Boulevard and continuing inbound to the Pease Street exit. This section of freeway had a sufficiently diverse traffic volume that 725 vehicles/h or less (light), 726-1225 vehicles/h (medium), and 1226 vehicles/h and more (heavy) were observed several times during a 24-h

period in one direction on both the three- and the four-lane sections.

Scheduling of Test Drivers

For scheduling purposes, the drivers from TTI were assigned to two groups of four drivers each and those from SDHPT were placed into three groups of four drivers each. Groups of drivers were assigned to volume conditions according to a Latin square design such that the sequence of runs was not the same for all drivers. This method of scheduling drivers provided a method of measuring any learning effect that might be associated with the sequence of administration and simultaneously attempted to equalize sequence effects across conditions.

Data-Collection and Reduction Method

The approach used to acquire the data consisted of an in situ instrumented-vehicle study in which distance was determined indirectly by recording vehicle speed and time required to complete each run. Vehicle speed was not fixed; however, a maximum speed of 55 mph was established for each driver. This maximum was established to coincide with the legal speed limit and for safety purposes. In a few instances this imposed maximum speed was violated. It was also felt that by maintaining a fixed speed, an additional loading factor would be placed on the driver. All drivers were instructed to drive in their usual manner to reduce the negative psychological effects of being a test driver.

For the maneuvering study, the independent variables were traffic volume (light, medium, and heavy) and number of lanes (three or four) on the freeway. Vehicles per hour on a per-lane basis was used as the criterion for establishment of traffic volume. The experimental design required that each driver negotiate three lane-change maneuvers for each traffic volume on both the three- and the four-lane sections of I-45 inbound.

The pushbutton on the experimenter's master control panel was used to synchronize both recording devices for each run. The button made the recording mechanism on the recorder deflect equivalent to an instantaneous 4-mph increase in speed and held the recording mechanism for channel 1 of the four-channel event recorder in the on position for as long as the button was depressed.

For an accurate synchronization, the button was depressed by the experimenter for a full 5 s. Any synchronization period of less than 5 s would be difficult to locate. The synchronization procedure was required because the paper speed of the analog recorder was slightly faster than the four-channel event recorder due to the differential in the power supplies. The analog recorder paper drive required 12 V DC and the four-channel event recorder required 110 V AC.

The events recorded on channels 2, 3, and 4 of the event recorder were manually input by the experimenter through the master control panel. In the maneuvering test, channel 2 recorded the length of time required to make the first lane change, channel 3 recorded the length of time required to make the second lane change, and channel 4 recorded the length of time required to make the third lane change.

The Rustrak analog recorder continuously recorded the speed of the vehicle while the master control switch was in the on position. In situations of rapid acceleration or deceleration, the recorder indicated every 4-mph differential in speed. In all other situations, 1-mph differentials in speed could be determined.

The procedure used to reduce the data from these two tapes included aligning both tapes by using the synchronization marks for each run. In this way, time and speed could be read directly from both tapes. After the tapes had been aligned, each lane change was marked off to isolate total time required and speed of the vehicle during the lane change. In situations where speed varied during the lane change, each variation equivalent to 1 mph or greater was subdivided, and the associated time was marked off corresponding to that speed level.

To determine the distance associated with each lane change, the time associated with that lane change and the vehicle speed were recorded. The procedure allows distances to be calculated by using the following formula:

Total distance = vehicle speed (mph) x 1.467
   x time (s),

where 1.467 is a constant to convert miles per hour to feet per second.

In situations where speed fluctuated, the distances traveled during each speed fluctuation were determined and then added together to obtain the total distance for that lane change.

A lane change began when the drivers signaled their intention to change lanes by turning on the directional signal and continued until all four wheels of the test vehicle were in the adjacent lane. To obtain the distance that the driver stayed in a lane, it was necessary to isolate the time and the speed of the vehicle between each lane change. This required a visual inspection of two channels on the event recorder tape. The time interval and speed of the vehicle between the end of one lane change and beginning of the next lane change were noted. After the time and the speed had been isolated, the procedure to determine lane-changing distance was used to determine distance in the lane. The response time of the experimenter was 0.2 ± 0.1 s in flipping the switches. The distance associated with this time (0.2 s) was subtracted from the driver's distance to eliminate the experimenter's response time. After the lane-changing distances and distance in each lane had been computed, a summation of these resulted in the total maneuvering distance for each run.

After an extensive analysis of the heavy-traffic maneuvering distance, it was determined that two separate subclassifications of heavy traffic would be necessary. These two subclassifications are heavy high-speed (HS) and heavy low-speed (LS) traffic. This decision was reached based on the large differences in maneuvering distance associated with the differences in speed. Speeds were arbitrarily classified in terms of those above and those below 35 mph so that the sample sizes of HS and LS groups would be as nearly equal as possible for statistical purposes.

After the distances required for each maneuvering run had been computed for all drivers, these distances were analyzed with regard to traffic volume and the number of lanes on the freeway.

After the distances had been tallied for each classification category, they were ranked from the shortest distance to the longest distance within categories. A cumulative distribution for each type was then determined. From this cumulative distribution the 25th-, 50th-, 75th-, 85th-, 90th-, 95th-, and 100th-percentile levels were determined.

The first analysis of variance was a 2x3x7 two-way classification repeated-measure design. This analysis of variance tested the threeand four-lane freeway under light, medium, and heavy LS traffic volumes for seven drivers. These seven drivers were selected because they had test runs on both the three-lane and the four-lane freeways in heavy LS traffic. The second analysis of variance was a 2x3x15 two-way classification repeated-measure design. This analysis of variance tested the three- and the four-lane freeways under light, medium, and heavy HS traffic volumes for 15 drivers. These 15 drivers had test runs on both the three- and the four-lane freeways in heavy HS traffic. Four drivers had at least one test run in both heavy LS and heavy HS traffic conditions. These drivers' test runs were used in both analyses of variance.

RESULTS AND DISCUSSION

The major emphasis of this study was to determine from actual driving tests the distances required for a driver to maneuver from the extreme left lane to the extreme right lane in light, medium, and heavy traffic. Lane changing in heavy traffic was a worst-case maneuver for drivers attempting to exit a freeway. It was also expected that there would be a significant difference in maneuvering distance associated with the number of lanes a driver must maneuver across. Tests were performed on a three-lane and a four-lane section of freeway, but results could be generalized to freeways with more lanes or fewer lanes.

The manner in which the maneuvering test was conducted creates a situation in which each driver may incur a learning effect. The maneuvering test required that each driver perform a number of maneuvers for each condition being investigated. Those maneuvers performed first could require a greater distance than those performed last, due to the driver's unfamiliarity with the test vehicle. A Spearman's rank correlation test was subsequently performed on each driver's test runs to determine whether there was any learning effect due to this ordering of the drivers. This test indicated there was no significant correlation (ρ_{12} = 0.254, ρ_{13} = 0.159, and ρ_{23} = 0.167) due to the order in which the drivers performed the test.

After all the distances had been determined, they were classified according to the number of lanes and traffic volumes investigated. The mean maneuvering distances were determined and are reported in Table 1. Inspection of these results indicates that maneuvering distance is affected by traffic volumes on both the three-lane and the four-lane sections of freeway although there is clearly not a linear increase in the maneuvering distance with increasing traffic volume. As predicted, medium traffic volume required greater distance than did light traffic volume. However, on both the three-lane and the four-lane sections, the drivers required as much or more maneuvering distance in medium traffic as in heavy traffic.

To determine the cause of this reversal, an analysis of the time required to maneuver and the speed of the vehicle during the maneuver was performed. It was suspected that those traveling at low speeds (20-25 mph) in heavy traffic would take longer to perform the lane-change maneuver than would those who were able to travel at higher speeds (45 mph) in heavy traffic. In other words, traffic speed as well as volume might be critical in lane-changing time and associated distance.

In order to perform this analysis, an arbitrary speed of 35 mph was selected as the cutoff for selecting those maneuvers classified as LS maneuvers, so that sample sizes associated with both speeds would be as close as possible for statistical purposes. The results of this analysis are given in Table 2. The data indicate that those traveling at low speeds required more time to perform the maneu-

ver than did those traveling at high speeds when both are traveling in heavy traffic.

It may be noted that when the HS and LS distances are compared, the distances are substantially greater for the HS group and the travel times, as expected, were substantially less. The drivers at

Table 1. Mean maneuvering distance during light, medium, and heavy traffic.

Condition	N	Mean Distance (ft)	Standard Deviation (ft)
Three-lane			
Light	56	925	270
Medium	56	1009	308
Heavy	59	1002	527
Four-lane			
Light	48	1205	382
Medium	57	1521	391
Heavy	63	1375	377

Note: The difference in the number of observations for each traffic volume was due primarily to unexpected changes in traffic volumes during the maneuvering test.

Table 2. Means of maneuvering time, speed, and distance for three and four lanes in light, medium, and heavy LS traffic.

Condition	N	Mean Time (s)	Mean Speed (mph)	Mean Distance (ft)
Low Speed (<35 mph)				
Three-lane				
Light	21	13	47	892
Medium	21	14	47	984
Heavy	14	35	19	809
Four-lane				
Light	17	16	45	1043
Medium	19	25	43	1538
Heavy	13	34	26	1178
High Speed (>35 mph)				
Three-lane				
Light	41	13	50	941
Medium	41	15	48	1024
Heavy	32	17	47	1164
Four-lane				
Light	35	17	47	1136
Medium	44	24	44	1529
Heavy	37	24	43	1453

Table 3. Analysis of variance: maneuvering distances associated with LS and HS light, medium, and heavy traffic.

Source	df	Mean Square	F-Statistic
Low Speed			
Blocks	6	87 238	1.27
Treatment			
A	1	1 224 363	17.76[a]
B	2	318 017	4.61[b]
AB	2	143 417	2.08
Residual	30	68 940	
Total	41		
High Speed			
Blocks	14	92 804	1.32
Treatment			
A	1	2 364 228	33.53[c]
B	2	694 666	9.85[c]
AB	2	194 674	2.76
Residual	70	70 501	
Total	89		

Note: A = number of lanes; B = volume.
[a] $p < 0.01$. [b] $p < 0.05$. [c] $p < 0.001$.

low speeds were apparently able to maneuver across lanes in a shorter distance because they accepted smaller gaps, whereas the HS drivers took a greater distance to maneuver. This analysis points out that the time required to complete the LS maneuver was between 2.1 and 4.1 (for the three-lane and the four-lane conditions, respectively) times greater than the time required in the HS maneuver. Speed, on the other hand, was between 2.5 and 1.7 (for the three-lane and the four-lane conditions, respectively) times slower during the LS maneuver than during the HS maneuver. This difference in time and speed resulted in the shorter distance during the LS maneuver.

Two repeated-measure analyses of variance were performed to determine whether there was a significant difference between the maneuvering distances associated with the LS light, medium, and heavy traffic conditions and the HS light, medium, and heavy traffic conditions (Table 3). The results of these two analyses of variance substantiated the original hypothesis that traffic volume has a significant effect on lane-maneuvering distance. As was pointed out earlier, maneuvers during LS heavy traffic require less distance than maneuvers during HS heavy traffic. This research has indicated that in certain situations speed more than volume affects maneuvering distance.

It was also originally hypothesized that for a given traffic volume, maneuvering distance would increase as the number of lanes increased. This hypothesis was based on the assumption that the addition of another lane would necessarily increase the distance required to maneuver across all the lanes.

Analyses of variance performed to determine whether the number of lanes affected lane-maneuvering distance are also presented in Table 3. Both of these analyses of variance indicate that the number of lanes had a significant effect on lane-maneuvering distance. Although only the three- and the four-lane conditions were investigated, these results might be generalized to mean that as the number of lanes increases, the lane-maneuvering distance increases, and as the number of lanes decreases, the maneuvering distance decreases.

SUMMARY OF CONCLUSIONS

In this in situ instrumented-vehicle study, a set of required distances was determined to be used to estimate advanced sign-placement distances. The following is a brief summary of the findings and conclusions of this research:

1. Maneuvering distances were affected by traffic volumes on both the three- and the four-lane freeways. As volume increased, maneuvering distance also increased.

2. After the mean maneuvering distance in medium traffic and heavy traffic on both the three- and the four-lane freeways had been computed, the mean maneuvering distance in medium traffic was larger than that in heavy traffic. An analysis of the time required to complete the maneuver indicated that the time required in heavy traffic was significantly longer than that in medium traffic. An analysis of the speed of the test vehicle indicated that during maneuvers in which the test vehicle was very slow (8-17 mph), the time required to complete the maneuver was very long and the maneuvering distance was substantially shorter than the maneuvering distances in heavy traffic at higher speeds (40-47 mph). It was therefore concluded that recommended distances should be based on two types of maneuvering in heavy traffic. These two types were heavy LS and heavy HS maneuvers.

3. Traffic volumes had an effect on both the speed of the test vehicle and the time required to complete the maneuver. In general, the speed of the test vehicle decreased as the traffic volume increased and the time required to complete the maneuver increased as the volume increased.

4. The number of lanes affected maneuvering distances. As the number of lanes increased, the maneuvering distance also increased. Therefore the sign-placement distance should take into account the maximum number of lane changes for the particular freeway.

5. The number of lanes affected both the time required to complete the maneuver and the speed of the test vehicle for maneuvers performed in light, medium, and heavy HS conditions. The number of lanes did not affect either speed or time in light, medium, and heavy LS conditions. This was due to large differences in both time and speed between the heavy LS and heavy HS conditions.

6. The lane-maneuvering distances (mean distance) from this research are as follows: (a) 917 ft on a three-lane freeway during light traffic, (b) 1004 ft on a three-lane freeway during medium traffic, (c) 809 ft on a three-lane freeway in heavy LS traffic, (d) 1164 ft on a three-lane freeway in heavy HS traffic, (e) 1090 ft on a four-lane freeway during light traffic, (f) 1534 ft on a four-lane freeway during medium traffic, (g) 1178 ft on a four-lane freeway in heavy LS traffic, and (h) 1453 ft on a four-lane freeway in heavy HS traffic.

REFERENCES

1. Manual on Uniform Traffic Control Devices for Streets and Highways. U.S. Government Printing Office, 1978.
2. D.J. Mace, R.S. Hostetter, and E.L. Seguin. Information Requirements for Exiting at Interchanges. HRB-Singer, Inc., State College, University Park, PA, Publ. 89211-F, Sept. 1967.
3. M. Levin. Some Investigations of the Freeway Lane Changing Process. Texas A&M Univ., College Station, Ph.D dissertation, 1970.
4. J.W. Eberhard. Driver Information Requirements and Acceptance Criteria. HRB, Highway Research Record 265, 1969, pp. 19-30.

Exponential Filtering of Traffic Data

PAUL ROSS

Real-time traffic-control systems commonly filter their input data with exponential filters. The filtering constant β can be objectively determined by recasting the filter as a predictor and choosing the β that would have optimized the predictions for some observed time series of data—mostly taken to be 24 h of input here. Results, by using 70 consecutive days of traffic volume counts from Toronto plus some Los Angeles and District of Columbia data, indicate that the optimal β can be approximated by $\beta = (1 - a) \exp^{(-bT)} + a$, where T is the data-aggregation period and a and b are constants that depend on traffic-peaking characteristics. The filter constant β can be allowed to change during the day; the algorithms of this kind that were investigated all entail only a modest increase in mean-square prediction error. Algorithms that reoptimize β over the most recent N data points do not appear practical, but a simple recursive algorithm that gives good accuracy is presented.

Measurement noise in the usual sense constitutes no more than a few percent of traffic volume counts. However, traffic volumes from successive short periods often differ substantially from each other so that it is convenient to think of the traffic volume itself as being made up of some meaningful signal plus meaningless noise. Thus, detector-measured traffic data (usually volumes or occupancies) constitute a noisy time series. Time-series analysis is a well-established subject of study (1); the methods of Box-Jenkins analysis (2,3) and Kalman filtering (4-6) are quite sophisticated and powerful techniques.

With respect to traffic measurements, virtually all the time-series analysis effort has been directed toward short-term traffic prediction. Polhemus (7,8) codified the prediction problem at a high level (unfortunately without any immediately applicable results). The problem has been presented in a Box-Jenkins framework (9-11) and some specific prediction problems have been formulated in terms of point processes (12-15). Breiman (16) proposed an algorithm that has some intuitive appeal and showed how to choose its parameters to minimize the maximum (not mean) square prediction error.

Real, on-line computer-controlled traffic systems have to make predictions at many locations, and their computing power is usually taxed. Consequently, the prediction and filtering algorithms that have been used in practice have been economical of storage and simple in execution; they tend to be based on intuitive models of traffic.

The prediction algorithm used in the second-generation urban traffic control system (UTCS) program has been described by Ganslaw (17) and by Sperry (18); the predictor in the third-generation UTCS program has been described by McShane, Lieberman, and Goldblatt (19) and by Lieberman and others (20). More effort has gone into devising such algorithms than testing them; the only convincing evaluations of practical prediction algorithms were reported by Kreer (21,22). None of the algorithms tested was convincingly better than simple time-of-day historical averages.

The separate but related filtering problem has been largely ignored. Only Houpt and others (23) describe the filtering of traffic data in more than a passing manner; the technique used is the powerful but cumbersome "extended Kalman filter."

In real traffic problems it is almost universal to use exponential filtering. Briefly, the use of

exponential filtering assumes that the data are generated by

$$a = a^* + noise \qquad (1)$$

where a is the observed traffic count and a* is an underlying true traffic volume, which varies only slowly from measurement to measurement. The noise is assumed to have a mean value of zero. The problem is to make a good estimate â of a*.

If it is true that a* varies slowly with time, then one intuitively wants to discount the older observations, and the simplest way to do so is by exponential weighting:

$$\hat{a}(t) = (1 - \beta) \sum_{i=1}^{t} \beta^{t-i} a(i) \qquad (2)$$

$$= (1 - \beta) a(t) + \beta \hat{a}(t - 1) \qquad (3)$$

where β is the filter constant.

This use of exponential filtering has many advantages. First, the calculation is recursive--as can be seen from Equation 3. The recursive character means that data-storage requirements are minimal; only the previously calculated estimate â(t - 1) and the value of the filter constant itself need be stored. Second, the calculation could hardly be simpler or quicker. And finally, this simple filter is optimal for the common MA(1) class of Box-Jenkins processes. For the purpose of this paper, the use of exponential filtering as defined above will not be questioned.

FILTERING CONSTANT β INDEPENDENT OF TIME OF DAY

Figure 1 shows a typical day of raw traffic counts and the resulting values after the data have been filtered with β = 0.2, 0.4, 0.6, and 0.8. [The principal data used in this paper were obtained in Toronto in the fall of 1973. That data set, the most extensive that I have located, consists of 5-min counts for each of eight detectors 24 h/day

for 77 consecutive days. For a description of this data set and some of its properties, see the report by Pignataro and others (24) and by McShane and Crowley (25). Figure 1--and all other cases where a single day is used in this paper--is from detector 1 on September 24, 1973, the first Tuesday of the study.]

Which value of the smoothing constant should be used appears to be a matter of personal taste--how well filtered one prefers the data. One finds such statements as "the constant (β) is typically 0.7 to insure that about five previous speed measurements make a significant contribution to the correct average" (26) and "a value of (β) in the range of 0.4-0.5 is generally quite satisfactory in reducing the (difference between the data and their filtered values) to low-variance white noise" (20).

However, a close look at the fundamentals of the problem indicates that one can establish an objective criterion for the value of β. The filtered data point â(t) from Equation 2 or 3 is the best available estimate of the signal a* and hence is the best prediction of the next datum a(t + 1). After data have been collected for an entire day, one can compare the predicted value â(t) with the actual value a(t + 1) and determine which value of β would have given the best estimate; such is the approach that will be used in this paper. The criterion of best will mean minimum mean-square error in the predictions made. Figure 2 shows how the mean-square prediction error depends on the value of β for the traffic data used in Figure 1. The best predictions occur when β = 0.424.

Location of this minimum point is most conveniently done by differentiating the mean-square error curve and locating the β that makes this derivative zero. The derivative of the total square error with respect to β can be written as follows:

$$dE(t)/d\beta = [dE(t-1)/d\beta] + [a(t) - \hat{a}(t-1)] \ (d/d\beta) \hat{a}(t-1) \qquad (4)$$

where

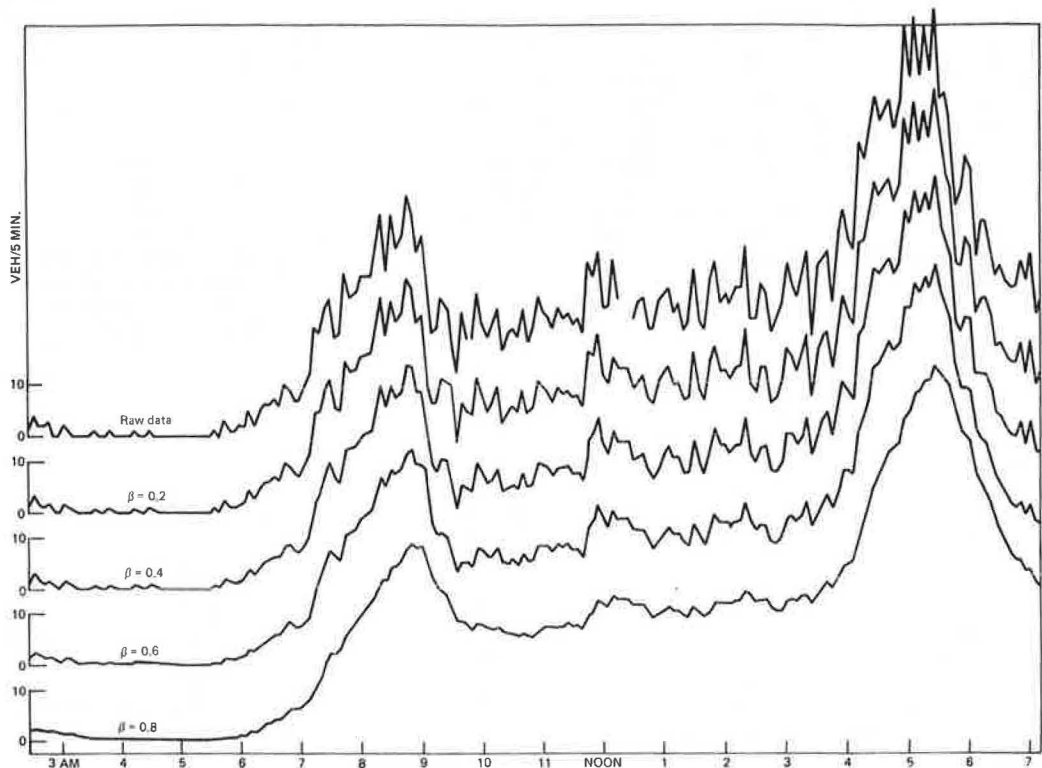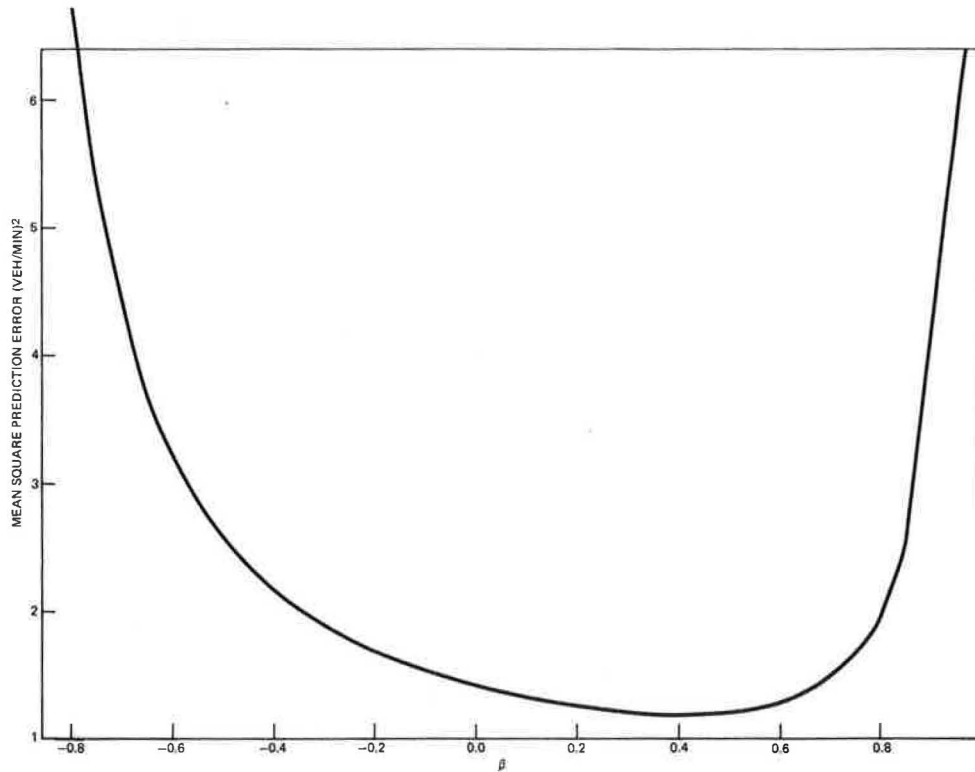Figure 1. Five-minute traffic-volume counts filtered with various values of filtering constant β.

Figure 2. Mean-square prediction error as function of filtering constant β.

$$\hat{a}(t - 1) = (1 - \beta) a(t - 1) + \beta \hat{a}(t - 2)$$

$$(d/d\beta) \hat{a}(t - 1) = -a(t-1) + \hat{a}(t - 2) + \beta (d/d\beta) \hat{a}(t - 2)$$

The optimal β is a root of $dE(t)/d\beta = 0$ from Equation 4. In practice there appears to be only one real root. [Since $dE(t)/d\beta$ is of odd degree in β, it must have at least one real root.] The complex roots complicate many root-finding methods; a bisection algorithm was used in this study. Equation 4 is recursive and therefore easy to program for electronic computers.

Figure 3 shows the optimal smoothing constants β for detector 1 during the first seven days of Toronto data. β is shown as a function of the period of aggregation. Intuitively, it seems that β ought to be large (nearly unity) for very small aggregation periods (because short periods of data are very noisy) and nearly zero for long periods; this seems approximately true for the Saturday and Sunday data. However, the weekday data definitely imply that the optimal β is negative for long aggregation periods. This reflects a violation of the assumption that the underlying signal varies only slowly from period to period; on weekdays the traffic volumes vary substantially (compared with the noise) when the aggregation period is 30 min or more.

The appearance of negative values for β does not prohibit the use of exponential smoothing even though the original assumption about the character of the process is violated for large aggregation periods; the actual prediction error (Figure 4) remains small even for large aggregation periods. The observed increase in prediction error with increased aggregation period is due to two separate effects--the model failure discussed above and the fact that there are fewer data periods when the periods are long. The increase in prediction error with the Saturday and Sunday data is almost wholly due to this latter effect.

A summary of all the available data is shown in Figure 5. Optimal smoothing constants were calculated for every detector for every day (except that the last seven days, which had a great deal of data missing, were deleted from the data set). The curve shown for each detector is the average of 50 daily curves (weekdays) or 20 daily curves (weekends). Also shown are two curves based on District of Columbia data that recorded each vehicle detection for 90 min (with two short breaks to remount the tapes). These arterial data appear to be of the same character as the Toronto weekday data.

The third set of curves in Figure 5 represents 10 detectors on a Los Angeles freeway. These 10 detectors are randomly selected from the four lanes in the northbound direction on the San Diego Freeway; the detector array spans 1 mile. These data contain individual vehicle actuations uninterrupted for 150 min.

All three sets of curves fit a family that can be written as follows:

$$\beta = (1 - a) \exp(-bT) + a \qquad (5)$$

where a and b are constants that depend on the peaking characteristics of the traffic and T is the data-aggregation period in minutes. The values of a and b shown below give a good visual fit to the curves in Figure 5:

Curve	a	b (min⁻¹)
Freeway	0.00	0.86
Arterial (Sat.-Sun.)	-0.05	0.13
Arterial (Mon.-Fri.)	-0.042	0.10

FILTERING CONSTANT β ALLOWED TO VARY DURING DAY

The foregoing discussion has assumed that a single, fixed value of the filtering constant must be used throughout the day. Intuition says that there may be substantial benefits in allowing β to vary. A

Figure 3. Optimal filtering constant β as function of aggregation period on seven consecutive days.
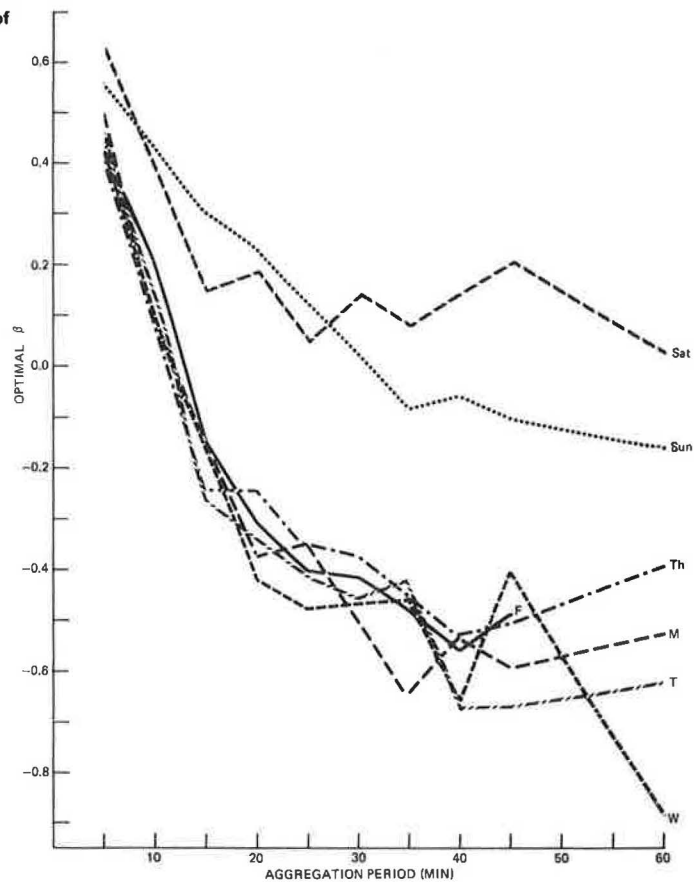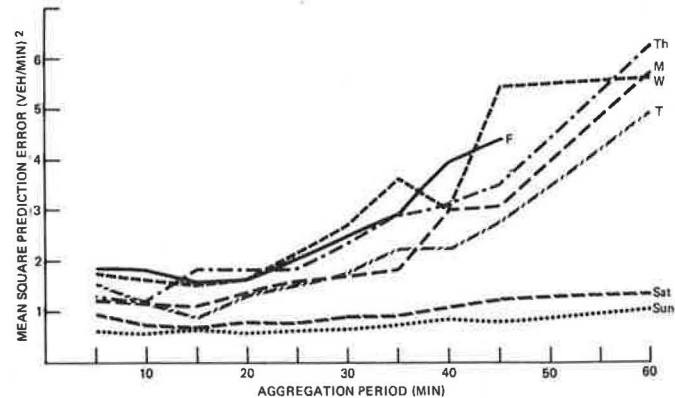

Figure 4. Mean-square prediction error as function of aggregation period by using same data as those in Figure 3.

properly chosen algorithm for the variation will, in effect, allow β to "tune" itself, obviating the need to guess in advance what the correct value of the smoothing constant will be. A variable β can conceivably even produce smaller prediction errors than using a single fixed value all day.

The obvious way to calculate such a variable smoothing constant is to recalculate the optimal β every period or two by using the most recent several data points. Figure 6 shows how some smoothing constants vary during the day. (Shortly after midnight, when fewer than the desired number of data points were available, the optimal calculation was limited to the available data only.)

Figure 7 shows how the mean-square prediction error depends on the number of data points used. As one would expect, the accuracy appears to be ap-

proaching a limit as the reoptimization includes more data. The limiting value of the mean-square error is about 2 percent worse than using the 24-h optimal smoothing constant throughout the day. Since the variable smoothing constant is based on the data available up to the time of the calculation only, it is not surprising that its limit is worse than that of the fixed 24-h optimal constant, which, in effect, is chosen with foreknowledge of the data yet to come.

One can evaluate the practicality of direct reoptimization by using the latest N points. Examination of Figure 6 indicates that the 5-min smoothing constant is likely to vary appreciably in 10 min, so that it will be necessary to recalculate β at least that often. How many of the latest data points must be used to produce acceptable accuracy

Figure 5. Average optimal filtering constants β for 8 arterial detectors in Toronto, 2 arterial detectors in Washington, and 10 freeway detectors in Los Angeles.
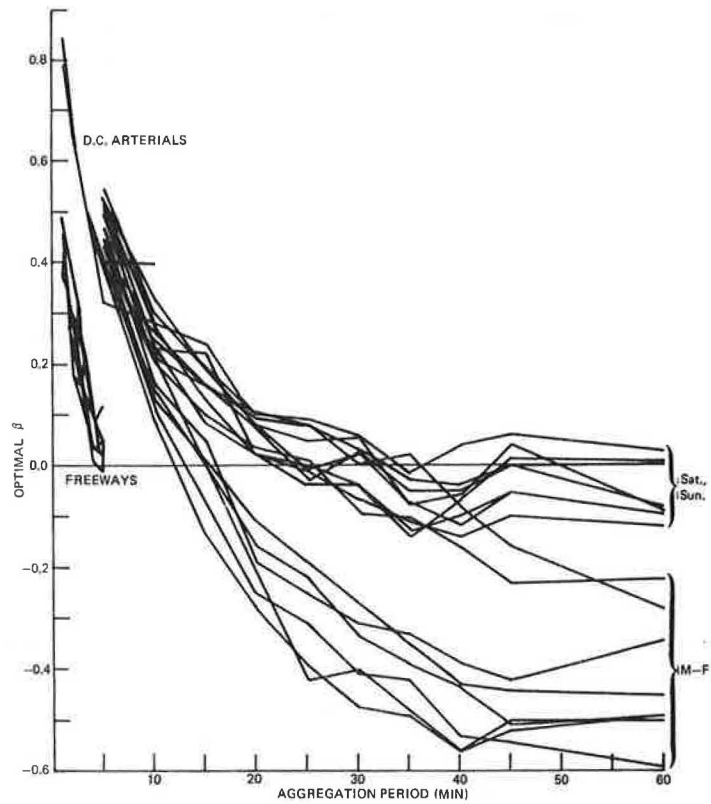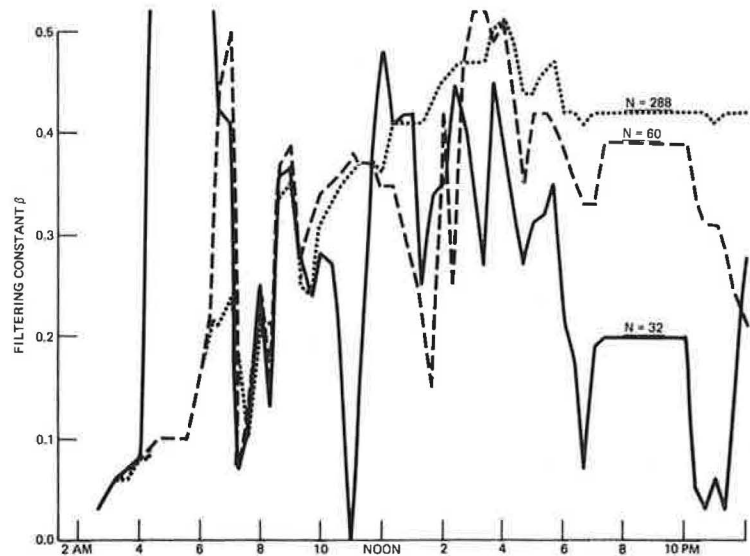

Figure 6. Variation in optimal β based on most recent 32, 60, and 288 observations as function of time of day.

is a somewhat subjective question; based on Figure 7, I would use at least N = 80 points. The use of 80 data points implies finding a root of a 159th-degree polynomial. If this must be done every 10 min for every detector in the system, it is clear that no exact reoptimization based on the latest N data points can be workable.

After abandoning exact reoptimization of β, one wonders if an approximate formulation can be found. Recursive algorithms are particularly attractive since they require only modest storage of data and preceding results and are generally simple in execution. Such an algorithm has been devised as follows.

At time t, the most recent prediction error is

$$a(t) - \hat{a}(t-1) = \beta_t [a(t-1) - \hat{a}(t-2)] - a(t-1) + a(t) \qquad (6)$$

If the value of β_t had been as follows,

$$\beta_t{}^* = [a(t-1) - a(t)]/[a(t-1) - \hat{a}(t-2)] \qquad (7)$$

the prediction would have been perfect, i.e., zero prediction error.

It is desired that the new filtering constant β_{t+1} be a linear combination of $\beta_t{}^*$ and β_t. There seems to be no logically imperative choice for the weights associated with $\beta_t{}^*$ and β_t; I have used, somewhat arbitrarily, the square prediction error in the (t - 1)st term and

Figure 7. Mean-square prediction error by using optimal filtering constant based on N floating data points.
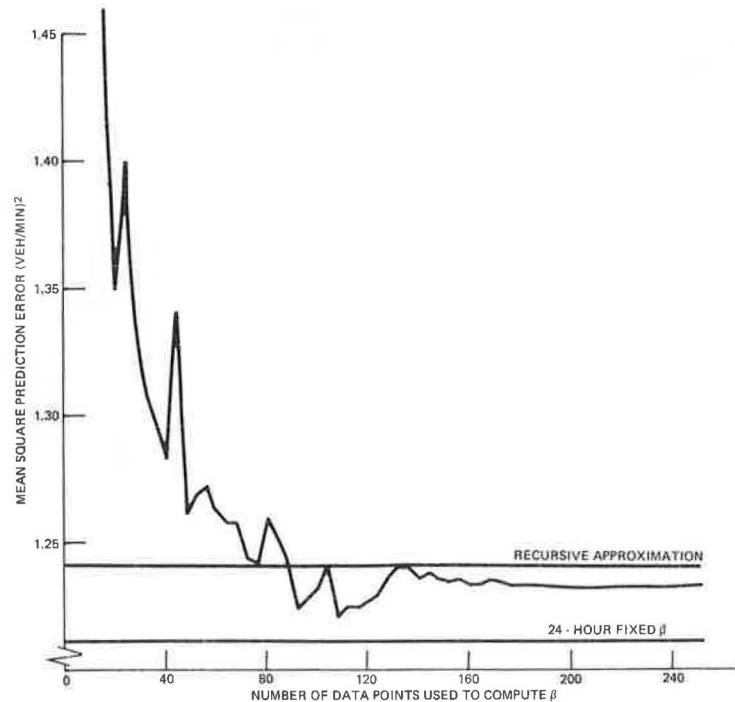
the total preceding square prediction error as the weights for β_t^* and β_t, respectively. This combination seems to work well; the mean-square prediction error is less than 0.5 percent worse than using all available data (i.e., up to 288 values) in a rigorous optimization when tested with the 5-min aggregation test case.

This algorithm reduces to three fairly elegant recursion equations:

$$Z_t = a(t-1) - \tilde{a}(t-2)$$

$$E_t = E_{t-1} + Z_t^2$$

$$\beta_{t+1} = \left\{ E_{t-1}\beta_t + Z_t [a(t-1) - a(t)] \right\}/E_t \qquad (8)$$

In the above equations, Z_t can be interpreted as the $(t-1)$st prediction error and E_t is the cumulative square prediction error through time t.

SUMMARY

Exponential filtering of traffic volume counts is simple, quick, and reasonably accurate. A constant value for the filter constant may be used with good accuracy; if so, the appropriate value of the filter constant depends on the aggregation period and the daily traffic variation as shown in Equation 5 and the tabulation below it.

The necessity of predicting the best value for the smoothing constant and updating it as conditions change can be obviated by using on-line updating of the filtering constant. Updating by exact reoptimization--even over short histories--is not feasible, but Equation 8 presents a simple recursive method for calculating the filter constant that gives good results.

REFERENCES

1. M.G. Kendall. Time-Series. Charles Griffin and Co., London, 1973.
2. C.R. Nelson. Applied Time Series Analysis for Managerial Forecasting. Holden-Day, San Francisco, 1973.
3. G.E.P. Box and G.M. Jenkins. Time Series Analy-
sis, Forecasting and Control. Holden-Day, San Francisco, 1977.
4. R.E. Kalman. A New Approach to Linear Filtering and Prediction Problems. Transactions of ASME: Journal of Basic Engineering, March 1960, pp. 35-45.
5. A.E. Bryson, Jr., and Y.-C. Ho. Applied Optimal Control. Blaisdell Publishing Co., Waltham, MA, 1969.
6. A.P. Sage and J.L. Melsa. Estimation Theory with Applications to Communications and Control. McGraw Hill, New York, 1971.
7. N.W. Polhemus. Statistical Analysis of Traffic Systems: A Time Series Approach. Department of Civil Engineering, Princeton Univ., Princeton, NJ, Ph.D. dissertation, 1976.
8. N.W. Polhemus. Time Series Analysis of Local Fluctuations in Traffic Parameters. Transportation Research, Vol. 10, 1976, pp. 311-317.
9. Data Sciences, Inc. Mathematical Models for Traffic Volume Prediction. TRW Systems, Houston, TX, Nov. 17, 1972.
10. Integrated Systems, Inc. A Traffic Volume Predictor for Networks. FHWA, April 20, 1972.
11. M. Levin and Y.-D. Tsao. On Forecasting Freeway Occupancies and Volumes. TRB, Transportation Research Record 773, 1980, pp. 47-49.
12. R.L. Lopez and P.K. Houpt. Estimation of Traffic Variables Using Point Processing Techniques. Research and Special Programs Administration, Office of Systems Engineering, U.S. Department of Transportation, Rept. DOT-TSC-RSPA-78-9, 1978.
13. J.S. Baras, W.S. Levine, A.J. Dorsey, and T.L. Lin. Advanced Filtering and Prediction Software for Urban Traffic Control Systems. Transportation Studies Center, Univ. of Maryland, College Park, 1977.
14. J.S. Baras and W.S. Levine. Estimation of Traffic Flow Parameters in Urban Traffic Networks. Proc., 1977 IEEE Annual Conference on Decision and Control, San Francisco, CA, pp. 428-433.

15. J.S. Baras, W.S. Levine, A.J. Dorsey, and T.L. Lin. Advanced Filtering and Prediction Software for Urban Traffic Control Systems. U.S. Department of Transportation, Feb. 1981.
16. L. Breiman. Predicting Input Flows. Systems Development Corp., Santa Monica, CA, Tech. Memorandum TM-4638/016/01, Sept. 1971.
17. M.J. Ganslaw. A 5-Minute Volume and Speed Predictor for the UTCS/BPS Second Generation Software. TRW Systems, Houston, TX, Sept. 1975.
18. Sperry Systems Management. Development of Traffic Logic for Optimizing Traffic Flow in an Intercity Corridor. FHWA, April 1, 1976.
19. W.R. McShane, E.B. Lieberman, and R. Goldblatt. Developing a Predictor for Highly Responsive System-Based Traffic Signal Control. TRB, Transportation Research Record 596, 1976, pp. 1-2.
20. E. Lieberman, W.R. McShane, R. Goldblatt, and D. Wicks. Variable Cycle Signal Timing Program: Volume 4--Prediction Algorithms, Software and Hardware Requirements, and Logical Flow Diagrams. KLD Associates, Inc., Huntington, NY, June 1974.
21. J.B. Kreer. Factors Affecting the Relative Performance of Traffic Responsive and Time-of-Day Traffic Signal Control. Transportation Research, Vol. 10, 1976, pp. 75-81.

22. J.B. Kreer. A Comparison of Predictor Algorithms for Computerized Traffic Control Systems. Traffic Engineering, Vol. 45, No. 4, April 1975, pp. 51-56.
23. P.K. Houpt, M. Athans, D.G. Orlahac, and W.J. Mitchell. Traffic Surveillance Data Processing in Urban Freeway Corridors Using Kalman Filter Techniques. Research and Special Projects Administration, U.S. Department of Transportation, Nov. 1978.
24. L.J. Pignataro, W.R. McShane, K.W. Crowley, B. Lee, and T.W. Casey. Traffic Control in Oversaturated Street Networks. NCHRP, Rept. 194, 1978.
25. W.R. McShane and K.W. Crowley. Regularity of Some Detector-Observed Arterial Traffic Volume Characteristics. TRB, Transportation Research Record 596, 1976, pp. 33-37.
26. D. Ghosh and C.H. Knapp. Estimation of Traffic Variables Using a Linear Model of Traffic Flow. Transportation Research, Vol. 12, No. 6, Dec. 1978, pp. 395-402.

Operational Effects of Two-Way Left-Turn Lanes on Two-Way Two-Lane Streets

PATRICK T. McCOY, JOHN L. BALLARD, AND YAHYA H. WIJAYA

The two-way left-turn lane (TWLTL) has been installed on two-way streets under a wide variety of conditions as a solution to the safety and operational problems caused by the conflict between midblock left turns and through traffic. Although the safety effectiveness of the TWLTL has been the subject of many studies, very few studies have been made of its operational effectiveness. Consequently, its effects on the efficiency of traffic flow have not been precisely measured. The objective of this study was to quantify the effects of a TWLTL on the efficiency of traffic flow on a two-way two-lane street. By using computer simulation models specifically developed and validated for the purpose of this study, traffic operations were simulated over a range of traffic volumes and driveway densities. From the outputs of these simulation runs, the reductions in stops and delays that result from a TWLTL were computed. Isograms of the stop and delay reductions were prepared to facilitate the use of the results of this study to evaluate the potential cost effectiveness of TWLTL installations.

The two-way left-turn lane (TWLTL) is recognized as a possible solution to the safety and operational problems on two-way streets that are caused by the conflict between midblock left turns and through traffic. The primary function of the TWLTL is to eliminate this conflict by removing the deceleration and storage of vehicles making these turns from the through lanes, thereby enabling through traffic to move past them without delay. However, the extent to which a TWLTL can improve the efficiency of traffic operations depends on the traffic volumes and density of driveways involved. Although the principle of the complex relationship between these factors and the operational effectiveness of the TWLTL is intuitively apparent, it has yet to be quantita-

tively expressed. Consequently, traffic engineers have not been able to precisely predict the amount of improvement in the efficiency of traffic operations that would result from the installation of a TWLTL.

An extensive review of the literature and nationwide survey of experience with the TWLTL were conducted by Nemeth (1) in developing guidelines for its application. This effort revealed that the TWLTL has been installed under a wide variety of conditions. In most cases, it was considered to have noticeably improved the quality of traffic flow. Numerous before-and-after accident evaluations were found that provided measures of the safety effectiveness of the TWLTL. But similar studies of its effect on the efficiency of the traffic were rare, and measures of the operational effectiveness of the TWLTL were not found.

Likewise, in developing guidelines for the control of access on arterial streets, Glennon and others (2) found that empirical data pertinent to the determination of the operational effectiveness of the TWLTL were lacking. This deficiency precluded the precise estimate of the delay-reduction potential of the TWLTL. And this in turn limited the specificity with which the conditions that warrant installation of a TWLTL could be defined.

In response to the need of traffic engineers to be able to more precisely predict the operational effectiveness of a TWLTL and more clearly define those circumstances that justify its installation, a

study of the operational effects of a TWLTL was conducted at the University of Nebraska-Lincoln. The objective of this study was to quantify the effects of a TWLTL on the efficiency of traffic flow on a two-way two-lane street. Two computer simulation models were developed and validated for this study. One of the models was used to simulate traffic operations on a two-way two-lane street with a TWLTL, and the other model was used to simulate traffic operations on a two-way two-lane street without a TWLTL. Traffic operations were simulated with both models over a range of traffic volumes and driveway densities. The outputs of these simulation runs were then compared to determine the reductions in stops and delays that resulted from the TWLTL.

This paper presents the procedure and findings of this study. Also presented is a brief description of the simulation models and their validation, and to facilitate the implementation of the results of the study, isograms of the stop and delay reductions provided by a TWLTL over the range of traffic volumes and driveway densities are included.

SIMULATION MODELS

The two computer simulation models developed in this study were written in the General Purpose Simulation System (GPSS) language (3). These models are basically the same, except that one is for a two-way two-lane street with a TWLTL and the other is for a two-way two-lane street without a TWLTL. A brief description of the input, logic, output, and validation of these models follows.

Input

The input to the models consists of two types of information--traffic characteristics and street geometry. The traffic characteristics input to the models are the volume and average speed of traffic in each direction and the percentage of the traffic volume turning left into each driveway on the street. Also, the arrival pattern of the traffic entering at each end of the street is specified. The models can generate random and nonrandom arrival patterns, so that the effects of traffic signals can be simulated by the models.

Because of the nature of the GPSS language, the street geometry is defined in terms of sections. Each lane on the street is divided lengthwise into 20-ft sections, and driveway locations on the street are defined by the numbers of the sections in which they are located. Also, input for each driveway in the model with the TWLTL is the section number of the farthest point upstream at which a vehicle turning left into the driveway can enter the TWLTL.
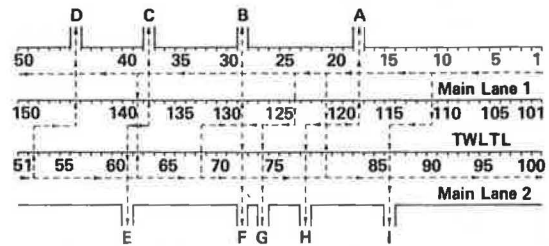
The geometry of a 1000-ft street segment with a TWLTL is illustrated in Figure 1. Each lane is divided into fifty 20-ft sections, which are numbered as follows:

    Lane 1: sections 1-50,
    Lane 2: sections 51-100, and
    TWLTL: sections 101-150.

The section numbers of the driveway locations and their corresponding TWLTL entry points that would be input to the model with the TWLTL are shown below:

Driveway	Lane No. Entered From	Driveway Location Section	TWLTL Entry Point Section
A	2	18	121
B	2	29	133
C	2	38	139
D	2	45	149

Figure 1. Geometry of 1000-ft street segment with TWLTL.

Driveway	Lane No. Entered From	Driveway Location Section	TWLTL Entry Point Section
E	1	61	139
F	1	72	124
G	1	74	124
H	1	78	121
I	1	86	111

In the case of a 1000-ft street segment without a TWLTL, sections 101-150 would not exist. Therefore, only the driveway location section numbers would be input to the model without a TWLTL.

Logic

In both models, traffic enters the street segment at either end in accordance with the traffic volumes and arrival pattern specified in the input. The course of any vehicle entering the segment will be one of two types: (a) traverse the entire length of the segment without turning left and exit at the other end or (b) traverse a portion of the segment and exit by turning left at one of the driveways. On entering the segment, the course taken by each vehicle is determined probabilistically in accordance with the left-turn percentages specified in the input.

Vehicles move through each section in the main lanes at the average speeds specified in the input and maintain at least 2-s headways. Thus, if a vehicle is stopped, the time required for it to traverse the next section is at least 2 s plus the travel time at the average traffic speed. Vehicles traversing the entire street segment remain in the main lanes and do not pass other vehicles in their lanes.

In the model without the TWLTL, turning vehicles also remain in the main lanes until they reach the driveways into which they turn. However, in the model with the TWLTL, a turning vehicle remains in the main lane until it reaches the entry point to the TWLTL, which is designated in the model input for the driveway into which it turns. The vehicle then moves from the main lane to the TWLTL. Once in the TWLTL, a vehicle moves ahead at a speed of 10 mph until it reaches the driveway into which it turns or until it is stopped by vehicles already in the TWLTL waiting to turn left.

If a turning vehicle reaches its entry point to the TWLTL and finds that the section is occupied by a left-turning vehicle from the other direction, it remains in the main lane and moves ahead until it finds an unoccupied section in the TWLTL upstream from the driveway into which it turns or until it reaches the driveway. If it reaches the driveway before it finds an empty section in the TWLTL, it turns left into the driveway from the main lane.

In both models, a turning vehicle must have an acceptable gap in the opposing traffic stream before

it can turn left. The required length of the gap is determined probabilistically in accordance with the left-turn gap-acceptance function derived by Gerlough and Wagner (4). If the required gap is available, the vehicle turns left. Otherwise, it waits until one is available. However, in the model without the TWLTL, if this wait exceeds 30 s, the attempt to turn is aborted and the vehicle traverses the entire length of the segment as if it were a through vehicle.

Output

The output from the models includes the following data:

1. Number of vehicles entering and exiting the segment,
2. Number of left turns attempted and completed,
3. Number of stops,
4. Travel time in the segment, and
5. Stopped-time delay.

The travel time, stops, and delay totals are output separately for through vehicles, turning vehicles, and all vehicles.

Validation

In order to validate the models, traffic flow on two two-way two-lane street segments (one with and one without a TWLTL) in Lincoln, Nebraska, was filmed. Those films were analyzed to determine the volumes, left-turn percentages, travel times, delays, and stops of the traffic on the two street segments. The traffic volumes, left-turn percentages, and geometries of the segments were then input to the models to simulate the traffic operations on them.

A series of t-tests comparing the simulation and observed mean delay times and number of stops indicated that there were no significant differences ($\alpha = 0.05$) between the simulation and observed mean values. In addition, during the conduct of this study, the models were used to simulate operations on segments that have a wide range of traffic characteristics and driveway densities, and in all cases the models gave consistent and reasonable results.

PROCEDURE

The operational effects of a TWLTL on a two-way two-lane street were determined in this study by a pairwise comparison of the outputs from the two models for identical traffic volumes and driveway densities. The two models were used to simulate traffic operations on a 1000-ft street segment, with and without a TWLTL, under three levels of traffic volume, left-turn volume, and driveway density. Simulation runs were made for all of the 27 possible combinations of these variable levels. The specific values used for these levels are given in Table 1. These values were selected as being comparable with the low, medium, and high levels of volumes and driveway density, which were used by Glennon and others (2) in developing guidelines for control of access on two-way two-lane arterial streets.

The traffic volumes given in Table 1 include left turns. And both the traffic volumes and left-turn volumes given are the volumes in each direction. Thus, the evaluation of the TWLTL in this study was for balanced traffic-flow conditions (i.e., the same traffic flow in each direction).

Also, the left-turn volumes are the total number of left turns made into all the driveways on one side of the 1000-ft street segment. In this study,

Table 1. Volume and driveway-density levels studied.

Level	Traffic Volume[a] (vph)	Left-Turn Volume[b] (vph/1000 ft)	Driveway Density[c] (no./mile)
Low	350	35	30
Medium	700	70	60
High	1000	105	90

[a]Volume in each direction, including left turns.
[b]Volume in each direction.
[c]Total number of driveways on both sides of street.

Figure 2. Driveway configurations.

(a) 30 Driveways Per Mile

(b) 60 Driveways Per Mile

(c) 90 Driveways Per Mile
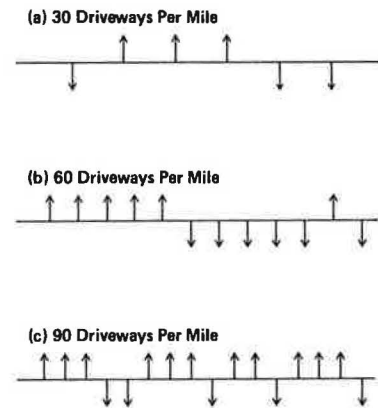
all the driveways on one side of the segment had the same volume of entering left turns. Therefore, for a given left-turn volume, the number of left turns entering each driveway was inversely proportional to the number of driveways.

The average running speeds used for each traffic-volume level were 35 mph for 350 vph, 30 mph for 700 vph, and 25 mph for 1000 vph. According to the Highway Capacity Manual (5), this speed-volume relationship was reasonable for a two-way two-lane urban arterial street.

Intuitively, the location of the driveways within the street segment has an effect on the efficiency of traffic operations. However, it was beyond the scope of this study to investigate the differences in traffic operations within driveway density levels. Instead, our primary concern was to examine the differences between driveway density levels. Therefore, only one configuration of driveway locations was evaluated for each density level. In each configuration, the driveways were evenly spaced throughout the segment. However, the side of the street on which each driveway was located was determined at random. The driveway configurations used are shown in Figure 2.

When the computer simulation runs were conducted with each model, the variability was controlled by selecting the random number variates so that the same traffic-flow and gap-acceptance sequence was always used for each driveway configuration. Therefore, for a given combination of traffic and left-turn volume levels, the differences in traffic operations were due only to the effects of the driveway configurations and the TWLTL.

Every simulation run was initialized by running the model for a few minutes to achieve system stability. Once stability was achieved, the model was run for 1 h of simulated time. Traffic operations data were collected and output for this hour.

FINDINGS

The reduction in stops and delay that results from

the installation of a TWLTL on a two-way two-lane street was computed by a pairwise comparison of the outputs from the two simulation models. The results of these computations were expressed in terms of the number of stops per hour and minutes of delay per hour that were eliminated by the installation of the TWLTL. These reductions are given in Tables 2 and 3. Although these two tables contain the same information, they are arranged differently. The reduction in stops and delay is given within driveway density in Table 2. In Table 3, it is given within traffic volume.

Examination of Table 2 reveals that in no case

Table 2. Reduction in stops and delay within driveway density.

Driveway Density[a] (no./mile)	Traffic Volume[b] (vph)	Left-Turn Volume (vph/1000 ft)[c]		
		35	70	105
Reduction in Stops (no./h)				
30	350	23	36	45
	700	98	157	290
	1000	186	612	982
60	350	0	14	29
	700	69	189	206
	1000	140	804	1216
90	350	18	27	48
	700	74	206	244
	1000	326	814	1630
Reduction in Delay (min/h)				
30	350	4.1	8.8	11.4
	700	13.7	16.8	43.8
	1000	19.4	44.2	79.8
60	350	0	3.8	9.0
	700	5.3	24.1	30.3
	1000	16.1	75.6	123.6
90	350	1.8	6.5	14.4
	700	6.9	30.2	37.6
	1000	47.3	83.6	271.1

[a]Total number of driveways on both sides of street.
[b]Volume in each direction, including left turns.
[c]Volume in each direction.

Table 3. Reduction in stops and delay within traffic volume.

Traffic Volume[a] (vph)	Driveway Density[b] (no./mile)	Left-Turn Volume (vph/1000 ft)[c]		
		35	70	105
Reduction in Stops (no./h)				
350	30	23	36	45
	60	0	14	29
	90	18	27	48
700	30	98	157	290
	60	69	189	206
	90	74	206	244
1000	30	186	612	982
	60	140	804	1216
	90	326	814	1630
Reduction in Delay (min/h)				
350	30	4.1	8.8	11.4
	60	0	3.8	9.0
	90	1.8	6.5	14.4
700	30	13.7	16.8	43.8
	60	5.3	24.1	30.3
	90	6.9	30.2	37.6
1000	30	19.4	44.2	79.8
	60	16.1	75.6	123.6
	90	47.3	83.6	271.1

[a]Volume in each direction, including left turns.
[b]Total number of driveways on both sides of street.
[c]Volume in each direction.

did the TWLTL increase stops and delay. In only one case, there was no reduction in stops and delay. Also, as expected, the size of this reduction increased within each level of driveway density as the traffic and left-turn volumes were increased. These increases in the reduction were greatest above the 700-vph traffic-volume level with more than 70 left turns per 1000 ft in each direction.

A review of Table 3 indicates that the effect of driveway density within each level of traffic volume was not consistent over the range of left-turn volumes. The reduction in stops and delay increased with driveway density only at the highest levels of traffic and left-turn volume. At the lower volume levels, these reductions were generally larger at the level of 30 driveways/mile than at 60 driveways/mile. This is probably because the average number of left turns entering each driveway at the level of 30 driveways/mile is twice that at 60 driveways/mile. However, when the driveway density was increased from 60 to 90 driveways/mile, these reductions are increased rather than decreased. One explanation of this apparent contradiction is the fact that unlike the levels of 30 and 60 driveways/mile, 90 driveways/mile did not have the same number of driveways on each side of the street segment. As illustrated in Figure 2, there were 11 driveways on one side and 5 on the other.

Thus, it is apparent that driveway configuration has an effect on the amount of the reduction in stops and delays that can be realized by a TWLTL. Therefore, application of the results of this study should be limited to street segments with driveway configurations that are similar--at least with respect to the number of driveways on each side of the segment--to those shown in Figure 2.

Another factor that should be remembered in using the results of this study is that the simulation model without the TWLTL assumed that the maximum length of time that any driver will wait to turn left was 30 s. Thus, in the model, when this wait exceeds 30 s, the turn was aborted. Although this assumption worked in the validation of the model, it was based on simulation stability requirements rather than on observed driver behavior. Wait limits much greater than 30 s caused the simulation to break down at the highest volume levels. The numbers of aborted left turns experienced in this study are given in Table 4.

The reduction in stops and delay determined in this study provides a basis for evaluating the effectiveness of a TWLTL on a two-way two-lane street from the standpoint of user costs, energy consumption, and air quality. The stops and delay reduction values are directly applicable to procedures for evaluating traffic engineering improvement, such

Table 4. Number of left turns aborted per hour.

Driveway Density[a] (no./mile)	Traffic Volume[b] (vph)	Left-Turn Volume (vph/1000 ft)[c]		
		35	70	105
30	350	5	6	18
	700	6	15	18
	1000	12	26	49
60	350	0	1	3
	700	2	9	11
	1000	9	31	50
90	350	2	6	13
	700	2	11	20
	1000	8	36	96

[a]Total number of driveways on both sides of street.
[b]Volume in each direction, including left turns.
[c]Volume in each direction.

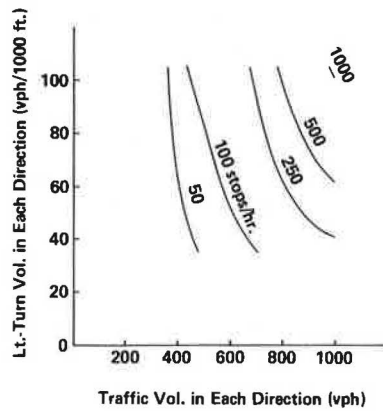**Figure 3. Reduction in stops: 30 driveways/mile.**

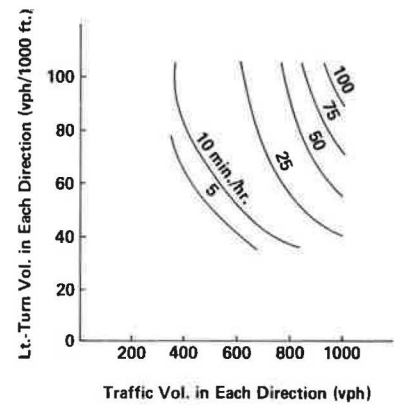**Figure 4. Reduction in stops: 60 driveways/mile.**

Figure 5. Reduction in stops: 90 driveways/mile.

Figure 6. Reduction in delay: 30 driveways/mile.
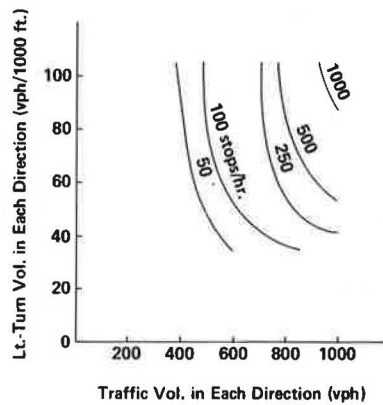

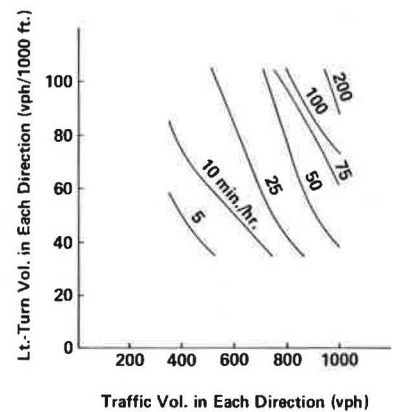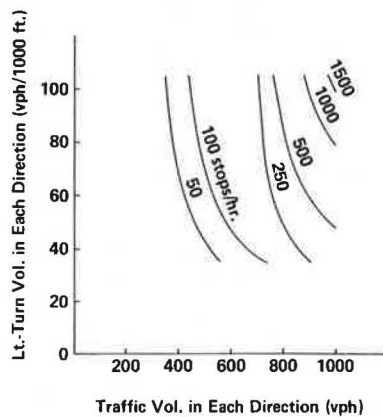**Figure 7. Reduction in delay: 60 driveways/mile.**

Figure 8. Reduction in delay: 90 driveways/mile.

as the procedure outlined by Dale (6). Therefore, to facilitate this application of the results of this study, isograms of the reduction in stops and delay were constructed from the data in Table 2. The stop-reduction isograms for the three levels of driveway density are shown in Figures 3, 4, and 5, and the delay-reduction isograms are shown in Figures 6, 7, and 8.

CONCLUSIONS

Based on the findings of this study, it was concluded that the installation of a TWLTL on a two-way two-lane street improves the efficiency of traffic operations over a wide range of traffic volumes, left-turn volumes, and driveway densities. Under balanced traffic-flow conditions, it is particularly effective at traffic volumes above 700 vph in each direction with more than 70 midblock left turns per 1000 ft from each direction.

The stop and delay reduction isograms that were developed in this study facilitate the quantitative evaluation of the operational effectiveness of a TWLTL under balanced traffic-flow conditions on a two-way two-lane street. Used within the context of a cost-effectiveness analysis, these isograms contribute to the identification of the circumstances under which the installation of a TWLTL on a two-way two-lane street would be justified.

However, this study is just a start. The need for further research is obvious. Additional studies need to be conducted for more levels of traffic volume, left-turn volume, and driveway density. Future studies should address unbalanced as well as balanced traffic-flow conditions, and the effects of driveway configuration need to be evaluated. Of course, similar research needs to be conducted to
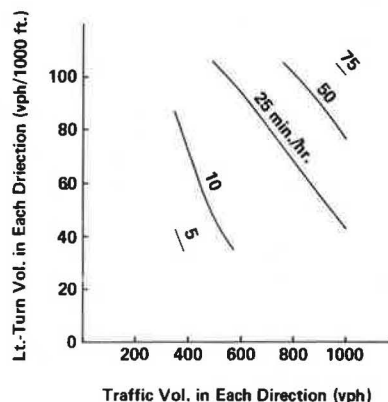
evaluate the operational effects of a TWLTL on two-way four-lane streets.

REFERENCES

1. Z.A. Nemeth. Development of Guidelines for the Application of Continuous Two-Way Left-Turn Median Lanes. Engineering Experiment Station, Ohio State Univ., Columbus, Final Rept. EES 470, July 1976.

2. J.C. Glennon, J.J. Valenta, B.A. Thorson, and J.A. Azzeh. Technical Guidelines for the Control of Direct Access to Arterial Highways--Volumes 1 and 2. FHWA, Repts. FHWA-RD-76-86 and FHWA-RD-76-87, Aug. 1975.
3. T.J. Schriber. Simulation Using GPSS. Wiley, New York, 1974.
4. D.L. Gerlough and F.A. Wagner. Improved Criteria for Traffic Signals at Individual Intersections. NCHRP, Rept. 32, 1967.
5. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
6. C.W. Dale. Procedure for Evaluating Traffic Engineering Improvements. ITE Journal, April 1981, pp. 39-46.

Effects of Trucks on Freeway Vehicle Headways Under Off-Peak Flow Conditions

WILEY D. CUNAGIN AND EDMUND CHIN-PING CHANG

The results of a study to determine the effects of the presence of heavy trucks on traffic flow in sections of freeway as an operational measure of total throughput capacity are presented. The variable used to evaluate truck impacts was time headway. Data were collected at two sites on the Houston, Texas, freeway system during off-peak flow conditions. After each observed headway had been classified as to types of vehicles involved in the interaction, various statistical tests were performed to analyze variations in headway due to headway type, lane width, and traffic volume. Headway type (i.e., the types of vehicles involved in the headway interaction) was shown to be the major determinant in length of the headway; those that involved trucks exhibited the greatest magnitude.

In recent years the construction of new highway facilities has not kept pace with the expansion of vehicular travel. In urban areas in particular, concern with measures to increase the efficiency of traffic operations has aroused increasing interest as the emphasis has shifted toward making the existing system work as well as it can. The diverse mixture of vehicle sizes, weights, and operating characteristics has become a potential limiting factor in trying to attain maximum efficiency and minimum accident experience from the highway system.

Of approximately 145 million motor vehicles in operation in this country today, nearly 7 million are trucks with empty gross vehicle weights of 10 000 lb or more. When these trucks are involved in accidents with the passenger cars in the traffic stream, the results can be startling. Although heavy trucks comprise less than 2 percent of the vehicle population, they were involved in accidents that accounted for almost 9 percent of all traffic fatalities in 1976. Of these, 91 percent were persons in other vehicles that conflicted with the trucks (1).

The problem is further complicated by an increasing polarization of the vehicle mix into very small cars and very large trucks. The trend toward smaller, more efficient passenger cars is undeniable. In 1963, automobiles made up 84.3 percent of the total vehicle fleet and included about 8 percent automobiles with registered vehicle weights of 3000 lb or less (2). By 1978, automobiles were down to 79 percent of the vehicle total but the small-car portion had risen to 22 percent. By 1990, the proportion of automobiles is expected to be 75 percent while more than 50 percent of those will have registered weights of less than 3000 lb (3). Unfortunately, the quest for more economical personal transportation vehicles has been pursued through methods that reduce the survivability of the passengers in an accident, since the smaller passenger cars generally are characterized by reduced track width, higher center of gravity, reduced horsepower, reduced weight, reduced structural integrity, and lower driver eye height.

Spurred by both demand for more fuel-efficient vehicles due to rapidly rising gasoline prices and mandatory standards set in the Energy Policy and Conservation Act of 1975, gains in mileage per gallon have been attained primarily by lowering horsepower and increasing ratios of weight to horsepower. These changes have tended to reduce acceleration rates and therefore the vehicle performance capabilities (3). A study by Woods and others (4) showed that although smaller vehicles accelerated adequately at low speeds, their acceleration capability at highway speed was substantially lower than that of full-size cars. Indeed, at 50 mph, more than 200 additional ft were required for the 85th-percentile small cars to pass another automobile. A recent study by the Institute for Highway Safety (5) showed relative injury rates on a normalized experience basis by make and model of automobile. The best vehicles from the standpoint of protecting occupants were full-size cars, and the worst were subcompacts or smaller. For example, drivers of a Honda Civic are three times as likely to be killed

or injured in an accident as drivers of an Oldsmobile Delta 88.

Concurrent with these developments, legal gross weights for permitted vehicles are expected to reach 120 000 lb in the near future (6). Many states have already increased the weight limits to 80 000 lb with allowable lengths of 65 ft and articulated configurations (7). Truck lengths are projected to reach 94 ft with three trailers pulled by a single tractor in 1990. The large trucks are operationally limited in their ability to stop, accelerate, corner, and maneuver relative to the performance of passenger cars and other smaller trucks. As more fuel-efficient trucks are introduced, the use of engines with low revolutions per minute and low friction, low-rolling-resistance tires, aerodynamic-drag-reduction devices, and low-parasitic-power-loss accessories and lubricant will further diminish their natural direct-braking capabilities (8).

In view of the projected increase of 30 percent more vehicles and increase in miles per vehicle as well as larger percentages of heavy trucks and small cars (3) and the demonstrated serious consequences of operational conflicts between these two vehicle types, it is worthwhile to consider the operational impacts of heavy trucks on the traffic stream.

The effect of grade on heavy-truck performance has long been a subject of concern (9-17). On two-lane roadways in particular (18), climbing lanes are often provided on steep and/or long grades. Current practice generally calls for climbing lanes (or vertical profile modification) when the truck speeds are expected to fall 15 mph below the average passenger-car speeds. A joint study by the Texas Transportation Institute and the Center for Highway Research (19) showed that the distance required for a full-size car to pass a 95-ft triple-bottom truck is about 330 ft more than that to pass a 65-ft double-bottom truck.

Lower-performance small cars will simply compound this problem. The operational effects of trucks on grades are addressed in the Highway Capacity Manual (20). In this guide, for a freeway on level terrain, one truck is the equivalent of two passenger cars. On rolling terrain, one truck is equivalent to four passenger cars generally, but more precise (and much higher) equivalents are given by percent and length of grade. Obviously, these guidelines must be reevaluated. Two current research studies sponsored by the Federal Highway Administration are addressing the relative performance of different types of cars and trucks and three recreational vehicles.

The operational problems of trucks on downgrades have also been addressed (21). Consequently, many states have implemented emergency escape ramps for runaway vehicles; these include both gravity ramps and arrestor beds.

The disparity in operational characteristics between heavy commercial vehicles and passenger cars is aggravated by geometric design practices, which often have not addressed the problem. Heavier, longer trucks on grades will introduce speed differentials that were not expected when the facilities were designed and may cause impatient motorists to attempt to pass in unsafe situations. Further, previously safe passing zones may no longer be adequate for lower-performance automobiles. Certainly, the design driver eye height of 3.75 ft is much higher than that which the majority of drivers enjoys (22,23).

OBJECTIVE

Although various aspects of freeway truck operations

have been examined elsewhere (24), the impact of heavy trucks on the capacity of urban freeways has not been satisfactorily ascertained. This paper presents the results of a study of heavy trucks on the freeway in Houston, Texas. A significant program of increasing the capacities of freeway bottlenecks has been conducted in this urban area by the Texas State Department of Highways and Public Transportation (TSDHPT) with support from the Texas Transportation Institute. The variable chosen to indicate traffic flow performance was time headway, defined as the time in seconds for the front bumper of two successive vehicles in the same lane to pass a single datum point.

DATA COLLECTION

Data were collected by photographic means. A total of 2 h of data was collected at each of two urban freeway locations in Houston, Texas. One study site was the westbound lanes of the Southwest Freeway (US-59) east of the loop IH-610 West Interchange. The other location was the eastbound lanes of IH-610 South at the new SH-288 interchange, which was under construction. Fifteen-minute traffic volumes were recorded by lane. Super-8 movie films were taken for one 5-min period in each of the quarter hours. The data were collected between the hours of 10:00 a.m. and noon at each site. The study was conducted at the Southwest Freeway sites on Tuesday, July 31, 1979, and at the IH-610 site on Wednesday, August 8, 1979. Physical measurements of roadway features such as lane widths and offsets from landmarks were obtained on site from each location. Although the film speed was set at 9 frames/s, a test car was driven through the study area during each study to confirm the camera speed.

MEASUREMENT TECHNIQUE

At both locations, the procedure followed was to set up the data-collection station above and to the right of the traveled way. In the case of the site on the Southwest Freeway, permission was obtained from the operator of a parking garage to set up on the top floor. At the IH-610 site, the contractor constructing the interchange and TSDHPT authorized setting up the data-collection site on an overpass. A Kodak XL55 Super-8 movie camera with a telephoto lens was used to collect the photographic data. Two research assistants collected the 15-min lane-count data on each of the five lanes on the Southwest Freeway and three lanes on IH-610 South. Physical measurements at each site were made about 6:00 a.m. on Sunday mornings following the studies. This was necessary due to the almost constantly high volumes on the Houston Freeway system. The test car was driven through the study area at a constant speed of 50 mph to obtain calibration data for the film speed.

ANALYSIS

The vehicle headways were obtained from the Super-8 movie films by using a time-lapse projector. The raw headway data were recorded in terms of frames between successive passages of a datum point on the projection screen. Conversion from frame headway to time headway was achieved by considering the test-vehicle data. From physical measurements of the lane striping, it was known that there was a distance of 260 ft between the ends of lane stripes at either end of the study area visible on the films. The test vehicle traversed this distance in 15.5 frames at 50 mph to yield the following frame rate:

Frame rate = 260 ft/{(15.5 frames)(50 mph)[1.467(ft·h)/(s·mile)]} = 0.229 s/frame.

The headways were classified into one of the following types:

Type 1: car following car (automobile/automobile),

Type 2: car following single-unit truck (automobile/truck),

Type 3: single-unit truck following car (truck/automobile),

Type 4: car following tractor with trailor (automobile/truck),

Type 5: tractor with trailer following car (truck/automobile), and

Type 6: truck following truck (truck/truck).

Headways greater than 9 s were not recorded since these were not considered to exhibit intervehicular interactions. All headways involving trucks were used, whatever type of leading and/or following vehicle was recorded, but not all automobile/automobile headways were recorded. Instead a representative sample of the automobile/automobile headways less than 9 s was selected. The resulting sample sizes by lane and headway type are shown in Table 1. The lane volumes by 15-min periods are shown in Table 2.

RESULTS

Duncan's multiple-range test was applied to compare mean headways for different headway types. The test results are shown below (note that mean headways

Table 1. Sample sizes by headway type and lane for Southwest Freeway and IH-610 data.

Lane No.	Headway Type					
	1	2	3	4	5	6
Southwest Freeway Data						
1	51	8	7	11	11	2
2	64	14	11	6	8	2
3	146	17	26	6	7	0
4	179	14	15	0	1	1
5	74	1	8	1	2	2
Total[a]	514	54	67	24	29	7
IH-610 Data						
1	197	14	20	26	29	2
2	154	15	24	13	18	6
3	60	5	11	1	2	3
Total[b]	411	34	55	40	49	11

[a]Total observations = 695.   [b]Total observations = 600.

Table 2. Lane volumes.

Quarter Hour	Southwest Freeway Lane					IH-610 Lane		
	1	2	3	4	5	1	2	3
1	258	305	330	318	150	324	136	116
2	231	272	297	280	128	324	125	108
3	230	300	325	301	127	305	143	111
4	267	299	350	304	143	299	460	212
5	274	344	345	334	165	324	306	109
6	302	330	387	377	181	356	322	142
7	317	274	380	421	182	321	316	115
8	321	189	403	380	223	301	332	150
Total	2200	2313	2817	2715	1299	2554	2140	1063

Note: Statistical analyses were performed after coding for each headway site number, quarter-hour period number, lane number, headway type, headway in seconds, lane width in feet, and quarter-hour total traffic volume.

Figure 1. Frequency histogram for automobile/automobile headways.
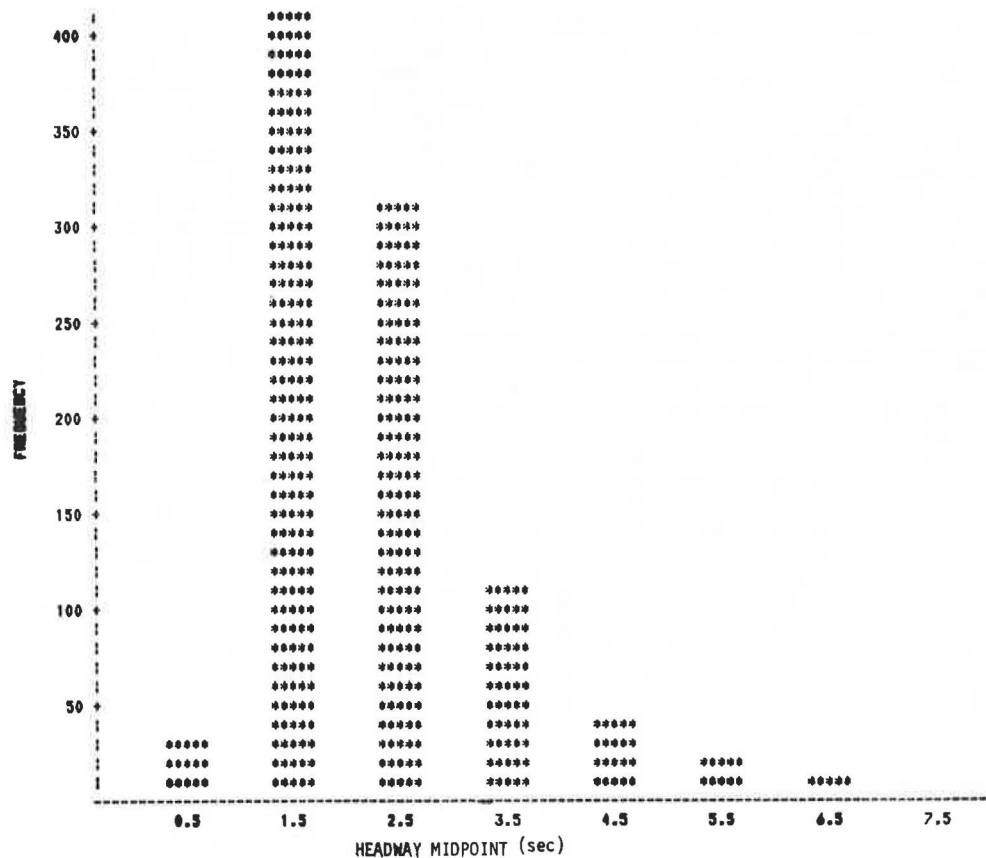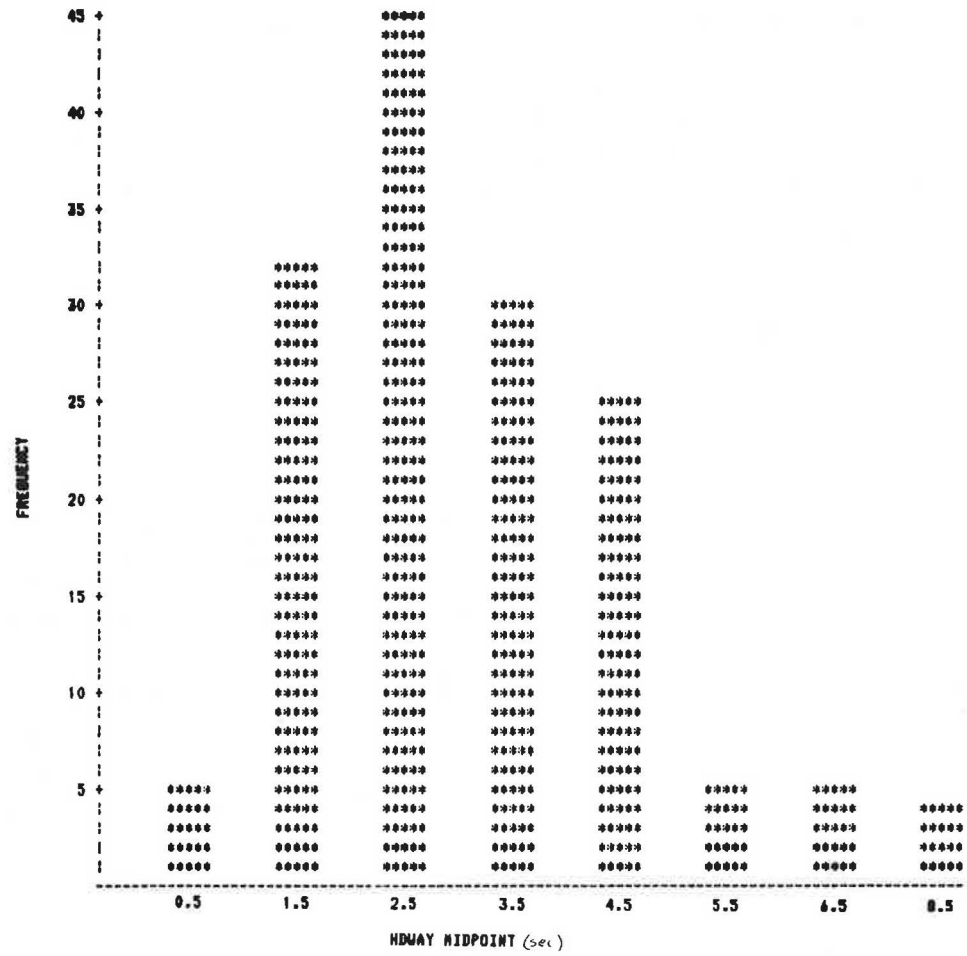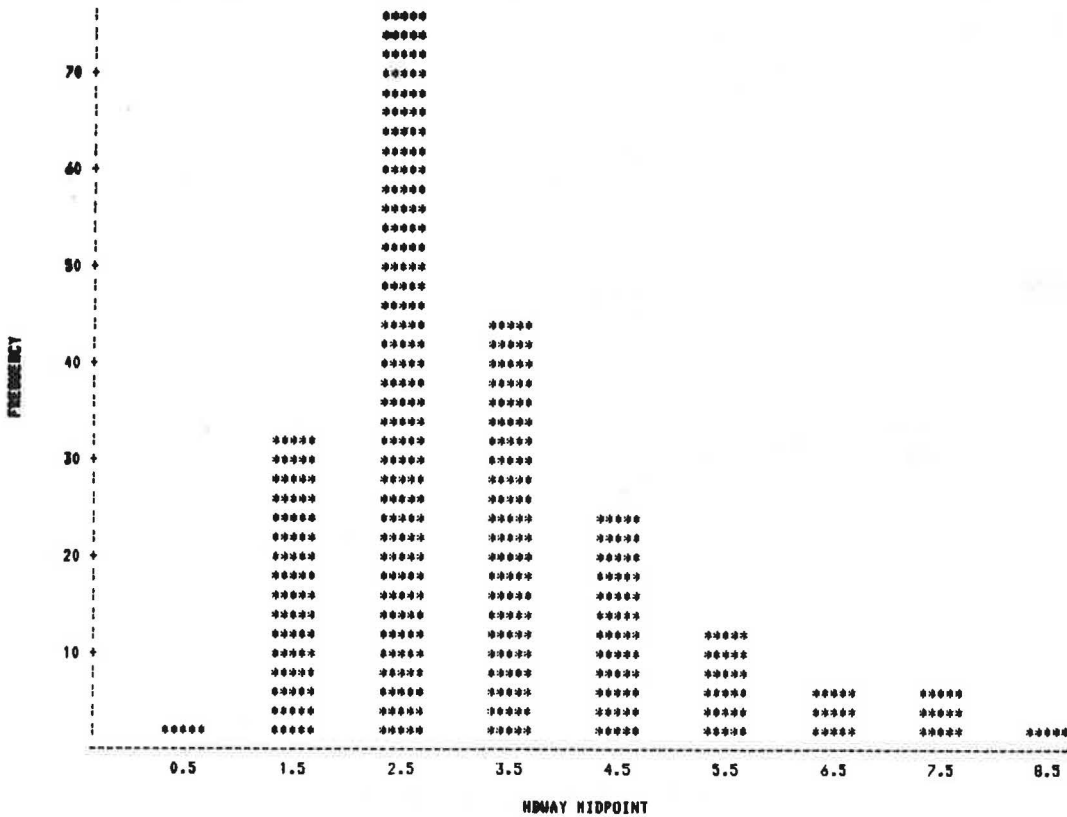
Figure 2. Frequency histogram for automobile/truck headways.

Figure 3. Frequency histogram for truck/automobile headways.

underlined as a group are not significantly different):

1. All freeway data (α-level = 0.05, degrees of freedom = 1287, mean square = 1.463 41):

Headway Type					
1	3	2	4	5	6
2.297	2.297	3.005	3.288	3.834	4.041

2. Southwest Freeway data (α-level = 0.05,

degrees of freedom = 687, mean square = 1.3343):

Headway Type					
1	3	2	4	5	6
2.383	2.971	3.147	3.416	3.864	4.237

3. IH-610 data (α-level = 0.05, degrees of freedom = 592, mean square = 1.5421):

Headway Type					
1	2	3	4	5	6
2.189	2.778	2.874	3.212	3.596	4.153

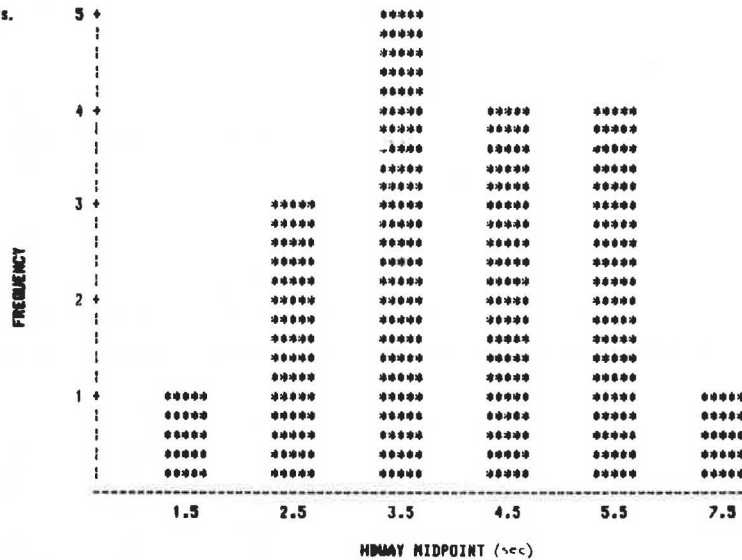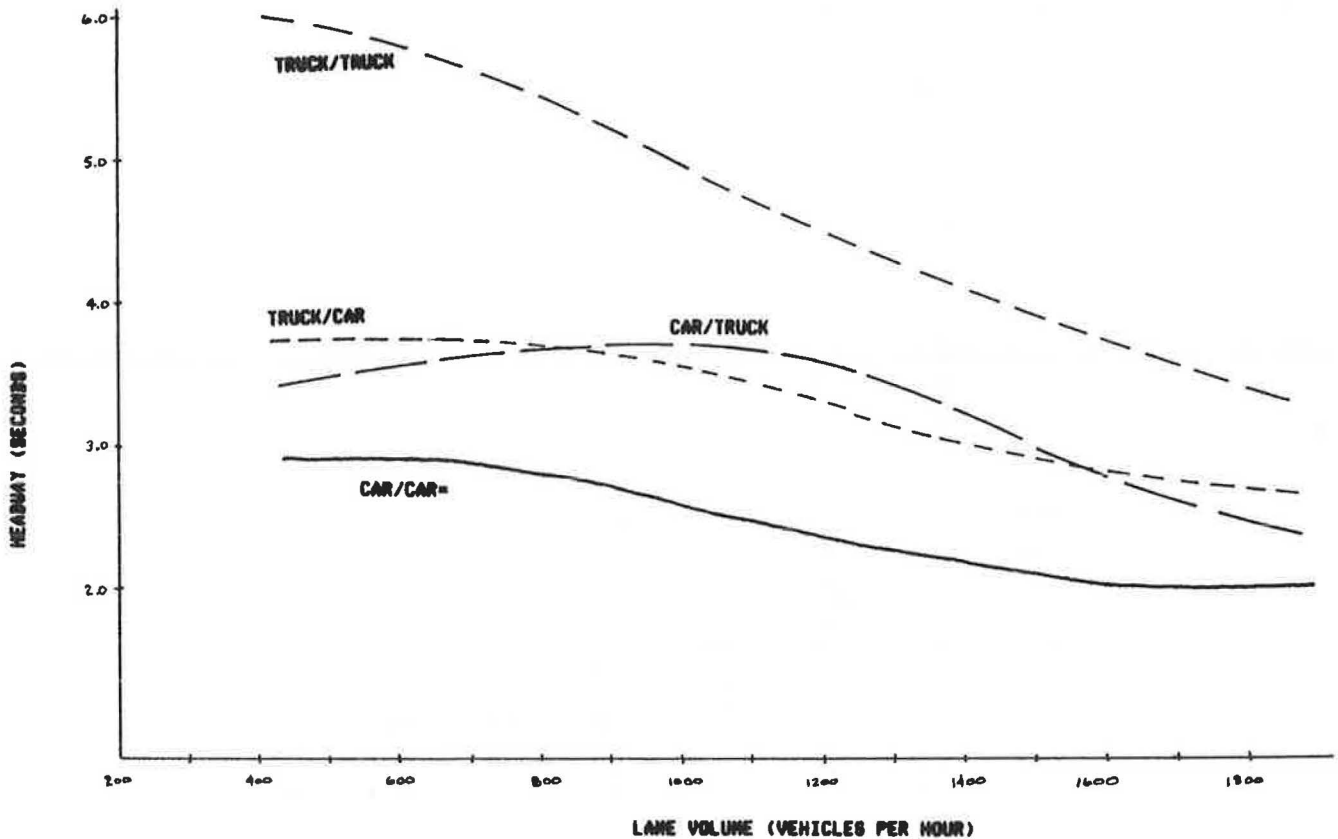Figure 4. Frequency histogram for truck/truck headways.

Figure 5. Headway means by headway type versus lane volume.

The results clearly show that the automobile/automobile headways are significantly less than all other headway types that involved trucks with automobiles or with each other. The type of truck also seemed to have an effect, since the mean headways of cars following trucks and of a single-unit truck following a car, taken as a group, are significantly less than those of tractors with semitrailers following either cars or other trucks. Examination of the histograms of the frequencies of headways by 1-s intervals for the automobile/automobile (Figure 1), automobile/truck (Figure 2), truck/automobile (Figure 3), and truck/truck (Figure 4) headway types reveals obvious differences in the shapes of the distributions. Indeed, the automobile/automobile headway type is skewed strongly to the right, whereas the automobile/truck and truck/automobile headway types are skewed only slightly to the right. The truck/truck headway type appears to be skewed slightly to the left.

Note that the headway data include only headways less than 9 s in length, whereas the traffic volume data include all vehicles passing during the study period.

Figure 5 is a plot of headway means by headway type versus lane volume. Although all headway types indicate a reduction in headway with increasing volume, the truck/truck interaction again appears to be the most profound.

CONCLUSIONS

The presence of trucks in the traffic stream is accompanied by an increase in the mean headway. Although this phenomenon is not critical at the flow rates observed in this study, the reduction in capacity (predicted by the Highway Capacity Manual) might become significant during the peak hours. Data collected in this effort were not sufficient to quantify the expected reduction in capacity due to heavy truck interactions during the peak period.

Contrary to complaints often expressed by automobile drivers, truck drivers did not appear to operate their vehicles unnecessarily close behind other vehicles. Indeed, they seemed to allow more room to the front than did automobile drivers.

ACKNOWLEDGMENT

REFERENCES

1. Improved Commercial Vehicle Conspicuity and Signaling Systems. FHWA, Rept. RFP NHTSA-0-A554, July 1979.
2. C. Pinnell. Changing Vehicle Mix on Our Highways and Attendant Problems. Presented at the 56th Annual Meeting, TRB, 1977.
3. C.V. Wootan. The Changing Vehicle Mix and Its Implications. Presented at Texas State Department of Highways and Public Transportation's Fifty-Third Annual Short Course, Nov. 1979.
4. G.D. Weaver, D.L. Woods, and others. Passing and No-Passing Zones: Signs, Markings, and Warrants. FHWA, Final Rept., Sept. 1978.
5. Crash Test Study. Institute for Highway Safety, National Highway Traffic Safety Administration, 1979.
6. Interagency Study of Post-1980 Goals for Commercial Motor Vehicles. FHWA, n.d.
7. Summary of Size and Weight Limits. American Trucking Associations, Inc., Washington, DC, 1978.
8. A Study of the Cost Benefits Derived from Using Retarders in Heavy Duty Commercial Vehicles. FHWA, Rept. RFP NHTSA-9A531, 1979.
9. A Policy on Geometric Design of Rural Highways. AASHTO, Washington, DC, 1965.
10. W.A. Stimpson and J.C. Glennon. Critical Review of Climbing-Lane Design Practices. HRB, Highway Research Record 371, 1971, pp. 1-11.
11. J.K. Whitfield. Runaway Truck Arresting Schemes. FHWA, in preparation.
12. J.C. Glennon. An Evaluation of Design Criteria for Operating Trucks Safely on Grades. HRB, Highway Research Record 312, 1970, pp. 93-112.
13. T.S. Huff and F.H. Scrivner. Simplified Climbing Lane Design Theory and Road-Test Results. HRB, Bull. 104, 1955.
14. R.E. Dunn. Motor Vehicle Performance on Ascending Grades. HRB, Bull. 104, 1955.
15. W.E. Willey. Truck Congestion on Uphill Grades. HRB, Bull. 10r, 1955.
16. R.M. Williston. Truck Deceleration Rate Study. Traffic Division, Connecticut Highway Department, Wethersfield, 1967.
17. J.E. Leisch and J.P. Leisch. New Concepts in Design-Speed Application. TRB, Transportation Research Record 631, 1977, pp. 4-14.
18. F.F. King and others. Seven Experimental Designs Addressing Problems of Safety and Capacity on Two-Lane Rural Highways, Volume 5. FHWA, Final Rept. DOT-TSC-FMWA-78-2-V, May 1978.
19. Effects of Heavy Trucks on Texas Highways. Center for Highway Research and Texas Transportation Institute, College Station, TX, Interim Rept. 1-8-78-231, Sept. 1978.
20. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
21. Emergency Escape Ramps for Runaway Heavy Vehicles. FHWA, Rept. FHWA-TS-79-201, 1979.
22. W.D. Cunagin and A.D. Abrahamson. Driver Eye Height: A Field Study. ITE Journal, May 1979.
23. A.D. Abrahamson. Driver Eye Height Measurement. Texas A&M Univ., College Station, Master's thesis, 1978.
24. J.B. Humphreys. Effect of Trucks on the Urban Freeway. HRB, Highway Research Record 308, 1970, pp. 32-75.

Estimation of Passenger-Car Equivalents of Trucks in Traffic Stream

MATTHEW J. HUBER

The passenger-car equivalent (PCE) of a truck represents the number of passenger cars (basic vehicles) displaced by each truck in the traffic stream under specific conditions of flow. A model is proposed for estimating PCE-values for vehicles under free-flowing, multilane conditions. Some measure of impedance as a function of traffic flow is used to relate two traffic streams—one that has trucks mixed with passenger cars and the other that has passenger cars only. PCE-values are related to the ratio between the volumes of the two streams at some common level of impedance. A deterministic model of traffic flow (Greenshields') is used to estimate the impedance-flow relationship. Three measures of impedance are considered, each of which will generate a separate PCE-value for a truck of given characteristics. PCE-values are also shown to relate to speed and length of subject vehicles and to vary with the proportion of trucks in the traffic stream.

The passenger-car equivalent (PCE) of a truck is introduced in the Highway Capacity Manual as follows ($\underline{1}$, p. 101):

> Trucks (defined for capacity purposes as cargo-carrying vehicles with dual tires on one or more axles) reduce the capacity of a highway in terms of total vehicles carried per hour. In effect, each truck displaces several passenger cars in the flow. The number of passenger cars that each dual-tired vehicle represents under specific conditions is termed the "passenger car equivalent" for those conditions.

The Highway Capacity Manual lists PCE-values for two categories of vehicles--trucks and intercity buses. PCE-values for a third category of vehicles, recreation vehicles, have been determined from field observations made on Canadian highways ($\underline{2},\underline{3}$).

It is evident that each of the three categories will include a wide range of vehicles. The truck category as now defined in the Highway Capacity Manual includes vehicles ranging from single-unit trucks with 6 tires to combination trucks with 18 or more tires. A single PCE-value for trucks does not adequately reflect the diverse characteristics of the many categories of vehicles that may be observed in the traffic stream.

The Federal Highway Administration is in the process of updating the national highway cost-allocation study. One factor in these cost-allocation studies is an analysis of the highway service capacity consumed by various classes of vehicles. To this end, the number of vehicle categories has been expanded to 15, as listed below [categories 1-13 are from a Federal Highway Administration report ($\underline{4}$, Table III-2.1); categories 14 and 15 are from a Voorhees report ($\underline{5}$)]:

1. Automobiles, large (15 ft and more);
2. Automobiles, small;
3. Motorcycles;
4. Buses;
5. Single-unit trucks, two axles, four tires;
6. Single-unit trucks, two axles, six tires;
7. Single-unit trucks, three or more axles;
8. Three-axle combination trucks;
9. 2S2 four-axle combination trucks;
10. Other four-axle combination trucks;
11. 3S2 five-axle combination trucks;
12. Other five-axle combination trucks;
13. Six-axle or more combination trucks;

Figure 1. Flow-impedance relationship.
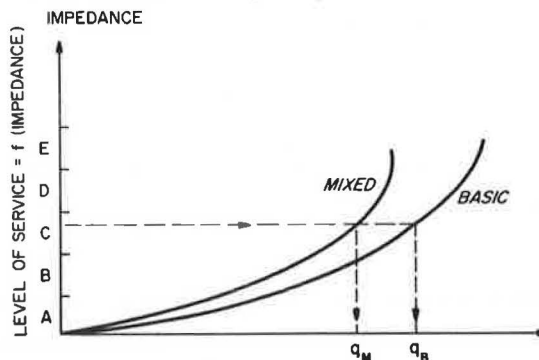
14. Noncommercial vans; and
15. Four- and six-tire recreational vehicles (campers, mobile homes, trailers).

There are currently (July 1981) nationwide studies under way that are being conducted to determine PCE-values for the different categories of vehicles in the following situations: urban arterials, urban freeways, rural two-lane two-way roadways, and rural freeways.

The analysis that follows has been made in order to anticipate the results that will follow from the nationwide studies listed above and to determine the underlying relationships between vehicle characteristics and the determination of PCE-values for the different categories of vehicles. A simple model has been used to represent steady-state traffic flow with and without trucks present, and the relationships between the resulting flows have been used to calculate the PCE-values.

FRAMEWORK FOR ESTIMATING PCE-VALUES

Consider the relationship between some measure of impedance along a length of roadway and the flow rate along that same roadway for two different traffic streams. The flow-impedance relationship is shown in Figure 1, in which the basic curve represents a stream consisting solely of basic vehicles (passenger cars) and the mixed curve represents a stream with proportion of trucks p and of basic vehicles $(1 - p)$. As the flow rate q increases, the impedance increases; the increase in impedance is at a greater rate for the mixed flow. The impedance in turn can be related to the level of service (LOS) on the roadway, where LOS A is the most desirable and LOS E is the least desirable.

For any given LOS (or impedance) it is possible to calculate corresponding flow rates q_B and q_M as shown. These flow rates for the basic and mixed streams will produce identical measures of LOS and can then be equated so that $q_B = (1 - p)q_M + pq_M(PCE)$. Solving for PCE, the result is

$$PCE = (1/p)[(q_B/q_M) - 1] + 1 \qquad (1)$$

Figure 2. Sample calculation of PCE-values.

$n_B = 6$

length

stop line

$n_M = 4$

$p = 0.25$

Figure 3. Greenshields model of traffic flow.

DENSITY   FLOW
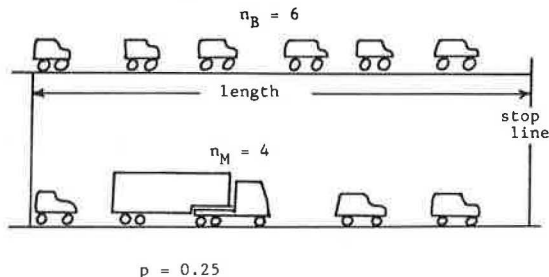
where

> PCE = passenger-car equivalent,
> p = proportion of trucks in mixed traffic flow, and
> q_B, q_M = flow rate at common LOS for basic and mixed traffic streams, respectively.

An example of the concept given in Equation 1 is shown in Figure 2, where a PCE-value is developed for a standing queue of vehicles as might be observed at a signalized intersection. In the first instance, there were six basic vehicles (n_B) observed over a length of roadway ℓ, while on an adjacent lane there were four vehicles (n_M) observed, one of which is a truck (p = 0.25), over the same length ℓ. The two queues, n_B and n_M, develop a common measure of length ℓ so that by reasoning similar to that of Equation 1, we calculate

$$PCE = (1/0.25)[(6/4) - 1] + 1 = 3.0.$$

There are several variables that may be used as a measure for the LOS or impedance shown on the vertical axis of Figure 1. A common measure is the average travel time [t(q)] over a length of roadway, where the travel time will increase as the flow q increases. For example, consider a single lane of a multilane urban roadway with a speed limit of 50 km/h (31 miles/h). At a flow rate q of 100 vehicles/h, the mean velocity is 50 km/h and the travel time t(q) will be 1/50 h/km or 1.200 min/km (1.931 min/mile). At 900 vehicles/h, the mean velocity is reduced to 40 km/h (24 miles/h) and the travel time t(q) will increase to 1/40 h/km or 1.500 min/km (2.414 min/mile).

An alternative measure is the time of occupancy or total travel time [T(q)], where T(q) is the volume q times the mean travel [t(q)]. For the same example cited above the time of occupancy becomes at q = 100 vehicles/h:

T(q) = 100 vehicles/h x 1/50 h/km = 2.00 vehicle-h/km-h,

and at q = 900 vehicles/h,

T(q) = 900 vehicles/h x 1/40 h/km = 22.500 vehicle-h/km-h.

The time of occupancy, in turn, is numerically equivalent to the density k, where density k = flow q ÷ speed u, or at 900 vehicles/h,

k = 900 vehicles/h ÷ 40 km/h = 22.500 vehicles/km.

Either of these two measures of impedance, the mean travel time t(q) or the density k [numerically equal to the time of occupancy T(q)], can be used to calculate PCE-values as suggested in Figure 1. The
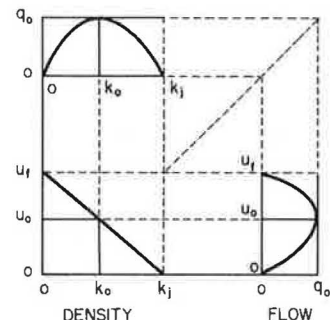
Greenshields (6) model of traffic flow, which assumes a straight-line relationship between density and velocity, is used to develop the interrelationships among the variables speed (u), density (k), and flow rate (q) for steady-state flow. These relationships are shown in Figure 3 where the parameters are jam density (k_j), free-flow speed (u_f), optimum density (k_0), optimum speed (u_0), and maximum flow (q_0).

FLOW PARAMETERS OF MIXED TRAFFIC

In a simplified case, mixed traffic is assumed to be made up of only two types of vehicles, basic vehicles with an effective length L_B and free-flow speed u_{FB} and trucks with an effective length L_T and free-flow velocity u_{FT}. The effective length of a vehicle is the distance the vehicle occupies when in a standing queue and is measured from the rear bumper of the preceding vehicle to the rear bumper of the subject vehicle. The free-flow speed is the speed of the vehicle when not influenced by other vehicles on the roadway.

The mixed-flow rate is the sum of the flow rate of basic vehicles plus the flow rate of trucks:

$$q_M = q_{MB} + q_{MT} \tag{2}$$

where

> q_M = flow rate of mixed vehicles,
> q_{MB} = flow rate of basic vehicles within mixed stream, and
> q_{MT} = flow rate of trucks within mixed stream.

The proportion p of trucks in the mixed traffic stream flow is as follows:

$$p = q_{MT}/q_M \tag{3}$$

The density of the mixed flow is the sum of the density of basic vehicles plus the density of trucks:

$$k_M = k_{MB} + k_{MT} \tag{4}$$

where

> k_M = density of mixed vehicles,
> k_{MB} = density of basic vehicles within mixed stream, and
> k_{MT} = density of trucks within mixed stream.

The proportion p' of trucks in the mixed traffic density is as follows:

$$p' = k_{MT}/K_M \tag{5}$$

The mean velocity of the mixed stream of traffic is the harmonic mean of the velocities of the basic vehicles and trucks:

$$u_M = 1 / \left\{ (p/u_{MT}) + [(1-p)/u_{MB}] \right\} \qquad (6)$$

where

u_M = mean velocity of mixed traffic stream,
u_{MB} = mean velocity of basic vehicles within mixed traffic stream,
u_{MT} = mean velocity of trucks within mixed traffic stream, and
p = proportion of trucks in mixed traffic stream flow.

The development of the proportion p' of trucks in the mixed traffic density follows from the relationship $q_O = k_O u_O$ where the subscript O refers to maximum (optimum) flow rate:

$$q_{OMT} = p q_{OM} = k_{OMT} u_{OMT}$$

$$q_{OMB} = (1-p) q_{OM} = k_{OMB} u_{OMB}$$

and

$$k_{OMT} = q_{OMT}/u_{OMT} = p q_{OM}/u_{OMT} \qquad (7)$$

$$k_{OMB} = q_{OMB}/u_{OMB} = (1-p) q_{OM}/u_{OMB} \qquad (8)$$

Equations 7 and 8 are substituted into Equation 5:

$$
\begin{aligned}
p' &= k_{OMT}/k_{OMT}/(k_{OMT} + k_{OMB}) \\
&= 1/[1 + (k_{OMB}/k_{OMT})] \\
&= 1/\left(1 + \left\{ [(1-p) q_{OM} \times u_{OMT}]/(p q_{OM} \times u_{OMB}) \right\} \right) \\
&= 1/\left\{ 1 + [(1-p) u_{OMT}/p u_{OMB}] \right\}
\end{aligned}
$$

Since, for the Greenshields model of traffic flow $u_O = u_f/2$, the final expression is as follows:

$$p' = 1/\left\{ 1 + [(1-p) u_{FT}/p u_{FB}] \right\} \qquad (9)$$

The jam density (k_j)—the number of stopped vehicles in a length of roadway—becomes, for basic vehicles only,

$$k_{jB} = L/L_B \qquad (10a)$$

and for mixed vehicles,

$$k_{jM} = L/[p' L_T + (1-p') L_B] \qquad (10b)$$

where L is the unit length of roadway [1000 m (5280 ft)] and L_T, L_B are the effective length of trucks and basic vehicles.

NUMERICAL EXAMPLE OF TRAFFIC FLOW

Consider a steady-state stream of traffic on a single lane of a multilane urban arterial with 10 percent trucks ($p = 0.10$). The free-flow velocity of basic vehicles u_{FB} is 48.280 km/h (30.0 miles/h) and of trucks u_{FT} is 32.187 km/h (20.0 miles/h). The effective length of basic vehicles L_B is 7.62 m (25 ft), and the effective length of trucks L_T is 22.86 m (75 ft).

The free-flow velocity of the mixed flow is found from Equation 6:

$$u_{FM} = 1/[(0.10/32.187) + (0.90/48.280)] = 45.981 \text{ km/h (28.571 miles/h)}.$$

Substituting in Equation 9,

$$p' = 1/\{1 + [(0.9/0.1)(32.187/48.280)]\} = 0.143.$$

The jam density of mixed flow (k_{jM}) is found by substituting in Equation 10b:

$$k_{jM} = 1000/[0.143(22.86) + 0.857(7.62)] = 102.071 \text{ vehicles/km (164.267 vehicles/mile).}$$

The optimum flow rate for mixed vehicles (q_{OM}) is found from the following relationship:

$$q_{OM} = (k_{jM}/2)(u_{FM}/2)$$

$$= (102.071/2)(45.981/2) = 1173.33 \text{ vehicles/h}$$

For a stream flow of basic vehicles only, the parameters are as follows:

u_{FB} = 48.280 km/h (30 miles/h),
k_{jB} = 1000/7.620 = 131.234 vehicles/km (211.200 vehicles/mile), and
q_{OB} = (48.280/2)(131.234/2) = 1584.000 vehicles/h.

The relationships between pairs of variables for basic vehicles only and for mixed vehicles are shown in Figures 4, 5, and 6. Figure 4 represents the velocity-density relationship, Figure 5 the flow-density relationship, and Figure 6 the flow-velocity relationship. The curves shown are based on the data used in the numerical example.

ASSUMPTION OF EQUAL AVERAGE TRAVEL TIME

It is assumed that a flow rate q_B of basic vehicles only will produce the same average travel time $t(q)_B$ as is produced by a flow rate q_M of mixed vehicles, so that $t(q)_B = t(q)_M$.

For any given length of roadway, this results in equal average velocities for the two traffic streams, so that, as shown in Figure 7, $u_B = u_M = u$. If we recall Equation 1 and note that $q_B = k_B u$ and $q_M = k_M u$, it follows that $q_B/q_M = k_B/k_M$. By similar triangles in Figure 7, $k_B = k_{jB}(u_{FB} - u)/u_{FB}$ and $k_M = k_{jM}(u_{FM} - u)/u_{FM}$, so that the following holds:

$$q_B/q_M = k_B/k_M = [(u_{FM}/u_{FB})(k_{jB}/k_{jM})][(u_{FB}-u)/(u_{FM}-u)] \qquad (11)$$

Case 1 is the general case where $u_{FM} < u_{FB}$; $k_{jM} < k_{jB}$, so from Equation 1:

$$PCE = (1/p)\left\{ [(u_{FM}/u_{FB})(k_{jB}/k_{jM})][(u_{FB}-u)/(u_{FM}-u)] - 1\right\} + 1 \qquad (12)$$

Case 2 is the special case where the truck is longer than the basic vehicle but has the same free-flow speed, so that $u_{FM} = u_{FB}$; $k_{jM} < k_{jB}$ and

$$PCE = (1/p)[(k_{jB}/k_{jM}) - 1] + 1 \qquad (13)$$

Equation 13 can be further reduced by substituting for the values of k_{jM} and k_{jB} given in Equations 10a and 10b and further noting that $p' = p$:

$$k_{jB}/k_{jM} = (L/L_B)[p L_T + (1-p)L_B]/L = (p L_T/L_B) + 1 - p$$

Substitution into Equation 13 gives the final result:

$$PCE = L_T/L_B \qquad (14)$$

which is a constant value over all ranges of p and for all volumes.

Case 3 is the special case where the truck is the same length as the basic vehicle but has a lesser free-flow velocity, $u_{FT} < u_{FB}$; $k_{jB} = k_{jT}$:

$$PCE = (1/p)\left\{ (u_{FM}/u_{FB})[(u_{FB}-u)/(u_{FM}-u)] - 1 \right\} + 1 \qquad (15)$$

By inspection of Figure 7 and Equation 12 it will be seen that the PCE is undefined for $u > u_{FM}$. For $u < u_{FM}$ and by using the Greenshields model of traffic flow,

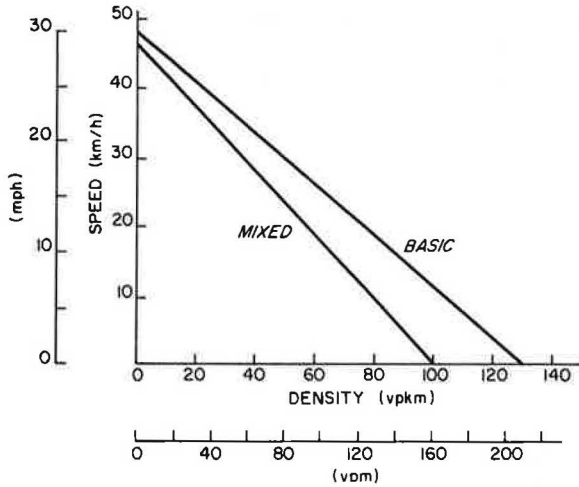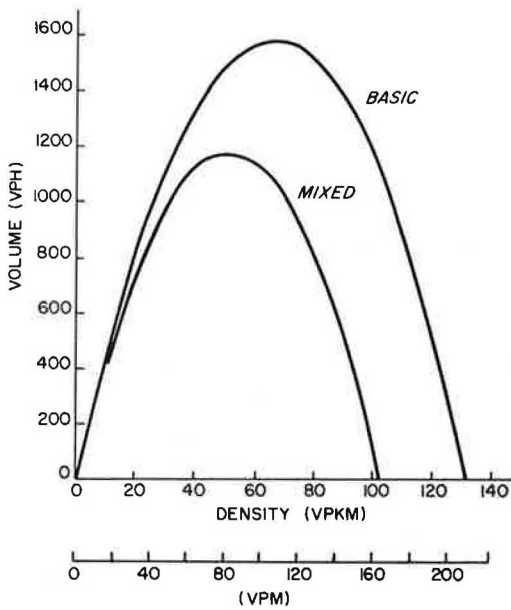Figure 4. Speed-density relationship: numerical example.

Figure 6. Flow-speed relationship: numerical example.

Figure 5. Flow-density relationship: numerical example.
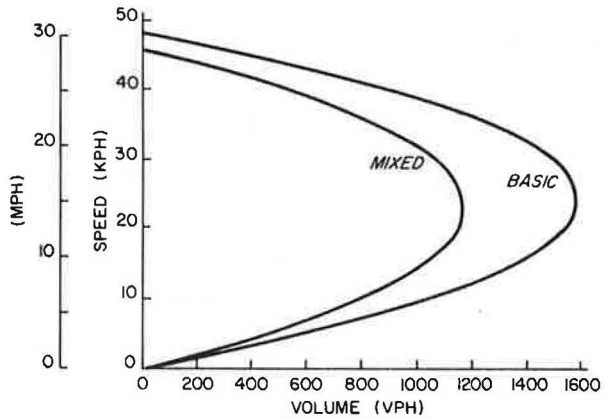
Figure 7. Determination of PCE-values by equal travel time.

$$u_B = u_M = (u_{FB}/2) [1 + (1 - y)^{1/2}] \qquad (16)$$

where $y = q_B/q_{OB}$ and other terms are as previously defined.

A numerical example incorporating the previously calculated data follows. Recall the earlier example of a traffic stream with 10 percent trucks with the following characteristics:

$u_{FM} = 45.98$ km/h (28.57 miles/h),
$k_{jM} = 102.07$ vehicles/km (164.27 vehicles/mile),
$q_{OM} = 1173.33$ vehicles/h,
$u_{FB} = 48.28$ km/h (30.00 miles/h),
$k_{jB} = 131.23$ vehicles/km (211.20 vehicles/mile), and
$q_{OB} = 1584.00$ vehicles/h.

Since $u_{FM} < u_{FB}$ and $k_{jM} < k_{jB}$, Equation 12 will apply. Consider the PCE-value when the flow in basic vehicles is 600 vehicles/h, $y = 600/1584 = 0.379$, and, by Equation 16, $u_B = u_M = 43.17$ km/h (26.82 miles/h).

If we substitute in Equation 12,

$PCE = (1/0.10)\{[(45.98/48.28)(131.23/102.07)] \times [(48.28 - 43.17)/(45.98 - 43.17)] - 1\} + 1 = 13.247$.

Each truck is equivalent to 13.247 basic vehicles if we are to satisfy the criterion that $u_B = u_M$. This is equivalent to substituting 269.7 mixed vehicles q_M (of which 10 percent are trucks) for 600.0 basic vehicles q_B, as demonstrated below:

$0.9q_M + 0.1q_M (13.247) = 600$,
$2.2247q_M = 600$,
$q_M = 269.7$.

Of particular interest is the PCE-value associated with low volumes on the mixed curve so that the mean speeds $u_B = u_M$ approach the free-flow speed of the mixed curve u_{FM}. For instance, at a flow rate q_M of 10 vehicles/h (10 percent trucks), $u_B = u_M = 45.88$ km/h. The equivalent flow rate on the basic-vehicle-only curve $q_B = 298.98$ vehicles/h, so that

$PCE = (1/0.10)[(298.98/10) - 1] + 1 = 289.98$.

At low volumes, the value of the PCE is at a maximum and decreases as the volume of basic vehicles increases. This pattern is shown in Figure 8 for a traffic flow with 10 percent trucks. Also shown in Figure 8 are PCE-values associated with 1

Figure 8. PCE versus volume by percentage of trucks, equal travel time.
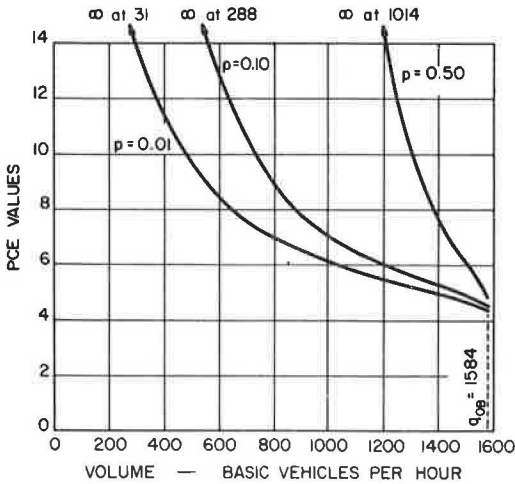
Figure 9. Determination of PCE-values by equal total travel time.
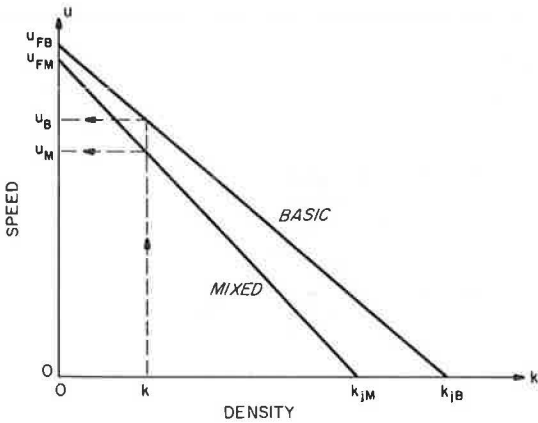
percent and 50 percent trucks in the traffic stream. As the percentage of trucks increases, the value of the PCE associated with a given volume of basic vehicles also increases as long as the free-flow velocity of trucks (u_{FT}) is less than the free-flow velocity of basic vehicles (u_{FB}).

ASSUMPTION OF EQUAL TOTAL TRAVEL TIME

It is assumed that a flow rate q_B of basic vehicles only will produce the same total travel time $T(q)_B$ as is produced by a flow rate q_M of mixed vehicles, so that $T(q)_B = T(q)_M$.

For any given length of roadway, this is equivalent to equal vehicle hours of occupancy per hour for the two traffic streams, and since $T(q)$ is numerically equal to density k, $k_B = k_M = k$, as shown in Figure 9. If we recall Equation 1 and note that $q_B = ku_B$ and $q_M = ku_M$, it follows that $q_B/q_M = u_B/u_M$.

By similar triangles in Figure 9,

$$u_B = u_{FB} (k_{jB} - k)/k_{jB}$$

and

$$u_M = u_{FM} (k_{jM} - k)/k_{jM}$$

so that

$$q_B/q_M = u_B/u_M = [(u_{FB}/u_{FM})(k_{jM}/k_{jB})] \ [(k_{jB} - k)/(k_{jM} - k)] \tag{17}$$

Case 1 is the general case where $u_{FM} < u_{FB}$; $k_{jM} < k_{jB}$, so from Equation 1:

$$PCE = (1/p)\left\{ [(u_{FB}/u_{FM})(k_{jM}/k_{jB})] \ [(k_{jB} - k) \right.$$
$$\left. \div (k_{jM} - k)] - 1 \right\} + 1 \tag{18}$$

Case 2 is the special case where the truck is longer than the basic vehicle but has the same free-flow speed, so $u_{FM} = u_{FB}$; $k_{jM} < k_{jB}$ and

$$PCE = (1/p)\left\{ (k_{jM}/k_{jB}) \ [(k_{jB} - k)/(k_{jM} - k)] - 1 \right\} + 1 \tag{19}$$

Case 3 is the special case where the truck is the same length as the basic vehicle but has a lesser free-flow velocity, $u_{FT} < u_{FB}$; $k_{jB} = k_{jT}$:

$$PCE = (1/p) \ [(u_{FB}/u_{FM}) - 1] + 1 \tag{20}$$

Equation 20 can be further reduced by substituting the value of u_{FM} from Equation 6 into Equation 20:

$$u_{FB}/u_{FM} = u_{FB} \left\{ (p/u_{FT}) + [(1 - p)/u_{FB}] \right\} = (pu_{FB}/u_{FT})$$
$$- (pu_{FB}/u_{FB}) + 1$$

Finally,

$$PCE = (1/p) \ [(pu_{FB}/u_{FT}) - (pu_{FB}/u_{FB}) + 1 - 1] + 1$$
$$= u_{FB}/u_{FT} \tag{21}$$

which is a constant value over all ranges of p and over all volumes.

Again, the Greenshields model of traffic flow is used for the interrelationships among flow rate (q), density (k), and velocity (u):

$$k_B = k_M = k_{jB}/2 \ [1 - (1 - y)^{1/2}] \tag{22}$$

A numerical example employing the same data as were used to illustrate the model of equal average travel time follows. Since for this example, $u_{FM} < u_{FB}$ and $k_{jM} < k_{jB}$, Equation 18 will apply. With a flow rate of 600 basic vehicles/h, $y = 600/1584 = 0.379$ (as before) and by Equation 22,

$$k_B = k_M = 13.90 \text{ vehicles/km (22.37 vehicles/mile)}.$$

By substituting in Equation 18,

$$PCE = (1/0.10)\{ [(48.28/45.98)(102.07/131.23)] \text{ x}$$
$$[(131.23 - 13.90)/(102.07 - 13.90)] - 1 \} + 1$$
$$= 1.868.$$

Each truck is equivalent to 1.868 basic vehicles if we are to satisfy the criterion that $T(q)_B = T(q)_M = (k_B = k_M)$. This is equivalent to substituting 552.1 mixed vehicles, q_M (of which 10 percent are trucks), for 600.0 basic vehicles, q_B. Recall that to maintain equal average times $t(q)_B = t(q)_M = (u_B = u_M)$, it was possible to substitute only 269.7 mixed vehicles (10 percent trucks) for 600.0 basic vehicles, so that PCE = 13.247.

From Figure 9 it will be observed that PCE is undefined for $k > k_{jM}$. This is in the "backward-bending" portion of the total-travel-time curve where LOS is F (stop-and-go conditions) and is not an area in which PCE-values are of concern for steady-state flow.

The relationship between PCE-values and volume is shown in Figure 10. PCE-values are lowest at low volumes and increase gradually until the maximum

mixed-flow rate q_{OM} is attained (point A on the curve) and then increase at a greater rate as the mixed-flow curve is operating on the backward-bending portion of the total-travel-time curve.

In Figure 10 curves are also shown for 1 percent and 50 percent trucks in the traffic stream. It is only after the equivalent flow rate on the mixed-vehicle curve has exceeded the point of maximum flow (corresponding to point A on the PCE curves) that there is a marked difference in PCE-value as related to percentage of trucks in the traffic stream.

The significance of the backward-bending point can be seen by reference to Figure 11, which shows the relationship between density k (total travel time) and flow for a basic-vehicle-only curve and a mixed curve with 10 percent trucks. Points B and B' are associated with an associate travel time of 13.90 vehicle-h/km-h (22.37 vehicle-h/mile-h). A mixed flow of 552.1 or 600 vehicles/h, basic vehicles only, will give this value of total travel time, and the PCE-value is 1.868.

Points A and A' are associated with the maximum flow q_{OM} on the mixed curve. At this point the total travel time is 51.04 vehicle-h/km-h (82.13 vehicle-h/mile-h). The associated volumes are 1173.3 and 1505.8 while the PCE-value is 3.833. Beyond point A' the flow for mixed vehicles becomes stop

and go with an associate decrease in volume. The volume associated with point A is noted on the PCE-versus-volume curves of Figure 10.

Points C and C' are associated with the maximum-flow q_{OB} on the basic-vehicles-only curve. The total travel time is 65.62 vehicle-h/km-h (105.60 vehicle-h/mile-h). The associated volume of basic vehicles has increased to 1584.0 vehicles/h while the mixed-vehicle flow has decreased to 1077.6 vehicles/h and the PCE-value has increased to 5.700.

ASSUMPTION OF EQUAL AVERAGE TRAVEL TIME FOR BASIC VEHICLES

The mixed stream of traffic is made up of two component steady-state flows. The first component is basic vehicles within the mixed stream with parameters u_{FMB} ($= u_{FB}$) and k_{jMB} [$=(1 - p')k_{jM}$]. The second component is trucks within the mixed stream with parameters u_{FMT} ($= u_{FT}$) and k_{jMT} ($= p'k_{kjM}$). The speed-density curves for the mixed flow and the two component flows are shown in Figure 12.

From Figure 12 and by the Greenshields relationship,

$$u_{MB} = (u_{FMB}/2) [1 + (1 - y_{MB})^{1/2}] \qquad (23a)$$

and

$$u_{MT} = (u_{FMT}/2) [1 + (1 - y_{MT})^{1/2}] \qquad (23b)$$

where

$$y_{MB} = (1 - p)q_M/(1 - p)q_{OM} = q_M/q_{OM} = y_{MT}$$

Speeds calculated from the traffic stream with 10 percent trucks as used in earlier examples are presented in Table 1 along with flow rates and PCE-values. The top section is based on the assumption that PCE is defined by two traffic flow rates, q_B and q_M, such that the mean velocities of individual vehicles within the traffic streams are equal ($u_B = u_M$). It will be observed that the speed of basic vehicles within the mixed stream (u_{MB}) is greater than the speed of the basic vehicles in the basic stream (u_B).

The middle part of Table 1 is based on the assumption that PCE is defined by two traffic flow rates, q_B and q_M, such that the total travel

Figure 10. PCE versus volume by percentage of trucks, equal total travel time.

Figure 11. Calculation of PCE-values with backward-bending curves.

Figure 12. Speed-density relationships for mixed flow and two component flows.
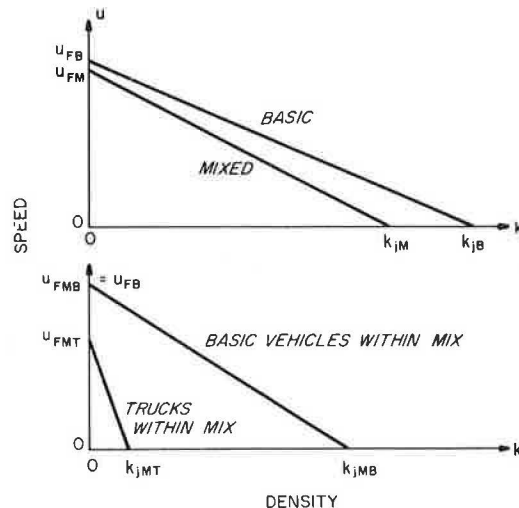
Table 1. Speed of vehicles within mixed flow.

q_B	u_B (km/h)	q_M	u_M (km/h)	u_{MB} (km/h)	u_{MT} (km/h)	PCE
Assumption of Equal Individual Travel Time						
288	45.98	0.56	45.98	48.28	32.19	5120.96
800	41.12	443.46	41.12	43.18	28.79	9.04
1584	24.14	1170.40	24.14	25.35	16.90	4.53
Assumption of Equal Total Travel Time						
288	45.98	270.36	43.16	45.32	30.21	1.65
800	41.12	724.02	37.22	39.07	26.06	2.05
1505	29.53	1173.33	23.03	24.17	16.11	3.83
Equal Individual Travel Time for Basic Vehicles						
288	45.98	213.33	43.79	45.98	30.66	4.50
800	41.12	592.59	39.17	41.12	27.42	4.50
1584	24.14	1173.33	22.99	24.14	16.09	4.50

Note: It is assumed that $L_T = 22.86$ m, $L_B = 7.62$ m, $p = 0.10$, $u_{FT} = 32.187$ km/h, and $u_{FB} = 48.280$ km/h. 1 km/h = 0.6 mile/h.

Figure 13. PCE-values by equal basic vehicle travel time.
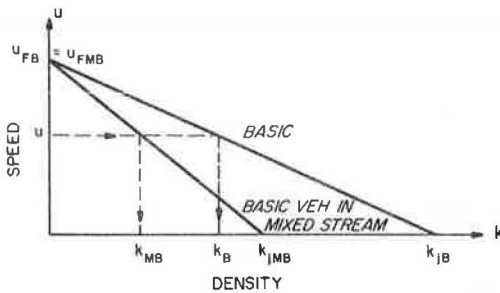
times per kilometer are equal. Eight hundred basic vehicles per hour will require 19.45 vehicle-h to traverse 1 km (0.6 mile) of roadway [800 x (1/41.12)], as will 724.02 mixed vehicles [724.02 x (1/37.22)]. The speed of basic vehicles within the mixed stream (u_{MB}) is always less than the speed of vehicles in the basic stream, and the difference increases with increasing volume (and increasing PCE-values).

St. John (7) has used mean speed of basic vehicles as the criterion for determining PCE-values. A diagram of the situation is shown in Figure 13 where $u_{FMB} = u_{FB}$ and $u_B = u_{MB} = u$. By Equation 1,

$$PCE = (1/p)[(q_B/q_M) - 1] + 1 \qquad (1)$$

where

$$q_M = q_{MB}/(1-p) = uk_{MB}/(1-p); q_B = uk_B$$

so that

$$q_M/q_B = (1-p)k_B/k_{MB}$$

By similar triangles in Figure 13,

$$k_{MB} = k_{jMB}(u_{FB} - u)/u_{FB}; k_B = k_{jB}(u_{FB} - u)/u_{FB}$$

so that

$$q_M/q_B = (1-p)k_{jB}/k_{jMB}$$

But $k_{jMB} = (1 - p')k_{jM}$ and

$$q_M/q_B = [(1 - p)/(1 - p')](k_{jB}/k_{jM}) \qquad (24)$$

By Equations 10a and 10b,

$$k_{jB}/k_{jM} = (L/L_B)\{[p'L_T + (1 - p')L_B]/L\} = [p'L_T + (1 - p')L_B]/L_B$$

Substitution in Equation 24 gives

$$q_B/q_M = (1 - p)\{[p'/(1 - p')](L_T + L_B)\}/L_B \qquad (25)$$

From Equation 9,

$$p'/(1 - p') = \{pu_{FB}/[pu_{FB} + (1 - p)u_{FT}]\}\{[pu_{FB} + (1 - p)u_{FT}]$$
$$\div (1 - p)u_{FT}\} = pu_{FB}/(1 - p)u_{FT}$$

Substituting in Equation 25,

$$q_B/q_M = [(pu_{FB}/u_{FT})(L_T/L_B)] + (1 - p)$$

so that by Equation 1,

$$PCE = (1/p)[(q_B/q_M) - 1] + 1 = (u_{FB}/u_{FT})(L_T/L_B) \qquad (26)$$

which is a constant value over all ranges of p and for all volumes.

Recall the numerical example previously cited in which the basic vehicle had an effective length of 7.62 m and free-flow velocity of 48.280 km/h and the truck was 22.86 m long and had 32.187 km/h velocity.

By Equation 26, the resulting PCE is calculated as follows:

$$(48.280/32.187) \times (22.86/7.62) = 4.50.$$

The speeds of vehicles within the traffic stream for these assumptions are shown in the bottom part of Table 1. Although the average velocity of all vehicles in the mixed flow (u_M) is less than the speed of basic vehicles only, the speed of the basic vehicles within the mixed flow (u_{MB}) is equal to u_B at all volume levels shown.

INFLUENCE OF VEHICLE SIZE AND SPEED ON CALCULATED PCE-VALUES

Table 2 contains a comparison of PCE-values for vehicles varying in effective length from 6.10 m (20 ft) to 22.86 m and free-flow velocity from 32.19 km/h to 56.33 km/h (35 miles/h). It will be recalled that the basic vehicle had an effective length of 7.62 m and a free-flow velocity of 48.28 km/h. There are 10 percent nonbasic vehicles in the mixed traffic stream and the flow of basic vehicles is 600 vehicles/h. The criterion for LOS is given as average travel time, so that $t(q)_B = t(q)_M$ ($u_B = u_M$).

Two entries are shown for each vehicle. The first is the PCE-value; the second is the volume of mixed vehicles that will produce the same mean travel time (or the reciprocal velocity) as is produced by 600 basic vehicles (V). As the size of the vehicle increases and the velocity decreases, the PCE-value increases; values range from a minimum of -0.23 for a short, fast vehicle to a maximum of 13.25 for a long, slow vehicle.

Of particular interest are the negative values of PCE in columns 1 and 3. Recall that the PCE-value represents the number of basic vehicles that are displaced by a nonbasic vehicle. For example, the truck 18.29 m (60 ft) long with a free-flow velocity of 40.23 km/h (25 miles/h) has a PCE-value of 5.01. The mixed traffic stream of 428.2 vehicles/h has the same mean travel time as 600 basic vehicles. The mixed stream contains 42.82 trucks and 385.38 basic vehicles. The 42.82 trucks have displaced 214.62 (600.00 - 385.3) basic vehicles.

Table 2. PCE-values by vehicle length and free-flow velocity, average travel time.

| Free-Flow Velocity (km/h) | Effective Length of Truck (m) | | | | | | | | | | | |
| | 6.10 | | 7.62 | | 9.14 | | 13.72 | | 18.29 | | 22.86 | |
	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V
56.33	-0.23	684.3	-0.08	672.4	0.08	660.9	0.54	628.6	1.01	599.4	1.48	572.8
48.28	0.80	612.2	1.00	600.0	1.20	588.2	1.80	555.6	2.40	526.3	3.00	500.0
40.23	2.75	510.7	3.03	498.7	3.31	487.2	4.16	455.8	5.01	428.2	5.86	403.7
32.19	7.81	357.0	8.30	346.8	8.80	337.1	10.28	311.2	11.76	289.0	13.25	269.7

Notes: It is assumed that $p = 0.10$, $q_B = 600$ vehicles/h; basic vehicle has effective length of 7.62 m and free-flow velocity of 48.28 km/h. 1 km/h = 0.6 mile/h; 1 m = 3.2 ft.
V = volume of mixed vehicles that will produce same mean travel time (or reciprocal velocity) as produced by 600 basic vehicles.

Table 3. PCE-values by vehicle length and free-flow velocity, equal total travel time.

| Free-Flow Velocity (km/h) | Effective Length of Truck (m) | | | | | | | | | | | |
| | 6.10 | | 7.62 | | 9.14 | | 13.72 | | 18.29 | | 22.86 | |
	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V
56.33	0.84	609.9	0.86	608.7	0.88	607.4	0.94	603.7	1.00	599.9	1.06	596.2
48.28	0.98	601.4	1.00	600.0	1.02	598.6	1.10	594.3	1.17	590.0	1.24	585.8
40.23	1.17	589.9	1.20	588.2	1.23	586.6	1.32	581.7	1.40	576.8	1.49	571.8
32.19	1.46	573.4	1.50	571.4	1.54	569.5	1.64	563.7	1.75	557.9	1.87	552.1

Notes: It is assumed that $p = 0.10$, $q_B = 600$ vehicles/h; basic vehicle has effective length of 7.62 m and free-flow velocity of 48.28 km/h. 1 km/h = 0.6 mile/h; 1 m = 3.2 ft.
V = volume of mixed vehicles that will produce same mean travel time (or reciprocal velocity) as produced by 600 basic vehicles.

Table 4. PCE-values by vehicle length and free-flow velocity, average travel time.

| Free-Flow Velocity (km/h) | Effective Length of Truck (m) | | | | | | | | | | | |
| | 6.10 | | 7.62 | | 9.14 | | 13.72 | | 18.29 | | 22.86 | |
	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V	PCE-Value	V
56.33	0.69	619.5	0.86	608.7	1.03	598.3	1.54	569.1	2.06	542.6	2.57	518.5
48.28	0.80	612.2	1.00	600.0	1.20	588.2	1.80	555.6	2.40	526.3	3.00	500.0
40.23	0.96	602.4	1.20	588.2	1.44	574.7	2.16	537.6	2.88	505.1	3.60	476.2
32.19	1.20	588.2	1.50	571.4	1.80	555.6	2.70	512.8	3.60	476.2	4.50	444.4

Notes: It is assumed that p and q_B are variable; basic vehicle has effective length of 7.62 m and free-flow velocity of 48.28 km/h. 1 km/h = 0.6 mile/h; m = 3.2 ft.
V = volume of mixed vehicles that will produce same mean travel time (or reciprocal velocity) as produced by 600 basic vehicles.

Conversely, the short (6.10-m) vehicle with a free-flow velocity of 56.33 km/h has a mixed traffic flow rate of 684.3. Of these vehicles, 68.43 are fast vehicles and 615.87 basic vehicles. The presence of the 68.43 fast vehicles decreases the average travel time sufficiently that an extra 15.87 basic vehicles are added to the traffic (rather than displaced); hence the PCE sign is negative. The ratio of added basic vehicles to nonbasic vehicles is 0.23 (15.87/68.43). If the number of basic vehicles in the mixed stream, $(1 - p)q_M$, is greater than the number of basic vehicles, q_B, the PCE-value will be negative.

Table 3 is similar to Table 2 except that the criterion for comparable LOS values is that the total travel time for all vehicles over a length of highway be equal for the two traffic streams, $T(q)_B = T(q)_M$ ($k_B = k_M$). The range of PCE-values is reduced but, again, increased PCE-values are associated with longer, slower vehicles. There are no negative PCE-values in Table 3, although, as is to be expected, there are PCE-values less than 1.0.

Table 4 is based on the criterion that $u_B = u_{MB}$. The PCE-values would be the same over all values of q_B, but the value of q_M in each cell is for the value $q_B = 600.0$.

DISCUSSION OF RESULTS

Three criteria for defining LOS have been considered. The first of these assumes equal mean travel time for two flows, mixed and basic vehicles. The introduction of slow-moving vehicles into the traffic stream will reduce average speed even at very low flow rates. A substantial number of basic vehicles can be expected to produce the same average speed. As a consequence, (a) PCE-values are undefined at very low volumes, and (b) PCE-values decrease as volumes increase.

It would be desirable to define PCE-values over all values of volume. It would also appear reasonable that at low volumes large, slow vehicles would have a minimum effect on traffic flow but that as volume increased, there would be greater interaction between vehicles, so the PCE-values should be increasing as volumes increase. Letting LOS be defined by $u_B = u_{MB}$ gives constant PCE-values over all volumes.

The definition of LOS by the criterion of total travel time satisfies the difficulties noted above and is recommended as a more desirable approach. In making this recommendation it should be recalled that the traffic flow model used (Greenshields') is a simple deterministic model that only partly repre-

sents observed traffic flow. Only two categories of
vehicles have been considered; in reality an analy-
sis of PCE-values should be based on a model that
considers three or more vehicle categories simul-
taneously.

Despite the shortcomings noted above, it is felt
that this analysis will provide some guidance to
those doing research to determine PCE-values.

ACKNOWLEDGMENT

This paper was prepared while I was with the Traffic
Systems Division, Office of Research and Develop-
ment, of the Federal Highway Administration on sab-
batical from the University of Minnesota. I am in-
debted to the Regents of the University of Minnesota
for making the leave possible and to the Federal
Highway Administration for the opportunity to engage
in the research program of that organization.

Discussion

A.D. St. John

Huber has performed a timely service by examining
alternative bases for the passenger-car equivalent
(PCE). The analyses detect undesirable characteris-
tics to be avoided and mandate a careful examination
of definitions and concepts. This discussion pre-
sents alternative points of view, describes a few of
the results from microscopic simulation models, and
suggests additional factors to be considered in
selecting bases for the PCE.

At the most fundamental level, Huber adheres to
past practice. The ideal or reference vehicle mix
is 100 percent passenger cars, and speed or its in-
verse, travel time, are examined in several forms as
best bases for equivalence between moving traffic
streams.

For the mixed traffic streams the paper analyzes
speed in two forms, which differ significantly. In
one form, the average speed is calculated with data
for all vehicles in the stream; the other form uses
speed data for passenger cars only. The paper se-
lects the all-vehicle form as preferable because of
the PCE characteristics estimated for it. I believe
that the choice here should depend more strongly on
the concept desired for equivalence between traffic
streams. For example, if the all-vehicle form is
used, a steep sustained upgrade would be calculated
to reduce service for a car-and-truck mix even if
the car speeds were not appreciably depressed by the
presence of low-speed trucks. Alternatively, if the
car-only form is used, there is no direct measure of
the speed depression exprienced by trucks due to the
upgrade.

Several facets of the 1965 Highway Capacity Man-
ual (1) suggest that car-only data were meant to be
used for the mixed stream. Operating speed must
have been selected for its sensitivity at low flow
rates, a sensitivity that would be distorted by in-
cluding trucks. Also, the definition implies that
operating speed is limited only by highway design
speed and interactions with other vehicles. However,
a passenger car is specified as the vehicle type
only in the definition for free-flow operating speed.

I agree with Huber's requirements that the defi-
nition for PCE and the PCE-values be well behaved
over the range of variables. It is also preferable
that the PCE-values not conflict with engineering
intuition. However, I question Huber's preference

that the PCE for a truck be small at low flow rates
and increase with flow rate when the percentage of
trucks is held constant.

A PCE that is essentially constant with flow rate
is desirable for two reasons. First, the constant
PCE economizes the field or model data needed and,
second, constant PCE implies fundamental relation-
ships that do not change in form between the car-
only and mixed flows. As an example of the latter
aspect, consider Figure 14, in which the operating
speed versus flow rate is sketched for a mixed flow
and for a flow of cars only. If the form of the
operating-speed function is constant, the curves of
Figure 14 will merge into the single curve of Figure
15 when the abscissa is normalized to volume/capac-
ity. The single, normalized curve is obtained only
if the PCE is essentially constant over flow rate.
The preservation of form in the single normalized
curve is convenient computationally and conceptually.

For currently defined PCE estimates there is evi-
dence both supporting and conflicting with the idea
that PCE is constant over flow rate. The tables in
the 1965 Highway Capacity Manual indicate only small
changes in PCE between high and low service levels.
An extensive collection of results from a micro-
scopic model of multilane flow (8,9) conformed with
normalized curves exemplifed in Figure 15. Although
PCE-values were not derived, normalized organization
indicated that they would be essentially constant
over flow rate. More recent results with the same
model (10) employed the free speeds observed under
the 55-mile/h (89-km/h) speed limit. These results
reveal cases in which PCE-values would diminish at
high flow rates; that is, speeds of cars are
depressed by impeding vehicles at low and intermedi-
ate flow rates (PCE > 1) but the approach to ca-
pacity is essentially unaffected (PCE $\approx$ 1).

Two considerations, absent in the paper, are
given here to complete the set of important con-

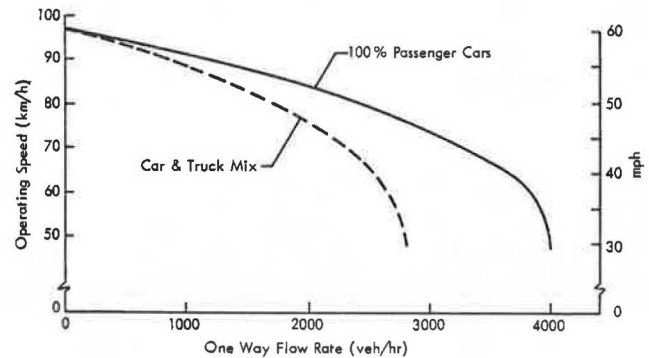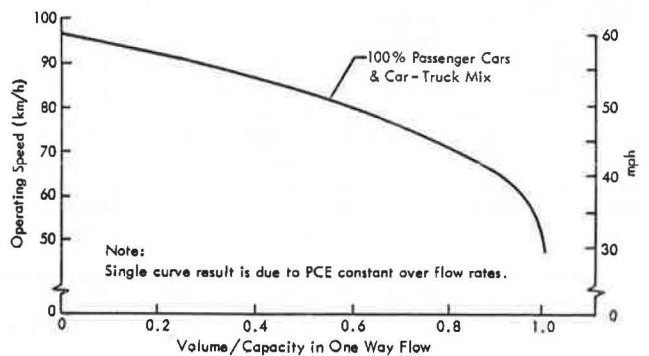Figure 14. Operating speed versus flow rate.

Figure 15. Operating speed versus volume/capacity ratio.

siderations. They are the field measurements and the nonlinear dependence of PCE-values on percentage of trucks.

The bases for the PCE should lead to field measurements that are feasible and economical. Operating speed exemplifies a measure that has desirable characteristics in equivalency but is difficult and expensive to measure in the field. Because it is not a central (average) measure, operating speed is at a disadvantage even in the analysis of results from models.

The nonlinear relationship between the PCE and percentage of trucks indicated by most available data constitutes a problem. Tables in the 1965 Highway Capacity Manual (1) and results from models (8,9,11) indicate that the PCE diminishes as the percentage of trucks increases. This is more than a computational inconvenience; it increases the field or model data required and complicates estimates for two or more types of impeding vehicles. A derivation (11) has been partly successful in establishing an analytical relationship between PCE and percentage of trucks. The results provide a measure that is invariant with percentage of trucks and a method for estimating the effects of two or more types of impeding vehicles. These possibilities should be considered in selecting the basis for PCE-values.

In summary, I agree with the aims of Huber's paper and with part of the conclusions. I suggest that more attention be directed to the fundamental concepts of equivalence, that it is desirable for the PCE to be constant over flow rate, and that additional considerations should be included in the selection of bases for equivalency. Also, final decisions should be based on extensive field data or results from comprehensive models. This does not detract from the value of simple models, however, since they frequently point to alternatives otherwise overlooked.

Randy Machemehl

The work by Huber is a significant contribution toward better characterization of the effects of commercial vehicles on traffic flow. The mathematical relationships developed through his work are understandable and reasonable. When viewed as a conceptual framework for PCE computation, the work is both interesting and useful. Huber notes several key weaknesses in his recommendations for further study.

These include the fact that only two vehicle categories--basic and nonbasic--are considered. Operational characteristics of the entire vehicle population and resulting effects on the traffic stream are known to vary widely. Recreational vehicles, for example, represent a vehicle class that generally has significantly different operational capabilities, and drivers of such vehicles often lack experience required to fully utilize the capabilities their vehicles possess. Variability in operational features of intracity-type commercial delivery vehicles versus over-the-road commercial trucks, local buses versus intercity commercial buses, and even small underpowered versus large passenger cars represents additional examples of the need for more than two vehicle classes. Huber notes that the Federal Highway Administration, as part of the national highway cost-allocation study, is in the process of developing PCE-values for as many as 15 categories of vehicles for four different roadway

classes. This effort should answer many questions regarding operational effects of the spectrum of vehicle classes.

The Greenshields model of traffic flow, which uses a linear relationship between density and velocity, is employed by Huber to develop interrelationships among speed, density, and flow rate. This model probably is a good approximation for many density-velocity conditions. However, it may not be completely sufficient for the entire velocity-density range. A nonlinear relationship might very likely improve the validity of PCE estimates near boundary values. The deterministic nature of this model is also problematic. As noted earlier, tremendous variability among the driver-vehicle population and their performance on various roadway types is indicative of the need for stochastic modeling. Characterization of this variability and its effect on traffic flow would likely be achieved most efficiently through use of stochastic representation.

The analyses presented in Huber's paper represent a contribution that will be of value to those investigating PCE techniques. It is both direct and easily understood and Huber is to be commended for his efforts.

Author's Closure

I am indebted to Machemehl and St. John for their thoughtful reading and discussion of this paper.

Machemehl is correct in pointing out that the modeling should be extended to permit analysis of several (more than two) categories of vehicles within the mixed stream. Preliminary investigation has shown that this is possible if one continues to use the Greenshields model of traffic flow. Incorporation of stochastic modeling would require extensive computer simulation but should provide further insight.

As St. John has noted, one can only make intuitive analyses in establishing a basis for determination of PCE-values as discussed in the 1965 Highway Capacity Manual. This paper represents an attempt to examine the alternative bases for PCE-values.

My preference that the PCE for a truck be small at low flow rates and increase with increased flow rates is predicated on the intuitive feeling that at low volumes there are few basic vehicles that can be influenced by a truck; the greater time and distance spacing tends to minimize intervehicular interference. As the flow rate increases, the opportunity for interaction between basic vehicles and trucks is increased with a subsequent increase in PCE-values.

I agree with St. John in his preference for a constant PCE-value because this will economize data requirements. His second contention, that a constant PCE implies fundamental relationships that do not change in form, is illustrated by reference to Figure 14. Implicit to this relationship is a common free-flow speed at extremely low flow rates. My analysis differs by considering that the car and truck mix has a lesser free-flow speed (because of the slower trucks) as illustrated in Figure 6. It is only when one uses as a basis $t(q)_{MB} = t(q)_B$ that the analysis will be similar to that suggested by St. John. (Conversion of the form given in Figure 13 to that given in Figure 14 is straightforward.)

Again, I thank both discussants for their helpful comments.

REFERENCES

1. Highway Capacity Manual. TRB, Special Rept. 87, 1965.
2. A. Werner and J.F. Morrall. Passenger Car Equivalencies of Trucks, Buses, and Recreational Vehicles for Two-Lane Rural Highways. TRB, Transportation Research Record 615, 1976, pp. 10–17.
3. A. Werner, J.F. Morrall, and G. Halls. Effect of Recreational Vehicles on Highway Capacity. Traffic Engineering, Vol. 45, No. 5, May 1975, pp. 20–25.
4. First Progress Report on the Federal Highway Cost Allocation Study. FHWA, March 1980.
5. Passenger Car Equivalence Analysis Framework: An Interim Report as Part of Quality of Flow and Determination of PCE on Urban Arterials. PRC Voorhees, McLean, VA, Jan. 1981.
6. B.D. Greenshields. A Study of Traffic Capacity. HRB Proc., Vol. 14, 1934, pp. 448–477.
7. A.D. St. John. Nonlinear Truck Factor for Two-Lane Highways. TRB, Transportation Research Record 615, 1976, pp. 49–53.
8. A.D. St. John, D.R. Kobett, and others. Traffic Simulation of the Design of Uniform Service Roads in Mountainous Terrain. Midwest Research Institute, Kansas City, MO, MRI Project 3029-E, Final Rept., Jan. 1970.
9. A.D. St. John, D.R. Kobett, and others. Freeway Design and Control Strategies as Affected by Trucks and Traffic Regulations. FHWA, Rept. FHWA-RD-75-42, Jan. 1975.
10. A.D. St. John and W.D. Glauz. Implications of Light-Weight, Low-Powered Future Vehicles in the Traffic Stream--Volume 2. Midwest Research Institute, Kansas City, MO, MRI Project 4562-S, Final Rept., Oct. 1981.
11. A.D. St. John and D.R. Kobett. Grade Effects on Traffic Flow Stability and Capacity. NCHRP, Rept. 185, 1978.

Model for Calculating Safe Passing Distances on Two-Lane Rural Roads

EDWARD B. LIEBERMAN

A model describing the kinematics of vehicle trajectories during the passing maneuver on two-lane rural roads is presented. This model is based on the hypothesis that there exists a point in the passing maneuver that can be identified as a critical position. At this point, the decision to complete the passing maneuver will provide the same factor of safety relative to an oncoming vehicle as will the decision to abort the maneuver. The model locates the critical position in terms of exogenous parameters. The results of a series of sensitivity studies conducted with the model are also presented. These results provide insight into those parameters that strongly influence the required sight distances. It is shown that the current sight-distance specifications of the American Association of State Highway and Transportation Officials may be inadequate from a safety standpoint, particularly for high-speed passing maneuvers and for passing vehicles that are low-powered subcompacts.

The calculation of passing sight distance as presented in the Blue Book of the American Association of State Highway and Transportation Officials (AASHTO) (1) is based on several simplifying assumptions. In this paper we examine the kinematics of the passing maneuver in greater detail and offer another point of view. The results obtained with this new model are compared with those detailed in the Blue Book (1); the implications of these comparisons are then discussed.

The benefits of an analytical model describing the passing maneuver on two-lane rural roads include the ability to identify those factors that play a role in determining safe passing sight distances. Furthermore, it is possible to conduct sensitivity studies to determine which of these factors are important relative to the others.

With the changing composition of the traffic stream--larger, faster, more powerful trucks mixing with smaller, lower, less powerful automobiles--such a model can be very useful in assessing the associated changes in safety margins provided by current sight-distance standards. It would also be possible to determine whether there is a need for changes in these standards or whether more positive forms of control are required to improve the safety characteristics of two-lane rural roads. Clearly, any change in these standards could also affect rural road capacity as well as operating speed.

OVERALL APPROACH

When a vehicle traveling on a rural road desires to pass an impeding vehicle, the driver must assess a large number of factors in deciding whether to attempt a passing maneuver. This assessment is a continuous one that extends, after the initial decision is made, throughout the passing maneuver.

This model is based on the hypothesis that there exists a point in the passing maneuver that can be identified as a critical position whenever an oncoming vehicle is in view. This critical position is defined as follows: At the critical position, the decision by the passing vehicle to complete the pass will afford it the same clearance relative to the oncoming vehicle as will the decision to abort the pass.

This implies that if a decision to abort the pass takes place downstream of the critical position (i.e., later in the passing maneuver), the clearance (and therefore the safety factor) relative to the oncoming vehicle will be less than if the passing vehicle completes the pass. The converse applies to a decision to complete the pass if made upstream of the critical position.

The determination of this critical position is a central issue in the development of the model. The

critical position is defined in terms of the longitudinal separation between the passing vehicle in the passing lane and the impeder vehicle in the normal lane.

It should be noted that this hypothesis applies even in the absence of an oncoming vehicle. In this case, the oncoming vehicle is replaced by the need for the passing vehicle to return to its normal lane at the terminus of the passing zone or by the possibility of the sudden appearance of an oncoming vehicle if the sight distance is limited. This paper addresses only the condition where the oncoming vehicle is in view; these other cases can be easily represented by suitable modification of the model.

Given the definition of the critical position, the approach taken is to consider that the complete/abort decision is made at this point in the maneuver. (This implies that the decision processes of motorists are accurate--an optimistic assumption.) On this basis, it is possible to locate the critical position for any combination of vehicle operating conditions and to determine the required sight distances.

Figure 1 is a schematic of the passing maneuver from the instant the critical position is attained (a) until the pass is completed (b) or aborted (c). A glossary of all terms used in Figure 1 and in the model formulation is given below:

$\bar{A}$ = average acceleration by passer vehicle to increase its speed from V to V + m (m/s^2),

A_{max} = maximum acceleration achievable at zero speed (m/s^2),

a = design value of abort maneuver deceleration $[m/(s \cdot s)]$,

C = clearance between passing and oncoming vehicles at completion of successful passing maneuver (m),

$\hat{C}$ = clearance between passing and oncoming ve-

hicles at completion of aborted passing maneuver (m),

D_1 = distance traveled by passer from start of passing maneuver to critical position (m),

$\hat{D}_1$ = distance traveled by impeder vehicle while passer is moving to critical position (m),

D_2 = distance traveled by impeder vehicle during successful passing maneuver from time critical position established (m),

$\hat{D}_2$ = distance traveled by impeder vehicle during aborted passing maneuver from time critical position established (m),

D_3 = distance traveled by passing vehicle from critical position to its return to original lane during its successful passing maneuver (m),

$\hat{D}_3$ = distance traveled by passing vehicle from critical position to its return to original lane during its aborted passing maneuver (m),

$d_{\bar{A}}$ = distance traversed by passer vehicle while accelerating (m),

G' = space headway between impeder and passing vehicles at start of passing maneuver (m),

G = space headway between passer and impeder vehicles at instant passer returns to normal lane (m) (in general, $G \neq G'$),

m = V_p - V = speed difference, passing versus impeder vehicles, at critical position (m/s),

S_c = sight distance (to oncoming vehicle) when passing vehicle is at critical position (m),

S_o = distance traveled by oncoming vehicle from time critical position is attained to end of passing maneuver (m),

T_1 = travel time from start of passing maneuver to attainment of critical position (s),

t = time for passing vehicle to return to its own lane from its critical position for completed passing maneuver (s),
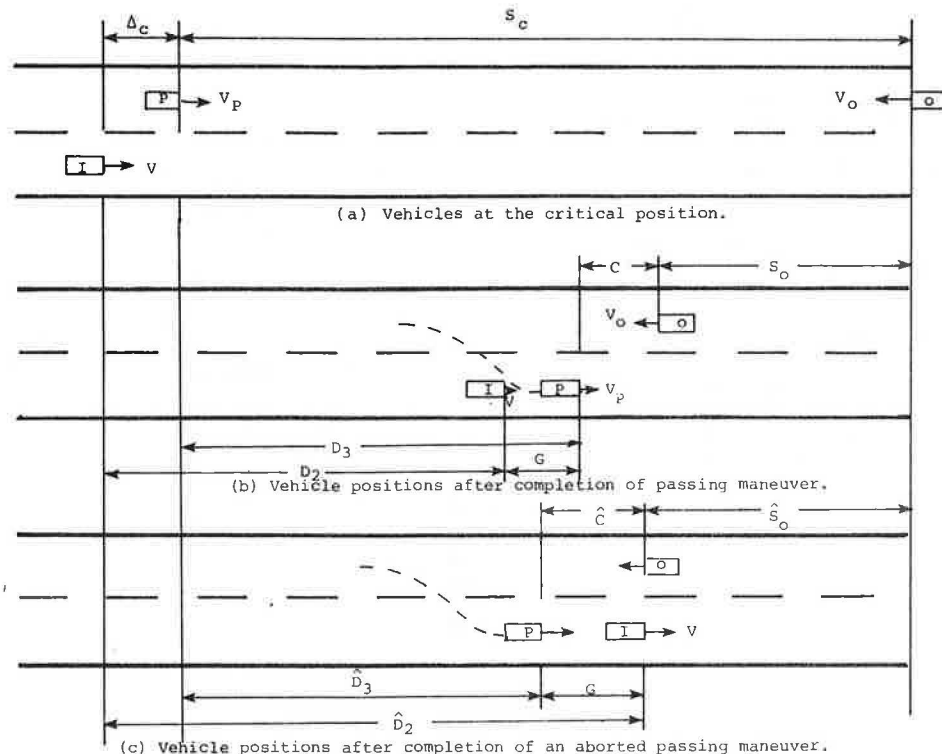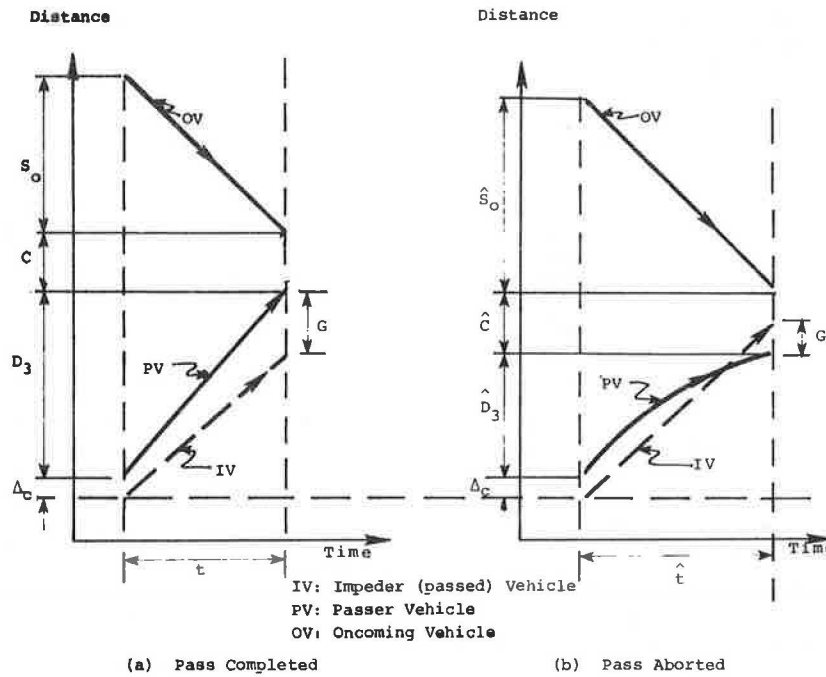
Figure 1. Passing scenarios.

(a) Vehicles at the critical position.

(b) Vehicle positions after completion of passing maneuver.

(c) Vehicle positions after completion of an aborted passing maneuver.

Figure 2. Vehicle trajectories.

IV: Impeder (passed) Vehicle
PV: Passer Vehicle
OV: Oncoming Vehicle

(a) Pass Completed   (b) Pass Aborted

t_1 = time passer vehicle moves at speed V_p to reach the critical position (s),

$\hat{t}$ = time for passing vehicle to return to its own lane from its critical position for aborted passing maneuver (s),

$t_{\overline{A}}$ = time passer vehicle spends accelerating to speed $V + m$ (s),

S.D. = required sight distance (m),

V = speed of impeder vehicle (m/s),

V_o = speed of oncoming vehicle (m/s),

V_p = speed of passing vehicle at its critical position (m/s),

V_{max} = maximum speed achievable at zero acceleration (m/s), and

Δ_c = distance that passing vehicle is downstream of impeder vehicle at critical position (m).

The formulation of the model proceeds in two parts, which are subsequently joined together:

1. Description of the passing and abort maneuvers from the critical position to the completion of these maneuvers and

2. Description of the passing maneuver from its inception until the critical position has been attained.

DEVELOPMENT OF MODEL

The trajectories of the vehicles of interest--passer, impeder, and oncomer--from the critical position to the completion of the maneuver are shown in Figure 2. In Figure 2a, the passer completes the maneuver; in Figure 2b, the passing maneuver is aborted. It is assumed that the passer has attained passing speed (V_p) by the time the critical position is reached and that the impeder and oncoming vehicles travel at constant speeds, V and V_o, respectively.

From Figure 2a it is seen that $D_3 + \Delta_c = Vt + G$. Here,

$$D_3 = (V + m)t$$

$$m = V_p - V \text{ (speed difference)}$$

Substituting, we obtain

$$(V + m)t + \Delta_c = Vt + G$$

or

$$\Delta_c = G - mt \qquad (1)$$

From Figure 2b, it is seen that $\hat{D}_3 = \hat{V}\hat{t} - G - \Delta_c$. Here,

$$\hat{D}_3 = (V + m)\hat{t} - (1/2)a\hat{t}^2$$

Substituting, we obtain

$$(V + m)\hat{t} - (1/2)a\hat{t}^2 = V\hat{t} - G - \Delta_c$$

or

$$1/2 a\hat{t}^2 = m\hat{t} + G + \Delta_c \qquad (2)$$

The sight distance from the passer vehicle to the oncoming vehicle when the former is at the critical position is

$$S_c = D_3 + C + S_o = \hat{D}_3 + \hat{C} + \hat{S}_o \qquad (3)$$

By definition of critical position, $C = \hat{C}$. Then, $D_3 + S_o = \hat{D}_3 + \hat{S}_o$. Substituting and assuming the oncoming vehicle speed to be $V_o = V$ yields

$$(V + m)t + Vt = (V + m)\hat{t} - 1/2 a\hat{t}^2 + V\hat{t}$$

or

$$(2V + m)(\hat{t} - t) = (1/2)a\hat{t}^2 \qquad (4)$$

Equating Equations 2 and 4 yields

$$(2V + m)(\hat{t} - t) = m\hat{t} + G + \Delta_c \qquad (5)$$

Substituting Equation 1 into Equation 5 yields

$$t = \hat{t} - (G/V) \qquad (6)$$

Substituting Equation 6 into Equation 4 and solving for $\hat{t}$ yields

$$\hat{t} = [2G(2V + m)/aV]^{\frac{1}{2}} \tag{7}$$

Solving Equations 7, 6, and 1 in that sequence yields the values of $\hat{t}$, t, and Δ_c, respectively. The calculation for the required sight distance from the critical position (S_c) follows immediately from Equation 3:

$$S_c = (2V + m)t + C \tag{8}$$

It is seen that, subject to reasonable assumptions, the critical position (Δ_c) is independent of the sight distance (S_c) but is dependent on the value of deceleration (a) that is acceptable to the motorist during any abort maneuver.

Figure 3 depicts the vehicle deployments from the start of the passing maneuver to the attainment of the critical position. During this period, the passer vehicle accelerates from its initial speed, assumed to be that of the impeder ("flying" passes

are not considered here), until it attains its passing speed (V_p) before or at the critical position.

This acceleration is a function of the vehicle speed, as shown in Figure 4. This function, in turn, depends on the type of vehicle, its weight-to-horsepower ratio, and other factors. To simplify the formulation yet retain the dependence of acceleration on vehicle speed, we will use an average acceleration ($\overline{A}$), calculated as follows:

$$\overline{A} = A_{max} \left\{ 1 - [(V + m/2)/V_{max}] \right\} \tag{9}$$

It follows that the time to attain passing speed (V + m) is

$$t_{\overline{A}} = m/\overline{A}$$

and

$$d_{\overline{A}} = Vt_{\overline{A}} + 1/2\overline{A}t_{\overline{A}}^2 = (m/\overline{A}) [V + (m/2)] \tag{10}$$

After attaining its passing speed, $V_p = V + m$, the passer vehicle travels for t_1 s until it

Figure 3. Start of passing maneuver.

(a)   Initial Position of Passer (p) and Passed (i) Vehicles

(b)   Deployment of Vehicles at the Critical Position
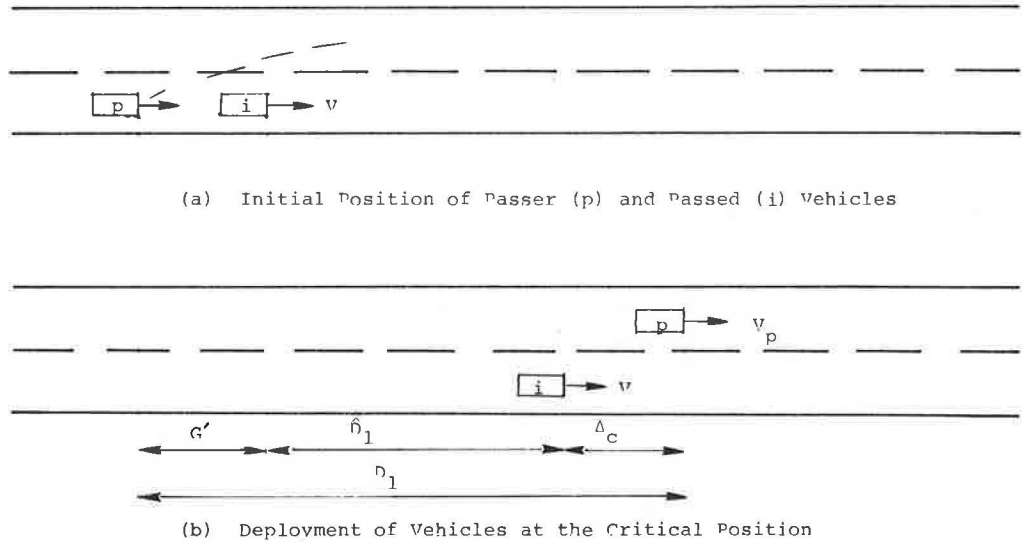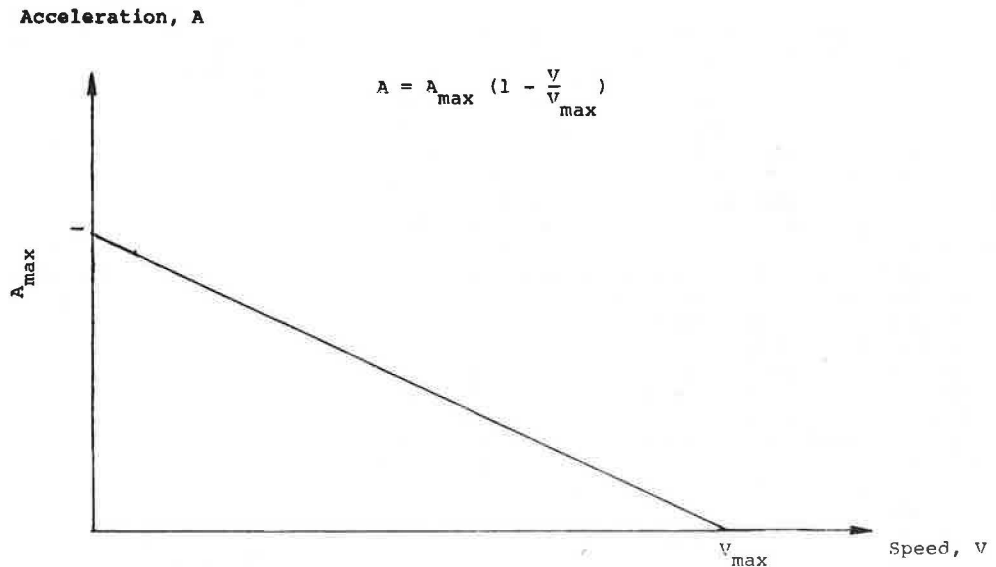
Figure 4. Dependence of vehicle acceleration on vehicle speed for specified vehicle type (level tangent).

Acceleration, A

$$A = A_{max} \left(1 - \frac{V}{V_{max}} \right)$$

reaches the critical position. Since the total travel time is T_1, then

$$t_1 = T_1 - t_{\bar{A}} = T_1 - (m/\bar{A}) \qquad (11)$$

Thus,

$$D_1 = d_{\bar{A}} + (V + m)t_1$$

From Figure 3 it is seen that

$$D_1 = G' + \hat{D}_1 + \Delta_c$$

Equating these two expressions for D_1 and recognizing that $\hat{D}_1 = VT_1$ yields the following:

$$d_{\bar{A}} + (V + m)[T_1 - (m/\bar{A})] = G' + VT_1 + \Delta_c$$

Substituting Equation 10 and solving for T_1 yields

$$T_1 = [(G' + \Delta_c)/m] + (m/2\bar{A}) \qquad (12)$$

Then

$$D_1 = G' + VT_1 + \Delta_c \qquad (13)$$

and the required sight distance is

$$S.D. = D_1 + S_c \qquad (14)$$

To solve this system, the passer vehicle type must be specified in order to obtain A_{max} and V_{max}, its operating characteristics. With these values ascertained, $\bar{A}$ may be found from Equation 9, then T_1 from Equation 12, where Δ_c is provided from Equation 1; the values of D_1 and of S.D. follow immediately by using Equations 13 and 14, where S_c is provided by Equation 8. Note that the critical position (Δ_c) is not dependent on the passer vehicle type and that the values of G and of G' depend on the vehicle speeds and lengths.

It should be emphasized that a major difference between this model and the approach used to develop the AASHTO specifications is that both the aborted and the completed passing maneuvers are considered here. (The analysis indicates that the abort time $\hat{t}$ always exceeds the passing time as measured from the critical position--see Equation 6.) The need to provide safe sight distances for both passing options is self-evident.

REPRESENTATIVE RESULTS

This formulation was programmed on a TI-59 calculator and a parameter study was undertaken.

Figure 5 displays the sensitivity of the critical position (Δ_c) with respect to the acceptable abort deceleration (a) for three values. It is seen that this sensitivity increases with speed (V) for any speed difference (m). Note that the critical position of the passing vehicle moves upstream of the impeder (i.e., Δ_c negative) as the speed difference (m) increases. To state it another way, the passer must decide earlier whether to abort, since the speed difference of the passer increases relative to that of the impeder.

Of particular interest is the great sensitivity of Δ_c with respect to the acceptable abort deceleration (a). The more this rate of deceleration decreases (which implies increased safety by virtue of easier vehicle handling and lower driver work load), the earlier the passer motorist must decide to abort. These relationships are intuitively satisfying.

Figure 6 displays the sensitivity of the critical sight distance (S_c) as measured from the critical

position with respect to impeder speed (V), speed difference (m), and acceptable deceleration (a). These values correspond to standard automobiles and the assumption of a 1.5-s headway for the calculation of the space headway (G).

As expected, the required critical sight distance (S_c) increases with speed (V). S_c is relatively insensitive to the speed difference (m), other factors being equal. This reflects the impact of two opposing factors: Higher value of m implies lower time to complete the pass, but the passer is farther upstream of the impeder at the critical position (as shown in Figure 5); thus, a longer distance is required to complete the pass.

Of particular interest is the sensitivity of S_c with acceptable abort deceleration (a). It is seen that the passing motorist must accept higher rates of deceleration in order to accept more passing opportunities when the available sight distance is limited. This sensitivity again demonstrates the need to consider the abort option when computing required sight distances.

The total sight distance (S.D.) is used as a basis for designing rural-road passing zones. To produce representative values of required passing sight distance by applying the model, a speed difference, m = 10 mph (16.1 km/h), is employed. (This value of m is used for the AASHTO specifications.) With this value of m, it is possible to estimate a reasonable value of abort deceleration (a), which provides computed passing sight distances that approximate the AASHTO sight-distance specifications. As shown in Figure 7, this value of acceptable abort deceleration is approximately 12 ft/s² (3.66 m/s² or 0.37 $\underline{g}$) over the range of speeds considered.

Examination of Figure 7 reveals that the AASHTO specifications for required passing sight distances are reasonable and conservative for impeder speeds up to 45 mph (72 km/h) for the case of a standard automobile passing an automobile. These results reflect the lower speeds of impeder vehicles in the 1940s when the empirical data supporting these specifications were gathered relative to speeds characteristic of today's traffic.

Recent observations of traffic on rural roads confirm that trucks (which often act as impeders) frequently travel at speeds well in excess of 50 mph (80 km/h). As shown in Figure 7, the model predicts that the required passing sight distances at these higher speeds are substantially greater than those specified currently by AASHTO.

It is necessary to consider the impact of the changing fleet composition characterized by smaller, low-powered automobiles and longer, high-powered trucks on the calculation of required passing sight distances. The model presented here permits the determination of required sight distances for different passing scenarios involving different types of vehicles. Specifically, the following passing scenarios are examined:

Scenario	Passing Vehicle	Impeder Vehicle
1	Standard automobile	Automobile
2	Standard automobile	Trailer truck [65 ft long (19.5 km)]
3	Subcompact automobile	Automobile
4	Subcompact automobile	Trailer truck

It is instructive to normalize all results with respect to the AASHTO specifications. In Figure 8, the horizontal axis represents the AASHTO specifications. Where the curves representing the results of this model extend above the axis, the AASHTO passing sight-distance specifications are inadequate; where

Figure 5. Sensitivity of critical position to impeder speed, speed difference, and acceptable abort deceleration (standard-sized automobile).

(a) V = 30 mph
1 mph = 1.61 km/hr.
1 ft = 0.3 m

V = Passed vehicle speed
m = Speed difference at C.P.
a = Acceptable abort deceleration

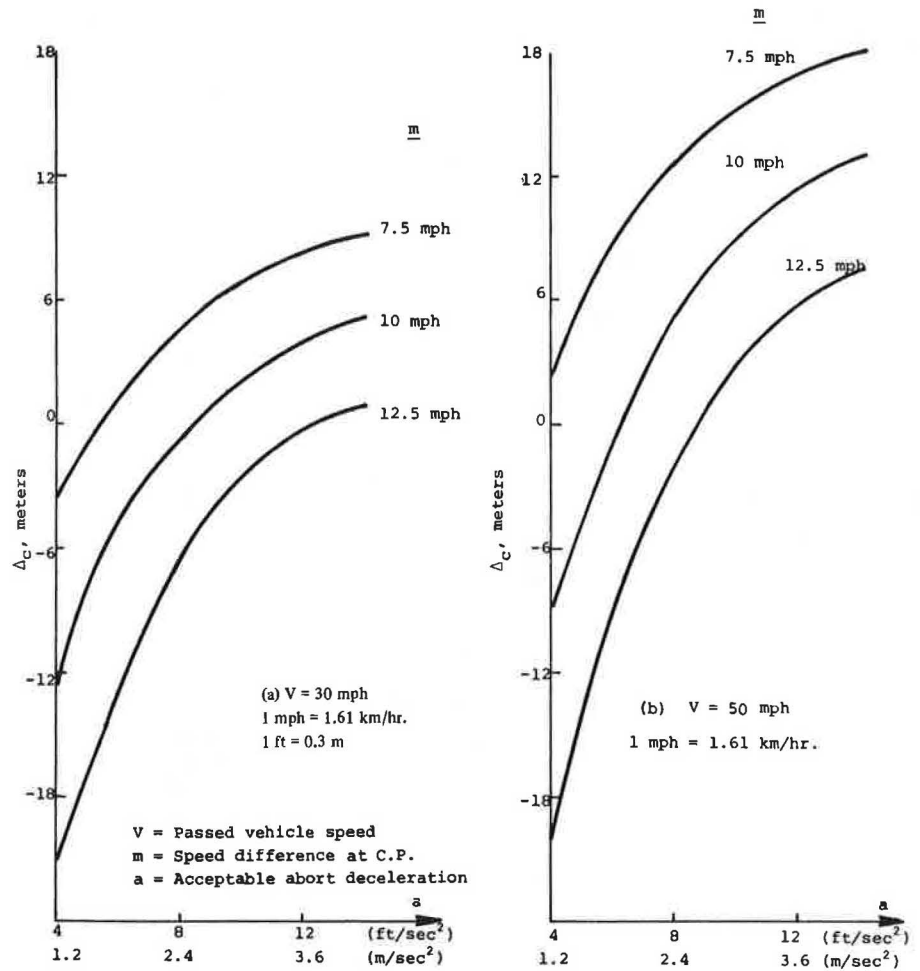(b) V = 50 mph
1 mph = 1.61 km/hr.

Figure 6. Critical sight distance (S_c) versus speed, speed difference, and abort deceleration.

Figure 7. Required sight distance.

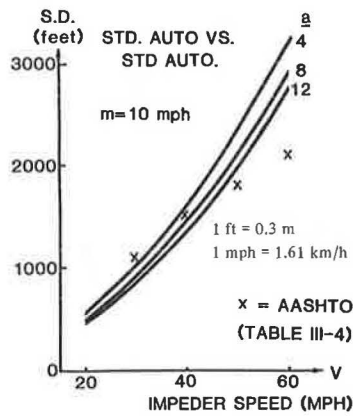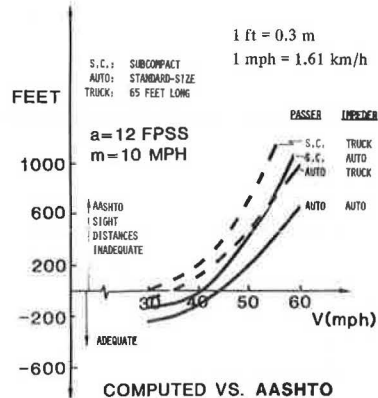more pronounced when the impeder is a truck or the passer is a subcompact automobile. As indicated in Figure 8, up to 1000 additional ft (300 km) of passing sight distance may be required relative to the current AASHTO specifications when the impeder vehicle is traveling at 55 mph (88 km/h). These results have been confirmed by empirical observation (2-4).

SUMMARY AND CONCLUSIONS

A formulation is presented that describes the passing maneuver on two-lane rural roads. A parameter study was undertaken that generated results describing the sensitivity of required passing sight distance with respect to the specified conditions. These results were compared with current AASHTO specifications for required passing sight distance. This comparison raises some questions concerning the adequacy of these AASHTO specifications when impeder speeds exceed 40 mph (64 km/h), particularly when subcompact automobiles and trucks are involved.

Based on these results, a review of the AASHTO specifications appears to be justified, particularly since the size of automobiles is projected to be reduced over the next decade. More conservative sight-distance requirements should enhance the safety of traffic roads. The impact of more restrictive passing zone delineation on roadway capacity may well increase the need to upgrade roadway geometrics or to improve control of passing operations. This is a problem area that deserves further study.

Figure 8. Assessment of AASHTO passing sight distance.

REFERENCES

1. A Policy on Geometric Design of Rural Highways (Blue Book). American Association of State Highway Officials, Washington, DC, 1965.
2. D.E. Peterson and R. Gull. Triple Trailer Evaluation in Utah. Utah Department of Transportation, Salt Lake City, Final Rept., 1975.
3. Report on Testing the Triple Trailer Combinations in Alberta. Alberta Department of Highways and Transport, Edmonton, Alberta, Canada, 1970.
4. R.J. Troutbeck. Overtaking Behavior on Australian Two-Lane Rural Highways. Australian Road Research Board, Nunawading, Victoria, Australia, ARRB Special Rept. 20, Sept. 1981.

the curves are below this axis, the AASHTO specifications are satisfactory.

Figure 8 compares the required sight distances estimated by this model for m = 10 mph and a = 12 ft/s² with those specified by AASHTO for the four scenarios defined above. As indicated, subcompact automobiles require a longer sight distance than do standard automobiles. Also, longer sight distances are required when the impeder is a truck than when the impeder is an automobile, which reflects the longer distance the passer must travel when the impeder is a truck.

The AASHTO specifications for passing sight distance are reasonable for speeds well below 44 mph (70 km/h) but appear to be increasingly inadequate (i.e., unsafe) at higher speeds. This inadequacy is

Performance Characteristics of Coal-Hauling Trucks in Mountainous Terrain

RONALD W. ECK, ABISHAI POLUS, AND KAI-CHU TSOU

Objectives of the study were to analyze and evaluate speed characteristics of coal trucks on two-lane mountain highways and to identify and quantify traffic performance limitations related to heavy trucks. To accomplish these objectives, spot speed data were collected on three coal-haul roads in West Virginia and a simulation analysis was performed by using the geometric characteristics of the same three sites as input. Upgrade and downgrade speed profiles were plotted for passenger cars and three classes of trucks. Truck speeds were significantly lower than passenger-car speeds on upgrades. On downgrades, the speed difference was not so pronounced and depended on whether trucks used braking or lower gears to reduce speed. Although simulated truck speeds showed good agreement with field data, simulated passenger-car speeds were uniformly higher than field speeds. This was attributed to the narrow roadway and rough pavement condition of the study sites. Geometric delay was significantly greater than traffic delay. Both types of delay were quantified for a variety of geometric and flow conditions. Acceleration noise, another traffic-flow parameter, was used as a measure of accident potential and stability of flows. Acceleration noise increased as volume increased on both upgrades and downgrades. Upgrade acceleration noise was greater than downgrade noise. Several practical applications of the results and directions for future research were presented.

The flow of traffic on mountain two-lane rural highways is adversely affected by the nonuniform performance capabilities of vehicles. These performance differences, particularly between loaded trucks and passenger vehicles, are most pronounced on steep upgrades and downgrades. The influence on flow is twofold. First, the reduction of highway capacity is significant on grades since, in effect, trucks take up the space of a larger number of passenger cars on grades than they do on level sections. Second, because of poorer truck performance on grades, there is increased likelihood for traffic instabilities, accident potential, and delays. This becomes even more noticeable with an increase in the proportion of low-performance cars, which in turn increases the variability both between and within vehicle groups.

Most highways in West Virginia and throughout Appalachia are narrow two-lane roads carrying significant amounts of large coal-hauling trucks, which cause the effects mentioned above. As coal assumes a greater role as an energy source, local citizens, planners, and engineers are devoting increased attention to evaluating the impacts of the growing number of coal trucks on the highway system. The study described in this paper examined the impact of coal trucks on traffic flow characteristics of two-lane rural highways.

Two primary objectives, or stages, were established for the study. The first objective was to analyze and evaluate speed characteristics of coal trucks on upgrades and downgrades and to compare the findings with those from previous research. The second objective was to identify and quantify other traffic performance limitations caused by or related to heavy trucks and to suggest possible measures to reduce these adverse effects.

In order to accomplish the first objective, speed data were collected on sections of three coal-haul roads in West Virginia. For the second stage of the research, a simulation analysis was performed, the input for which was collected during the first stage.

PREVIOUS STUDIES

A few previous studies have been conducted on performance limitations of coal or other heavy trucks on grades, although these limitations are widely accepted. Upgrade performance is primarily influenced by engine capabilities, specifically the ratio of weight to horsepower. This is suggested by the Policy on Geometric Design of Rural Highways (1) of the American Association of State Highway Officials (AASHO), which presents some speed-distance curves for upgrades based on an earlier flow study conducted by Huff and Scrivner (2). The Arizona Highway Department conducted another early study (3) on congestion caused by trucks on mountainous uphill grades. It was concluded that the total delay was dependent on the topographical features of the individual hill, the alignment, sight distance, roadway width, and percentage of grade. Walton and Lee (4) presented findings of a study in Texas where speed-distance curves were developed for a range of grade profiles; these were applicable to the evaluation of the need for and design of climbing lanes for trucks.

Downgrade performance is affected by a complexity of components, which may include the length and steepness of the grade as well as the previous and the following grades, sight distance, and driver skill and attitude. The AASHO guide (1) suggests that, compared with level operation, heavy vehicles on downgrades show an increase in speed for grades up to about 5 percent and a decrease in speed for grades of about 7 percent or steeper. A California study (5) dealt with downhill speeds of heavy vehicles; no distinction was made between loaded and empty trucks. Analysis of field observations showed a distinct difference in the behavior of trucks on long downgrades as compared with that on shorter downgrades. For long grades, trucks were observed to slow down to a crawl and maintain that speed until not far from the bottom of the grade. For short, steep grades, truck speeds were slower near the summit but increased uniformly down the grade. A more recent study (6) analyzed speed characteristics of heavy vehicles on downgrades. It was found that loaded trucks reduce their speed considerably at the beginning of a downgrade. The amount of reduction was shown to be related to the length and slope of the downgrade; the second variable contributed exponentially to the increase in speed gradient.

Some studies have developed simulation models to study traffic flows. One such study on two-lane mountain highways conducted by the Midwest Research Institute (7) suggested a measure for rating grades by analyzing truck performance on grades with a simulation technique. St. John and Kobett (8) improved and adjusted the simulation model to duplicate the characteristics of mixed (cars and trucks) flows on grades. Botha and May (9) used this same model to determine an optimum length and location of a climbing lane on a specific upgrade and to determine general guidelines for the most cost-effective location of climbing lanes when several upgrades are considered. It was determined that the travel-time benefit obtained from a climbing lane is most sensitive to gradient.

St. John's detailed model was adopted for use in the second stage of this study to examine the impacts of coal trucks on flow on mountain two-lane

Table 1. Characteristics of three study sites in Preston County, West Virginia.

Site No.	Site Location	Average Slope[a] (%)	Overall Length (m)	Peak-Hour Volume	Trucks (%)		
					T_1	T_2	T_3
1	WV-7, Cascade	5.10	866	380	5.5	3.3	3.0
2	WV-7, Kane's Creek	6.41	970	280	4.8	6.1	2.5
3	WV-26, Kingwood	4.98	1150	250	3.2	6.9	1.6

[a]Computed as weighted (by length) average of individual subsection slopes on each grade.

Table 2. Truck characteristics used in study.

Vehicle Type	Description	Avg GVW (kg)	Manufacturers' Suggested Avg Weight-to-Horsepower Ratio (kg/hp)	Avg Length (m)
T_1	Single-unit truck, three axles or less	21 000	95	10.7
T_2	Single-unit truck, four axles	28 000	165	10.7
T_3	Trucks with five axles or more, including semitrailer combinations	36 000	170	16.7

Figure 1. Field data collected at site 1, WV-7 near Cascade.

(a) Geometric Profile

(b) Speed Profiles

highway sections. The simulation model is described in detail elsewhere (8).

DATA COLLECTION

By using a radar device, spot speed observations were taken at five points on each of three sections of mountain roads in Preston County, West Virginia. Characteristics of the three study sections are presented in Table 1; the truck types shown will be described below. These three sites were selected to be two-lane rural grades on a tangent section, where the approach roadway was approximately level. This was to permit study of traffic characteristics on upgrades and downgrades where speed was not limited by horizontal curvature. Other criteria in the selection of sites were the presence of a significant number of coal trucks, proximity to West Virginia University, and close physical resemblance to one another in terms of design speed, pavement width, and shoulder width.

Since heavily loaded trucks have significantly lower performance characteristics compared with empty trucks, they have the most severe impacts on traffic flow both in the upgrade and downgrade directions. For the purpose of this study, it was decided to examine mainly the loaded trucks, although comparisons were made with performance of other vehicles, particularly passenger cars. The loaded trucks were primarily hauling coal, although some were carrying limestone and some were tank trucks transporting petroleum products. All trucks were classified according to number of axles, while the average gross vehicle weight (GVW), ratio of weight to horsepower, and average length were obtained from the literature, since time and resource constraints precluded field determination of these quantities. A summary of truck characteristics is presented in Table 2.

FIELD SPEED CHARACTERISTICS

The speed data collected were sorted by the three truck types described above in addition to passenger cars. Based on visual observations, it was concluded that most of the four and five-axle trucks (types T_2 and T_3) were coal-hauling vehicles originating at coal mines near the study sites and
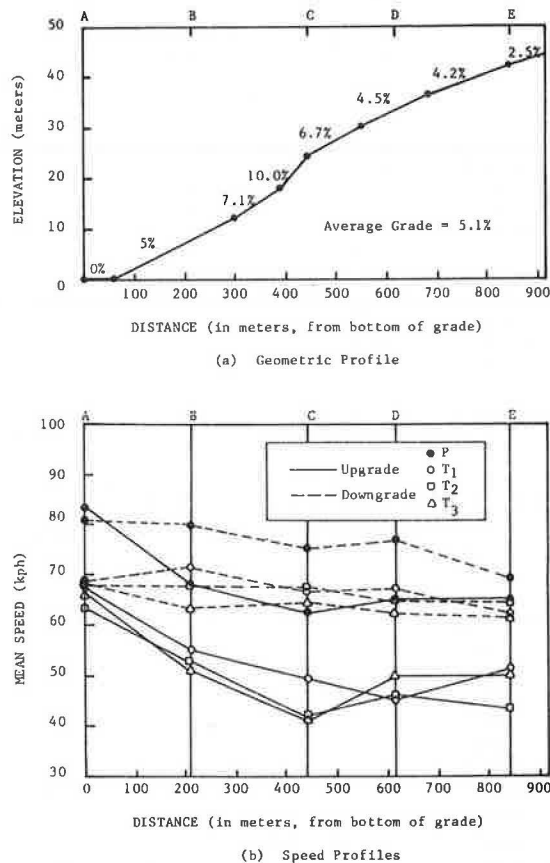
destined for a power plant or rail-loading facility.

Figure 1 presents the mean spot speed profiles of the four types of vehicles for site 1 for upgrade and downgrade directions. At the foot of the grade (station A) the mean speed of the passenger car (P) was about 16.1 km/h higher than those of all three classes of loaded trucks. There was no significant difference between truck speeds at station A. Table 3 presents an analysis of speed differences for different vehicle types for the five stations at site 1. A continuous decrease in speeds was noted for station C, from which a slight increase in speed was noted. At station E, located at the top of the upgrade, the relative speed difference remained the same, although the single-unit four-axle trucks (T_2) had the lowest mean speed. T_2 speeds were significantly lower than those of passenger vehicles and five-axle trucks (T_3), as shown in Table 3. The four-axle trucks were found to have the lowest speed at the top of the grade on all three sites. The vast majority of these trucks were coal-hauling trucks, and it is suspected that overloading was a prime reason for their relatively poor performance.

Figure 1 also presents the mean speed profiles for downgrade vehicles at site 1. For trucks, there

Table 3. Comparison of mean speeds for different vehicle types on site 1 upgrade.

Vehicle Type	Station									
	A		B		C		D		E	
	t-Statistic	Significance[a]	t-Statistic	Significance[a]	t-Statistic	Significance[a]	t-Statistic	Significance[a]	t-Statistic	Significance[a]
P_1 vs T_1	−4.867	+	−4.085	+	−5.927	+	−7.031	+	−5.198	+
P_1 vs T_2	−12.315	+	−13.010	+	−11.821	+	−9.882	+	−11.037	+
P_1 vs T_3	−6.695	+	−4.268	+	−8.729	+	−5.703	+	−5.816	+
T_1 vs T_2	0.900	−	0.445	−	1.866	+	0.223	−	1.895	+
T_2 vs T_3	0.951	−	0.347	−	0.120	−	0.742	−	1.655	−

[a]The plus sign means significant at the 5 percent level; the minus sign means not significant at the 5 percent level.

Figure 2. CDF of upgrade and downgrade mean spot speeds at selected points on site 2, WV-7, Kane's Creek.
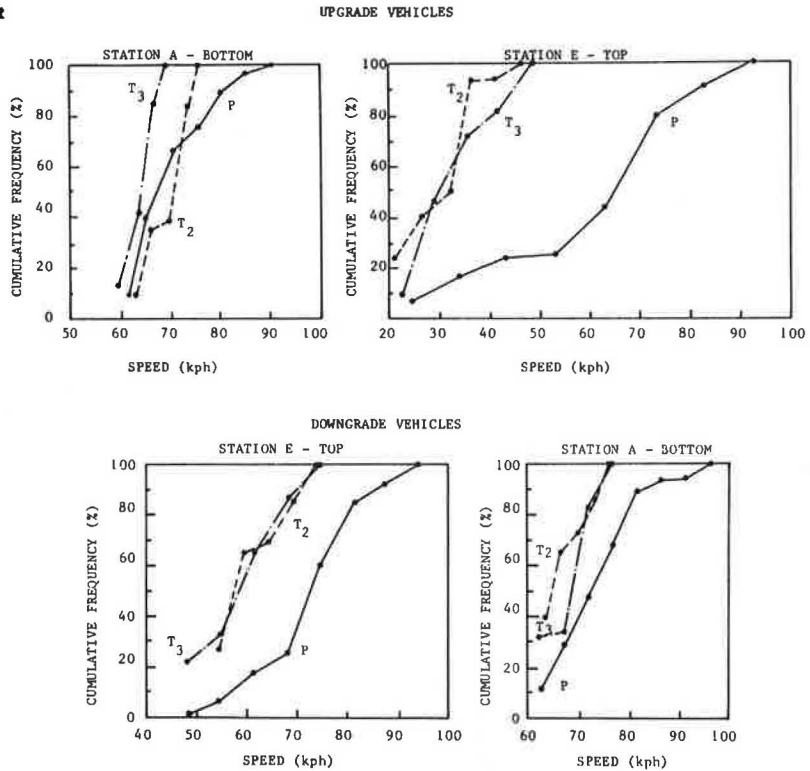
was relatively little acceleration proceeding downgrade. By the time the steepest part of the grade (station C) was reached, loaded-truck speeds were only about 3.2 km/h greater than speeds at the top of the grade. Passenger-car speeds, on the other hand, increased until the grade steepened significantly. Prior to entrance of the −10 percent section, braking reduced vehicle speed slightly. At the bottom, loaded-truck speeds were essentially equal but about 11.3 km/h slower than those of passenger cars. The widest variation between passenger-car and loaded-truck speeds occurred at this point. An analysis of speed differences similar to the one performed for the upgrade direction found that the differences between passenger cars and all types of loaded trucks were significant at the 5 percent level. Differences between T_2 and the other two loaded-truck types were not significant for this downgrade or for sites 2 and 3.

Cumulative density functions (CDFs) of mean spot speed for all three sites were examined. Figure 2 presents typical examples of the CDFs for site 2, upgrade and downgrade directions, at stations A and E. It can be seen that the cumulative distributions followed the typical S-shaped curve. Slopes of the

curves were very steep at the entrances to the grades, which indicated uniform speeds within each vehicle category. However, on the grades themselves, the curves were flatter, which indicated more variable speeds as the grades affected vehicle performance.

The downgrade CDFs portrayed relatively low variability in speeds between truck types. At most observation stations, loaded-truck curves tracked one another very closely. Passenger-car curves demonstrated slightly higher speeds at all observation stations. This is what one would expect intuitively; i.e., passenger cars are fastest, followed by loaded trucks where drivers use both gears and brakes on relatively long and steep downgrades.

SIMULATION SPEED RESULTS

For further evaluation of coal-truck performance and impacts, a previously developed simulation model, which has been documented (8) and tested in Kansas, was used. However, it was necessary to check the validity of the model when applied to the narrow mountain roads of West Virginia.

A preliminary validity test had been conducted in

a previous study in West Virginia (10). The study determined that the model's space mean speeds closely resemble actual space mean speeds. Further checks were made by using a sensitivity analysis, a procedure in which the input data are altered to see what effect they have on the results. Changing model input data did in fact produce simulated traffic situations that agreed with what would be expected intuitively. Since the model proved satisfactory for further analysis, the previously collected geometric and flow data were used as an input to the simulation model. Output from the model contained several prominent results. These results are discussed in the next two sections of this paper.

Figure 3 presents a typical example of mean field spot speeds versus mean simulation spot speeds as obtained for site 3. Two striking features may be observed. Simulation truck speeds and field truck speeds tended to be relatively similar without any major observable differences. This statement was valid for the other sites as well. Field and simulation performances of passenger cars, however, were statistically different; the simulation results were significantly higher. There are at least two possible reasons for the phenomenon. The study-site roadways were narrow (about 2.7-m lanes with almost no shoulders) and had rough surfaces (due to frequent patching). Both these factors could cause passenger-car speeds to be reduced below those predicted by the simulation model. However, truck speeds, since they were lower, were not so sensitive to these adverse geometric conditions.

It is also possible that, due to downsizing, passenger-car characteristics have changed since the simulation model was written in the late 1970s. The researchers observed a substantial number of four-cylinder subcompact cars on the study sections. These vehicles would be expected to perform more poorly than standard passenger cars on upgrades. However, this does not explain the fact that field speeds were lower than simulated speeds on the downgrade.

The second noticeable feature was the reduction in truck speeds versus the increase in passenger-car speeds. This phenomenon was attributed to a long horizontal curve located near the beginning of the grade. When the upgrade was encountered, truck speeds dropped off rapidly; loaded coal trucks (T_2) were most affected, since their speeds dropped about 24 km/h. Passenger cars, however, displayed some acceleration ability, which in fact increased the instability in terms of rear-end collision hazard between passenger vehicles and trucks. Similar situations, in which a horizontal curve precedes a steep upgrade, may call for certain geometric measures such as provision of climbing lanes that will attract trucks and reduce flow hazard and friction between heavy vehicles and passenger cars.

Findings of simulated speeds versus field speeds on downgrades for site 3 are also presented in Figure 3. Again, except for passenger vehicles, the similarities between the simulation model and the field data were noticeable. The reasoning here is similar to that discussed above where the narrow lanes and shoulders reduced passenger-car speeds considerably.

TRAVEL-TIME DELAY

When a passenger vehicle approaches a slower, heavier vehicle on an upgrade, it may or may not encounter some travel delay, depending on the geometric conditions and volume and speed of opposing traffic. When there is no opposing traffic in the adjacent lane and when the geometrics of the site, such as sight distance, are satisfactory, usually no

Figure 3. Field versus simulated mean speed profiles for different vehicle types at site 3, WV-26, Kingwood.
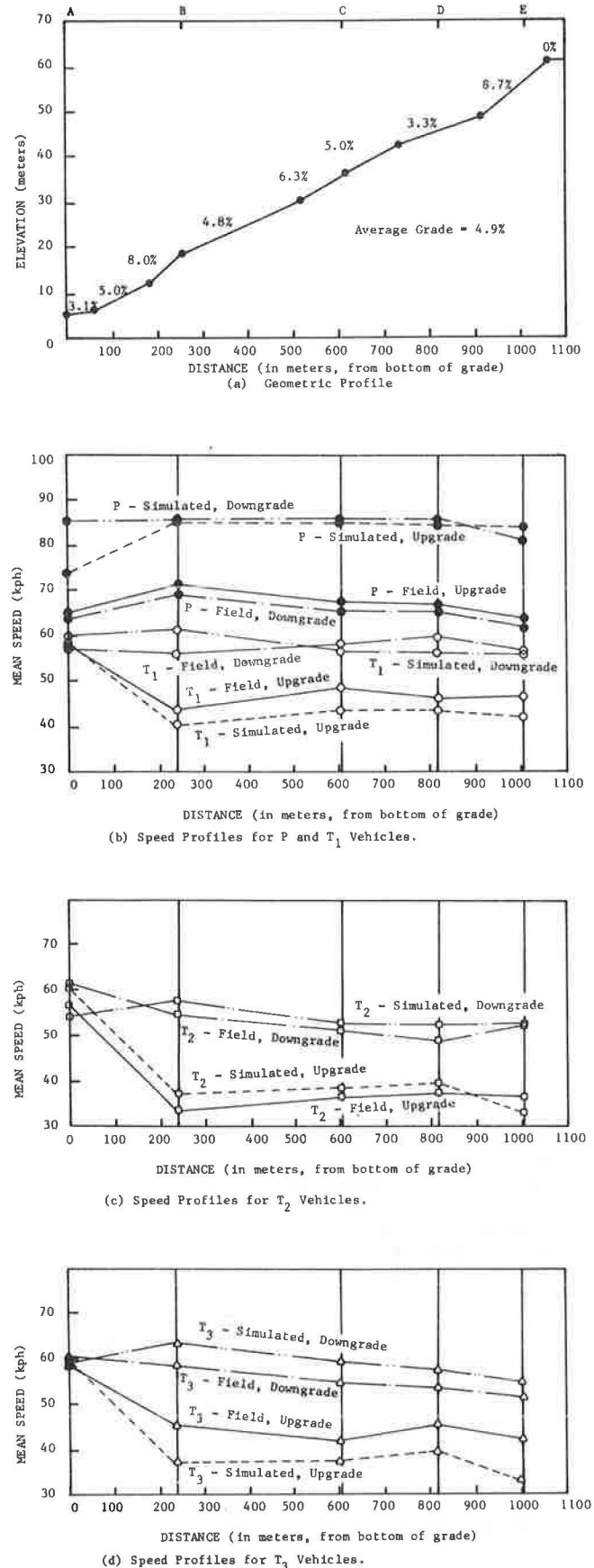

(a) Geometric Profile

(b) Speed Profiles for P and T_1 Vehicles.

(c) Speed Profiles for T_2 Vehicles.

(d) Speed Profiles for T_3 Vehicles.

Figure 4. Speed profiles of simulated passenger cars delayed on site-1 upgrade.

Legend:
- ○ Free Flow Speeds
- △ Reducing Speed for Slow-Moving Truck
- ● Delayed Speeds

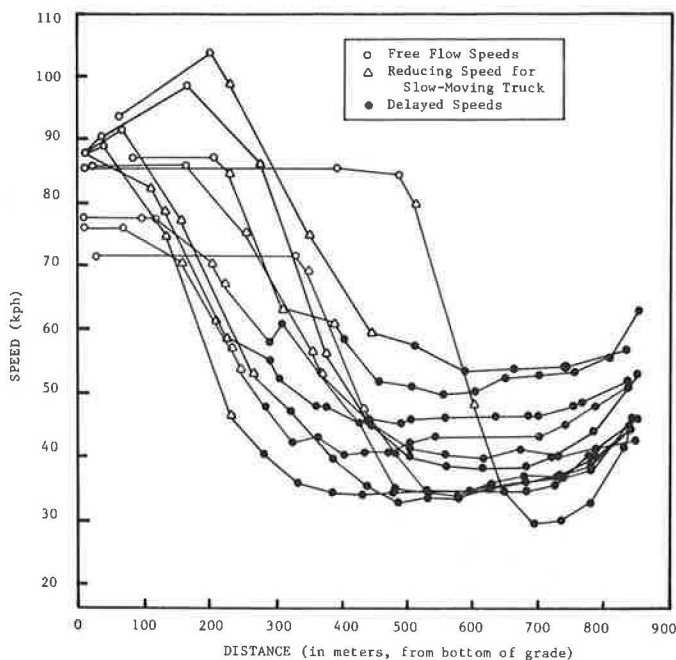SPEED (kph) vs DISTANCE (in meters, from bottom of grade)

Table 4. Geometric and travel-delay rates for site 1.

Direction of Travel	Geometric Delay (s/km)			Hourly Volume	Avg Traffic Delay (s/km)	
	T_1	T_2	T_3		All Trucks	Passenger Cars
Upgrade	47.0	55.9	59.3	150	10.6	6.3
				200	12.0	8.2
				250	1.7	9.9
				300	1.0	12.0
Downgrade	1.1	0.3	0.1	150	2.8	2.2
				200	3.9	1.6
				250	2.1	2.2
				300	2.7	2.8

reduction in speed and no delay are detected. If, however, an opposing vehicle appears, the impeded vehicle has to reduce its speed and continue to trail the slower-moving trucks. The total time delay rate in this study was defined as travel time per kilometer distance, whereas speed was reduced due to both roadway geometry and presence of other traffic. Delay due to traffic was obtained by subtracting geometric delay from total time delay.

The development and magnitude of this impedence are presented in Figure 4, which is based on the simulation model and demonstrates the sudden drop in speeds of originally free-flowing passenger cars as they encounter trucks on the upgrade and indicates their travel at slow speed behind the trucks. (In Figure 4, each curve represents a speed profile of one passenger car.) Further quantification of the delay data is presented in Table 4, where the geometric and traffic components of the travel-time delays are presented both for the upgrade and the downgrade of site 1. The following conclusions can be drawn: (a) the upgrade delay, due to both geometry and traffic, is considerably higher than the downgrade delay, regardless of traffic volume; (b) passenger-car traffic delay increases proportionally with volume; (c) passenger-car geometric delay is essentially zero for this site both for the upgrade and the downgrade; and (d) truck geometric delay is significant for the upgrade (47-59 s/km); downgrade

geometric delay does not seem to be significant at this site. Sites 2 and 3, however, present greater downgrade geometry delay--about 18 and 22 s/km, respectively. This is attributable to steep subsections on each grade that caused trucks to gear down to a crawl to preserve brakes on the grade.

ACCELERATION NOISE ON GRADES

The acceleration-noise concept has been used as a measure of accident potential and stability of flows and is defined (11) by the standard deviation of the driver's acceleration about the mean acceleration over a given section of highway. The total acceleration noise of a vehicle in traffic is a superposition of its natural noise (the acceleration noise at level section and very low traffic volume) and its response to various geometric and flow conditions. Table 5 presents the simulation findings of the acceleration noise for the three sites in both the upgrade and the downgrade directions. As would be expected, the acceleration noise generally increased with volume. As the traffic volume increased from 150 to 300 vph, the acceleration noise increased by 22-56 percent for the upgrades; acceleration noise increased by 9-53 percent for the downgrades. The influence of the volume increase seemed to be less pronounced at site 1, where section length was relatively short and the grade was not too steep. For site 2, where the grade was the longest and steepest, the volume increase generated the most instability. Another noticeable conclusion to be drawn is that the magnitude of the upgrade acceleration noise was greater than the downgrade noise, which means that the upgrade disturbances to the stability of the flow were greater. One can note, however, that some noise above the normal level section noise was attributable to downgrades as well, which has been found (11) to be on the order of 0.04 m/s². This was especially noticeable at sites 2 and 3, where trucks maintained crawl speed over part of the downgrade.

CONCLUSIONS AND RECOMMENDATIONS

This study analyzed speed and flow characteristics

Table 5. Acceleration noise for mixed vehicles for different flow conditions on three study sites (simulated results).

Site	Traffic Direction	Noise (m/s^2) at 150 vph	Flow Rate (vph)					
			200		250		300	
			Noise (m/s^2)	Percent Increase[a]	Noise (m/s^2)	Percent Increase[a]	Noise (m/s^2)	Percent Increase[a]
1	Upgrade	0.349	0.440	+25.8	0.386	+10.6	0.444	+27.2
	Downgrade	0.133	0.132	−0.0	0.138	−4.1	0.145	+8.9
2	Upgrade	0.306	0.370	+20.8	0.610	+99.3	0.477	+55.8
	Downgrade	0.134	0.435	+224.9	0.237	+76.6	0.204	+52.1
3	Upgrade	0.287	0.328	+14.2	0.345	+20.4	0.352	+22.6
	Downgrade	0.211	0.235	+11.2	0.311	+47.1	0.324	+53.5

[a]Relative to acceleration noise for 150 vph.

of heavy vehicles and passenger cars on three two-lane mountain highway sites in West Virginia. The study was based on analysis of gathered field data as well as further evaluation of results obtained by a simulation model calibrated with input geometric and flow data obtained at the study sites. Several conclusions drawn from this study will be discussed below.

Heavy-truck speeds, such as those of coal trucks, were significantly lower than passenger-car speeds on upgrades. On downgrades, the difference between heavy and light trucks was not so pronounced as on upgrades. On downgrades with relatively steep subsections, loaded trucks may reach crawl speed. Passenger cars following such trucks can experience significant delay.

Travel-time delay is composed of geometric delay caused by sharp curvature, steep grades, and inadequate pavement width and traffic delay caused by slow-moving vehicles where there are no opportunities to pass. The geometric delay on the West Virginia mountain roads studied was significantly greater than the traffic delay. Traffic delay, however, was more significant for passenger cars than for trucks. Acceleration noise increased with volume increases on both upgrades and downgrades. The upgrade noise was greater than the downgrade noise.

Although several of the conclusions cited above are not new, the research results presented here are significant because they quantify the impact of trucks of the type used to haul coal on traffic flow on two-lane mountain roads. Results should be of particular interest to highway planners evaluating the transportation effects of proposed mine openings. However, the data have wider applicability beyond coal transportation. For example, the information could be used by highway engineers concerned with evaluating alternative highway improvements such as rerouting trucks, improved horizontal and vertical geometry, or the addition of truck climbing lanes.

Relative to truck climbing lanes, the research results have shown that consideration for introduction of climbing lanes on upgrades should include some additional factors beyond traffic volume and steepness and length of grade. Such factors should include the horizontal alignment, particularly the presence of sharp horizontal curves at or next to the bottom of the grade, and the sight distance.

Climbing-lane criteria have been based on ratios of truck weight to horsepower of 400 to 1 and a speed difference of 24 km/h. According to Weaver, Wootan, and Woods (12), recent trends have considered weight/horsepower ratios of 275:1 since truck engine sizes have been increased with little increase in weight. However, such criteria probably will not be applicable in the future as engine performance is reduced to meet fuel constraints. Studies like the one described in this paper may be used to provide baseline data for analyzing the impact of these lower-performance trucks on traffic flow. In addition to the compatibility problems between different classes of vehicles, the data might be useful to examine problems created by performance limitations of a class of vehicles, e.g., automobiles, as the proportion of lower-performance subcompact cars increases.

Results have also shown that the maintenance conditions of rural highways, especially those that contain a relatively large volume of heavy trucks, can have a definite influence on traffic flow. Maintenance is important, not only to preserve the pavement from a structural standpoint but also to maintain a smooth and uniform flow, especially for passenger cars.

Another general area in which the research results may be applied and in which additional work is needed concerns the implications of geometric design on motor vehicle fuel consumption. There is a growing concern nationwide in identifying geometric design features and ways in which they can be altered without adversely affecting safety or level of service. For example, this study has indicated that geometric features such as insufficient passing opportunities on two-lane rural highways, sharp horizontal curves, or narrow roadways can cause speed changes. These design features affect motor vehicle fuel consumption, usually adversely, in the case of unnecessary speed changes. By combining specific vehicle fuel consumption data with geometric and traffic flow information, engineers and planners could estimate the energy impacts of different geometric design alternatives. In this way, decision-makers could give energy conservation as much or greater consideration as safety and travel time in the design and operation of the highway system.

A final recommendation for further research is that additional field and simulation analyses be made under a greater variety of geometric conditions, particularly on longer and steeper grades. More work is needed on the simulation model so that it portrays the mountain highways of Appalachia better. Development of prediction models for such measures as accident rates, instability potentials, or delays based on geometric or volume characteristics of existing or proposed highways would be appropriate. In this regard, a previously mentioned study (6) developed a model of mean approach speed gradient incorporating length and slope of downgrade. The contribution of the average slope of the downgrade was exponential in nature and its contribution to average approach speed gradient was more significant than that of length of the downgrade. Speed gradients for the three downgrades examined in this study were plotted on a graphical presentation of the equation (not shown here). The points fit the previously developed model very well. Such good agreement suggests that further study of longer and steeper grades would be worthwhile.

REFERENCES

1. A Policy on Geometric Design of Rural Highways. AASHO, Washington, DC, 1965.
2. J.S. Huff and F.H. Scrivner. Simplified Climbing Lane Design: Theory and Road Test Results. HRB, Bull. 104, 1955.
3. Arizona Highway Department. Report on Congestion Caused by Trucks on Uphill Mountain Grades. Division of Economics, U.S. Bureau of Public Roads, Oct. 1953.
4. C.N. Walton and E.C. Lee. Characteristics of Trucks Operating on Grades. TRB, Transportation Research Record 631, 1977, pp. 23-30.
5. G.H. Webb. Downhill Truck Speeds. California Division of Highways, Sacramento, Traffic Bull. 1, July 1961.
6. A. Polus, J. Craus, and I. Grinberg. Downgrade Speed Characteristics of Heavy Vehicles. Transportation Engineering Journal of ASCE, Vol. 107, No. TE 2, March 1981, pp. 143-152.
7. D.R. Kobett and others. Traffic Simulation for Design of Uniform Service Roads in Mountainous Terrain--Volume 2: Description and Validation of the Simulation. Midwest Research Institute, Kansas City, MO, 1970.
8. A.D. St. John and D.R. Kobett. Grade Effects on Traffic Flow Stability and Capacity. NCHRP, Rept. 185, 1978.
9. J.L. Botha and A.D. May. A Decision-Making Framework for the Evaluation of Climbing Lanes on Two-Way Rural Roads. Univ. of California, Berkeley, Res. Rept. UCB-ITS-RR-80-8, July 1980, 29 pp.
10. J. Halkias. Impact of Coal Trucks on Delay and Safety of Other Highway Users. Department of Civil Engineering, West Virginia Univ., Morgantown, M.S. thesis, 1980, 149 pp.
11. D.L. Gerlough and M.J. Huber. Traffic Flow Theory. TRB, Special Rept. 165, 1975.
12. G.D. Weaver, C.V. Wootan, and D.L. Woods. Highway Design Criteria: Problems and Projections. Proc., Specialty Conference Implementing Highway Safety Improvements, American Society of Civil Engineers, San Diego, CA, March 12-14, 1980, pp. 38-53.