TRANSPORTATION RESEARCH RECORD 682

Urban System Operation and Freeways

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1978

Transportation Research Record 682
Price $6.40

modes
1 highway transportation
2 public transit

subject areas
12 planning
54 operations and traffic control
55 traffic flow, capacity, and measurements

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

Notice
The papers in this Record have been reviewed by and accepted for publication by knowledgeable persons other than the authors according to procedures approved by a Report Review Committee consisting of members of the National Academy of Sciences, the National Academy of Engineering, and the Institute of Medicine.

The views expressed in these papers are those of the authors and do not necessarily reflect those of the sponsoring committee, the Transportation Research Board, the National Academy of Sciences or the sponsors of TRB activities.

To eliminate a backlog of publications and to make possible earlier, more timely publication of reports given at its meetings, the Transportation Research Board has, for a trial period, adopted less stringent editorial standards for certain classes of published material. The new standards apply only to papers and reports that are clearly attributed to specific authors and that have been accepted for publication after committee review for technical content. Within broad limits, the syntax and style of the published version of these reports are those of the author(s).

The papers in this Record were treated according to the new standards.

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Urban system operation and freeways.

(Transportation research record; 682)
Seventeen reports prepared for the 57th annual meeting of the Transportation Research Board.
Includes bibliographical references.
1. Traffic engineering–Congresses. 2. Express highways–Congresses. 3. Urban transportation–Congresses. I. Title. II. Series.
TE7.H5 no. 682 [HE332] 380.5'08s [388.4'11] 79-15545
ISBN 0-309-02827-2

Sponsorship of the Papers in This Transportation Research Record

GROUP 1–TRANSPORTATION SYSTEMS PLANNING AND ADMINISTRATION
E. Wilson Campbell, New York State Department of Transportation, chairman

Environmental Quality and Conservation of Resources Section
Warren B. Lovejoy, Port Authority of New York and New Jersey, chairman

Committee on Transportation and Air Quality
Earl C. Shirley, California Department of Transportation, chairman
John B. Chernisky, Federal Highway Administration, secretary
John Kenneth Bell, Earl H. Bowman, Norman L. Cooper, Walter F. Dabberdt, Diane L. Donley, Denis E. Donnelly, John H. Gastler, John Hawley, Ronald C. Hilfiker, Joel L. Horowitz, Kenneth E. Jones, Thomas P. Kozlowski, Jerry A. Kurtzweg, William C. Lockett, Francis J. Mason, Leonard O. Myrup, Carlton Thomas Nash, Kenneth E. Noll, Fedele L. Palmieri, Robert M. Patterson, Steven D. Reynolds, David R. Souten, N. Thomas Stephens, Lawrence R. Taylor

GROUP 3–OPERATION AND MAINTENANCE OF TRANSPORTATION FACILITIES
Adolf D. May, University of California, Berkeley, chairman

Facilities and Operations Section

Committee on Urban System Operations
Carlton C. Robinson, Highway Users Federation, chairman
D. E. Bergstrom, Kenneth W. Crowley, Ralph Gakenheimer, Joseph M. Goodman, David N. Goss, David D. Grayson, Marble J. Hensley, Sr., Joel L. Horowitz, Dan W. Hoyt, Herbert S. Levinson, Gary E. Maring, William R. McGrath, Chester F. Phillips, Anne Marie Zerega

Committee on Parking and Terminals
Raymond H. Ellis, Peat, Marwick, Mitchell and Company, chairman
Stephen G. Petersen, Consultant, Gaithersburg, Maryland, secretary
Frank E. Barker, Dalton L. Baugh, Sr., George K. Benn, Harvey B. Boutwell, Paul C. Box, Robert G. Bundy, John P. Cavallero, Jr., William D. Heath, James M. Hunnicutt, Delbert F. Karmeier, Walter H. King, Herbert S. Levinson, Sven Lindqvist, Norene Dann Martin, Harry F. Orr, Chester F. Phillips, Harry B. Skinner, Jason C. Yu

Committee on Freeway Operations
Robert S. Foote, Port Authority of New York and New Jersey, chairman
Joseph M. McDermott, Illinois Department of Transportation, secretary
Patrick J. Athol, John L. Barker, Elmer N. Burns, Glen C. Carlson, E. Ryerson Case, Kenneth G. Courage, Conrad L. Dudek, Walter M. Dunn, Jr., J. Glenn Ebersole, Jr., John W. Erdman, Joseph W. Hess, E. Vinson Hoddinott, Donald L. Hoffman, Richard L. Hollinger, Daniel J. Mayer, William R. McCasland, Richard J. Murphy, Ann Muzyka, Robert A. Reiss, Frederick D. Rieger, David H. Roper, Ronald C. Sonntag, A. Taragin, Joseph Treiterer, Joseph A. Wattleworth, Jr., Paul M. Weckesser

K. B. Johns and Stephen E. Blake, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units and officers and members are as of December 31, 1977.

Contents

Alternative Objectives in Arterial-Traffic Management

Paul P. Jovanis, Institute of Transportation Studies, University of California, Berkeley
Adolf D. May, Department of Civil Engineering, University of California, Berkeley

An arterial traffic management model has been developed by modifying an existing model—TRANSYT 6. The new model—TRANSYT 6C—includes fuel-consumption and vehicle-emission impacts, spatial and modal demand responses, and an extended multiple-objective function for traffic-signal timing. Four traffic management strategies were evaluated in a variety of operating environments. The timing of traffic signals to minimize passenger delay and the use of exclusive bus lanes are the two strategies that best meet the objectives of increased passenger mobility and minimal environmental impacts. A series of sensitivity tests was conducted to determine the effects of different objective functions for traffic-signal timing. Minimizing passenger delay and minimizing fuel consumption were the most effective single-impact objective functions evaluated. Tests in which the length was varied indicate a trade-off in impacts: short cycle lengths minimize total delay and vehicle emissions, and longer cycle lengths tend to decrease priority-vehicle delay, total stops, and fuel consumption.

Transportation engineers can no longer build transportation systems by reacting to crises. They must manage the development of a transportation system by considering the optimum use of natural resources and the protection of the environment. As the construction of new facilities becomes less feasible, the better operation of existing ones is a key to better system management. Some of the questions that should be asked are, "In what way is one strategy better than another? What measures should be used in system evaluation? How does the impact selected influence the resulting management strategy? How do traffic-intensity changes affect management strategies and control parameters?"

Furthermore, the impacts may vary with time. The strategy evaluation should consider impacts at three time periods: short term (immediate), longer term (after demand shift but before demand growth), and long term (with demand growth).

Researchers at the University of California, Berkeley, have been considering these and other issues, as part of a project on managing the future evolution of the urban transportation system. One of the five task groups participating in the project—the traffic management group—has developed and tested a series of models that can be used to evaluate the impact of traffic management strategies.

One of the primary areas of concern was the improved management of arterial street operations. The major aim of the research was to improve passenger mobility while also benefiting society through lower fuel use, decreased air pollution, and more efficient travel patterns.

The development and application of a traffic model for arterial streets is given below. A series of sensitivity tests are described that examine the applicability of different strategies in varying operating environments and by using different objective functions.

ARTERIAL MODEL DEVELOPMENT

A number of existing arterial-traffic models were surveyed to determine their suitability for use as a traffic management tool. By selecting an existing model, our major effort could then be directed toward the study of impacts and demand consequences, rather than model evolution. The investigation led to the selection of TRANSYT (1) and specifically to the most recent version, TRANSYT 6.

TRANSYT 6 Model

The TRANSYT model provides a macroscopic simulation of traffic flow. Vehicles are represented as platoon shapes that change as the vehicles proceed through signals and disperse along a route. The arterial is represented as a series of nodes (intersections) connected by a series of links. The model requires, as input, the flows, journey times, saturation flows, and link lengths for all intersections in the study section. It produces, as output, traffic performance measures for each link such as time spent, distance traveled, uniform and random delay, number of stops, and percentage saturation. The individual link values are summed to arrive at system performance measures.

The model has been widely used in studies of traffic-signal timing (2). The signal splits and offsets are changed sequentially as specified by the user. The optimization is a hill-climbing technique that searches the response surface for a minimum value of a performance index. The index is the sum of delay plus a weighted value of stops. TRANSYT has been used and calibrated at several sites, which adds to its credibility as a traffic performance model and enhances its attractiveness as a traffic management tool.

In addition to its widespread use, TRANSYT has several other desirable characteristics. The model provides for representation of different classes of vehicles (buses and automobiles) that can use either common or separate stop-lines. Bus flows can be designated by specifying a bus free-flow speed and an average bus stop time on each link. The speed and stop time are used to compute an equation for bus platoon dispersal that is separate from automobiles. The program also has a feature allowing the weighting of the delay on a link by a specific multiplier. By weighting the delays by the average passenger loads of the vehicles, estimates of delay in passenger hours are obtained for bus and automobile links.

Model Modifications

The major modification of the TRANSYT 6 model proceeded in two phases. First, fuel-consumption and vehicle-emission impacts and spatial and modal demand responses were incorporated to give the TRANSYT 6B model. Second, the fuel-consumption and vehicle-emissions computations were refined and the performance index was modified to consider a broader range of variables than just delay and stops; this gave the TRANSYT 6C model.

Additional Impacts

Fuel-consumption estimates were obtained from tables developed by Claffey (3). Three classes of vehicles are considered: automobiles, gasoline-powered trucks, and diesel-powered trucks and buses. Two driving aspects are considered: in motion and stopped. Corrections are applied for roadway gradient, curvature, surface condition, and acceleration and deceleration at signals. Additional computations are made for buses to estimate the amount of fuel consumed while stopped to pick up or discharge passengers. The traffic performance measures computed by TRANSYT and an average acceleration-deceleration rate are used to enter tables to obtain fuel-consumption estimates for all links.

A simplified version of a vehicle-emissions model was developed from a report by Kunselman (4). It was assumed that the amount of a particular pollutant emitted can be computed by multiplying the emission factor for that pollutant (for each driving aspect) by the extent of driving done in each aspect. The model form is therefore

$$Q_j(R,T) = \sum_k A_{jk} D_k(R,T) \qquad (1)$$

where

$Q_j(R, T)$ = quantity of pollutant j emitted by all vehicles in region R during time period T,
A_{jk} = amount of pollutant emitted per unit of driving aspect k (emission rate), and
$D_k(R, T)$ = extent of driving in aspect k.

[The pollutants used in the model are carbon monoxide (CO), hydrocarbons (HC), and nitrous oxides (NO_x).]

Three driving aspects—cruise, idle, and acceleration and deceleration—are considered. The times in cruise and idle were computed directly from the TRANSYT output and the acceleration-deceleration value was obtained from the estimated number of stops and an average acceleration-deceleration profile. Again, special computations are performed for bus links to account for stopping to process passengers.

Demand Responses

Traffic management strategies can result in a diversion of traffic over space, mode, or time, or in a change in the rate of trip making. The inclusion of demand responses is essential to the evaluation of alternative strategies. The amount of each type of shift depends on several factors (e.g., trip purpose, change in travel time, trip length, and availability of alternative routes and modes) and can vary substantially from site to site.

Submodels for modal and spatial shifts have been added to TRANSYT. Both models use the general formula that a response is a function of a stimulus times a sensitivity. The stimulus in both cases is the change in vehicle travel time computed by TRANSYT. The sensitivity is user specified and is applied to the demand shifts. Although both spatial and modal shifts can vary, the total passenger demand is assumed to remain fixed for the arterial corridor.

The spatial-shift submodel diverts traffic from or to the arterial being studied, depending on the change in vehicle travel time that results from the management strategy. By specifying sensitivity values from zero to four, the conditions of parallel routes are changed from no parallel routes available to an unlimited number available or to those available having very large unused capacity. The spatial shift is computed iteratively until a new traffic equilibrium is reached.

Modal choice is also a function of trip characteristics, trip-maker characteristics, and transportation system characteristics. The modal-shift submodel relates modal shift to changes in in-vehicle travel time only. The model was developed from a logit formulation used in the travel demand forecasting project (5). A range of sensitivities from zero to three specify very poor to excellent bus service. Modal-shift elasticities derived from Train's data correspond to the sensitivity value.

Both spatial and modal demand-response sensitivity values were applied to groups of links called segments. The segments are defined by the user and should correspond to the average trip length in each direction along the study arterial. The sensitivities must be specified for each segment in the study section.

By varying the values of the modal and spatial sensitivities, different operating environments can be simulated. The flow and design conditions on the arterial remain the same, and the impacts can be evaluated in a variety of operating environments. The spatial shift is always performed before the modal shift because it is assumed that drivers will seek alternative routes before changing modes.

Performance Index

The performance index used to optimize signal timing was modified to include delay, stops, and fuel consumption for priority and nonpriority vehicles, and total emissions of each of the three pollutants. The complete performance-index (PI) equation is thus

$$PI = \sum_{i=1}^{n} [(k_1 d_i)_{NP} + (k_2 s_i)_{NP} + (k_3 f_i)_{NP} + (k_4 d_i)_P + (k_5 s_i)_P + (k_5 f_i)_P + k_7 CO_i + k_8 NO_i + k_9 HC_i] \qquad (2)$$

where

i = link i,
n = number of links,
$k_1, k_2, \ldots, k_9$ = weighting factors,
d_i = delay on link i (vehicle-h),
s_i = stops on link i (stops/s),
f_i = fuel consumed on link i (L),
CO_i = carbon monoxide emitted on link i (kg),
NO_i = nitrous oxide emitted on link i (kg), and
HC_i = hydrocarbons emitted on link i (kg), and

NP and P refer to nonpriority and priority vehicles respectively. The user selects the coefficient values used in the performance index. An impact may be excluded from consideration by setting its coefficient value to zero. The impacts may also be combined by selecting different weights for the coefficients. For example, a user may set signals to minimize total delay by setting k_1 and k_4 to 1 and all other coefficients to zero. Alternatively, signals may be set to minimize fuel by setting k_3 and k_6 to 1 and all other coefficients to zero.

MODEL APPLICATION

The model was applied to a 4.94-km (2.75-mile) section of San Pablo Avenue in Berkeley, California. Travel in the area is heavily automobile dependent. Bus service on San Pablo Avenue is considered very good (modal shift sensitivity of three). The very good rating implies not only that headways are short, but also that frequently used origins and destinations are served by the buses. Although the arterial is not currently congested during any part of the day, it is an important link in the local network and data for study were readily available. Previous tests of TRANSYT 6B had been performed in a heavily saturated environment (6), so a less congested facility was desirable for further study.

Establishment of Base Conditions

The evening peak hour was chosen for study. The study section was described by inputting the following data to the model: (a) bus flows average 13-18 vehicles/h in both directions; (b) bus occupancy averages 30 passengers; (c) automobile occupancy averages 1.2 passengers; (d) the vehicle mix is approximately 2 percent trucks and buses, 50 percent of which are diesel; (e) the roadway is straight and level; (f) the street carries two-way traffic, has 3 lanes in each direction on a 32.5-m (74-ft) width, and has no parking on either side; and (g) the performance index seeks to minimize delay to both priority and nonpriority vehicles. For purposes of demand response, an average trip length of 4.94 km was assumed; i.e., only two segments, one in each direction, were used for demand response. The directional split along the study section was approximately 60:40, and the predominant flow was northbound.

Traffic Management Strategies

The following strategies were tested by using the TRANSYT 6C model.

1. Traffic-signal optimization on a vehicle basis: Each automobile and bus are equally weighted;
2. Traffic-signal optimization on a passenger basis: The contribution to the performance index of the delays on a link is weighted by the average passenger occupancy for that link; thus, bus delays are weighted by 30 and automobiles delays by 1.2;
3. Exclusive bus lane—an exclusive bus lane was conceptualized in the curb lane in both inbound and outbound directions: Capacities are adjusted appropriately and after simulating conditions with the exclusive lanes, the signals are optimized on a passenger basis (as in 2 above) for the new flow conditions; and
4. Reversible lane—a reversible lane was added in the peak direction (taken away from the nonpeak direction) by adjusting the link capacities along the study section: The signals are then optimized on a vehicle basis (as in 1 above) for the new flow conditions.

Demand Response

In describing the demand-response analyses that follow, it is helpful to identify three states of traffic conditions:

1. Existing conditions—these are the conditions represented by the street design, signal timing, and flow conditions as input to the model: They are indicative of traffic flow at the time of input data collection;
2. Short term—these results represent the conditions that will exist after the management strategy has been implemented, but before any potential spatial or modal shift occurs: They may be interpreted in two ways—they will exist for the first few days of operation before a new traffic equilibrium is reached or no parallel route or alternative bus service is available and these results may remain for an extended time; and
3. Longer term—these conditions exist after demand shift occurs and a new traffic equilibrium has been established: They represent changes in travel behavior over space and mode, but total passenger demand remains fixed; in this application of the model, a spatial-shift value of four was used and a modal-shift value of three (the spatial value of four indicates that many parallel routes are available in the area, and the modal-shift value of three indicates that very good bus service is available).

All comparisons discussed in the sections below are with respect to the existing conditions. When applicable, notes will be made of comparisons between strategies.

Results of Strategies

Table 1 summarizes the four strategies tested; the results are divided into short-term and longer term consequences. The impacts that result from signal optimization are almost the same in the short term whether the optimization is on a passenger or a vehicle basis. Both strategies show small improvements in overall time spent, fuel consumption, and vehicle emissions. The major difference is the improvement in bus time spent and bus fuel consumption that results from passenger-optimized signals.

In the longer term, passenger-based optimization results in a 30 percent increase in productivity rather than the 14 percent that results from the vehicle-based signal optimization. In the passenger-optimized strategy, the arterial handles the increased traffic with only small increases in time spent, fuel consumption, and vehicle emissions. The difference in the increase in productivity between the two runs is due to the passenger weight for delay used in the passenger-based signal optimization. Because buses run only on San Pablo Avenue, the passenger-based optimization allocates more green time to San Pablo itself. The greater green time on San Pablo allows more vehicles to be added during the spatial shift than does the vehicle-based optimization (which more evenly allocates green to all approaches). In the vehicle-based signal-timing optimization, there are few permanent impacts because of the 14 percent increase in productivity. The overall trend in the longer term results is to trade quality for quantity.

The strategy of a reversible lane combined with vehicle-based signal optimization had essentially the same systemwide results as the vehicle-based signal optimization alone, considering both short-term and longer term consequences. The differences were that the productivity increase for the reversible lane was 12 percent (rather than 14 percent) and no improvement in bus fuel consumption was evident.

The exclusive bus lane, in the short term, resulted in a small decrease in passenger time spent and sizable decreases in bus time spent and bus fuel consumption. Because of the retiming of signals after inclusion of the bus lane, the other impacts remain the same as the existing conditions. This indicates that significant excess capacity was available during the existing conditions. In the longer term, fuel consumption and vehicle emissions increased by 4 percent and productivity increased 12 percent. The substantial difference between this and the other strategies is that bus time spent and bus fuel consumption maintained their short-term improvements. Very small modal shift did occur with this strategy.

In summary, in the short term, almost all strategies resulted in small improvements in travel time, fuel consumption, and air pollution. In the longer term (be-

Table 1. Impacts of strategies.

Strategy	Improvement Over Base Conditions (%) Time Spent ST	Time Spent LT	Fuel Consumption ST	Fuel Consumption LT	Vehicle Emissions ST	Vehicle Emissions LT	Modal Shift ST	Modal Shift LT	Productivity ST	Productivity LT	Bus Time Spent ST	Bus Time Spent LT	Bus Fuel Consumption ST	Bus Fuel Consumption LT
Signal optimization: vehicle basis	-2	0	-2	0	-4	0	-	0	-	+14	0	0	-2	+1
Signal optimization: passenger basis	-2	+1	-3	+3	-3	+4	-	0	-	+30	-2	0	-5	+2
Reversible lanes	-2	0	-2	0	-4	-1	-	0	-	+12	0	+1	0	0
Bus lanes	-1	0	+1	+4	+1	+4	-	+1	-	+12	-3	-4	-9	-10

Note: ST = short-term results and LT = longer term results.

cause of traffic on accessible parallel routes), traffic was attracted to San Pablo Avenue, which resulted in significant increases in productivity (12 to 30 percent) but the return of travel times, fuel consumption, and vehicle emissions to their existing-condition values. Thus, only the bus-lane strategy maintained the short-term savings in bus time spent and bus fuel consumption into the longer term.

Because they best satisfied the objectives of improved environmental conditions and increased passenger mobility, signal timing on a passenger basis and exclusive bus lanes were the strategies selected for further testing.

SENSITIVITY ANALYSES

Several sensitivity tests were performed to determine the variations in results from the base conditions. Only passenger-based signal optimization and exclusive bus lanes were considered for further testing because they best met the criteria of improved environmental impact and improved passenger mobility.

First, the signals were optimized by using the following additional objective functions: (a) priority-vehicle delay only, (b) total stops only, (c) total fuel consumed, and (d) total vehicle emissions. Second, the same series of tests was run for the exclusive-bus-lanes strategy except that priority-vehicle delay was not tested, based on the results of the first series of tests. Third, another series of tests varied the spatial and modal sensitivity values. By varying these sensitivities, the study section was conceptually moved to different operating environments. Finally, a last series of tests evaluated impacts at different levels of traffic intensity and with different cycle lengths.

Signal Optimization for Different Objective Functions

By modifying the performance index, signal settings can be optimized to minimize total passenger delay, bus-passenger delay, number of stops, fuel consumption, and vehicle emissions (or combinations thereof). The new performance index equation thus allows a rather detailed evaluation of impacts and the consequences of different impact objectives. The base conditions discussed above (that had an objective function of minimizing total passenger delay) and four additional computer runs were analyzed; the results are shown in Table 2. Only the short-term consequences were analyzed because the demand responses of the different performance indexes were not of immediate interest.

The objective functions of minimizing total passenger delay and minimizing fuel consumption appear to be the best single-impact strategies because they resulted in improvements in all impacts. The trade-off between these two objective functions is dependent on the weight given to the different impacts. Minimizing total passenger delay gives larger savings in total passenger delay and priority-vehicle delay, and minimizing fuel consumption gives larger savings in stops and small improvements in fuel consumption and vehicle emissions.

The objective functions of minimizing vehicle emissions and minimizing number of stops appear to be the next-best single-impact strategies because they result in improvements in three or four of the impacts. Minimizing vehicle emissions results in general impact improvements except for the 12 percent increase in priority-vehicle delay. This illustrates a clear trade-off between environmental issues and passenger-movement efficiency. Minimizing stops has the disadvantage of causing severe congestion on isolated links, thus increasing total delay and total vehicle emissions.

The objective function of minimizing bus-passenger delay appears of less interest because it resulted in increases in fuel consumption, vehicle emissions, and total passenger delay. Due to the structure of the study section (priority vehicles on San Pablo Avenue only), minimizing priority-vehicle delay results in the maximum amount of signal green being allocated to San Pablo Avenue. Severe penalties to cross-street traffic were the primary cause of the impact increases. Although bus delay was substantially reduced (57 percent), the objective function of minimizing bus delay was not considered suitable for the problem under study and was thus eliminated from further testing.

Bus-Lane Strategy: Different Objective Functions for Signal Timing

By modifying the performance index, the bus-lane strategy was evaluated at traffic signal settings that would minimize different impacts. The objective functions tested were (a) minimize total stops, (b) minimize total fuel consumption, and (c) minimize total vehicle emissions. The results of these three tests as well as the base condition results (minimize total delay) are shown in Table 3.

The objective functions of minimizing total delay and minimizing total fuel consumption appear to be the best single-impact objective functions for the bus-lane strategy because they both decrease three or four of the impacts. Total delay is reduced the most in the delay-based optimization, and stops and priority-vehicle delay have the largest reductions when fuel consumption is optimized. Optimizing vehicle emissions results in a small decrease in emissions and moderate decreases in other impacts; therefore, it is the second choice among the single-impact objective functions.

Optimizing stops again results in very large increases in impacts because of oversaturation on particular links of the study section. The over-saturated links experience large increases in the random component of delay. The TRANSYT model calculates stops from the uniform delay only; therefore, large increases in random delay are not reflected in higher numbers of stops. Caution must be observed if stops are used as a single-impact objective function.

Sensitivity to Different Operating Environments

To more fully evaluate the passenger-optimized signal-timing and exclusive-bus-lane strategies, tests were conducted in an additional operating environment. The new environment was simulated by changing the spatial-shift sensitivity value from four to zero (no parallel routes available) and leaving the modal-shift sensitivity at three (very good bus service). The short-term and long-term results of both strategies in the new operating environment are shown in Table 4. The short-term results are the same as the base conditions; therefore, only the longer term results are discussed.

For the strategy of optimizing signals on a passenger basis, the short-term improvements continue into the longer term. There is no productivity increase because traffic is not attracted from parallel routes (there are none). There is little modal shift because the travel time saving on the arterial for nonpriority vehicles is approximately the same as for priority vehicles. Therefore, there is no stimulus to change mode. Thus, the improvements in all impacts are extended to the longer term, but no productivity increase or modal shift is evidenced.

Table 2. Impacts of different signal-timing objectives.

Impact	Minimize Total Delay	Minimize Priority-Vehicle Delay	Minimize Stops	Minimize Fuel Consumption	Minimize Vehicle Emissions
Total delay					
Before study, passenger-h	114.42	114.42	114.42	114.42	114.42
After study, passenger-h	102.90	265.48	180.03	106.34	110.05
Difference, passenger-h	11.52	151.06	65.61	8.08	4.37
Difference, %	-10.1	+132	+57	-7.1	-3.8
Priority-vehicle delay					
Before study, passenger-h	21.47	21.47	21.47	21.47	21.47
After study, passenger-h	17.63	10.13	20.49	18.26	24.00
Difference, passenger-h	3.84	11.34	0.98	3.21	2.53
Difference, %	-17.9	-53	-4.6	-15.0	+11.8
Stops					
Before study, number/km	2.28	2.28	2.28	2.28	2.28
After study, number/km	2.08	2.28	1.92	1.97	2.04
Difference, number/km	0.20	0	0.36	0.31	0.24
Difference, %	-8.5	0	-15.7	-13.5	-10.4
Fuel consumption					
Before study, L	1179.03	313.05	1179.03	1179.03	1179.03
After study, L	1147.76	468.72	1158.97	1140.57	1153.52
Difference, L	31.27	155.67	20.06	38.46	25.51
Difference, %	-2.7	+50	-1.7	-3.3	-2.2
Vehicle emissions					
Before study, kg	313.05	1179.03	313.05	313.05	313.05
After study, kg	303.44	1204.73	367.14	302.18	299.81
Difference, kg	9.61	25.70	53.09	10.87	13.24
Difference, %	-3.1	+2.2	+17.0	-3.5	-4.2

Notes: 1 km = 0.6 mile, 1 L = 0.26 gal, and 1 kg = 2.2 lb.
All differences are absolute values; negative percentage differences indicate savings relative to existing conditions.

Table 3. Impacts of bus-lane strategies with different signal-timing objectives.

Impact	Minimize Total Delay	Minimize Stops	Minimize Fuel Consumption	Minimize Vehicle Emissions
Total delay				
Before study, passenger-h	114.42	114.42	114.42	114.42
After study, passenger-h	107.66	473.33	112.32	112.05
Difference, passenger-h	6.76	358.91	2.1	2.05
Difference, %	-5.9	313.7	-1.8	-2.1
Priority-vehicle delay				
Before study, passenger-h	1.55	1.55	1.55	1.55
After study, passenger-h	1.52	1.27	1.47	1.49
Difference, passenger-h	0.03	0.28	0.08	0.06
Difference, %	-2.0	-18.3	-5.2	-3.8
Stops				
Before study, number/km	1179.03	1179.03	1179.03	1179.03
After study, number/km	1185.42	1711.24	1176.34	1187.32
Difference, number/km	6.39	532.21	2.69	8.29
Difference, %	+0.5	+31.1	-0.2	+0.7
Fuel consumption				
Before study, L	313.05	313.05	313.05	313.05
After study, L	314.70	638.80	318.72	312.09
Difference, L	1.65	325.75	5.67	0.86
Difference, %	+0.5	+104.1	1.8	-0.3
Vehicle emissions				
Before study, kg	21.47	21.47	21.47	21.47
After study, kg	14.52	19.29	13.67	17.59
Difference, kg	6.95	2.18	7.80	3.88
Difference, %	-32.4	-10.2	-36.3	-18.1

Notes: 1 km = 0.6 mile, 1 L = 0.26 gal, and 1 kg = 2.2 lb.
All differences are absolute values; negative percentage differences indicate savings relative to existing conditions.

The bus-lane tests in the new operating environment resulted in larger savings in bus time spent and bus fuel consumption in the longer term than did signal timing alterations alone. By more efficient signal timing, all other impacts remain the same as the existing conditions despite the loss of roadway space to the exclusive use of buses. Because the non-priority-vehicle travel-time savings were comparable to the travel-time savings of priority vehicles, no modal shift occurred. In these operating environment conditions, the bus lane is a very effective strategy to aid bus operation without severe environmental impacts.

In comparison with the base conditions, the tests in the new operating environment had very different longer term results. Although there is an increase in the quantity of flow over that in the base condition, it is at the expense of a slight decrease in the quality; the new operating environment results in maintaining quality improvements but there is no increased productivity.

Sensitivity to Different Cycle Lengths and Traffic Intensities

Examination of the minimum green times required for pedestrian movements indicated that a 50-s cycle was the shortest that could be used for the study section. To select a cycle length that was equally spaced, but greater than the existing cycle length, a 90-s cycle was chosen. Each of the three cycle lengths were tested under three flow conditions: (a) existing flows, (b) existing flows increased by 20 percent, and (c) existing flows increased by 40 percent. The test results are compared for the same set of flow conditions so that the passenger kilometers of travel are the same within each group. As with other tests, all comparisons are made with respect to the existing conditions for the given flow level. At each cycle length, the signals were optimized to minimize passenger hours of delay.

1. Existing flows: The tests that used existing flows indicate a clear trade-off (see Table 5). At the shortest cycle length (50 s), total delay and vehicle emissions show the greatest savings. Fuel consumption is only slightly improved, however, and stops are slightly increased.

The 70-s cycle, in this case, has much better results for priority-vehicle delay and stops and somewhat larger savings in fuel consumption than the 50-s cycle. The 90-s cycle improved stops by 12.6 percent and fuel by 2.7 percent, but the other impacts were not improved as much as by the 50- or 70-s cycles. The results indicate that, up to a maximum, priority-vehicle delay, stops, and fuel consumption are improved by cycles that are longer than the minimum and total delay and vehicle emissions are minimized by using short cycle lengths.

2. Tests that used increased flows: Both the 20 and 40 percent increases in flow yielded very similar results when compared with the existing flow conditions (see Tables 6 and 7). The shortest cycle length yielded minimum values of total delay and vehicle emissions. The 70-s cycle yielded the best results for priority-vehicle delay and close to minimum values for fuel consumption. By using the 90-s cycle, stops and fuel consumption are minimized, but total delay, priority-vehicle delay and vehicle emissions are significantly greater than minimum. The trade-off again appears to be one of vehicle-emission savings and total delay savings at short cycle lengths versus fuel-consumption, stops, and priority-vehicle-delay savings at longer cycle lengths.

3. Summary of cycle-length and flow-intensity tests: The trends in impact improvement as traffic intensity increases are interesting. It should be noted that a 20 percent increase in all flows results in a study section that is still undersaturated. It is only when the flows are increased by 40 percent that some links approach saturation and the arterial becomes fairly congested.

One would think that signal timing would be most important when congestion is eminent. The trends shown in Table 7 support this hypothesis. There are improvements for all cycle lengths for all impacts (except stops for the 50-s cycle). At lower traffic-intensity conditions, the 90-s cycle resulted in increases in total delays and no improvement in vehicle emissions.

It is clear then that, for a wide range of traffic intensities, selection of an appropriate cycle length and timing plan can significantly change the impacts that result. Short cycle lengths result in the least delays and vehicle emissions. Longer cycle lengths decrease stops, fuel consumption, and priority-vehicle delay. At high traffic intensities, selection of the appropriate cycle length and signal-timing plan become very important.

CONCLUSIONS

Strategies

1. Signal optimization to minimize vehicle delays: In the short term, there were moderate benefits in terms of delay, fuel-consumption, and vehicle-emission savings. Buses were only marginally aided, however. The policy appears useful for decreasing overall impacts but does not result in significant direct benefits to buses. The weakness of the strategy is its emphasis on vehicle movement, rather than passenger mobility.

2. Signal optimization to minimize passenger delays: The short-term results of this strategy are generally similar to those of the vehicle-delay case. An added benefit is the additional saving in bus time spent and bus fuel consumption. Tests that used short and long cycle lengths indicate a clear trade-off common to all flow levels. Shorter cycle lengths yield larger savings in total delays and vehicle emissions, and longer cycle lengths generally yield large savings in priority-vehicle delay, stops, and fuel consumption.

The longer term results require additional discussion. Because of the improvement in time spent (2 percent), there was an increase in productivity of 30 percent. The magnitude of the productivity increase is primarily due to the undersaturated conditions that exist on San Pablo Avenue. Large increases in flow are needed to cause small increases in time spent. One can certainly question, on behavioral grounds, whether a 30 percent diversion of traffic will occur for a 2 percent saving in time spent. In this series of tests, the spatial-shift estimates appear to be too large. Tests conducted in operating environments that were more congested initially resulted in smaller spatial-shift estimates that were much more behaviorally reasonable (6).

3. Reversible lanes: There were very few direct benefits to buses under any operating conditions. The improvement in overall impacts was similar to that of vehicle-delay optimization. This strategy is not recommended for the conditions existing at the study site, but could be appropriately used in situations that

Table 4. Impact sensitivity in operating environment that has very good bus service and no alternative routes.

	Improvement Over Base Conditions (%)													
	Time Spent		Fuel Consumption		Vehicle Emissions		Modal Shift		Productivity		Bus Time Spent		Bus Fuel Consumption	
Strategy	ST	LT	ST	LT	ST	LT	ST	LT	ST	LT	ST	LT	ST	LT
Signal optimization: passenger basis	-2	-2	-3	-3	-3	-3	-	0	-	0	-2	-2	-5	-5
Bus lanes	-1	0	+1	0	+1	0	-	0	-	0	-3	-3	-9	-10

Note: ST = short-term results and LT = longer term results.

Table 5. Effects of different cycle lengths on base-flow conditions.

Impact	50-s Cycle	70-s Cycle	90-s Cycle
Total delay			
Before study, passenger-h	114.42	114.42	114.42
After study, passenger-h	91.99	102.90	118.91
Difference, passenger-h	23.43	11.52	4.49
Difference, %	-19.6	-10.1	+3.9
Priority-vehicle delay			
Before study, passenger-h	21.47	21.47	21.47
After study, passenger-h	18.90	17.63	20.38
Difference, passenger-h	2.57	3.84	1.09
Difference, %	-11.9	-17.9	-5.1
Stops			
Before study, number/km	1.73	1.73	1.73
After study, number/km	1.74	1.58	1.51
Difference, number/km	0.01	0.15	0.22
Difference, %	+0.7	-8.7	-12.7
Fuel consumption			
Before study, L	1179.03	1179.03	1179.03
After study, L	1166.65	1147.76	1146.97
Difference, L	12.38	31.27	32.06
Difference, %	-1.0	-2.7	-2.7
Vehicle emissions			
Before study, kg	313.05	313.05	313.05
After study, kg	297.86	303.44	313.25
Difference, kg	15.19	9.61	0.20
Difference, %	-4.9	-3.1	0

Notes: 1 km = 0.6 mile, 1 L = 0.26 gal, and 1 kg = 2.2 lb.
All differences are absolute values; negative percentage differences indicate savings relative to existing conditions.

Table 6. Effects of different cycle lengths for 20 percent increase in traffic flows.

Impact	50-s Cycle	70-s Cycle	90-s Cycle
Total delay			
Before study, passenger-h	147.95	147.95	147.95
After study, passenger-h	119.02	133.23	153.34
Difference, passenger-h	28.93	14.72	5.39
Difference, %	-19.5	-9.9	+3.6
Priority-vehicle delay			
Before study, passenger-h	27.63	27.63	27.63
After study, passenger-h	25.50	23.94	26.73
Difference, passenger-h	2.13	3.69	0.90
Difference, %	-7.7	-15.4	-3.3
Stops			
Before study, number/km	1.82	1.82	1.82
After study, number/km	1.84	1.70	1.61
Difference, number/km	0.02	0.12	0.21
Difference, %	+1.4	-6.5	-11.7
Fuel consumption			
Before study, L	1452.68	1452.68	1452.68
After study, L	1438.72	1416.73	1413.32
Difference, L	13.96	35.95	39.36
Difference, %	-1	-2.5	-2.7
Vehicle emissions			
Before study, kg	388.66	388.66	388.66
After study, kg	366.63	375.81	388.10
Difference, kg	22.03	12.85	0.56
Difference, %	-5.7	-3.3	+0.1

Notes: 1 km = 0.6 mile, 1 L = 0.26 gal, and 1 kg = 2.2 lb.
All differences are absolute values; negative percentage differences indicate savings relative to existing conditions.

Table 7. Effects of different cycle lengths for 40 percent increase in traffic flows.

Impact	50-s Cycle	70-s Cycle	90-s Cycle
Total delay			
Before study, passenger-h	199.92	199.92	199.92
After study, passenger-h	155.13	170.70	195.20
Difference, passenger-h	44.79	29.13	4.72
Difference, %	-22.4	-14.6	-2.4
Priority-vehicle delay			
Before study, passenger-h	36.25	36.25	36.25
After study, passenger-h	32.36	30.59	35.27
Difference, passenger-h	3.89	5.66	0.98
Difference, %	-10.7	-15.6	-2.7
Stops			
Before study, number/km	1.92	1.92	1.92
After study, number/km	1.96	1.83	1.74
Difference, number/km	0.04	0.09	0.18
Difference, %	+2.3	-4.6	-9.4
Fuel consumption			
Before study, L	1751.70	1751.70	1751.70
After study, L	1730.12	1704.16	1699.77
Difference, L	21.58	47.54	51.93
Difference, %	-1.2	-2.7	-3.0
Vehicle emissions			
Before study, kg	481.53	481.53	481.53
After study, kg	444.79	456.02	470.89
Difference, kg	36.74	25.51	10.64
Difference, %	-7.6	-5.3	-2.2

Notes: 1 km = 0.6 mile, 1 L = 0.26 gal, and 1 kg = 2.2 lb.
All differences are absolute values; negative percentage differences indicate savings relative to existing conditions.

have unbalanced flows and heavy congestion.

4. Bus lanes: For the conditions of the study site, bus lanes resulted in generally favorable results. In the short term, buses were greatly aided and retiming the signals enabled nonpriority vehicles to operate as they did in the existing conditions. There were benefits of larger savings in bus time spent and bus fuel consumption, but the environmental benefits of decreased overall fuel consumption and vehicle emissions were lost. Tests with alternative signal-timing techniques indicate that some of the overall environmental benefits may be recovered, but not as much as by using other strategies.

The longer term results of the bus-lane strategy indicate that savings to buses are maintained even though productivity on the route increases. No significant modal shift occurred in either of the two operating environments because of the lack of a travel-time stimulus.

Objective Functions in Signal Timing

1. Total delay: The objective of minimizing total delays (in passenger hours) minimized the delay for the study section and generally decreased all other impacts. Priority-vehicle delay decreased due to the passenger weights applied to the links. Because it gen-

erally results in improvements in all impacts, minimizing passenger delays is one of the two highest ranked objective functions.

2. Priority-vehicle delay: Minimizing priority-vehicle delay resulted in allocating all available green time to the study arterial. Large penalties were thus accrued to vehicles approaching on a cross street. More thorough tests should be performed by using this objective before further guidelines are recommended. Tests on arterials that have intersecting bus routes or are in a network should prove illuminating.

3. Stops: Because the TRANSYT model computes stops from uniform delays only, the use of this objective alone may result in oversaturation of some intersections; thus, its use as a single-impact objective is not recommended. Stops may, however, be used in combination with delay as in the earlier TRANSYT 6 model—this gives a better balanced objective function.

4. Fuel consumption: Using fuel consumption as a single-impact objective generally minimizes fuel consumption for the study section. Stops are nearly commensurably decreased because the amount of fuel consumed is heavily dependent on the number of stops. Both total delay and priority-vehicle delay are decreased, but not as much as in the total-delay-based optimization. Because of its overall favorable impacts, minimizing fuel consumption is the other choice for highest ranking single-impact objective.

5. Vehicle emissions: Most of the tests indicated very small changes in vehicle emissions. Part of the problem may be due to the fact that, for the test section, most vehicle emissions result from cruise, not delay. In general, this objective function is more closely related to vehicle flows than to passenger flows. Therefore, impact savings for priority vehicles are usually sacrificed for better overall vehicle performance. This objective function may be helpful in decreasing the vehicle-emission impacts of strategies in operating environments in which vehicle emissions are a severe problem.

6. Extensions: The tests of alternative single-impact objective functions are a first step in a more thorough analysis of the potentially powerful multiple-objective function. Combinations of impacts can be used, and different weights can be assigned to each. An example of this would be to assign equivalent monetary costs to each impact and then set signals to minimize a net cost.

ACKNOWLEDGMENT

This research was supported by the Office of University Research, U.S. Department of Transportation.

REFERENCES

1. D. I. Robertson. TRANSYT: A Traffic Network Study Tool. U.K. Road Research Laboratory, Crowthorne, England, Rept. LR 253, 1969.
2. A. D. May. Some Fixed-Time Signal-Control Computer Programs. Proc., Traffic Control Symposium, Monte Carlo, Monaco, Sept. 1974.
3. P. J. Claffey. Running Cost of Motor Vehicles as Affected by Road Design and Traffic. NCHRP, Rept. 111, 1971.
4. P. Kunselman and others. Automobile Exhaust Emission Modal Analysis Model. U.S. Environmental Protection Agency, Rept. 460/3-74-005, 1974.
5. K. Train. TDFP Demand Model. Travel-Demand Forecasting Project, Institute of Transportation Studies, Univ. of California, Berkeley, Internal Rept., 1976.
6. P. P. Jovanis, A. J. Kruger, and A. D. May. Analysis of Traffic Management Impacts and Demand Responses on Surface Streets and Freeways. Institute of Transportation Studies, Univ. of California, Berkeley, 1977.

Publication of this paper sponsored by Committee on Urban System Operations.

Optimization of Large Traffic Systems

Stanley B. Gershwin, Electronics Systems Laboratory, Massachusetts Institute of Technology
Paul Ross and Nathan Gartner, Office of Research, Federal Highway Administration
John D. C. Little, Operations Research Center, Massachusetts Institute of Technology

A program for optimizing most of the significant traffic engineering quantities while allowing for the fact that drivers can and do change their routes and modes of travel is proposed. The program requires the person origin-destination table as input and gives the best traffic-signal settings, reserved-lane assignments, and ramp-metering rates as outputs. The optimization can be for minimum person time, for minimum fuel consumption, or for a combination of both. The optimization was coded by using arbitrary but reasonable models for the components and run for a hypothetical network. Sample results are given. The example problems described required from 2.5 min (system optimizing) to 3 min (user optimizing) central processing unit time per complete iteration.

All of the signal-optimization methods that are now available use fixed traffic demands—i.e., they assume that changes in the signal settings will not produce significant changes in the actual traffic. This assumption is made not only in those methods that optimize isolated signals, e.g., Webster's method (1), but also in the computer methods that optimize entire networks of signals simultaneously, such as TRANSYT (2, 3, 4), SIGOP (5), and MITROP (6, 7, 8). However, it is clear that signal settings do have an effect on traffic assignments (9, 10).

The research described in this paper is an attempt to examine the feasibility of carrying traffic optimization one or two steps beyond this level and to produce a technique that will allow the traffic engineer to determine the correct signal settings, freeway ramp-metering rates, and reserved-lane assignments while taking into account the fact that automobile drivers will change their routes and passengers will change their modes of travel. Travelers can also change their trip frequencies and destinations in response to traffic conditions and traffic engineers can make other changes to the system (such as intersection improvements), but we have not included such factors in the optimization model. The procedure we describe here is similar to others (11, 12); however, we provide numerous options and we believe our traffic equilibrium technique may be more accurate. Other authors have combined demand studies with assignment (13, 14, 15) and with modal split (16, 17, 18).

This research is an extension of that reported elsewhere (19). It is only a feasibility study, not a finished and polished program; the intent was to determine whether such large-scale, comprehensive optimizations are possible, not to actually produce a consumer product. Because of this emphasis on the development of a pilot program, there are a few caveats that should be remembered when examining the numerical results:

1. The component submodels for traffic assignment, signal optimization, and modal split have not been calibrated or validated. The models that were used are reasonable and have been fairly well justified in the literature but they are not meant to be authoritative; research is continuing in all these areas.
2. The computer code used for the numerical examples is not in a polished and finished form. The coding was done to examine the feasibility of the concept; it does not operate as smoothly and conveniently as a finished program would and may even contain minor bugs.

OVERALL OPTIMIZATION PLAN

The overall optimization plan is a fairly simple iteration among three separate subprograms—traffic assignment, signal optimization, and modal split. These three subprograms can use adaptations of established and well-known computer packages [such as UROAD (20) for the traffic assignment and TRANSYT for the signal optimization]. However, because of the iterative nature of the optimization, it is important that these routines execute very quickly; therefore, they should be written especially for this use. In the experiments carried out here, specially written subprograms were used, except that the MITROP signal-optimization program was also used in some runs.

Traffic Assignment

The most crucial element of the optimization is the traffic-assignment program. In addition to the usual feature of estimating the average traffic on all links of a network from fixed origin-destination (O-D) tables, the traffic assignment must be able to make these assignments on the basis of two different criteria—system optimization and user optimization. System-optimized traffic means that all traffic takes the route that is best for the system as a whole. That is, some drivers go out of their way and take longer or more expensive (or both) routes to avoid congesting certain streets so that other drivers will have much shorter or cheaper (or both) trips. In this way, the total time spent by all travelers in the system is a minimum, but such altruism is rare among real automobile drivers. Nevertheless, it is important that the traffic-assignment components have the ability to make such assignments for reasons that will be explained below.

User-optimized traffic assignments are the products of programs that assume that each individual driver does as well as he or she individually can; i.e., any other route that can be chosen takes at least as long as the one actually chosen. It follows that all routes that carry any traffic between a given O-D pair have the same trip times and the routes that are not used take even longer. There is reason to believe that actual traffic approximately distributes itself according to this user-optimization principle. The principles of user and system optimization, in the context of traffic, are described by Wardrop (21).

Signal Optimization

This traffic-assignment package can be combined with a signal-optimization package to optimize the signal settings in a traffic network while allowing for the ability of traffic to change its route distribution. In Figure 1, the traffic-assignment program shown in box 2 iterates back and forth with the signal optimization in box 3 until the traffic flows cease to change appreciably from iteration to iteration and the stopping criterion (box 4) is satisfied. The signal-optimization program should accept the average traffic volume on the streets as input and produce the best signal settings as output. In addition, it should provide output that allows the traffic-assignment program (box 2) to reflect the effects of signal synchronization and coordination. Of the available signal-network optimization packages, only MITROP has this ability.

Modal Split

Travelers will change their mode of travel if they perceive a benefit in doing so. In stochastic systems such as traffic, this perception takes a relatively long time but modal shift is nevertheless a real phenomenon. Changes in the mode of travel can be included in this optimization by adding another interation loop, as shown in Figure 2. Box 2 represents the entire program of Figure 1; it requires only an O-D table (broken down by mode of travel) and it gives, as output, the correct traffic flows, signal settings, and travel times for each O-D pair by mode. (We assume that the traffic-assignment subprogram can recognize several different kinds of traffic—e.g., automobiles, car pools, and buses—and evaluate the total travel time associated with trips by each mode.) Box 3, the modal-split determination, uses these travel times for each mode to estimate what fraction of persons will travel by each mode for each O-D pair. Any suitable program will serve for the modal-split module (box 3) if it can take the information provided by box 2 (travel times by mode for each O-D pair) and produce a reasonable estimate of the fractions of persons choosing each mode.

This three-way iteration converges to an optimal solution, but we are unable to prove whether this optimum is global. In the test cases, convergence was quite rapid at first, coming to within 1 or 2 percent of the eventual modal split in two or three iterations, but after that was relatively slow.

Florian (22) has solved the modal-split problem (for two modes, automobile and public transit, only) by a different iterative procedure. He guesses the transit travel times and solves an elastic demand problem in the automobile sector. This leads to new transit travel

Figure 1. Flow chart for calculation of signal settings and assignments.

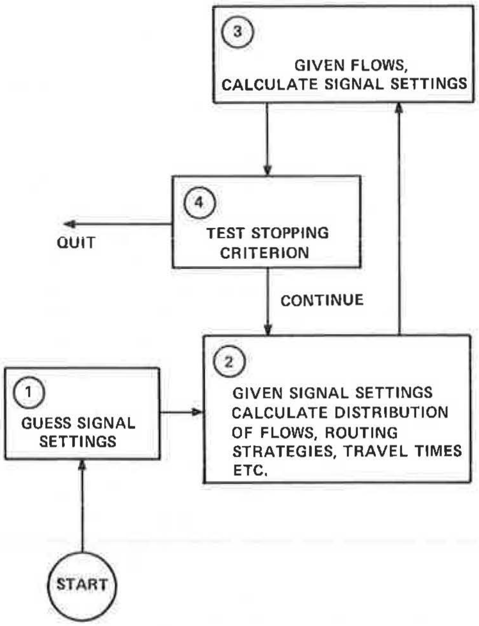
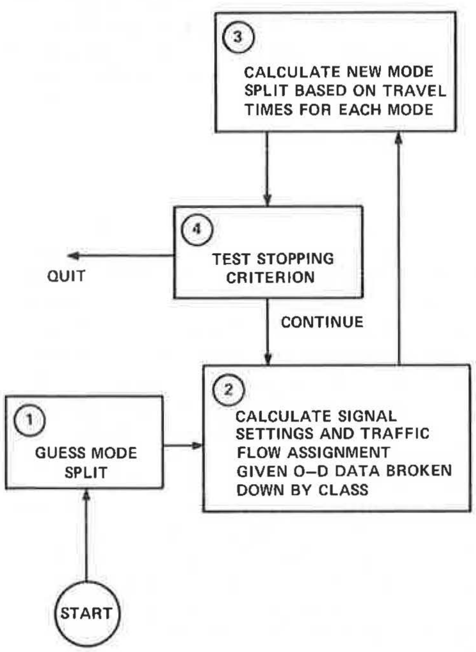


Figure 2. Flow chart for calculation of modal split.



times, and the procedure is then repeated.

Reserved Lanes

If, for example, we run the optimization of Figure 1 by using the traffic assignment in the system-optimizing mode, the program might assign only buses to a certain street and require automobiles and car pools to take longer routes to their destinations. This might be the only way to achieve the minimum person delay in the system. Of course, in real, user-optimizing traffic, ordinary automobiles would crowd onto this street and degrade the overall performance of the system unless the street was reserved exclusively for buses.

Lane reservations are made by an extension of this principle. The network geometry must be entered into the computer such that each lane that is a candidate for reserved status is coded as a separate link and suitable nodes are located so that traffic can enter and leave at reasonable locations. A link-and-node representation of a hypothetical freeway corridor is shown in Figure 3. Link 36, which joins nodes 15 and 29, is a (possibly) reserved lane. Traffic entering the freeway at node 16 cannot cross the regular lanes (link 37) until further downstream at node 29. Computer runs are made by using the traffic assignment in both the user- and system-optimizing modes, and the lanes where a significant change in the mix of traffic occurs are reserved for the traffic component that dominates in the system-optimized mode. (This is not a true optimization procedure; it only indicates that the option should be evaluated.) This comparison could be automated, but there are so many unquantifiable value judgments involved that we have chosen to leave the lane-reservation process as a manual judgment based on the program output and local knowledge.

Ramp Metering

The determination of optimal ramp-metering rates requires a compound optimization procedure of the type described by Gartner and others (23).

DETAILS OF ACTUAL COMPONENTS

To test computationally the concepts outlined above, we have constructed actual computer programs. Because this research was in the nature of a pilot study, no great effort was made to achieve ultimate accuracy and perfection in the programming but all components were at least reasonable. The parameters used were, in some cases, only reasonable guesses.

Traffic Assignment

The traffic-assignment subprogram is the heart of the optimization; it contains the traffic model and an optimization subroutine.

The traffic model must give the average travel time on each link while allowing for the presence of several different types of vehicles and several different kinds of streets (e.g., arterials, freeways, and freeway ramps). If the traffic model gives these travel times accurately and simply, the optimization portion will have little difficulty in assigning traffic to the various streets so that the total objective function is a minimum (system optimum) or the trip times by alternative paths are equal (user optimum). Because this traffic-assignment subprogram is used many times during the course of the overall program, it is important that the optimization routine operate with exceptional speed.

The traffic model itself has two components: the constraints and the objective function. The constraints describe the relations among the variables of the system. In the model developed for this study, the constraints were

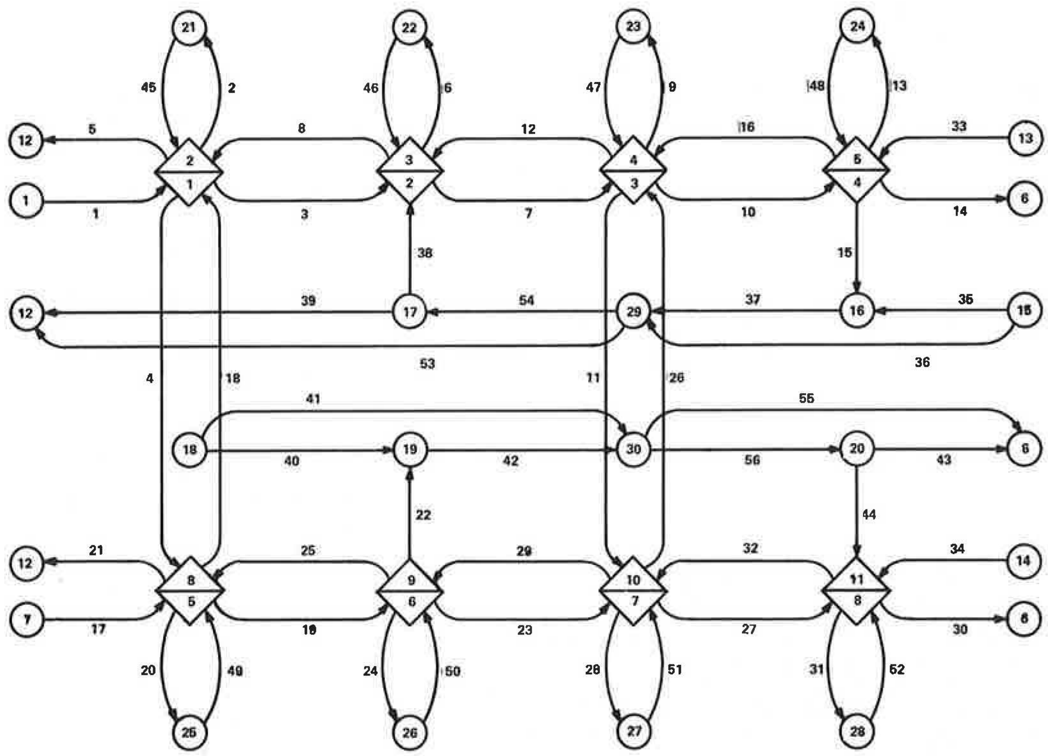
1. Nonnegative flows

$$\phi_{ij}^{(n)} \geq 0 \qquad (1)$$

for all i, j, n;

2. Net outflow from source node k equals O-D volume

Figure 3. Network that has freeway lanes separated.



$$\sum_{\substack{\text{links } \beta \\ \text{leaving} \\ \text{node k}}} \phi_{\beta j}{}^{(n)} - \sum_{\substack{\text{links } \alpha \\ \text{entering} \\ \text{node k}}} \phi_{\alpha j}{}^{(n)} = r_{kj}{}^{(n)} \qquad (j \neq k); \tag{2}$$

3. Net inflow to destination node k equals the sum of attractions

$$\sum_{\substack{\text{links } \alpha \\ \text{entering} \\ \text{node k}}} \phi_{\alpha k}{}^{(n)} = \sum_{m \neq k} r_{mk}{}^{(n)} \tag{3}$$

and, at transfer node k, net outflow equals net inflow

$$\sum_{\substack{\text{links } \beta \\ \text{leaving} \\ \text{node k}}} \phi_{\beta j}{}^{(n)} - \sum_{\substack{\text{links } \alpha \\ \text{entering} \\ \text{node k}}} \phi_{\alpha j}{}^{(n)} = 0 \qquad (j \neq k); \tag{3a}$$

4. Total passenger flow (P_i) defined on link i

$$P_i = \sum_j [\phi_{ij}{}^{(1)} + 2.5\, \phi_{ij}{}^{(2)} + 25\, \phi_{ij}{}^{(3)}]; \tag{4}$$

and

5. Total passenger-automobile-equivalent volume defined on link i

$$\phi_i = \sum_j [\phi_{ij}{}^{(1)} + \phi_{ij}{}^{(2)} + 3\phi_{ij}{}^{(3)}] \tag{5}$$

where

$\phi_{ij}{}^{(1)}$ = passenger-automobile flow (vehicles/h) on link i destined for node j,
$\phi_{ij}{}^{(2)}$ = car-pool flow (vehicles/h),
$\phi_{ij}{}^{(3)}$ = bus flow (vehicles/h), and
$r_{kj}{}^{(n)}$ = number of vehicles per hour of mode n originating at node k and destined for node j.

In constraint 4, we have assumed that private passenger automobiles carry one occupant, car pools average 2.5 occupants, and buses average 25 passengers. In constraint 5, we have assumed that buses have the same effect on traffic as three ordinary automobiles or car pools.

Objective Function: Time

The objective function (T) is the sum of the person travel times on all links

$$T = \sum_i P_i \tau_i \tag{6}$$

where P_i = total number of persons per hour using link i as defined by constraint 4 and τ_i = average travel time on link i. Different formulas were used for these travel times depending on whether link i represented a freeway, a freeway entry ramp, or an arterial street.

On freeway links, the average travel time depended only on the congestion on the link

$$\tau = l[1/v_0 + (0.3k_{max} - 1/v_0)(\phi/s)^6] \tag{7}$$

where

l = length of link,
k_{max} = jam density,
s = capacity, and
v_0 = free speed on link.

On freeway entrance ramps, there is one delay at the metering signal and another at the point of merge with the freeway. A reasonable model for the metering delay can be derived from the theory for (M/D/1) queues as

$$[2 - (\phi/M)]/2M[1 - (\phi/M)] \tag{8}$$

where M = metering rate and ϕ = link flow (24). (This metering delay was not included in the computer program whose results are described below.)

Merge delay occurs on the ramp when an automobile waits for an acceptable gap in the freeway flow. It is reasonable to assume that the automobiles on the ramp wait for a gap of length D; thus, we can calculate the average merge delay from the queuing theory for (M/M/1) queues as

$$\{\phi'/[\exp(\rho'D) - 1] - \phi\}^{-1} \quad (9)$$

where ϕ' = freeway curb-lane flow that inhibited the merge process and ρ' = vehicular density in that lane. If we combine transit time, metering delay, and merge delay, the average travel time on a freeway entrance ramp is

$$\tau = t_0 + \{[2 - (\phi/M)]/2M[1 - (\phi/M)]\} + \{[\phi'/\exp(\rho'D) - 1] - \phi\}^{-1} \quad (10)$$

On arterials, there are significant delays at signals. The results of the optimization will be most accurate if the arterial-delay function includes, in detail, the effects of cycle length, split, and offset, but this accuracy will be at the expense of a considerable number of iterations between the traffic-assignment package and the signal-optimization package. At the other extreme, if the arterial-delay function ignores these variables, only one iteration will occur in each step and the overall program will execute with considerably increased speed. Although it is not absolutely necessary, logical consistency suggests that this arterial-delay function in the traffic-assignment module should be the same as that used in the signal-optimization module.

We used two different levels of detail in the arterial-delay function. The simpler arterial-delay function was taken from Webster:

$$\tau = t_0 + 0.45\left(\{c(1 - g)^2/[1 - (\phi/s)]\} + [\phi/gs(gs - \phi)]\right) \quad (11a)$$

where c = signal cycle length and g = fraction of the cycle that the link experiences a green display. The use of this function implies local optimization only, i.e., at each signal independently.

The more detailed arterial-delay function was chosen to be compatible with MITROP. (The MITROP program uses piecewise linear approximations to this curve.) This function takes account of the network effects of signal synchronization and coordination.

$$\tau = t_0 + \{\gamma^2/2p[1 - (\phi c/ps)]\} + \phi/(gs - \phi) \times \{1.25[(gs - \phi)/g^2s^2c]^{\frac{1}{2}} + 2.25[(\phi/gs) \times (gs - \phi)/g^2s^2c] + 0.008(\phi/gs)^3\} \quad (11b)$$

where γ = platoon arrival time at the downstream signal and is a function of the offset on the link and p = platoon length on the link. Both of these values are available from the MITROP signal-network-optimization program; this function could not be used if the values of γ and p were not produced by the signal-optimization program.

Objective Function: Fuel Consumption

We included fuel consumption in the system optimal objective function; vehicle emissions can be included in a similar manner. The total objective function was written in the form

$$W_tT + W_fF \quad (12)$$

where

W_t and W_f = user-specified parameters,
T = total travel time in the system as defined in Equation 6, and
F = total fuel consumed in the system.

The total fuel is the sum of the fuel consumed on the different links where, again, different formulas must be used to evaluate the fuel consumed on different kinds of links.

On freeways, the fuel consumed on link i (ξ_f) was taken to be

$$\xi_f = l[\phi^{(1)}G^{(1)}(v) + \phi^{(2)}G^{(2)}(v) + \phi^{(3)}G^{(3)}(v)] \quad (13)$$

where $G^{(n)}(v)$ = fuel consumption (L/km) of vehicles of mode n for the average speed v on the link. The fuel-consumption functions have been investigated by Claffey (25). We have used polynomial approximations to his data:

$$G^{(1)} = G^{(2)} = 0.285\,468 - (9.852\,41 \times 10^{-3})v + (1.864\,59 \times 10^{-4})v^2 - (1.465\,63 \times 10^{-6})v^3 + (4.413\,77 \times 10^{-9})v^4 \quad (14a)$$

$$G^{(3)} = 0.304\,520 - (1.229\,78 \times 10^{-2})v + (3.010\,67 \times 10^{-4})v^2 - (2.736\,72 \times 10^{-6})v^3 + (9.507\,19 \times 10^{-9})v^4 \quad (14b)$$

Because Claffey did not publish results for buses, Equation 14b was fitted to his results for two-axle, six-tired trucks.

On freeway entrance ramps, the principal effects are decelerating to a stop and then accelerating back to speed plus the idling time at the metering signal and the merge point. The total fuel consumed on entry ramps (ξ_{er}) was taken to be

$$\xi_{er} = \sum_n \phi^{(n)}\left[H^{(n)}(v) + I^{(n)}\{[2 - (\phi/M)]/2M[1 - (\phi/M)]\} + \left(I^{(n)}\{\phi'/[\exp(\phi' t^*) - 1]\} - \phi\right)^{-1}\right] \quad (15)$$

where the fuel consumed during the deceleration-acceleration cycle was

$$H^{(1)} = H^{(2)} = -1.827\,92 \times 10^{-2} + (1.862\,66 \times 10^{-3})v - (3.646\,46 \times 10^{-5})v^2 + (3.910\,79 \times 10^{-7})v^3 - (1.433\,24 \times 10^{-9})v^4 \quad (16a)$$

and

$$H^{(3)} = -4.008\,24 \times 10^{-5} + (3.433\,29 \times 10^{-4})v + (1.068\,39 \times 10^{-5})v^2 + (7.827\,85 \times 10^{-8})v^3 - (1.766\,52 \times 10^{-9})v^4 \quad (16b)$$

where $H^{(n)}(v)$ = fuel consumption (L/km) of vehicles of mode n for the average speed v during the deceleration-acceleration cycle and the fuel consumed during idling was

$$I^{(1)} = I^{(2)} = 2.19 \text{ L/h} \quad (17a)$$

and

$$I^{(3)} = 2.46 \text{ L/h} \quad (17b)$$

On signalized arterials, we took the fuel consumption (ξ_a) to be

Figure 4. Person O-D table for example problems.

	TO 6	12	21	22	23	24	25	26	28
FROM 1	870								225
7	1005					135			
13		420		240			510		
14		579	112					135	
15		5250					135	135	
18	5850								
21	537		---						225
22				---			135		
23					---				180
24		315				---			
25	372				135		---		
26	315					90		---	
27				135					
28					90				---

$$\xi_a = (k_1 l + k_2 \tau)\,[\phi^{(1)} + \phi^{(2)} + 1.523\phi^{(3)}] \quad (18)$$

where k_1 = 0.1020 L/vehicle-km and k_2 = 0.000976 L/vehicle-s. This is based on the results of Evans and Herman (26). The 1.523 factor for buses is $G^{(3)}$ (48 km/h)/$G^{(1)}$ (48 km/h).

Optimization Portion of Traffic Assignment

We used the Cantor-Gerla algorithm (27) for the optimization portion of the test traffic-assignment routine; other traffic-assignment optimization routines, such as those based on the Frank-Wolfe method or the convex-simplex method, could also be used. [These methods have been reviewed by Gartner (28).] An experiment that used a novel extension of user optimization to nonseparable cost functions had promising results. The master step in the system-optimization algorithm can be thought of as finding a solution to a set of nonlinear equations and inequalities—the Kuhn-Tucker conditions (29). Suppose instead, we satisfy a different set of nonlinear equations and inequalities: We find flows to equate travel times on paths defined by the extremals already generated. If this procedure converges, it creates flows that equate travel times on paths used between origins and destinations—the user-optimization solution.

Signal-Optimization Package

We incorporated the MITROP optimization program into some of the runs; all components operated well, and good signal progressions were established. However, most of the computational work used a very simple signal-optimization procedure. Green time was assigned to each signal approach in proportion to the degree of saturation on that approach. This ignores the effects of cycle length, signal offset, and platoon formation but, nevertheless, provides a fast and easy solution; it is adequate when the signals are far enough apart so that platoons dissipate in traveling between them (30, 31). It is, of course, inadequate where the effects of signal synchronization and coordination are important. Because, in this method, the concepts of offset and platoon length are not used, values for the γ and p variables required in Equation 11b are not available; only the Webster delay formula (Equation 11a) could be used. All the results reported in this paper use the Webster delay function; cycle lengths were fixed at 60 s.

Modal Split

There does not appear to be a consensus as to what factors affect the choice of mode of travel. In the absence of such a consensus, it is proper to choose the simplest reasonable way of representing modal choice. We chose the well-known logit model (22, 32)

$$r_{ij}^{(n)} = [R_{ij}/w^{(n)}]\exp[-\alpha^{(n)} - \beta^{(n)}T_{ij}^{(n)}] \Big/ \sum_n \exp[-\alpha^{(n)} - \beta^{(n)}T_{ij}^{(n)}] \quad (19)$$

where

Mode	α	β
Automobile	0.00	0.010
Car pool	0.15	0.011
Bus	0.25	0.013

R_{ij} = total number of persons that will travel from origin i to destination j by all modes (the person O-D table),

$w^{(n)}$ = average occupancies of the modes (taken as 1, 2.5, and 25 in these examples and in constraint 4), and

$T_{ij}^{(n)}$ = trip time (seconds from origin i to destination j) via mode n.

The α and β parameters are arbitrary choices that seemed reasonable.

This closed form for the modal-choice function is attractive because of its computational efficiency.

RESULTS OF SAMPLE RUNS

To demonstrate the feasibility of the concept of optimization of traffic systems, the test programs described above were run by using the network shown in Figure 3 and the person O-D data shown in Figure 4. Both the network and the O-D table are hypothetical constructs chosen to illustrate typical conditions.

Optimization in the user-optimizing mode in which persons are free to change mode of travel approximates the way real traffic would occur in the network. If the hourly traffic volumes are as shown, 553.36 person-h would have been spent in the network and 2323 L of gasoline would have been consumed each hour.

The same O-D table was also run in the fuel system-optimizing mode with free modal choice. That is, each person chose his or her mode of travel by its apparent utility to himself or herself but took the best route to conserve fuel. Although real people do not choose their routes this way, it provided a sort of limit on the possible fuel consumption and served as the basis from which reserved lanes were determined. In this run, 2155 L of fuel were consumed; no combination of traffic improvements alone would be likely to reduce long-term fuel consumption below this value. To achieve lower fuel consumption, either coercive travel restrictions or changes in the average fuel consumption of vehicles or changes in the public perception of the utility of different modes would be required.

Examination of the traffic volumes in this system-optimal assignment showed that buses were not being assigned to the outer two lanes of some freeway links and that automobiles and car pools were restricted from the inner lane of the freeways. One freeway entrance ramp was being used only by buses, and the other was not being used at all. Consequently, a third run was made in which drivers were allowed to choose their own routes (user optimum), but there were extensive restrictions on what types of vehicles could use which links. Only buses were allowed on the innermost freeway lane (links 36, 41, 53, and 55). Entrance ramp 22 was closed to all traffic, and entrance ramp 15 was closed to all but buses. Buses were prohibited from the outer lanes on freeway links 35 and 39 because they were expected to use the inner lane (links 36 and 53). Modal splits were fixed at the values that resulted from the run that represented the situation that would be expected just after these lane restrictions were introduced and before travelers had an opportunity to change modes. Quite dramatic fuel savings resulted from the new traffic restrictions; even some travel time would have been saved as indicated below (1 L = 0.264 gal).

Run Conditions	Person-Time in Network (person-h/h)	Total Fuel Consumed (L/h)
Present value (user optimum, no reserved lanes, modal shifts allowed)	553.36	2323
Best possible fuel consumption; no changes in network (system optimal, fuel basis, modal shift allowed)	507.33	2155
Expected after reserved lanes introduced (user optimal, no modal shift)	530.42	2084
Expected eventually (user optimal, modal shift allowed)	507.69	2151

Travelers can be expected to continue to use their old mode of travel only until they perceive an advantage in changing. We have assumed that, after lane restrictions are instituted, it will take only a few days for the traffic to establish new routes but several months for the travelers to arrive at a new modal-choice equilibrium. Because the mode choices are based on personal utility, they will not necessarily improve overall fuel use. In fact, when we allowed travelers to choose their new modes of travel in the last run, the total fuel consumption returned to the optimal value from the second run. (The difference between these values is within the limits of the convergence tests; i.e., they are essentially equal.) In general, we would expect the new fuel consumption with lane restrictions to be above the optimum value; the lane restrictions used here turned out to be unexpectedly efficacious.

The differences between the third and fourth runs appear to be counterintuitive. When the reserved bus lanes are first introduced (the third run), in effect, a large part of the corridor capacity is lost; the most pronounced effect is that automobiles on the freeway are forced to move at slow, fuel-efficient speeds.

Many travelers whose trips lie along the length of the corridor switch to buses, which reduces the freeway congestion somewhat in the fourth run. The freeway itself becomes less congested, and automobiles move at higher, less efficient speeds; however, the shift to buses produces a small net saving of fuel on the freeway itself, although elsewhere the network remains congested and both automobiles and buses moving across the corridor are delayed. When both automobiles and buses are delayed, there is a strong shift from buses to automobiles. The net effect for the whole network is a decrease of 8 buses/h; this increased use of automobiles for trips across the corridor more than cancels the fuel savings along the corridor.

It is not meaningful to compare the eventual fuel consumption with the artificially low fuel consumption of the third run. That consumption could be maintained only if travelers were enjoined from changing modes of travel. The correct comparison is between the first and fourth runs and shows that the introduction of bus lanes has saved not only fuel but also time.

CONCLUSIONS AND PROSPECTS

This project has shown that a partial traffic-engineering optimization of traffic systems is feasible. Such a computer program should be quite useful for indicating the overall effects of contemplated changes and for highlighting points where improvements would be most effective. We feel that this has been the first attempt to look at second-order effects in the traffic system.

Vehicle emissions could be evaluated by methods analogous to those outlined here for fuel consumption. The resulting program would evaluate the total person time, fuel consumed, and vehicle emissions produced in a network and indicate what can be done to reduce any one or any combination of these factors (27, 33).

The algorithms that were used in this study seem suitable for large-scale use, but careful attention to clean, error-free programming and the minimization of execution time and memory are required. The utility of the Cantor-Gerla algorithm and the overall iterative scheme has been shown, but there are both obvious and subtle ways to improve their efficiency. (For example, the first few iterations—those through the traffic assignment and the signal optimization—require only approximate solutions, but the programs, nevertheless, produce fully accurate results based on the current approximate input. Much quicker, less accurate functions could be used for these first few iterations.) The example problems described in this paper required from 2.5 min (system optimizing) to 3 min (user optimizing) central processing unit (CPU) time per complete iteration. (When there are lane restrictions, as used in the second and third runs, the network is considerably simpler, and a complete modal-split iteration required only 0.38 min of CPU time.) We believe that clean and efficient programming will reduce these running times by an order of magnitude, which will make it possible to analyze corridors that have approximately 150 intersections and 500 links within reasonable computer time and storage requirements (perhaps three complete iterations in 30 min of CPU time and 250 K of core storage on an IBM 360/65).

REFERENCES

1. F. V. Webster. Traffic Signal Settings. U.K. Road Research Laboratory, Crowthorne, England, Technical Paper 39, 1957.
2. D. I. Robertson. TRANSYT: A Traffic Network Study Tool. U.K. Road Research Laboratory, Crowthorne, England, Rept. LR 253, 1969.
3. D. I. Robertson. TRANSYT Method for Area Traffic Control. Traffic Engineering and Control, Vol. 2, No. 6, Oct. 1969.
4. F. A. Wagner. TRANSYT Method for Network Optimization. Proc., ITE, 42nd Annual Meeting, Sept. 26, 1972.
5. SIGOP: Traffic Signal Optimization Program—A Computer Program to Calculate Optimum Coordination in a Grid Network of Synchronized Traffic Signals. Traffic Research Corporation, New York, Sept. 1966; NTIS, Springfield, VA, PB 173 738.
6. N. H. Gartner, J. D. C. Little, and H. Gabbay. MITROP: A Computer Program for Simultaneous Optimization of Offsets, Splits, and Cycle Times. Traffic Engineering and Control, Aug.-Sept. 1976, pp. 355-359.
7. N. H. Gartner, J. D. C. Little, and H. Gabbay.

Optimization of Traffic Signal Settings by Mixed-Integer Linear Programming—Part 1: The Network Coordination Problem and Part 2: The Network Synchronization Problem. Transportation Science, Vol. 9, 1975, pp. 321-363.
8. N. H. Gartner, J. D. C. Little, and H. Gabbay. Simultaneous Optimization of Offsets, Splits, and Cycle Time. TRB, Transportation Research Record 596, 1976, pp. 6-15.
9. M. J. Maher and R. Akcelik. The Redistributional Effects of an Area Traffic Control Policy. Traffic Engineering and Control, Sept. 1975.
10. N. H. Gartner. Area Traffic Control and Network Equilibrium. In Traffic Equilibrium Methods (M. A. Florian, ed.), Springer-Verlag, New York, 1976.
11. J. A. Charlesworth. The Calculation of Mutually Consistent Signal Settings and Traffic Assignment for a Signal-Controlled Road Network. Proc., 7th International Symposium on Transportation and Traffic Theory, Kyoto, Japan, July 1977, pp. 545-561.
12. R. E. Allsop and J. A. Charlesworth. Traffic in a Signal-Controlled Road Network: An Example of Different Signal Timings Inducing Different Routings. Traffic Engineering and Control, Vol. 18, May 1977, pp. 262-264.
13. M. Beckman, C. B. McGuire, and C. Winston. Studies in the Economics of Transportation, Yale Univ. Press, New Haven, CT, 1956.
14. M. Florian and S. Nguyen. A Method for Computing Network Equilibrium With Elastic Demands. Transportation Science, Vol. 8, No. 4, Nov. 1974.
15. M. Florian, S. Nguyen, and J. Ferland. On the Combined Distribution-Assignment of Traffic. Transportation Science, Vol. 9, No. 1, Feb. 1975.
16. M. R. Wigan and T. J. G. Bamford. An Equilibrium Model of Bus and Car Travel Over a Road Network. U.K. Transport and Road Research Laboratory, Crowthorne, England, Rept. LR 559, 1973.
17. J. C. Yu and R. K. Giguere. Optimizing Transportation Facility Arrangements. Transportation Engineering Journal, Nov. 1973.
18. D. Ghahrman, A. K. C. Wong, and T. Au. Perturbations of a Multimodal Network Model for Urban Transportation Planning. Trans., IEEE, Systems, Man, and Cybernetics, Vol. SMC-4, No. 3, May 1974.
19. S. B. Gershwin, J. D. C. Little, and N. H. Gartner. Computer Assisted Traffic Engineering Using Assignment, Optimal Signal Setting, and Modal Split. Research and Special Programs Administration, U.S. Department of Transportation, Rept. DOT-TSC-RSTA-78-10, 1978.
20. E. R. Ruiter. Network Equilibrium Capabilities for the UMTA Transportation Planning System. In Traffic Equilibrium Methods, Montreal, 1974, Springer-Verlag, Berlin, 1976, pp. 183-199.
21. J. G. Wardrop. Some Theoretical Aspects of Road Traffic Research. Proc., Institute of Civil Engineers, Pt. 2, 1952.
22. M. Florian. A Traffic Equilibrium Model of Travel by Car and Public Transit Modes. Transportation Science, Vol. 2, No. 2, May 1977, pp. 166-179.
23. N. H. Gartner, S. B. Gershwin, and J. D. C. Little. Supply Modeling for Urban Traffic Management. Paper presented at the International Symposium on Travel Supply Models, Univ. of Montreal, Nov. 1977.
24. H. A. Taha. Operations Research: An Introduction. Macmillan, New York, 1971.
25. P. J. Claffey. Running Costs of Motor Vehicles as Affected by Road Design and Traffic. NCHRP, Rept. 111, 1971.
26. L. Evans and R. Herman. A Simplified Approach to Calculations of Fuel Consumption in Urban Traffic Systems. Traffic Engineering and Control, Aug.-Sept. 1976.
27. D. G. Cantor and M. Gerla. Optimal Routing in a Packet-Switched Computer Network. Trans., IEEE, Computers, Vol. C-23, No. 10, Oct. 1974, pp. 1062-1069.
28. N. H. Gartner. Analysis and Control of Transportation Networks by Frank-Wolfe Decomposition. Proc., 7th International Symposium on Transportation and Traffic Theory, Kyoto, Japan, July 1977, pp. 591-623.
29. H. W. Kuhn and A. W. Tucker. Nonlinear Programming. Proc., 2nd International Berkeley Symposium on Mathematical Statistics and Probability (J. Neyman, ed.), Univ. of California Press, Berkeley, 1951.
30. F. V. Webster and J. A. Hillier. Traffic Control Systems. In Traffic Engineering Practice (E. Davies, ed.), Spon's Civil Engineering Series, London, 1968.
31. J. E. Baerwald. Traffic Engineering Handbook. Institute of Traffic Engineers, 1965.
32. P. R. Stopher and A. H. Meyburg. Urban Transportation Modelling and Planning. Lexington Books, Lexington, MA, 1975.
33. S. L. Cohen and G. Euler. Signal Cycle Length and Fuel Consumption and Emissions. TRB, Transportation Research Record 667, 1978, pp. 41-48.

Publication of this paper sponsored by Committee on Urban System Operations.

Public Reaction to Priority Lane for Buses and Car Pools in Miami

Kenneth G. Courage, Thomas H. Culpepper, * Charles E. Wallace, and Joseph A. Wattleworth, Transportation Research Center, University of Florida

A series of user surveys was conducted in conjunction with an evaluation of a bus-and-car-pool priority lane on I-95 in Miami. The groups surveyed included (a) express bus passengers, (b) express bus drivers, (c) car poolers, (d) other I-95 motorists, and (e) motorists using the I-95 corridor. Analyses of the survey results indicated that the bus-and-car-pool priority-lane concept was well received by all road-user groups. Despite some concerns about specific operational or physical aspects of the system, no group expressed dissatisfaction with it, and those who were qualified to use the priority lane were strongly supportive. Specific survey findings included the following: (a) more than half of all groups (50 to 94 percent) felt that the priority-lane system should remain in operation, (b) the existence of the priority lane was ranked second in importance among all system features by both bus passengers and car poolers, (c) all groups indicated a high degree of concern over use of the lane by nonqualified motorists, and (d) the existence of the priority lane was well established in the minds of all road-user groups surveyed, but the recognition and understanding of the reserved-lane diamond symbol was low.

Peak-period traffic congestion is a problem common to most urban areas in the United States and, because of growing public concern over the environmental and social impacts of roadway construction, transportation engineers have been exploring ways to reduce this congestion by increasing the efficiency of existing facilities. One aspect of this approach is the use of incentives to travel in high-occupancy vehicles (HOVs), which can reduce the congestion on and increase the passenger-carrying capability of a facility. Typical of these incentives is preferential treatment for HOVs in the form of exclusive rights-of-way, reserved arterial and freeway lanes, and priority-entry control systems on freeway ramps.

In 1973, the Florida Department of Transportation, the U. S. Department of Transportation, and several agencies in metropolitan Dade County embarked on a program to reduce the recurring peak-period congestion in the I-95 corridor, shown in Figure 1, and to demonstrate the feasibility of several priority-treatment techniques. This demonstration project included (a) several bus-priority strategies on NW 7th Avenue, (b) reserved lanes for buses and car pools on I-95, (c) the Orange Streaker express bus service, and (d) the Golden Glades park-and-ride lot. These treatments are evaluated elsewhere (1, 2, 3).

SCOPE

This paper presents an evaluation of the responses of several road-user groups who participated in questionnaire surveys related to the bus-and-car-pool priority operation on I-95. The road-user groups surveyed included (a) express bus passengers, (b) car poolers, (c) motorists using the I-95 corridor, (d) express bus drivers, and (e) other I-95 motorists.

In some cases, the surveys dealt with broad aspects of the bus-car-pool project; a detailed treatment of these results have been given by Wattleworth and others (1). The discussion presented here is limited to those parts of the survey that addressed the operational aspects of the exclusive lane on I-95.

BUS-PASSENGER SURVEY

A questionnaire survey of bus passengers was carried out in connection with the Orange Streaker project. The survey forms were distributed to the passengers as they boarded the bus and collected from them at the end of the trip. Five questions related to the exclusive lane on I-95 were included to assess the following topics:

1. Relative preference: the exclusive lane on the freeway versus the reserved lane and other strategies previously used by the express buses on NW 7th Avenue,
2. Reaction to certain physical and operational features of the exclusive lane,
3. Relative importance of the exclusive lane compared with other system elements,
4. Estimated time savings, and
5. Overall reaction to the exclusive-lane concept.

A total of 838 responses were received, and the following results were obtained.

Relative Preference for Freeway System

This response, as shown in Figure 2, was overwhelmingly favorable to the exclusive lane on the freeway. Three-fourths of the respondents rated the freeway lane far superior to the arterial system, 92 percent expressed some degree of preference for the freeway lane, and less than 1 percent preferred the NW 7th Avenue system.

Reaction to Physical and Operational Features

To assess the reactions of the bus passengers to the various physical and operational features of the I-95 system, they were asked to indicate which, if any, of the following items caused them to feel unsafe or uncomfortable:

1. Speeds in the exclusive lane that were higher than those in adjacent lanes,
2. Lack of an inside shoulder,
3. Buses that operated too close to the concrete barrier wall,
4. Buses that had to change lanes too many times to enter and exit the exclusive lane, or
5. Other.

The responses to this question indicated that each problem category generated discomfort in about 5 percent of the respondents. A total of 18 percent indicated concern for at least one of the problems mentioned. The other category drew a 9 percent response. A frequent complaint in this category referred to the excessive weaving activity into and out of the exclusive lane by other drivers. Many of the complaints were not specifically related to the exclusive-lane operation (e.g., bus drivers that were too aggressive).

Estimated Time Savings

The bus passengers were asked to estimate the amount of time they saved by traveling in the exclusive lane. Of the 838 responses, 64 percent indicated some degree of perceived time savings, 13 percent saw no saving, and 23 percent offered no opinion. The distribution of the perceived time saving is shown in Figure 3.

Figure 1. Project corridor: I-95 and NW 7th Avenue.

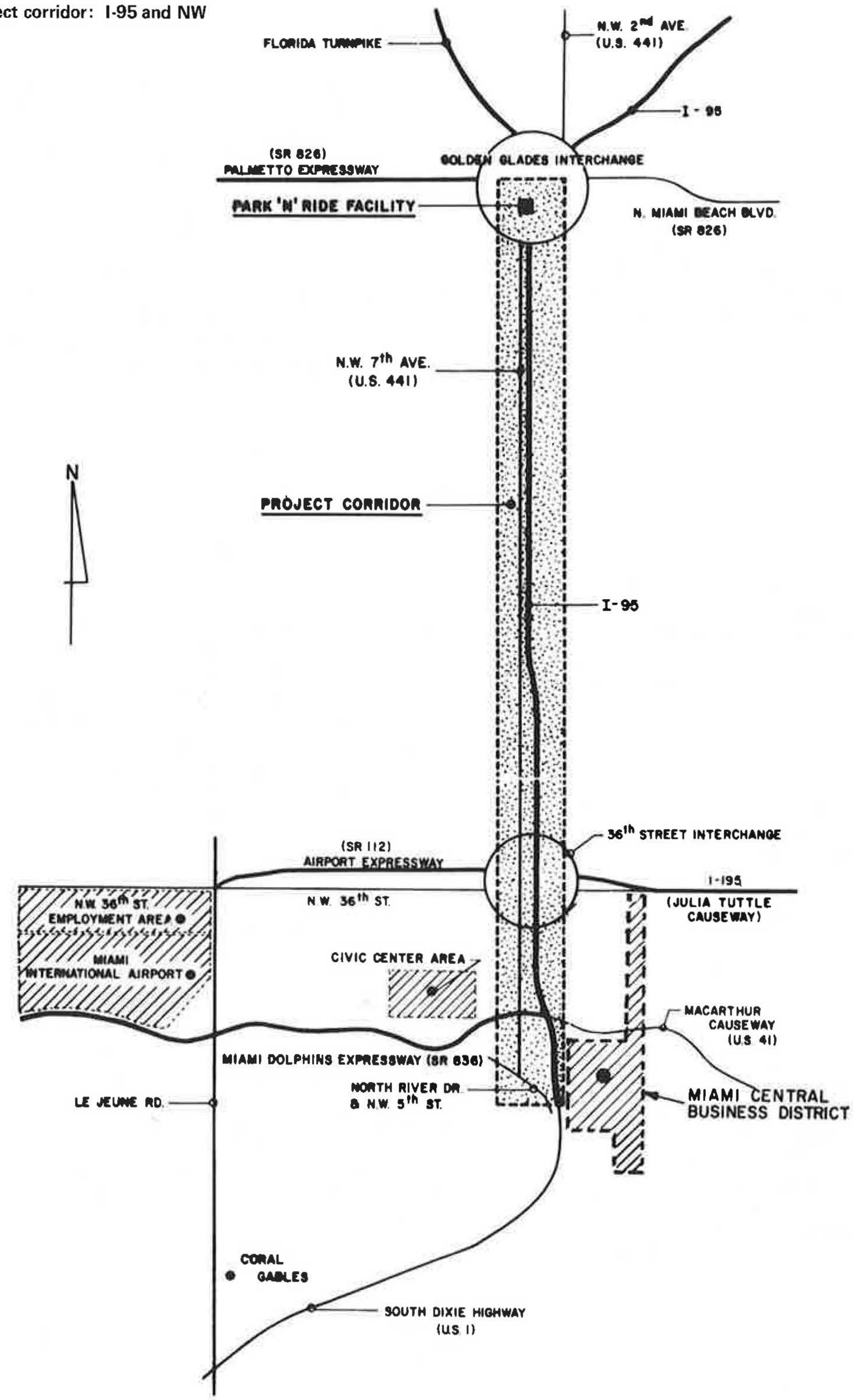
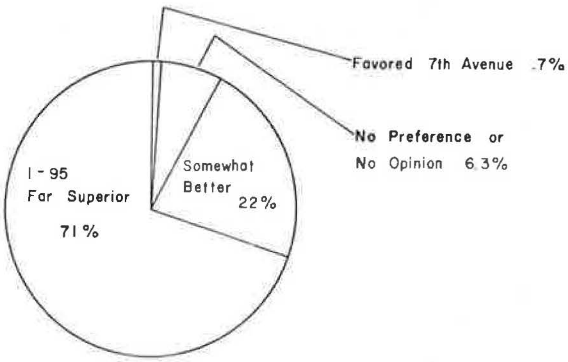


Figure 2. Relative preference of bus passengers: I-95 bus-and-car-pool lane versus NW 7th Avenue bus-priority system.



Relative Importance of Exclusive Lane

To gain an insight into the degree of importance to the respondents of the various project features, each respondent was asked to indicate the degree of importance of (a) the express bus service, (b) the exclusive lane on I-95, (c) the park-and-ride lot, and (d) the comfort of the buses. The responses to this question are summarized below.

Feature	Very Important (%)	Somewhat Important (%)	Not Important (%)	No Opinion (%)
Express bus service	89	9	1	1
Exclusive lane	74	21	4	1
Park-and-ride lot	65	20	12	3
Bus comfort	56	37	5	2

These results show that, as perceived by bus passengers, the exclusive lane ranked second in importance. The only feature considered more important was the provision of the bus service itself, and both the park-and-ride facility and the comfort of the buses were considered to be of lesser significance.

Overall Reaction to Exclusive-Lane Concept

To evaluate the overall reaction to the exclusive lane, the bus riders were asked to indicate whether or not they felt that this concept should

1. Remain on I-95,
2. Be installed on all highly congested freeways, or
3. Be installed on all urban freeways.

The responses to this question, as summarized in Figure 4, were highly favorable. However, this would be anticipated here because the respondents were all receiving some benefit from the exclusive lane as bus passengers. Of the total response, 94 percent felt that the express lane should remain on I-95, and only 2 percent felt that it should be discontinued. The degree of enthusiasm for extending the concept to other facilities was also reasonably high: 57 percent favored its extension to all congested freeways and 48 percent favored its extension to all urban freeways.

CAR-POOL SURVEY

A separate survey was conducted among car-pool participants by distributing a mail-back questionnaire to each occupant of every vehicle that had two or more persons leaving the Golden Glades parking facility during a selected morning peak period. A total of 42 responses were returned. The questionnaire dealt with several aspects of the Orange Streaker demonstration project and included five specific questions regarding the exclusive-lane operation on I-95. These were similar to those asked of the express bus passengers, except that the car poolers were asked about their use of the exclusive lane, rather than about their preference for the freeway or arterial system.

Use of Exclusive Lane

Of the 42 respondents, 66 percent indicated that they normally used the exclusive lane, 31 percent indicated that they normally drive in the general lanes on I-95, and 3 percent indicated that they did not use I-95.

Reaction to Physical and Operational Features

Because certain physical and operational features of the exclusive lane were felt to be potential sources of discomfort, each respondent was asked to identify areas of particular concern. The responses to this question are summarized below:

Feature	Percentage of Respondents
Accessibility of lanes	38
Lack of shoulder	33
Speed differential	29
Proximity of barrier	10
Other	43
Any of the above	71

Other-category comments dealt primarily with the abuse of the exclusive lane by other drivers.

By comparing these results with the corresponding responses from the express bus passengers, it can be seen that the degree of concern is considerably higher among automobile occupants. For example, 71 percent of the car poolers expressed concern over at least one item as opposed to 18 percent of the bus passengers. This increased concern is probably attributable to the fact that the automobile occupants interact more closely with these problems than do the bus passengers.

Estimated Time Savings

The exclusive-lane users were also asked to estimate the amount of time they saved by using the exclusive lane. The resulting distribution of estimated time savings is shown in Figure 5. The mean estimated saving was 12.6 min, which, when compared with measured saving of approximately 3 min, indicates that the perceived saving is substantially greater than the actual saving. [The measured time difference, which was determined by moving-vehicle studies, is discussed in greater detail by Courage and others (2).]

Relative Importance of Exclusive Lane

This topic was addressed by use of two questions. First, the car poolers were asked to rate the importance of the various physical elements of the I-95 system as (a) very important, (b) somewhat important, (c) not important, or (d) no opinion. The re-

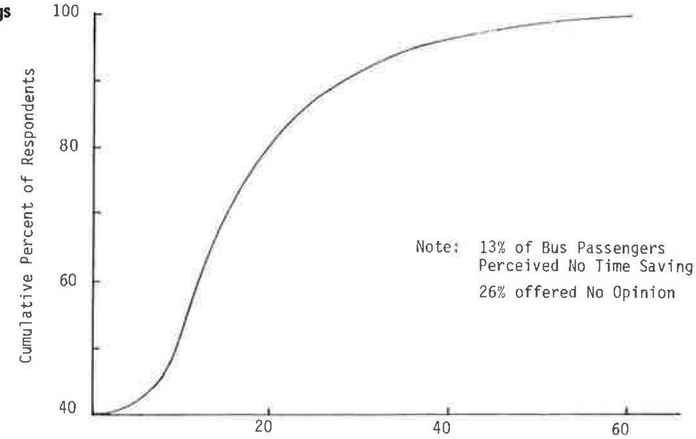
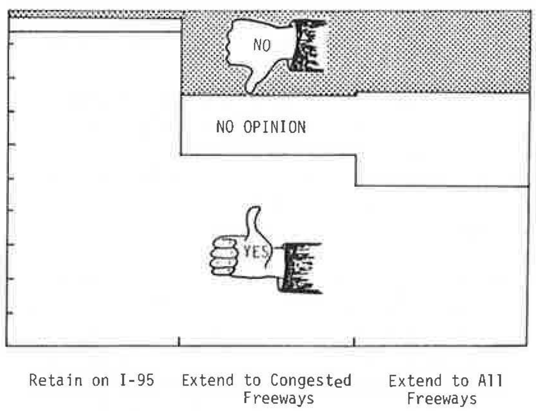


Figure 3. Distribution of bus-passenger perceived travel-time savings due to exclusive lane.

Figure 4. Bus-passenger attitudes toward retention and extension of bus-and-car-pool priority lanes.



sponses to this question are summarized below.

Feature	Very Important (%)	Somewhat Important (%)	Not Important (%)	No Opinion (%)
Exclusive lane	71	18	5	5
Park-and-ride lot	71	26	3	6
Flyover	72	11	6	11

It is observed that the three physical attributes of the system (the park-and-ride lot, the flyover, and the exclusive lane) were regarded as being of more or less equal importance by car poolers. Approximately 70 percent regarded each attribute as being very important, and less than 6 percent suggested that any attribute was not important.

Second, the car poolers were asked to indicate the factor–(a) cost of driving alone, (b) lack of available automobile, (c) lack of parking facilities, (d) cost of parking, (e) time advantage of exclusive lane, (f) concern for energy conservation, and (g) other–that most influenced their decision to car pool. The responses to this question, illustrated in Figure 6, indicate that about one-half of the respondents were motivated toward car pools primarily by cost factors and about one-fourth were motivated by time savings due to the exclusive lane. This suggests that, in the case of the I-95 system, cost factors were approximately twice as important to car-pool formation as the potential time saving (this is despite the relatively large perceived time savings as shown in Figure 5).

Overall Reaction to Exclusive-Lane Concept

To assess their overall reaction to the exclusive lane on I-95, car poolers were asked their opinion regarding the continuation and future expansion of the exclusive-lane system. The responses to this question, illustrated in Figure 7, were similar to those obtained in the bus passenger survey. A high proportion (86 percent) favored continuation of the reserved lane on I-95 and more than half favored extension of the concept to other congested freeways (60 percent) and and to all urban freeways (53 percent). The convergence of responses to this question is interesting when compared with the responses of bus riders (Figure 4). It again reflects, first, the more intimate concerns of the car poolers, but the stronger sense of direct benefit of the bus passengers.

HOME INTERVIEW SURVEY

A telephone survey was conducted by using a sample of 1903 persons observed using either I-95 or alternative routes within the project corridor. Because this survey dealt with the overall demonstration project, it was not possible to address the I-95 exclusive-lane features in detail. However, questions were included to assess the following areas:

1. Awareness of the existence of the exclusive lane,
2. Extent of exclusive-lane use, and
3. Overall reaction to the exclusive-lane concept.

[A complete analysis of the results of this survey is given by Long and others (4).] The survey respondents were separated into two categories:

1. Project trip respondents: those who indicated a trip origin in the market area and a destination in one of the service areas served by the Orange Streaker and
2. Nonproject trip respondents: those who did not qualify as project trip respondents by the above definition.

Awareness of Exclusive Lane

Of the project trip respondents, 99.4 percent indicated that they were aware of the exclusive lane on I-95 and, of the nonproject trip respondents, 86 percent indicated a knowledge of the exclusive lane. This suggests a very high awareness among both

Figure 5. Distribution of estimated travel-time savings by car poolers who use priority lane on I-95.

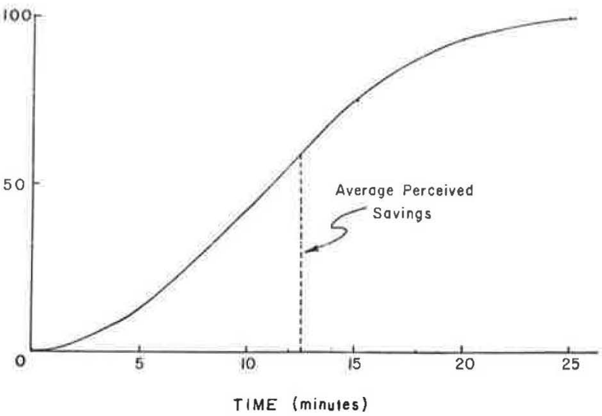


Figure 6. Primary factor in decision to car pool.

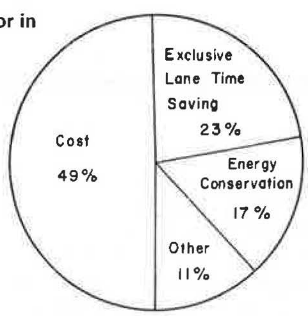
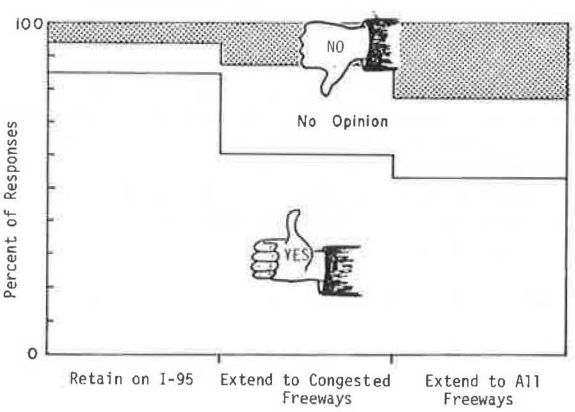


Figure 7. Car-pooler attitudes toward retention and extension of bus-and-car-pool priority lanes.



respondent categories, although the project trip respondents demonstrated a predictably higher awareness than the nonproject trip respondents.

Use of Exclusive Lane

In both respondent categories, those who indicated that they were aware of the exclusive lane and used I-95 were asked whether or not they used the exclusive lane. This question was also put to all project trip respondents who indicated that they car pooled at least occasionally; however, those who indicated that they drove alone were not asked whether they used the exclusive lane. This was done to eliminate the possible fear of self-incrimination. All nonproject trip respondents were questioned on exclusive-lane use, because they were not asked about their car-pooling habits.

Thirty-five percent of the project trip respondents reported using the exclusive lane, although only 12 percent of the nonproject trip respondents indicated such use. These figures are not directly comparable because of the elimination of single-occupant users from the project-trip-respondents sample. Furthermore, it cannot be inferred from the 35 percent use by project trip car poolers that the nonuse rate was 65 percent because, at the time of the survey, two-occupant car pools were not permitted to use the exclusive lane.

Overall Reaction to Exclusive-Lane Concept

The respondents to the home interview survey were also asked about their overall acceptance of the exclusive-lane concept. The responses to this question are summarized in Table 1 and statistical comparisons support the following inferences:

1. There were no significant differences in the responses of the project trip and the nonproject trip respondents.
2. Persons who used I-95 were more favorable toward the concept as evidenced by a smaller proportion of negative responses.
3. Persons who did not use I-95 were less certain about their opinion as evidenced by the larger proportion of "don't know" responses.

Inspection of Table 1 also indicates that the actual differences among the various groups were relatively small. The opinions of the road users as a whole can, therefore, be represented by the combined aggregate of all the groups, as shown in Table 1 and in Figure 8. By comparing these results with those for the car poolers and the bus passengers, we can see that, although the overall response patterns are similar, the general road user is about 40 percent less inclined to favor the continuation or extension (or both) of the exclusive-lane concept.

BUS DRIVER SURVEY

The operators of the Orange Streaker buses were also surveyed to determine their reaction to the exclusive bus-and-car-pool lane. A total of 117 drivers participated; their questionnaire study focused on the following topics:

1. Magnitude of problems created by physical and operational features of the exclusive lane and
2. Overall reaction to the exclusive-lane concept.

Physical and Operational Features

The drivers were asked to rate the severity of problems associated with potentially troublesome features of the exclusive lane. These included

1. Hazards caused by sudden or unexpected stops in the priority lane,
2. Hazards caused by other vehicles cutting into the priority lane,
3. Lower speeds of traffic in adjacent lane,
4. Nearness of the concrete barrier wall,
5. Violations of the priority lane restrictions, and
6. Delays caused by other traffic using the priority lane.

Table 1. Corridor-user attitude toward retention and extension of exclusive-lane concept.

	Project Trip Respondents (%)						Nonproject Trip Respondents (%)						Combined Sample (%)		
	I-95 Users			I-95 Nonusers			Non-I-95 Users			I-95 Nonusers					
Response	Yes	No	?	Yes	No	?	Yes	No	?	Yes	No	?	Yes	No	?
Retain on I-95	50	32	18	50	27	23	50	32	18	50	27	23	50	29	21
Extend to congested freeways	40	35	25	38	32	30	40	35	25	39	32	29	39	34	27
Extend to all freeways	31	39	30	30	38	32	31	39	30	31	39	30	31	29	30

Figure 8. General road-user attitudes toward retention and extension of bus-and-car-pool priority lanes.

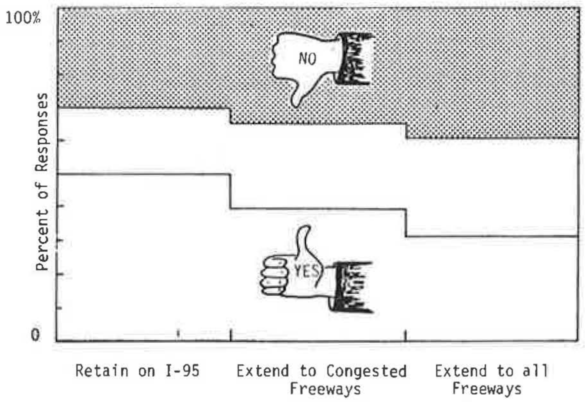


Figure 10. Bus driver attitudes toward retention and extension of bus-and-car-pool priority lanes.

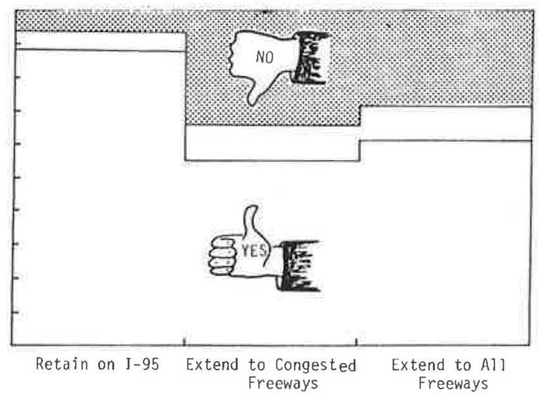


Figure 9. Bus driver perceptions of severity of exclusive-lane problems.

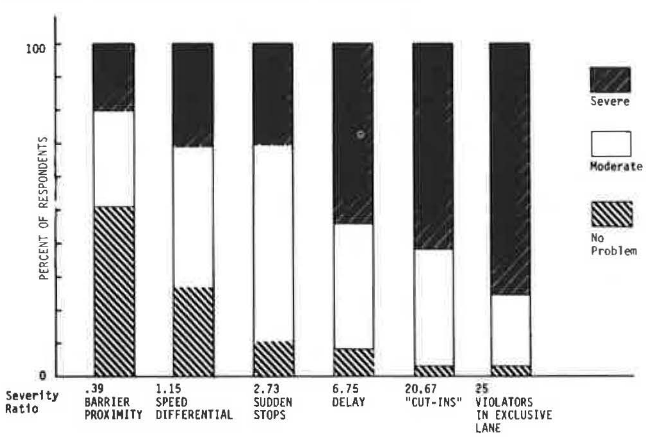
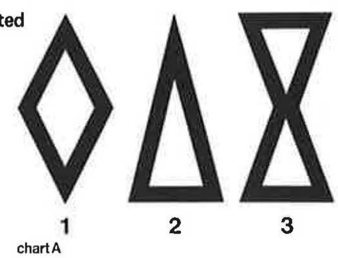


Figure 11. Symbols presented to drivers in diamond recognition survey.



The responses to this question are shown in Figure 9. The various features are rank ordered by inspection. The proximity of the concrete median barrier was of least concern to the operators; only 18 percent indicated this to be a severe problem and 50 percent indicated it to be no problem. At the other end of the scale, violation of the exclusive lane was viewed as a serious problem by 75 percent of the respondents and only 4 percent indicated that no problem was experienced with violators.

The magnitude of a particular problem can also be expressed conveniently in terms of the ratio of the severe-problem to the no-problem responses. This measure is defined for purposes of this study as the severity ratio and is also shown in Figure 9. In only one case (proximity of the median barrier) was the severity ratio less than 1.0. In all other cases, more bus operators rated the problem severe than nonexistent. The highest ratio (25) was that of the exclusive-lane violators. Based on these severity ratios, it can be concluded that the bus operators were quite concerned about the potential problems associated with the use of the exclusive lane. This concern was greatest in problem areas involving misuse of the lane by other drivers (e.g., violators and weaving). The inherent physical and operational characteristics such as proximity of the median barrier, speed differential, and turbulence were viewed with less concern. However, there was no prescribed nor measured definition of the word "problem", and the results should be viewed accordingly.

Overall Reaction to Exclusive-Lane Concept

To determine the overall reaction to the exclusive lane, each bus operator was asked to indicate a preference for continuation or extension (or both) of this concept. The response to this question is shown in Figure 10. A strong preference for continuation of the I-95 system is evident; 89 percent of the responses were affirmative, and 7 percent were negative. Enthusiasm for extension of the concept to other facilities followed the same response pattern as the previous groups; more than half of the express bus drivers favored implementing additional exclusive-lane systems, and more favored the extension of the concept to all facilities than to congested facilities only. This response is not internally consistent and may suggest

Figure 12. Degree of recognition of symbols.

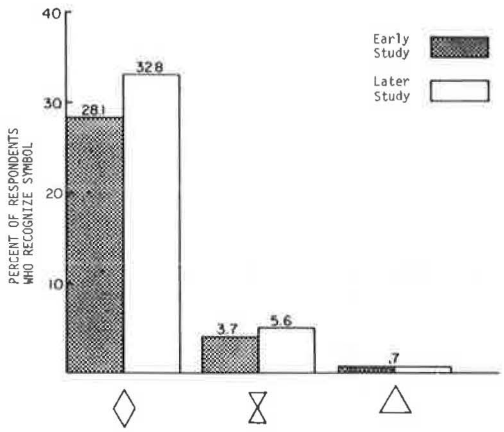
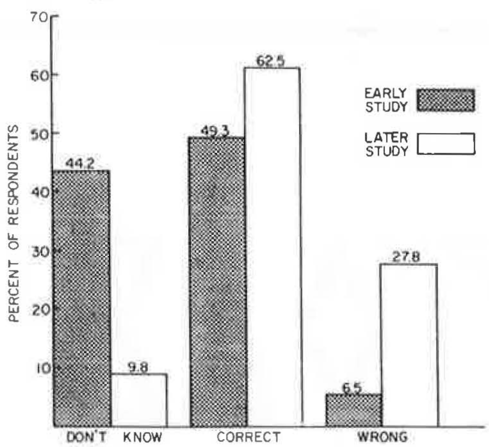


Figure 13. Degree of identification of meaning of diamond symbol.



misinterpretation of the question. The bus driver group had fewer no-opinion responses than any other road-user group. This suggests that the attitudes of the bus drivers to the system were more strongly developed because of their greater familiarity with the system.

DIAMOND SYMBOL SURVEY

The purpose of this survey was to assess the degree of motorist recognition and understanding of the reserved-lane symbol as a traffic control device. The survey was first conducted immediately after the diamond symbol had been implemented on the exclusive-lane signs and as pavement markings and then repeated approximately 3 months later.

The survey approach was to interview motorists leaving the freeway at two selected exit ramps during the afternoon peak period. The drivers were approached while they were stopped in a queue for the downstream traffic signal and were shown a chart depicting three different symbols. These symbols, shown in Figure 11, were

1. The diamond symbol used on the freeway to identify the exclusive lane,
2. An elongated triangle of approximately the same proportions as the diamond, and
3. An angular hourglass symbol of approximately the same proportions as the diamond.

Neither the hourglass nor the triangle were used on the freeway for traffic control purposes, and the charts were changed periodically to eliminate any bias that might be caused by the order of presentation. As the chart was being shown, each driver was asked the following questions:

1. "Where did you get on I-95 for this trip?"
2. " Did you notice any of these shapes being used as traffic symbols on the freeway?"
3. "What does this symbol mean to you?" (Indicating the selected symbol or the diamond.)

By observation, each vehicle was categorized by number of occupants and by area of residence (a) local (Dade and Broward counties), (b) nonlocal Florida, (c) out of state, and (d) rental. The descriptive parameters for these surveys are given below.

Parameter	Early Survey	Later Survey
No. of responses	437	341
Survey location breakdown, %		
62nd Street exit	46	50
135th Street exit	54	50
Proportion from Dade and Broward counties (local drivers), %	95	93
Average occupancy (passengers per vehicle)	1.50	1.46

Recognition of Diamond Symbol

The degree of recognition of the various symbols is illustrated in Figure 12. Nearly one-third of the drivers recognized the diamond symbol in both studies, although recognition was increased in the later study by approximately 15 percent. This change is statistically significant at the 95 percent level, although the actual number who recognized the diamond was surprisingly low considering the degree of exposure. At the time of the later study, the symbol was visible at approximately 100 locations on the pavement and 40 locations on overhead signs. The diamond was, however, recognized by significantly more drivers than the two fictitious symbols. The fictitious symbols combined received less than 20 percent the degree of recognition of the diamond.

Meaning of Diamond Symbol

Those motorists who recognized the diamond were also asked to identify its meaning. The results of this question are summarized in Figure 13 and show that the proportion of motorists who said they did not know the meaning of the symbol dropped substantially (from 44 percent in the early study to 10 percent in the later study) and that the proportion giving the correct meaning increased from 49 to 62 percent. However, the proportion giving an incorrect answer increased from 6 percent to 28 percent.

Effects of Study Location, Level of Occupancy, and Trip Length

As noted above, these surveys were conducted at two freeway exits: (a) 62nd Street, which was the first exit in the priority section, and (b) 135th Street, which was near the downstream end of the priority section. The sample was also stratified by level of occupancy as car pools (three or more occupants) versus noncar pools and by length of trip in the priority section. A summary of the effects of these parameters on the survey responses is given below (percentages may not add because not all respondents answered completely).

Parameter	Recognized (no. of persons)		Wrong (percentage of yes respondents)	Meaning (percentage of yes respondents)	
	No	Yes		Don't Know	Correct
Location					
62nd Street	265	104	21	32	47
135th Street	282	135	14	23	63
Occupancy					
Noncar pool	483	207	18	29	53
Car pool	63	32	13	13	74
Trip length					
< 4.8 km (3 miles)	288	114	21	30	49
> 4.8 km (3 miles)	259	125	14	24	62

Analyses of these factors indicate the following:

1. The study location did not influence the degree of recognition of the diamond symbol as a traffic control device.
2. The drivers leaving the freeway at the 135th Street exit who recognized the diamond symbol were more familiar with its meaning than those who left at 62nd Street (99 percent level of significance).
3. The vehicle occupancy did not affect the degree of recognition of the diamond symbol as a traffic control device.
4. Drivers of vehicles that qualified to use the exclusive lane who recognized the diamond symbol were more familiar with its meaning than drivers of noncar-pool vehicles (95 percent level of significance).
5. The trip length did not affect the degree of recognition of the diamond symbol as a traffic control device.
6. Drivers who had longer trip lengths and who recognized the diamond symbol were more familiar with its meaning than drivers who had shorter trip lengths (99 percent level of significance).

CONCLUSIONS

As a result of these user surveys, the following conclusions are offered regarding the public reaction to the bus-and-car-pool priority lanes on I-95.

System Acceptance

The bus-and-car-pool priority concept was well accepted by all road-user groups surveyed. Although one-third of the general road users expressed opposition to the priority lane, only 2 percent of the bus passengers felt that the operation should be discontinued. More than half of each group (50 and 94 percent, respectively) felt that the system should remain on I-95, and a positive attitude was expressed toward the extension of this concept to other facilities.

Bus passengers rated the exclusive lane second in importance among all the system elements. The provision of the bus service itself was the only feature that was rated as being of greater importance. The I-95 system was favored over the NW 7th Avenue bus-priority system by 93 percent of the bus passengers.

Car poolers also rated the express lane second in importance. In this case, the reduced cost of car pooling was considered to be the most important benefit. Both the bus passengers and car poolers tended to overestimate the time savings due to the exclusive lane by a significant amount.

Operational Problems

All groups surveyed indicated a high degree of concern over the abuse of the priority lane by violators. This was rated as the most significant problem by each user group surveyed. Car poolers expressed more concern about the physical and operational features than did bus passengers. Bus drivers generally expressed a greater degree of concern than either of these groups, but their concerns were concentrated more on the operational features, rather than on the physical elements.

Familiarity With System

The existence of the bus-and-car-pool priority lane was well established in the minds of the user groups who were surveyed, but knowledge of the use and meaning of the diamond symbol was low. More than 99 percent of the corridor users who had an origin in the market area and a destination in one of the service areas were familiar with the exclusive lane. On the other hand, only one-third of the freeway motorists recognized the diamond symbol as a traffic control device, and only one-third of these were able to correctly indicate its meaning. Drivers who had greater exposure to the symbol and drivers of qualified car pools demonstrated a higher degree of recognition of it.

Considering the overwhelming general awareness of the project, this suggests that, if a special symbol is to be used to identify a project, the symbol and its significance from a traffic engineering (and enforcement) point of view should be well publicized.

General Reaction to Bus-and-Car-Pool Priority System

These conclusions support the overall conclusion that the I-95 bus-and-car-pool systems demonstration project was well received by the road-user community. Despite some concern about particular operational and physical aspects of the system, no group expressed dissatisfaction with the system, and most who were able to afford themselves the opportunity to use the system were strongly supportive. Thus, exclusive lanes for buses and car pools can be a successful technique for improving highway transportation efficiency and receive strong support from the public.

ACKNOWLEDGMENTS

We wish to acknowledge the support of the Florida Department of Transportation, the Federal Highway Administration, and the Urban Mass Transportation Administration, U. S. Department of Transportation, who sponsored this research. All opinions, findings, and conclusions expressed in this paper are ours and not necessarily those of the U. S. or Florida departments of transportation.

REFERENCES

1. J. A. Wattleworth and others. Evaluation of the NW 7th Avenue Express Bus and Bus Priority Systems. Transportation Research Center, Univ. of Florida, Gainesville, Sept. 1977.
2. K. G. Courage and others. Traffic Control of Car Pools and Buses on Priority Lanes. Transportation Research Center, Univ. of

Florida, Gainesville; Federal Highway Administration, Rept. FHWA-RD-77-148, Aug. 1977.
3. J.A. Wattleworth and others. Evaluation of the I-95 Express Bus and High-Occupancy-Vehicle Priority Systems. Transportation Research Center, Univ. of Florida, Gainesville, Jan. 1978.
4. G. Long and others. Evaluation of Characteristics of Users and Nonusers of the I-95 Express Bus/Car Pool System. Transportation Research Center, Univ. of Florida, Gainesville, Dec. 1977.

Publication of this paper sponsored by Committee on Urban System Operations.

**Mr. Culpepper is currently with the Department of Civil Engineering, Auburn University.*

Role of Parking in Transportation System Management

Michael J. Demetsky, University of Virginia and Virginia Highway and Transportation Research Council, Charlottesville
Martin R. Parker, Jr., Virginia Highway and Transportation Research Council, Charlottesville

The appropriateness is examined of including parking management strategies in the transportation system management (short-range) component of the transportation plan for an urban region. The probable effects of parking management schemes are described and evaluated with respect to short-range objectives to determine the compatibility between the two. A definition that applies to the total set of parking management options is given and tested against the results of a survey of 173 cities in the United States and a review of the literature. Parking-control strategies can be divided on various bases: (a) whether they are supply- or cost-related controls, (b) whether they are intended to reduce automobile travel in selected areas or to make the highway more efficient, or (c) whether they can be implemented within the short-range element or the long-range component of the transportation plan. Considerable public, political, and business opposition to restrictions on parking in urban areas was found. Public support for parking controls that alter travel behavior must be developed gradually in association with areawide planning objectives. The majority of the parking measures were found to be long-range planning elements rather than transportation system management components.

When the Urban Mass Transportation Administration (UMTA) and the Federal Highway Administration (FHWA) issued a set of joint regulations in September 1975 (40 Federal Register, 42 976-42 979, 1975), a new dimension of the urban transportation planning process was formally established as the transportation system management (TSM) element. To comply with these regulations, the transportation plan for an urban region must include a transportation system management, or short-range, element in addition to the long-range element that had been the only requirement. Specifically, this TSM element requires the following:

1. Provision for the short-range transportation needs of the urban area by making efficient use of existing transportation resources and provision for the movement of people in an efficient manner and
2. Identification of traffic engineering, public transportation, regulatory, pricing management, operational, and other improvements to the existing urban transportation system, but not including new transportation facilities or major changes in existing facilities.

The long-range element addresses the future transportation needs of the area and should identify new transportation policies and facilities or major changes in existing facilities by location and modes to be implemented.

"Parking management" is the term currently used to describe the management and organization of parking as an element of TSM and is thus implied to be a consideration in the transportation planning process. This interpretation places a new burden on the transportation planner, because he or she is now responsible for considering the effect of the parking component of the transportation system on travel behavior, rather than merely considering parking needs in response to automobile travel demand. Thus, the supply of parking must be viewed in terms of the actions that affect the spatial and temporal allocations of spaces in addition to the traditional view in terms of the capacity or number of spaces provided at activity centers.

Conceptually, the role of parking management in improving the efficiency of urban transportation can be shown by examining the types of effects that it imposes. For example, the most notable result of a decrease in the supply of parking is a reduction in automobile trips, which implies some or all of the following:

1. An increase in automobile-occupancy levels,
2. A decrease in person trips,
3. Faster travel times for the remaining trips and a decrease in delays,
4. An increase in transit use,
5. A reduction in air pollution, and
6. Lower ambient noise levels.

Other effects can also be listed, but basically all strategies fall into either of two groups: those that reduce automobile travel and those that improve the efficiency of the highway network and make automobile travel more desirable. The impacts of the former are considerably more difficult to foresee than those of the latter because they are more likely to require major changes in travel behavior. The strategies in the latter group are basically traffic-engineering improvements to existing facilities, and they will not have widespread effects on travel behavior. Thus, there appear to be both short-run and long-range components of parking management policy that imply localized and areawide programs respectively.

Soon after the FHWA and UMTA planning regulations were issued, serious problems of interpretation arose. Primarily, the regulations did not make clear whether the long list of TSM actions they contained was a shopping list from which a locality might choose appropriate activities or whether every urban area would be required to undertake every suggested action or justify the failure to do so (1). Accordingly, controversial issues soon surrounded the TSM planning process, and it became imperative that each of the recommended options be clearly defined and understood before being written into the TSM element. Subsequent guidelines issued in the Federal-Aid Highway Program Manual (volume 4, chapter 4) have aided in clarifying the planning process; however, the effects of various strategies have not been properly demonstrated.

PURPOSE AND SCOPE

This paper develops a definition of parking management that relates to its proposed role as an element of TSM. Alternative parking management strategies are identified, their effects on the transportation system are evaluated, and the validity of the definition is tested. The analysis uses survey data and reported experiences. The selected strategies are considered in terms of their suitability as supply- or cost-related measures, whether their purpose is to reduce automobile travel in selected areas or to make the highway more efficient, and whether they can be implemented within the TSM or the long-range element of the transportation plan.

DEFINING PARKING MANAGEMENT

The definition of parking management that is derived here is based on consideration of the types of actions that are viewed as components of parking management and of their probable impacts on the transportation system. UMTA and FHWA have identified the following TSM strategies for managing and controlling parking:

1. Elimination of on-street parking, especially during peak periods;
2. Regulation of the number and price of public and private parking spaces;
3. Favoring of short-term users over all-day commuters in the provision of parking;
4. Provision of fringe and transportation-corridor parking to facilitate transfer to transit and other high-occupancy vehicles; and
5. Strict enforcement of parking restrictions.

By using these basic strategies as a focus, a literature review was conducted to define a listing of practical parking-control options. The results are shown below.

Parking Management Strategy	Support Strategy (public transportation incentives)
Supply	Improvements in transit service
Short-term on-street parking only	Demand-responsive transit
No on-street parking	Subscription services
Strict enforcement of parking regulations	Park-and-ride lots
Reserved parking for priority vehicles	Bicycle facilities
Restricted parking time at all facilities	Promotion of transit use
Residential parking permits	Staggered work hours
Freeze on number of parking spaces	Exclusive bus lanes
Limitations on parking garage construction	Peripheral parking
Zoning law limits (limitations on number of spaces allowed by zone)	Automobile-free zones
Price	Priority treatment for high-occupancy vehicles
High rates for single-occupancy vehicles	
Discriminatory hourly rates (low rates for short-term and high rates for long-term parkers)	
Increases in all parking rates	
Reductions in costs for priority vehicles	
Parking taxes on users	
Stall taxes on parking garage owners	

There are 15 basic means for managing parking and 11 complementary actions that may be required to sustain mobility when restraints are placed on the parking supply.

Measures of the effectiveness of parking-control methods that can be obtained from basic traffic and planning data include the following: measured traffic volumes, number of parkers, automobile-occupancy rates, transit-ridership levels, noise levels, air-contaminant levels, vehicle kilometers of travel, energy consumption, travel times and delays, and costs. Parking management strategies can be examined for their compatibility with TSM objectives and compared with one another or with other TSM strategies.

Because of the variety of actions that can affect the transportation system through the parking element and the issues that have been raised, the following definition of the basic ideas associated with parking management is now proposed.

A parking management strategy is a measure taken to alter the supply or cost of parking to either reduce automobile travel in a selected area or to make the operation of the urban highway system more efficient. Some of these actions can be implemented in the short-term as TSM projects, but others are dependent on institutional changes that can occur only over a considerable period of time.

This definition should establish the role of parking in meeting the transportation planning requirements specified by UMTA and FHWA.

METHODOLOGY

To study the appropriate role of parking management in the transportation plan for urban areas, associated experiences and attitudes were examined through a review of the literature and a nationwide survey. A questionnaire was designed to solicit opinions on parking management strategies from city officials throughout the United States. The purpose was to determine the types of parking controls in use, the reasons for their selection, the basis for their evaluation, and the problems encountered in implementing them.

Questionnaires were distributed to transportation officials in 458 U.S. cities. Every urban area that has a population greater than 100 000 was included, as well as a number of randomly selected cities that have populations between 10 000 and 100 000. A total of 173 questionnaires were returned by cities in 47 states, the District of Columbia, and Puerto Rico, for a return rate of 38 percent. The distribution and return of the questionnaires by population of the cities is given below.

Population of City (1970 census)	Mailed (no.)	Returned No.	Returned Percent
< 50 000	123	42	34.1
50 000-99 999	173	59	34.1
100 000-500 000	133	58	43.6
> 500 000	29	14	48.3
Total	458	173	37.8

Table 1. Reasons for not considering strategy for implementation.

Strategy	Prohibited: By Local Ordinance	Prohibited: By State Law	Cannot Be Enforced	Ineffective	Not Applicable to Problem	Opposition: Public	Opposition: Political	Opposition: Business	Funds Not Available	Other	Total
Supply											
Short-term on-street parking only	0	0	4	4	30	22	13	14	3	4	94
No on-street parking	2	0	2	3	36	28	17	26	3	3	120
Strict enforcement of parking regulations	0	0	1	1	19	6	5	1	6	3	42
Reserved parking for priority vehicles	3	1	9	6	56	20	13	17	4	5	134
Restricted parking time at all facilities	3	0	7	9	54	30	20	32	5	7	167
Residential parking permits	5	3	8	6	59	25	20	10	4	5	145
Freeze on number of parking spaces	5	0	1	6	62	31	28	38	4	8	183
Limitations on parking garage construction	4	2	3	6	67	25	26	31	7	8	179
Zoning law limits	11	3	1	8	65	23	24	25	5	7	172
Total	33	9	36	49	448	210	166	194	41	50	1236
Price											
High rates for single-occupancy vehicles	3	2	19	10	59	52	31	37	10	8	231
Discriminatory hourly rates	1	0	6	8	53	34	21	22	5	7	157
Increases in all parking rates	3	0	0	6	46	31	21	25	4	6	142
Reductions in costs for priority vehicles	2	1	10	10	64	21	15	15	7	6	151
Parking taxes on users	3	5	3	6	57	41	31	32	6	5	189
Stall taxes on parking garage owners	3	4	0	6	77	23	28	38	5	8	192
Total	15	12	38	46	356	202	147	169	37	40	1062

Table 2. Reasons for considering but rejecting strategy for implementation.

Strategy	Prohibited: By Local Ordinance	Prohibited: By State Law	Cannot Be Enforced	Ineffective	Not Applicable to Problem	Opposition: Public	Opposition: Political	Opposition: Business	Opposition: Parking Enterprise	Funds Not Available	Other	Total
Supply												
Short-term on-street parking only	0	0	1	0	1	3	2	4	0	0	0	11
No on-street parking	0	1	2	0	1	9	9	14	1	0	0	37
Strict enforcement of parking regulations	0	0	1	0	0	1	1	3	0	1	0	7
Reserved parking for priority vehicles	2	0	3	0	0	0	0	0	0	0	0	5
Restricted parking time at all facilities	0	0	0	0	1	1	0	3	1	0	0	6
Residential parking permits	4	4	2	2	2	4	2	0	0	0	1	21
Freeze on number of parking spaces	0	0	1	1	2	2	3	3	1	0	0	13
Limitations on parking garage construction	1	0	0	0	0	1	3	4	2	0	0	11
Zoning law limits	2	0	0	0	0	0	1	1	0	0	0	4
Total	9	5	10	3	7	21	21	32	5	1	1	115
Price												
High rates for single-occupancy vehicles	0	0	1	0	1	1	2	1	1	0	0	7
Discriminatory hourly rates	0	0	0	1	3	2	2	1	1	0	0	10
Increases in all parking rates	0	0	0	0	0	6	7	6	2	0	0	21
Reductions in costs for priority vehicles	0	0	0	0	1	0	1	2	0	0	0	4
Parking taxes on users	1	2	2	1	0	2	1	2	2	0	0	13
Stall taxes on parking garage owners	0	0	0	1	1	1	2	5	2	0	0	12
Total	1	2	3	3	6	12	15	17	8	0	0	67

This distribution indicates that each population category was proportionally represented by the responses ($\chi^2 = 2.19$, $p > 0.01$, df = 3).

Summaries of the questionnaire responses are shown in Tables 1, 2, and 3. Approximately 40 percent of the respondents made comments in the space provided.

The following discussion integrates the survey results with observations documented in the literature. An overview of the data is given first to identify the status of parking management strategies in general; then, the individual parking control strategies are evaluated.

ANALYSIS

Status of Parking Management: An Overview

Although parking management is widely practiced in many European cities, the development of plans for U.S. cities has been limited. Only five U.S. cities that responded to the survey have official plans for parking management, which shows that its time as a vital element of transportation system management has not yet arrived. However, 27 cities (16 percent) reported that a parking management study had been conducted and 38 (22 percent) indicated that studies were under way. This recent activity indicates a new awareness of the role of parking as an element of the transportation system and is most likely attributable to recent Environmental Protection Agency (EPA) emphasis on parking management. A review of the published studies on parking indicated that, for the most part, parking management strategies are being used at spot locations on an individual basis to increase capacity or improve safety. The parking management plans for the five cities mentioned above include implementation of parking controls on an areawide basis.

The use or nonuse of parking strategies is interpreted

Table 3. Reasons for implementing strategy.

Strategy	Improve Traffic Flow	Reduce Congestion	Improve Air Quality	Reduce Noise Level	Reduce Energy Consumption	Increase Use of Transit	Increase Automobile Occupancy	Reduce Accident Hazards	Other	Total
Supply										
Short-term on-street parking only	10	11	4	3	2	2	4	3	30	69
No on-street parking	34	23	7	3	5	7	2	13	7	101
Strict enforcement of parking regulations	35	27	4	4	4	7	4	12	32	129
Reserved parking for priority vehicles	7	7	1	1	2	5	5	6	9	43
Restricted parking time at all facilities	6	7	1	2	1	4	2	1	3	27
Residential parking permits	2	6	4	3	3	4	4	1	8	35
Freeze on number of parking spaces	1	3	3	1	3	3	2	1	0	17
Limitations on parking garage construction	1	1	2	0	1	2	1	0	2	10
Zoning law limits	7	7	3	0	3	4	1	5	1	31
Total	103	92	29	17	24	38	25	42	92	462
Price										
High rates for single-occupancy vehicles	2	1	1	0	1	1	1	1	0	8
Discriminatory hourly rates	8	7	3	2	3	5	3	3	3	37
Increases in all parking rates	5	7	1	0	1	6	3	1	10	34
Reductions in parking costs for priority vehicles	3	4	1	0	2	0	1	1	1	13
Parking taxes on users	2	1	1	1	1	1	2	0	5	14
Stall taxes on parking garage owners	1	0	0	0	0	0	0	0	0	1
Total	21	20	7	3	8	13	10	6	19	107

as an indicator of the role parking is being assigned as a TSM strategy. For example, 71 percent of the respondents had not considered using any of the 15 strategies listed in the survey. As shown in Table 1, the primary reasons cited for not using parking strategies were public, political, and business opposition (47 percent); however, 35 percent of the respondents felt that the controls were not applicable to their transportation problems. Although 71 percent of the respondents had not considered using parking controls, only 4 percent had considered and rejected the strategies. As shown in Table 2, the primary reasons cited for rejecting the strategies were opposition on the part of the public, political personages, and business enterprises.

Approximately 12 percent of the respondents indicated that one or more of the 15 parking management strategies were in use in their cities. Approximately 11 percent indicated that the strategies were being considered, while 2 percent of the controls had been programmed for implementation. The strategies in use were being employed primarily to improve traffic flow and reduce congestion as shown in Table 3. Of the parking controls that have been implemented, few have been evaluated for effectiveness. Only 47 percent of the cities that had implemented a parking strategy indicated that data had been collected for evaluation purposes. Eleven percent of the respondents stated that the effectiveness of their strategies had not been measured, and 16 percent reported that their evaluations were limited to engineering judgments. Thus, it is apparent that, in many cities, the effects of using various controls on an areawide basis are not known.

Aside from public, political, and business opposition and the question of the applicability of parking controls to a city's problem, the most significant reason for not implementing parking controls is probably that there are few perceived transportation-related problems in U.S. cities. As shown below, traffic congestion was rated as a major problem in 15 cities and conditions such as accidents and air pollution were felt to be major problems in only a few cities.

Category	No Problem	Minor Problem	Problem	Considerable Problem	Major Problem
Traffic congestion	12	53	69	17	15
Traffic accidents	3	56	74	23	8
Air pollution	43	66	32	16	8
Traffic noise	34	90	33	7	1
Inefficient use of energy	9	37	70	35	11
Other	0	0	1	3	9

As the magnitudes of some of these problems increase, it is possible that a proportional need for parking controls will also increase.

Generally, few legal problems were reported to have been encountered in using parking strategies. Twenty-seven cities had needed changes in the state and city codes to permit implementation of a strategy. In 5 cities, litigation and legislative action was under way to allow implementation of a parking control.

A list of factors that have been cited in the literature as impediments to the application of parking controls is given below.

1. There is a surplus of public and private parking in most cities. Nationwide, 92.7 percent of commuter automobile trips park free.
2. Most parking strategies that can be implemented are already in use.
3. Most parking facilities in large cities are privately owned and not subject to government regulation.
4. Public, political, and business interests generally oppose parking regulations.
5. Increases in parking fees, even when coupled with free bus service, have had little effect on the choice of travel mode.
6. Parking fees in cities are low, averaging $1.75/d.
7. Strict enforcement of parking regulations is not economically feasible.
8. Evaluation of the effectiveness of various parking strategies has been limited.

These cited situations are a result of policies that set no limit on the parking supply in an area and made little effort to control automobile travel. Such an atmosphere contributed to the recision of the parking regulations promulgated by EPA to promote conformance to the air quality standards set by the Clean Air Act of 1970 before the regulations were implemented.

In the following section, the results of the recent survey are reviewed in terms of specific parking-control actions.

Parking-Control Strategies

Parking management strategies can be divided into those controls that affect the supply of parking and those that regulate the price of parking. Of the 15 parking strategies addressed in the survey, 9 were associated with supply and 6 with price, as shown above. This classification recognizes that parking can be used to influence air quality, energy conservation, traffic congestion, land use, and other social and environmental factors.

Supply Controls

Supply controls are those parking management strategies that affect the restriction, removal, or reallocation of parking spaces to increase roadway capacity or to alter automobile-use patterns. The measures that restrict parking include the provision of short-term on-street parking, restrictions on long-term on-street parking, and the issuance of residential parking permits. These measures are aimed at discouraging commuters from tying up parking spaces all day and hence require alternative travel arrangements such as selective off-street parking and public transportation. The results of the questionnaire show that short-term parking is one of the most widespread traffic management strategies used in the United States. Fifty-nine cities reported using this technique, and 3 others reported plans for implementation. A definite policy that restricts long-term parking discourages work trips by automobile to selected areas. The most prominent substitute objective is induced transit use. According to the survey, 17 cities had implemented restrictions on parking time. The majority of cities (91) had not considered this strategy; their reasons are given in Table 1.

In large cities where the central business district is close to residential areas, commuters use the residential streets for parking and restrict the availability of spaces for residents. To provide residential parking, permits can be issued to residents and the streets signed to give these persons priority over commuters. The results of the survey showed that residential-permit parking controls are being used in five cities and have been programmed for implementation in four other areas. The reasons are given in Table 3. It is significant that, when the Virginia Supreme Court ruled the residential-permit system unconstitutional, Arlington County appealed the decision to the U.S. Supreme Court, which ruled in favor of allowing cities to issue residential permits. Thus, this strategy may be more widely used now that its legal status has been resolved.

To be effective, all of the restrictions on parking require strict enforcement, but its cost is a major problem. The results of the survey show that the enforcement of existing parking regulations was seen by the respondents to be the most frequently used parking management strategy. As shown in Table 3, strict enforcement of regulations was felt to improve traffic flow and reduce congestion and accident hazards.

The elimination of on-street parking improves vehicular flow by increasing the capacity of the roadway. This removal of parking could encourage commuters to use public transportation or car pools; however, it also eliminates space for short-term users. The loss of parking is seen in many instances to have a detrimental effect on business. Also, because the proportion of curb spaces in large cities is less than it is in small cities, removal of on-street parking will have a greater impact in small cities than in large ones. The survey showed that provisions for eliminating on-street parking had been made in 23 cities and programmed for implementation in 7 others.

An easy way to encourage car and van pooling is to provide preferential parking for high-occupancy vehicles. This strategy is a relatively new concept in parking management, and an evaluation of the impacts could not be found in the literature. In the survey, 12 cities reported using reserved parking for priority vehicles; reasons are shown in Table 3.

Another group of controls that reallocate the parking supply of an area can have a much more profound impact than those discussed above. These measures include freezing the number of spaces in areas of high trip density, limiting the construction of parking garages, and adjusting zoning laws to limit the number of spaces. A limitation on the number of spaces for parking facilities and a restriction on the construction of garages will not have an immediate impact but can have long-range effects; i.e., as the activity level increases, alternative means of transportation will have to be provided.

Based on the survey responses, Honolulu; High Point, North Carolina; and Portland, Oregon, were the only cities that had placed a freeze on the number of parking spaces. The reasons for this action are given in Table 3. Cambridge, Massachusetts, had programmed the measure for implementation. One hundred and four cities have not considered this action for the reasons given in Table 1.

High Point, Portland, New York City, and Bethlehem, Pennsylvania, had imposed a limit on the construction of parking garages, and Cambridge is also planning to implement this strategy. The reasons for limiting the construction of garages are given in Table 3. The reasons given by 107 cities for not considering this measure are given in Table 1. Although 103 cities had not considered zoning to limit the number of parking spaces, 7 had developed zoning plans; the reasons given are listed in Table 3.

The supply controls can be summarized in a hierarchical order, from those that restrict the use of existing spaces to those that alter the allocation of parking in a region. The former have localized impacts on traffic, and the latter have long-term effects on travel behavior.

Price Controls

Price controls are parking management measures that attempt to provide an improved transportation system through the use of selective pricing mechanisms. The strategies in this classification include high parking rates for single-occupancy vehicles, low rates for short-term parkers coupled with high rates for long-term users, a general increase in all parking rates, and taxes on users and operators of parking facilities.

To promote more efficient use of the urban street system, single-occupancy vehicles could be charged a higher rate for parking. One major problem with this proposal is that, nationwide, more than 90 percent of urban commuters do not pay for parking. Enforcement would also be a major problem.

The survey indicated that none of the cities responding to the questionnaire was charging higher rates for single-occupancy vehicles; however, the city of Hartford, Connecticut, had developed plans for implementing the strategy. The reasons given for these actions are listed in Table 3. More than 91 percent (112) of the cities had not considered this proposal; the reasons are noted in Table 1.

One method that has been suggested for reducing peak-period congestion and attracting shoppers and other

Table 4. Current status of parking management strategies.

Strategy	Planning Element				Perceived Purpose*		
	Long Range	TSM	Local Spot Improvements	Areawide Policy	Improve Traffic Flow	Reduce Accidents	Other
Supply							
Short-term on-street parking only		X	X		X		
Elimination of on-street parking	X		X		X	X	
Strict enforcement of parking regulations		X	X		X	X	
Reserved parking for priority vehicles	X			X	X		
Restricted parking time at all facilities	X			X	X		
Residential parking permits		X	X				Provide parking for residents
Freeze on number of spaces	X			X			
Limitations on parking garage construction	X			X			
Zoning law limits	X			X	X		
Price							
High rates for single-occupancy vehicles	X		X				
Discriminatory hourly rates	X	X	X		X		
Increases in all parking rates	X	X		X	X		
Reductions in costs for priority vehicles	X			X	X		
Parking taxes on users	X			X			
Stall taxes on parking garage owners	X			X			

*None of the cities related any of the strategies to the perceived purposes of improving the environment and conserving energy.

short-term parkers to the central business district is to charge low rates for short-term users and high rates for long-term users. Because many long-term parkers are commuters who do not pay for parking, this strategy would have limited application in most U.S. cities. The results of the questionnaire showed that seven cities had employed this strategy and three others had programmed it for implementation.

Installing parking meters or increasing the rates of on- and off-street parking can have a pronounced effect on the use of parking spaces. The results of a study conducted in a 4.4-km^2 (1.7-$mile^2$) section of downtown London indicated that, when meter rates were doubled or quadrupled, peak-period arrivals decreased 37 percent and, because of the increased availability of parking space, off-peak trips increased 49 percent. However, caution must be used in increasing parking rates because the short-term user may decide to shop in the outlying fringe areas rather than pay a high parking rate and this could lead to an economic decline of the downtown area. Based on the survey data, eight cities had increased their parking rates and four others were planning similar strategies. Preferential treatment for priority (high-occupancy) vehicles with regard to parking has had limited implementation. Anniston, Alabama; Dade County, Florida; and Baltimore reported that they had reduced parking costs for priority vehicles; the reasons are given in Table 3.

Another way to regulate parking in a congested area is to impose taxes on parking spaces. In October 1970, the city of San Francisco imposed a 2-year, 25 percent tax on all paid parking. In 1972, the tax was reduced to 10 percent at the insistence of the chamber of commerce because of decreasing sales. The parking industry estimated a 31 percent loss of revenue because of the tax. Although the data were inconclusive, only a 2 percent reduction in vehicle kilometers of travel were attributed to this strategy.

The survey responses showed that the cities of Washington, D.C., and Baltimore had also imposed a parking tax on users, and similar taxes were programmed for implementation in Lorain, Ohio, and Salem, Oregon; the reasons are given in Table 3. A parking tax had not been considered in 82 percent (100) of the cities; the reasons are given in Table 1.

None of the cities that responded to the questionnaire were using or considering the implementation of a parking-stall tax on garage owners. As shown in Table 1, the primary reason that the measure had not been considered was that the respondents felt it to be inapplicable to their problems.

Complementary Strategies

It is implied above that limiting the number of parking spaces or increasing the price should have an effect on the choice of travel mode. Therefore, if parking management strategies are used to regulate the supply or price of parking in a downtown area, alternative means of travel will have to be provided to maintain mobility and accessibility. To determine the extent to which complementary or support strategies were being used in cities, a list of 11 of these actions was included in the questionnaire survey. These strategies included improving transit service, providing bicycle facilities, and staggering work hours. It was originally anticipated that the identification of support strategies that had been implemented in conjunction with parking controls would be useful for determining the elements of a successful parking management policy. But because of the lack of experience with parking-control strategies in U.S. cities, this analysis could not be made. Generally, however, the cited support strategies were being used independent of, and to a much larger extent than, parking management strategies. Complementary strategies were used primarily to increase the use of transit, to reduce congestion, and to improve traffic flow. One of the major reasons given for not considering complementary strategies was the lack of funding, which indicates that complementary strategies may be more capital intensive than parking management controls.

Summary of Parking Management Strategies

At present, few parking management strategies are being used in U.S. cities. Thus, the potential benefits of applying parking strategies in conjunction with other TSM techniques have not been demonstrated. Specific strategies that have been implemented in many cities include (a) providing short-term parking, (b) eliminating on-street parking, and (c) strictly enforcing parking regulations. Although many other strategies appear to have potential benefits and some can be implemented within a short period at low cost, the results of the survey indicate that there is considerable oppostion to an active parking policy. Parking strategies, however, may have public support when implemented gradually in accordance with local needs and goals.

Based on the survey results, the status of the use of 15 parking strategies is given in Table 4. As shown, only a few of the controls are directly related to TSM planning. For example, allowing short-term on-street parking is a non-capital-intensive improvement that is well suited to use as a TSM strategy. This improvement is being used in local corridors and busy intersections or blocks to improve traffic flow. However, certain parking strategies have both short-range TSM and long-range planning aspects. For example, street parking can be eliminated at little cost, and this will improve local traffic flows and reduce accidents. However, when substantial amounts of on-street parking are eliminated, most cities have to provide additional off-street parking. This procedure is usually expensive, ranging from the purchase of land to the construction of parking garages, and takes considerable time to plan and implement. Thus, this strategy is more suited to use as a long-range planning element.

The majority of parking strategies are, in fact, long-range planning elements. For example, a freeze on the number of parking spaces within a city would take considerable planning to implement to ensure that mobility and the economic life of the affected area were maintained. These improvements, although not necessarily capital intensive, would require complementary strategies that, for the most part, would be expensive.

CONCLUSIONS

The environmental, economic, and energy issues of recent years have caused new emphasis to be placed on using parking management strategies to reduce automobile travel and increase the use of public transportation. Based on a review of the literature and the results of a survey questionnaire of U.S. cities, the following conclusions are offered.

1. Parking management strategies are not widely used on an areawide basis.
2. Parking strategies in the realm of TSM actions include (a) providing short-term parking, (b) eliminating on-street parking, and (c) strictly enforcing parking regulations. Most other strategies can have a dramatic effect on an urban area and are being implemented carefully over a long period of time. Thus, some parking strategies are short-range in nature and others are more appropriately part of the long-range element.
3. Parking strategies generally have diverse effects on an urban area. For example, encouraging short-term on-street parking may attract shopping trips and help revitalize the central business district. However, this will increase overall vehicle kilometers of travel, which will counter attempts to reduce air pollution and conserve energy.
4. There have been very few attempts to evaluate the effectiveness of parking controls, and little is known about the interrelationships between parking management strategies and supporting services.
5. There are few legal problems associated with implementing parking controls; however, public, political, and business opposition act as a deterrent to their implementation.
6. The use of parking policies to improve air quality, conserve energy, or attain other national goals may not achieve beneficial results in many urban areas, but could produce strong public opposition and cause economic decline of the central city.
7. In most cities, there is no perceived local problem that would require the extensive use of parking regulations to limit automobile travel.
8. Parking management policies may have beneficial effects in urban communities if they are applied gradually to alleviate local problems and promote achievement of local planning goals.

ACKNOWLEDGMENT

This study was conducted at the Virginia Highway and Transportation Research Council in Charlottesville. We are indebted to the sponsors—the Virginia Department of Highways and Transportation and the Federal Highway Administration. We especially thank the 173 city officials who returned the questionnaire.

The opinions, findings, and conclusions expressed in this paper are ours and not necessarily those of the sponsoring agencies.

REFERENCE

1. Transportation System Management. TRB, Special Rept. 172, 1977.

Publication of this paper sponsored by Committee on Parking and Terminals.

Freeway Incident-Detection Algorithms Based on Decision Trees With States

Harold J. Payne,* ESSCOR, San Diego
Samuel C. Tignor, Offices of Research and Development, Federal Highway Administration, Fairbank, Virginia

Incident-detection algorithms are a part of an overall freeway-traffic management system. These algorithms provide indications of the probable presence of freeway incidents by processing electronic surveillance data. In this paper, a class of algorithms that are designed to discriminate patterns in the data peculiar to incidents are described. The generic structure of these algorithms is the decision tree with states, the states corresponding to distinct traffic conditions. Ways to calibrate algorithm thresholds are described and applied to the algorithms. Performance evaluations based on traffic data from the Los Angeles system are presented.

An important function of freeway traffic management is the detection of and response to freeway incidents. Past research (1, 2) and operating experience have demonstrated that the detection of incidents can be automated through the use of specific incident-detection algorithms that operate on electronic surveillance data to produce indicators of the probable presence of an incident. The previously developed algorithms, however, are less effective than is desirable for operational use because they generate a high level of false alarms. This paper describes several new algorithms that are based on a generalization of the structure of the California algorithm, the decision tree with states, and that have been demonstrated to have significantly improved performance through evaluations based on a large amount of real surveillance data.

The results reported here represent a portion of the results obtained in the study; they are more fully reported elsewhere (3, 4, 5, 6, 7, 8) and have also been summarized (9).

The study was empirical in nature, being based entirely on a large amount of data obtained from the Los Angeles and Minneapolis freeway surveillance systems. In all, more than 10 000 000 vehicle-sensor crossings and approximately 150 incident events were used in the developments and evaluations. This is by far the most extensive data base ever used in research on incident-detection algorithms. Although more detailed data were available from the Los Angeles system, the data used in the algorithms were 20- and 30-s occupancies and volumes, averaged over all lanes at a station.

The research involved, first, an evaluation of previously developed algorithms (3). This evaluation indicated the superiority of occupancy-based algorithms, especially those that use occupancy values at adjacent stations on the freeway. Algorithms based on double exponential smoothing were among the best found, but the multiplicity of incident indications they generated was deemed to be an operational disadvantage. Subsequent evaluations (5) established that the California algorithm is the best of these, in that its detection-false alarm performance is best and, moreover, the incident indication was consistently associated with the actual location of the incident. This result contradicts earlier findings (1). A detailed examination of these earlier findings and our findings, however, showed that the bases for the conclusions were different. Generally, our findings emphasized detection rate performance at much lower false-alarm rates.

The results of these evaluations were used for the principal portion of the study, that of developing and evaluating new algorithms (which is the subject of this paper). The development of new algorithms initially followed two paths. The first, which was based on a complex, time-series approach (5), did not provide effective algorithms for the heavier traffic regimes that were our principal concern, but it does hold promise for light-to-moderate traffic regimes. The second path was based on extensions to the structure of the California algorithm.

Further attention was given to the effects of geometrics, sensor configurations, malfunctioning sensors, and weather and to means for identifying the lane of the incident (5).

The next section of this paper describes and analyzes the California algorithm and introduces the decision tree with states as a generic structure for several new incident-detection algorithms. Following this, we define the measures of performance, and detection and false-alarm rates and describe how the multiple thresholds in an algorithm can be calibrated to give an optimum trade-off curve in terms of these measures.

The several new algorithms and one old one are introduced after a discussion of the patterns in traffic data whose identification played a major role in the development of the new algorithms. Performance evaluations are presented for these algorithms.

CALIFORNIA ALGORITHM

As a preliminary to the introduction of decision trees with states as a general structure for incident-detection algorithms, we will describe the structure of the California algorithm and how its elements relate to patterns in traffic data. Of all previously defined incident-detection algorithms, the California algorithm is clearly superior (3, 5). Furthermore, the California algorithm, as compared with other algorithms, has the unique characteristic that three functions, rather than a single one, of the traffic data are used and that these functions (or features as we will henceforth refer to them) are selected to distinguish, in combination, patterns in the traffic data specifically related to incidents.

The several features—OCC, DOCC, OCCDF, OCCRDF, and DOCCTD—that are used in the California and the other algorithms to be described here are defined below [station indexes (i) increase in the direction of travel].

Feature	Description	Definition
OCC(i, t)	Occupancy at station i, for time interval t (percent)	
DOCC(i, t)	Downstream occupancy	OCC(i + 1, t)
OCCDF(i, t)	Spatial difference in occupancies	OCC(i, t) - OCC(i + 1, t)
OCCRDF(i, t)	Relative spatial difference in occupancies	OCCDF(i, t)/OCC(i, t)
DOCCTD(i, t)	Relative temporal difference in downstream occupancy	[OCC(i + 1, t - 2) - OCC(i + 1, t)] / OCC(i + 1, t - 2)

All features involve only the use of occupancy, measured as the average over all instrumented lanes at a single location on the freeway and over a 1-min interval.

Algorithms based on features involving volume and the volume-to-occupancy ratio were also investigated but were generally found to be inferior (3).

The California algorithm is described in Figure 1. This algorithm, as originally defined, used in place of DOCCTD a similar feature in which OCC (i+ 1, t - 5) replaced OCC (i+ 1, t - 2) as it appears in the definition of DOCCTD given above. The latter version is sometimes referred to as the modified California algorithm. It is seen that three features are used that have three corresponding thresholds. This algorithm is an example of a decision tree. The algorithm is executed in a sequence of steps, starting at the top, or root node (see Figure 1). The first test is made, and a branch to the left (OCCDF $\geq T_1$) or right (OCCDF $< T_1$) is followed. In this instance, a branch to the right yields a final designation of incident-free conditions, which are coded here as a state value of 0. Otherwise, subsequent tests are made in a similar manner, until a final designation of incident or incident-free is made.

This algorithm is executed for each adjacent pair of sensor stations, i.e., for sections of the freeway bounded by the sensor stations, at regular intervals of 20 s (as in Los Angeles), 30 s (as in Chicago and Minneapolis), or 1 min (as in our study).

The structure of the California algorithm derives from a consideration of the pattern of traffic data that typically arises when an incident occurs. An example of this pattern is given in Table 1, which uses actual values of occupancy taken from the Los Angeles system. The pattern that develops can be explained on a theoretical basis, but it is sufficiently evident that here we appeal only to driving experience and simple facts. The incident reduces the capacity of the freeway at the site of incident. If the resultant capacity is less than the volume of traffic upstream, congestion builds up (i.e., queueing develops); the boundary of the congested region propagates in an upstream direction on the freeway at a typical speed of 8 to 16 km/h (5 to 10 mph) (although this speed depends on the particular values of the capacity at the site and the upstream volume). The congested area includes the upstream sensor stations in an orderly sequence following the occurrence of the incident.

Downstream of the site of the incident, the freeway is cleared of traffic; the boundary of the cleared region propagates downstream at a speed that may be as high as 80 km/h (50 mph).

Thus, an incident creates, after some interval of time, a significant difference in occupancy values at the sensor stations bounding the site of the incident. The features OCCDF and OCCRDF are intended to measure this effect. Two features are used (rather than OCCRDF alone, for example) to avoid problems in light traffic where OCCRDF might be briefly large due to normal fluctuations in traffic data.

However, other normal conditions can also produce significant differences in occupancy values. Geometric bottlenecks are a frequent source of such differences. The occurrence of an incident can often be distinguished by the fact that the downstream occupancy decreases rather abruptly, so that the feature DOCCTD would have, briefly, an unusually large value. This is the third feature used in the California algorithm.

Figure 1. California algorithm.

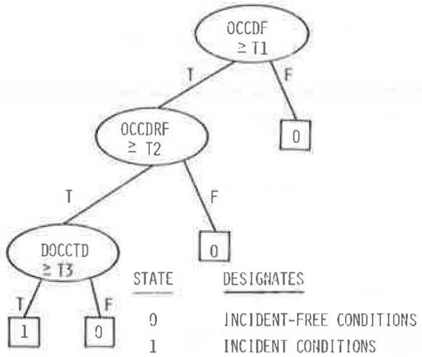


Table 1. Occupancy values: incident data set 74051501, Santa Monica Freeway eastbound.

Time	1-Min Occupancy at Station*						
	21	22	23	24	25	26	27
7:05	15	17	15	16	16	17	15
7:06	15	18	13	13	15	16	15
7:07	16	16	15	15	15	15	14
7:08	14	17	17	15	17	15	15
7:09	15	17	17	16	16	15	16
7:10	16	18	18	19	15	15	14
7:11	17	17	19	17	16	16	15
7:12	18	19	15	17	18	15	15
7:13	15	19	17	16	20	18	20
7:14	16	17	17	17	18	18	25
7:15	18	19	18	16	17	20	21
7:16	18	17	17	18	19	15	23
7:17	17	21	17	19	21	14	17
7:18	14	20	19	17	43	10	19
7:19	15	21	20	30	33	10	11
7:20	14	18	18	47	32	10	13
7:21	14	16	34	30	29	9	11
7:22	16	14	30	38	37	12	10
7:23	15	19	21	37	32	10	10
7:24	15	27	20	46	33	11	9
7:25	17	37	27	35	36	11	11
7:26	25	32	43	30	32	12	11
7:27	26	20	26	33	40	11	12
7:28	44	20	20	42	44	12	10
7:29	37	38	20	40	30	15	11
7:30	27	42	28	41	25	14	14
7:31	26	32	37	30	27	14	14
7:32	34	25	29	29	37	11	12
7:33	30	22	33	43	27	14	12
7:34	28	35	37	32	24	13	13
7:35	26	44	38	32	22	12	12
7:36	35	28	28	29	22	14	13
7:37	54	33	27	30	29	12	14
7:38	35	42	23	29	30	12	13
7:39	30	37	21	41	29	13	13
7:40	49	31	24	36	40	14	13

Note: Incident occurred at 7:15:40 between stations 25 (upstream) and 26 (downstream).
*Station indexes increase in direction of travel.

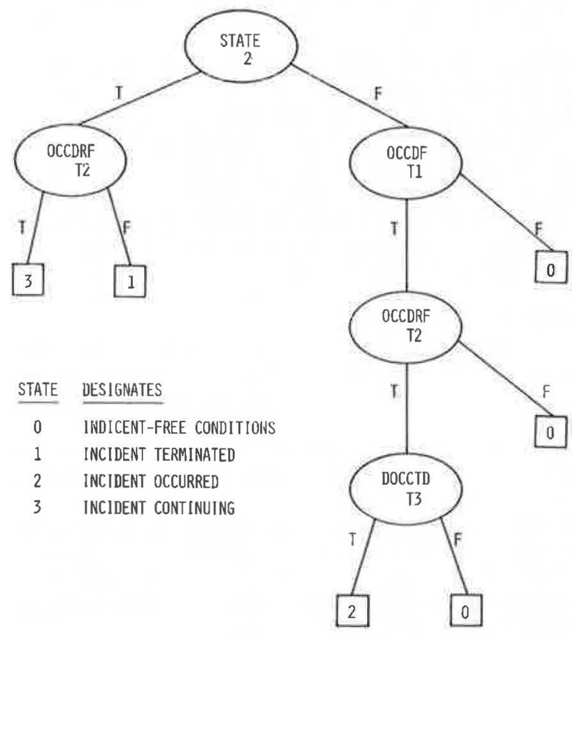
DECISION TREES WITH STATES

In freeway operations, it is desirable not only to detect the occurrence of an incident, but also to know when it has terminated. A simple means for accomplishing this is to augment the California algorithm in the manner suggested in Figure 2. Note that we now make use of four state values: 0 corresponds to incident-free conditions, 1 corresponds to termination, 2 corresponds to the initial detection of the incident, and 3 corresponds to the continued presence of incident conditions. The state values are retained for each section and used in the subsequent application of the algorithm. For example, Figure 2 shows that the first test in the algorithm is state ≥ 2. This is a decision tree with states.

We have found this general structure to be useful in a variety of ways; many more state values can be identified to provide a more refined classification of the traffic condition.

Any algorithm based on a decision tree can be defined by a data structure and a general algorithm for running down the tree. As a result, such algorithms are very easy to implement and extremely fast to execute. De-

Figure 2. Refinement of California algorithm (algorithm 2): example of decision tree with states.



tails of these matters are beyond the scope of this paper, but can be found elsewhere (5, 6, 7).

PERFORMANCE MEASURES AND THRESHOLD CALIBRATION

Up to this point, we have been concerned with features and the algorithm structure. Completion of the specification of an algorithm requires specification of the set of thresholds, e.g., T_1, T_2, and T_3 in Figure 1. As previously implemented, the thresholds for the California algorithm were determined by a trial-and-error process. A commonly used set is $T_1 = 8$, $T_2 = 0.5$, and $T_3 = 0.15$. Our purpose here is to describe a systematic threshold-calibration technique. This technique is designed to yield optimal performance as measured by false-alarm and detection rates.

Four possibilities arise when an incident-detection algorithm is executed, as indicated below.

Condition Indicated by Algorithm	Actual Condition: Incident-free	Incident
Incident-free	–	Missed detection
Incident	False alarm	Detection

By accumulating the results over a number of tests and defining N_F = total number of tests performed by the algorithm and N_{FA} = total number of false-alarm signals generated by the algorithm, we can compute the false-alarm rate as

$$\alpha = \text{false-alarm rate (percent)} \equiv 100 \times N_{FA}/N_F \tag{1}$$

Next, by defining N_I = number of incidents and N_D = number of incidents detected, we have

$$\beta = \text{detection rate (percent)} \equiv 100 \times N_D/N_I \tag{2}$$

To determine that a valid detection has been made, we must define the spatial and temporal deviations allowable between the actual incident event and the detection event. Our assessments required detections in a time interval beginning 5 min before and ending 20 min after the estimated time of occurrence and corresponding to the actual section on which the incident occurred or the next downstream section.

Other measures are also of interest, e.g., the mean time to detect but, in the interest of brevity, we shall confine our attention in this paper to the measures defined above. Information pertaining to the mean time to detect can be found elsewhere (5).

The false-alarm and detection rates clearly depend on the choice of threshold set, which we will denote as a whole by T. Suppose there is a requirement on the detection rate; if an allowable threshold set T is one for which $\beta(T) \geq y$, y is the minimum detection rate specified. Of all such threshold sets, the best one is that which minimizes the false-alarm rate. This leads to consideration of the problem

$$\min_T \{\alpha(T)|\beta(T) \geq y\}$$

The solution of this problem yields a threshold set $T^*(y)$, a detection rate $\beta[T^*(y)] \geq y$, and a false-alarm rate $\alpha[T^*(y)]$. By varying the parameter (y), one obtains corresponding threshold sets and false-alarm rates. Threshold calibration is thus reduced to a sequence of constrained minimization problems, where the functions involved [$\alpha(T)$ and $\beta(T)$] are defined by application of the algorithm to the available data. This procedure was mechanized in our study and used extensively to produce threshold sets. Details of the procedures and software are available elsewhere (5, 7).

Note the inherent trade-off: as the requirement on the detection rate is increased, the corresponding best false-alarm rate is also increased. The results of the calibration are therefore in the form of a trade-off curve.

PATTERNS IN TRAFFIC DATA

In the course of developing and evaluating incident-detection algorithms, we have discovered that certain types of patterns in the traffic data appear repeatedly and have developed an understanding of traffic conditions that provides rational explanations for the appearance of such patterns. In this section, we describe and rationalize these patterns so that the directions used in the development of the algorithms can be more easily understood.

Incident Patterns

Depending on the nature of an incident and the traffic conditions prevailing at the time, the pattern that the traffic data develop in the presence of that incident may be one of five types.

The first type is the most distinctive in that it is the most easily discriminated from patterns associated with incident-free conditions. This pattern occurs when the capacity at the site of the incident is less than the volume of oncoming traffic so that a queue develops upstream. Simultaneously, a region of light traffic develops downstream. An example of this pattern is illustrated by the occupancy data given in Table 1. This pattern is clearest when traffic is flowing freely before the incident occurs.

The second type of pattern occurs when the prevailing traffic condition is freely flowing but the impact of the incident is less severe (for example, as might result from a lane blockage that yields a capacity at the site of the incident that is greater than the volume of oncoming traffic). This situation is more difficult to distinguish from certain incident-free patterns and, therefore, may not result in a detection.

The third type of pattern occurs, again, in freely flowing traffic, but the impact of the incident is not noticeable in the traffic data. This may occur when the incident is a disabled vehicle in the median. Incident-detection algorithms cannot be expected to detect incidents of this sort.

The fourth type is one that occurs in heavy traffic when the capacity at the incident site is less than the volume (and the capacity) of traffic downstream. This difference leads, generally, to a clearance in the region downstream of the incident. Here the traffic pattern evolves rather slowly, and a distinguishable pattern develops only after several minutes. Obviously, a very severe incident in which several lanes were blocked would result in rapid development of the pattern, but the more typical situation is one involving a slowly developing pattern.

The fifth type occurs in heavy traffic in which the capacity at the site of the incident is not less than the downstream volume. The effects are then localized and are not noticeable in the traffic data. Incident-detection algorithms cannot be expected to detect incidents of this sort.

Patterns in Incident-Free Traffic That Tend to Produce False Alarms

Four types of patterns occur under incident-free conditions that are similar to incident patterns and therefore tend to produce false alarms. The first of these is that related to malfunctioning detectors (5).

The second type arises in heavy traffic in which individual vehicles experience significant speed variations. This phenomenon shows up in traffic data in the form of compression waves that propagate in a direction counter to the flow of traffic. Several such waves can be seen in the occupancy data given in Table 2. An examination of these data shows significant station-to-station differences in occupancies of the same magnitude as is seen in patterns related to incidents. This pattern is the most significant contributor to false alarms for algorithms identified before our work (at least as measured with respect to the Los Angeles data).

The third is associated with abnormal geometrics, for example, that which is found at freeway-to-freeway interchanges.

The fourth is associated with bottlenecks (for example, at locations that have a substantial volume of onramp traffic). These bottleneck locations are such that the total demand exceeds the capacity of the freeway. Examples of these latter two types of patterns are given elsewhere (5).

Consequences for Algorithm Development

It should be evident from this disucssion that effective incident-detection algorithms require more than an identification of a discontinuity in the traffic data—there are numerous such occurrences in incident-free data. The algorithms we have developed explicitly account for the differences in the detailed nature of the discontinuities found in the traffic stream under incident and incident-free conditions. Our developments have led to algorithms that can always detect the first type of incident pattern and can sometimes detect the second and fourth. At the same time, certain algorithms have been developed that are invulnerable to compression waves.

Incidents in light traffic are not generally detected by the algorithms we have developed. Tignor (10) has investigated the application of single-exponential smoothing to the detection in this regime. Another potentially effective means for detecting such incidents involves the use of traffic correlation and is discussed elsewhere (5).

SEVERAL NEW ALGORITHMS

The algorithm depicted in Figure 2 is a slight modification of the California algorithm. We now wish to turn to descriptions of several new algorithms that have been developed to discriminate more effectively, based on a qualitative assessment of patterns in traffic data, and that have, in fact, been found to have superior performance. In all, our research yielded 10 algorithms (including, essentially, the California algorithm and 2 simple variants). Our discussion here will be limited to a discussion of two of these, which for consistency with the references cited are denoted as algorithm 7 and algorithm 8.

Many—but not all—disturbances in incident-free traffic are short-lived and, although they may produce an incident signal, the associated incident-continuing state value does not last long (if it is produced at all). This is in contrast to the majority of incidents that produce a discontinuity in the traffic stream that generally lasts at least several minutes. Thus, it has been suggested that improved performance might be obtained by requiring that the discontinuity persist for a period of time.

The state feature of the binary-tree structure we have adopted provides a convenient way to include a persistence requirement. In its simplest form, one sets the state value equal to, for example, 1 when a discontinuity is first detected and then signals an incident (state value equal to, for example, 2) if the next test (for that station or section) indicates that the discontinuity has persisted.

Algorithm 7, which is illustrated in Figure 3 and has the same state values as algorithm 2, incorporates this persistence requirement and involves replacement of DOCCTD by DOCC, the occupancy at the downstream station. This choice is based on the observations that (a) the most common cause of false alarms in the California algorithm is a compression wave that moves in a direction counter to the flow of the traffic and (b) in heavy traffic that has compression waves, the downstream occupancy rarely drops below 20 percent (in the Los Angeles data), whereas incidents generally produce downstream occupancies substantially less than 20 percent.

The application of the threshold-calibration methodology described above yielded the performance measures and threshold sets given in Table 3.

Compression waves in heavy traffic are a principal source of false alarms, and some success in eliminating such false alarms is obtained by using algorithm 7. Further efforts to improve performance in heavy traffic involved attempting to account for the regularity in the pattern of traffic associated with compression waves.

Consider the data given in Table 2 for an incident-free data set. As we have noted, compression waves are manifested by sudden large increases in occupancy that move through the traffic stream in a direction counter to the direction of travel. Attempts were made to account for this pattern through the use of correlation analysis (5), but the patterns were not sufficiently regular for this technique to be successful.

Therefore, we considered a more gross way to account for the observed patterns. As Table 2 shows, a large increase at one station is typically followed some 2 to 5 min later by a large increase at the next station upstream. The typical station spacing in Los Angeles is 0.8 km (0.5 mile), corresponding to a shock wave speed of 9.7 to 24 km/h (6 to 15 mph). The fact that a compression wave has passed over a station can be captured by the simple test shown in Figure 4, where $T_5 = 30$ and $T_2 = -0.250$ have been found to be effective.

Of course, incidents also produce patterns that would yield positive results for this test. In heavy traffic that has compression waves, one would expect that a positive result for this test would be followed, typically within 5 min, by a positive result at the next station upstream. Thus, incidents are distinguished by the absence of a compression wave at the downstream station in the previous 5 min.

Algorithm 8, which is shown in Figure 5 and has state values as defined below, uses the state feature in an unusual way.

State	Designates
0	Incident-free
1	Compression wave downstream in this minute
2	Compression wave downstream 2 min ago
3	Compression wave downstream 3 min ago
4	Compression wave downstream 4 min ago
5	Compression wave downstream 5 min ago
6	Tentative incident
7	Incident confirmed
8	Incident continuing

Essentially, this algorithm suppresses incident detection at any station for a period of 5 min after detection of a compression wave at the downstream station. This algorithm, which has both a persistence requirement and a continuing incident state, may appear to be com-

Table 2. Occupancy values: incident-free data set 74090454, San Diego Freeway southbound.

	1-Min Occupancy at Station[a]						
Time	32	31	30	29	28	27	26
7:10	15	20	20	18	22	26	22
7:11	13	21	19	18	22	28	26
7:12	16	19	20	19	21	32	26
7:13	13	18	16	18	30	25	25
7:14	14	22	17	18	25	23	24
7:15	14	20	20	26	44	29	26
7:16	14	18	18	25	34	26	24
7:17	13	21	19	36	26	21	25
7:18	14	24	21	48	29	25	21
7:19	16	26	32	28	31	26	25
7:20	21	24	47	19	26	39[b]	23
7:21	14	26	32	27	26	21	22
7:22	14	52	32	22	29	19	23
7:23	14	27	23	20	50[b]	18	24
7:24	13	26	21	21	30	22	26
7:25	24	21	22	62[b]	23	26	24
7:26	39	20	23	38	23	28	23
7:27	23	21	65[b]	29	22	30	23
7:28	26	24	43	28	23	23	25
7:29	31	26	26	29	22	30	23
7:30	30	60[b]	22	35	22	24	23
7:31	31	41	21	30	17	26	24
7:32	37	29	27	26	23	18	26
7:33	50[b]	26	35	22	37	22	24
7:34	53	22	31	21	29	26	26
7:35	48	21	32	21	25	22	23
7:36	29	28	33	39	21	24	23
7:37	37	33	28	26	22	30	27
7:38	38	29	44	21	20	23	24
7:39	40	25	38	21	21	20	27
7:40	53	23	43	19	30	23	24
7:41	37	47	44	22	36	26	23
7:42	41	30	42	23	38	28	26
7:43	38	26	38	21	31	22	25
7:44	56	24	29	33	29	23	22
7:45	64	25	24	38	27	27	24

[a] Station indexes increase in direction of travel.
[b] Large-occupancy value indicative of compression wave.

Table 3. Thresholds and performance results.

Algorithm	T_1	T_2	T_3	T_4	T_5[a]	Detection Rate (%)	False-Alarm Rate (%)
2	5.4	0.325	0.011			82	1.341
	6.8	0.307	0.056			71	0.883
	15.0	0.335	-0.050			61	0.346
	9.9	0.552	-1.112			51	0.169
	9.5	0.629	-1.746			41	0.064
	21.3	0.646	-2.080			31	0.026
	29.9	0.685	-1.959			20	0.011
7	8.1	0.313	16.8			59	0.134
	12.9	0.360	16.6			51	0.050
	13.1	0.358	15.8			49	0.043
	9.6	0.359	12.3			41	0.029
	13.1	0.393	12.5			37	0.017
	21.6	0.301	13.9			31	0.006
	26.6	0.322	13.4			20	0.004
8	10.2	-0.443	0.312	28.8	30	61	0.177
	13.1	-0.296	0.309	15.9	30	51	0.038
	18.1	-0.310	0.356	18.5	30	41	0.024
	5.2	-0.401	0.590	27.9	30	31	0.010
	24.4	-0.392	0.579	13.0	30	20	0.003

[a] Held fixed.

plex because it involves 21 decision nodes, but its components have straightforward interpretations (5).

Application of the threshold calibration methodology yielded the performance measures and threshold sets given in Table 3.

ONE OLD ALGORITHM

To compare the new algorithms with a popular approach to the construction of incident-detection algorithms—double exponential smoothing—we include here a brief description of one further algorithm, designated algorithm 11 (5). This was the best performing of this type of algorithm that we considered. The basis for this type of algorithm is the smoothing of surveillance data, e.g., OCCDF(t), according to

$$S_1(t) = \alpha \mathrm{OCCDF}(t) + (1 - \alpha)\, S_1(t-1) \qquad (3)$$

and

$$S_2(t) = \alpha S_1(t) + (1 - \alpha)\, S_2(t) \qquad (4)$$

These derived functions, $S_1(t)$ and $S_2(t)$, are used to provide a forecast, e.g., of OCCDF(t), and the accumulated forecast error is then used as the basis for the algorithm.

Results of calibration and evaluation of algorithm 11 are given below.

Threshold	Detection Rate (%)	False-Alarm Rate (%)
-3.29	71	0.705
-3.88	61	0.373
-4.52	51	0.200
-5.06	41	0.126
-6.14	31	0.048
-7.52	20	0.021

COMPARISON OF ALGORITHM PERFORMANCE

Algorithms 2 and 11 are compared in Figure 6 on the bases of detection and false-alarm rates. At higher false-alarm rates, algorithm 11 has a slight advantage; at lower false-alarm rates, suitable for most operational purposes, algorithm 2 is clearly superior.

Figure 3. Decision tree for algorithm 7.

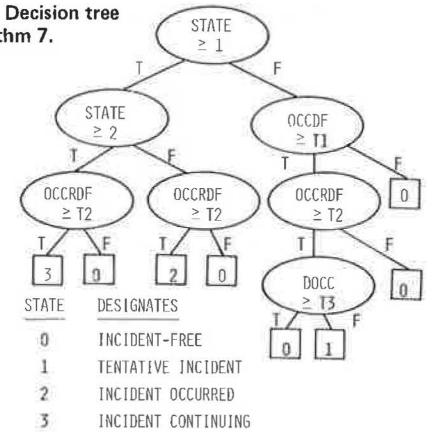
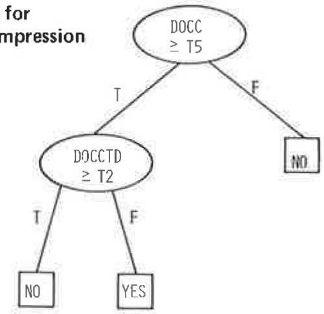


Figure 4. Test for presence of compression wave.



To provide an independent evaluation of algorithm performance, the ability of algorithms 2, 7, and 8 to detect incidents in an expanded data base involving 118 incidents was tested. (A data base with 50 incidents was used for the calibration). The results are shown in Figure 7. Algorithm 8 is generally superior, because of its superior handling of compression waves (present in much of the Los Angeles data); algorithm 7 gives somewhat poorer performance (but it is much simpler); and the California algorithm, algorithm 2, generally gave the poorest performance.

These relative performances held up when tested by using data taken from the Minneapolis system, thus providing evidence for the transferability of the research results to other freeway surveillance systems.

CONCLUSIONS

In this paper we have discussed the advantages associated with using (a) the persistence and (b) the shock-

Figure 5. Decision tree for an equivalent form of algorithm 8.

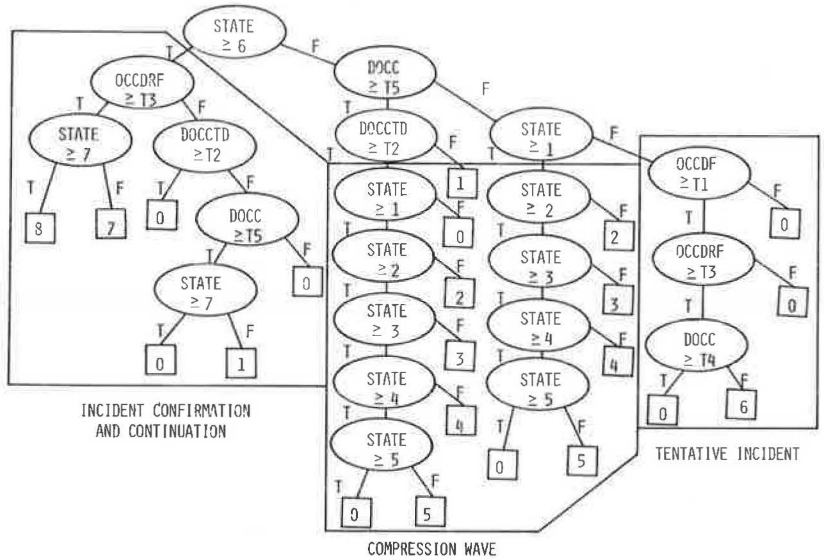


Figure 6. Comparison of performance: algorithms 2 and 11.

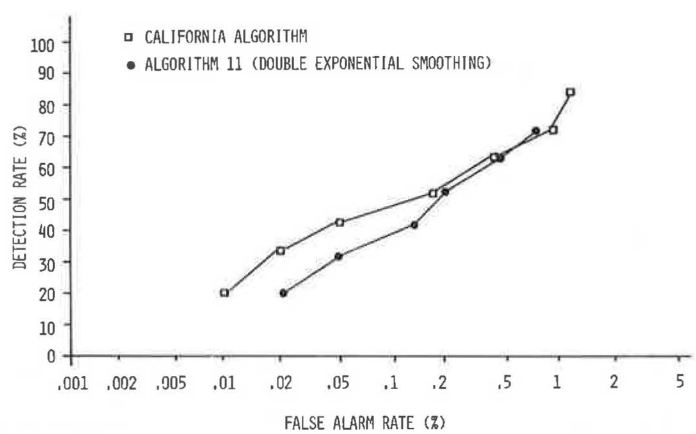
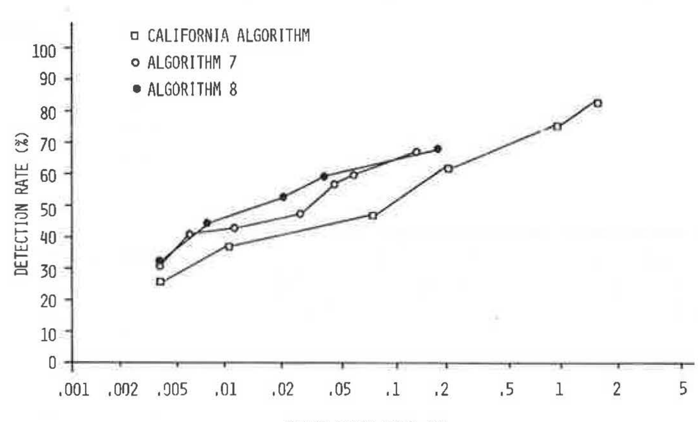


Figure 7. Comparison of performance: algorithms 2, 7, and 8.



wave detection capabilities to limit the occurrence of false alarms. These capabilities were built into two new incident-detection algorithms that are identified as algorithm 7 and algorithm 8; they have been shown to provide performance that is superior to that of the most commonly, previously used algorithm—the California algorithm.

Both algorithms 7 and 8 are primarily intended for use during moderate to heavy traffic (as are most previously developed incident-detection algorithms). Detection of incidents under low-volume conditions has been and continues to be an unresolved problem. Although it is sometimes important to be able to detect incidents at times of low volume, the greater need is for good incident-detection algorithms in moderate-to-heavy traffic conditions. It is during these times that traffic management is seriously hampered by undetected incidents. These algorithms can be implemented by using FORTRAN (6).

In the course of this work, we also identified an alternative approach for the construction of incident-detection algorithms for low-volume applications. This approach is based on the use of a traffic-correlation analysis (5). Based on preliminary results, it appears that traffic correlation is most consistent in the light-to-moderate volume regime at speeds of about 80 km/h (50 mph). However, additional work will be required before this approach can be implemented.

We also investigated ways to identify the lane(s) in which an incident had occurred (5). The lane location is important from the point of view of communicating specific information to drivers via changeable message signs or radio and also from the point of view of using appropriate ramp-control strategies for managing the freeway disturbance. The lane-location algorithm, which was effectively tested on the data base available, also appears to be potentially useful in reducing the frequency of false alarms in bottlenecks. The actual use of the lane-location algorithm for this slightly different purpose has not been attempted to date; however, we believe it offers considerable merit and should be considered in future research on this subject.

ACKNOWLEDGMENT

The research described in this report was supported by the Office of Research, Federal Highway Administration.

REFERENCES

1. A. R. Cook and D. E. Cleveland. Detection of Freeway Capacity-Reducing Incidents by Traffic-Stream Measurements. TRB, Transportation Research Record 495, 1974, pp. 1-11.
2. C. L. Dudek, C. J. Messer, and N. B. Nuckles. Incident Detection on Urban Freeways. TRB, Transportation Research Record 495, 1974, pp. 12-24.
3. H. J. Payne, D. N. Goodwin, and M. D. Teener. Evaluation of Existing Incident-Detection Algorithms. Federal Highway Administration, Rept. FHWA-RD-75-39, Feb. 1975; NTIS, Springfield, VA, PB241-883/AS.
4. H. J. Payne. Development and Testing of Incident-Detection Algorithms: Vol. 1, Summary of Results. Federal Highway Administration, Rept. FHWA-RD-76-19, April 1976.
5. H. J. Payne, E. D. Helfenbein, and H. C. Knobel. Development and Testing of Incident-Detection Algorithms: Vol. 2, Research Methodology and Detailed Results. Federal Highway Administration, Rept. FHWA-RD-76-20, April 1976.
6. H. J. Payne and H. C. Knobel. Development and Testing of Incident-Detection Algorithms: Vol. 3, User Guidelines. Federal Highway Administration, Rept. FHWA-RD-76-21, April 1976.
7. H. C. Knobel and E. D. Helfenbein. Development and Testing of Incident-Detection Algorithms: Vol. 4, Program Documentation for Evaluation, Calibration, and Implementation. Federal Highway Administration, Rept. FHWA-RD-76-22, April 1976.
8. H. J. Payne and E. D. Helfenbein. Freeway Surveillance Data. Federal Highway Administration, Rept. FHWA-RD-76-74, April 1976.
9. S. C. Tignor and H. J. Payne. Improved Incident-Detection Algorithms. Public Roads, June 1977, pp. 32-40.
10. S. C. Tignor. Operational Analysis and Improvements for Freeway Moving-Merge Systems. Department of Civil Engineering, Univ. of Michigan, PhD dissertation, 1974.

Publication of this paper sponsored by Committee on Freeway Operations.

**Dr. Payne was with the Technology Service Corporation, Santa Monica, California, when this research was conducted.*

Field Evaluation of Messages for Real-Time Diversion of Freeway Traffic for Special Events

Conrad L. Dudek, Graeme D. Weaver, Donald R. Hatcher, and Stephen H. Richards, Texas Transportation Institute, Texas A&M University

This paper presents the results of special-event route-diversion studies conducted in Dallas to evaluate 14 primary-candidate real-time messages that had resulted from extensive human-factors laboratory studies. The messages were displayed on matrix signs located on the freeway. The re-

sults of the study showed that all messages tested were effective in diverting freeway traffic to an arterial alternative route. The messages were equally effective before events that had fixed starting times (e.g., football games). Before events for which the arrival times were not set but spread over long periods of time (e.g., state fairs), some of the messages produced higher diversion rates than others, although all the messages resulted in high diversion rates.

Special events at places such as stadiums attract large volumes of traffic and create significant congestion at the site and on the adjacent freeways. Almost every driver who has attended a ball game at a major stadium, for example, has experienced considerable delay while stopped in lengthy queues of traffic. Often, less congested alternative routes are available and, if some of the approaching freeway traffic can be diverted to those routes, freeway congestion can be reduced.

As a part of a study of the human-factors requirements for real-time motorist information displays, the Texas Transportation Institute, in cooperation with the city of Dallas, conducted route-diversion studies to evaluate primary candidate messages that had resulted from extensive human-factors laboratory studies. The studies were conducted on three weekends when special events were held at the Dallas Fair Park:

1. July 4, 1976—annual fireworks display and special bicentennial-year activities,
2. October 9, 1976—the annual football game between the Universities of Texas and Oklahoma and also opening day of the annual Texas State Fair and October 10, 1976—the annual Texas State Fair, and
3. January 1, 1977—the annual Cotton Bowl football game.

The results of the July 4th studies were presented in an earlier paper (1). The objectives of that study were to evaluate two basic messages by using two lamp-matrix changeable-message signs. The results clearly showed that the matrix-sign messages greatly influenced route diversion to the alternative route. The October 9-10th and the January 1st studies were designed to evaluate several other candidate messages.

BACKGROUND

Fair Park Characteristics

Fair Park, a 97.5-hm^2 (240-acre) area located in south central Dallas, houses the Cotton Bowl and the permanent buildings, facilities, and midway for the annual Texas State Fair, in addition to several cultural buildings such as a music hall and museums. The Cotton Bowl, which has a seating capacity of 73 000 people, is a stadium in which many collegiate football games are played and is often used for other events and exhibitions.

The location of Fair Park relative to the freeway system is shown in Figure 1. The primary access route to the park and its associated parking facilities for traffic originating from the north, west, and south is provided by two exit ramps (Second Avenue and Haskell Avenue exits) from I-30 (R. L. Thornton Freeway). Traffic from US-75 (Central Expressway), I-35E (Stemmons Freeway), and the Dallas-Fort Worth Turnpike must connect with I-30 and then exit via either Second Avenue or Haskell Avenue. Queues of traffic exiting to Fair Park often extend 3.2 km (2 miles) back during peak demand conditions and affect operations on not only I-30, but also on US-75 and I-35E. This is a common occurrence before such football games as the Cotton Bowl game on New Year's Day or the Oklahoma-Texas game in October. Each of these games attracts more than 73 000 people. In addition, the October game coincides with opening day at the Texas State Fair, which attracts an additional 225 000 people to Fair Park on the same day.

Primary and Alternative Routes

The primary route, Central Expressway, and the alternative route, Fitzhugh Avenue, are shown in Figure 2. The expressway from Fitzhugh Avenue to I-30 is a six-lane facility that has a direct two-lane ramp connection to eastbound I-30. Fitzhugh Avenue is a two-way arterial that, about 0.4 km (0.25 mile) east of the expressway, becomes half of a pair of one-way routes (Bennett Avenue is the other). The street narrows from three to two lanes approximately 1.6 km (1 mile) east of the expressway. It has eight traffic signals between the expressway and Fair Park. For the studies, the traffic signals at the Fitzhugh-expressway diamond interchange were timed to favor the expected left-turn demand, and those along Fitzhugh Avenue were timed and coordinated to provide steady progression. Parking was prohibited along Fitzhugh Avenue to provide a minimum of two operating lanes throughout the complete alternative route.

Matrix-Sign Characteristics

The portable matrix signs are illustrated in Figure 3. The signs can display messages on two lines by using 46-cm (18-in) high characters. A computer located in an environmental cabinet on the front side of the trailer provides almost infinite capability for message selection and display. The messages can be displayed in a stationary mode or can be flashed or alternated with other messages. Message-display commands are made through a teletype available on each sign trailer. Each sign system can be connected to the regular line power or to a generator.

Matrix-Sign Locations

The trailer-mounted lamp-matrix signs were positioned on the overcrossing structures above the southbound lanes of Central Expressway at two locations upstream from the Fitzhugh Avenue exit ramp. Sign locations were selected with regard to several factors: (a) availability of overcrossing structures containing U-turn bays, (b) adequate tangent sight distance to the sign, and (c) adequate distance for lane-changing maneuvers in response to the sign message.

For the October 9-10th and January 1st studies, the matrix signs were positioned at the Lovers Lane and Yale Boulevard overcrosses, approximately 4.17 and 3.04 km (2.6 and 1.9 miles) north of the Fitzhugh Avenue exit ramp respectively.

Static Trailblazer Signs Along Alternative Route

The trailblazer signs used for the July 4th studies were 61 × 61-cm (24 × 24-in) sheeted aluminum. The sign background was dark brown, the legend "Fair Park" and the arrows were white, and the FP logo was orange and yellow with a black border. The legend, "Fair Park," was in 10.4-cm (4-in) series C letters. The supplementary panels, white with a brown legend, carried 6.4-cm (2.5-in) series C lettering.

After the July 4th studies, it was decided to enlarge the logo signs to 91.5 × 91.5 cm (36 × 36 in) and make the supplementary panels larger also. The revised signs carried 10.4-cm series C letters on the upper supplementary panels. The first four trailblazer signs

Figure 1. Dallas Fair Park.

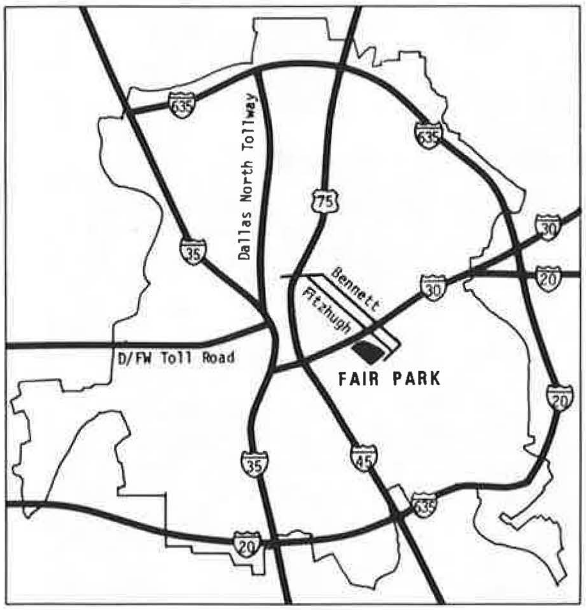
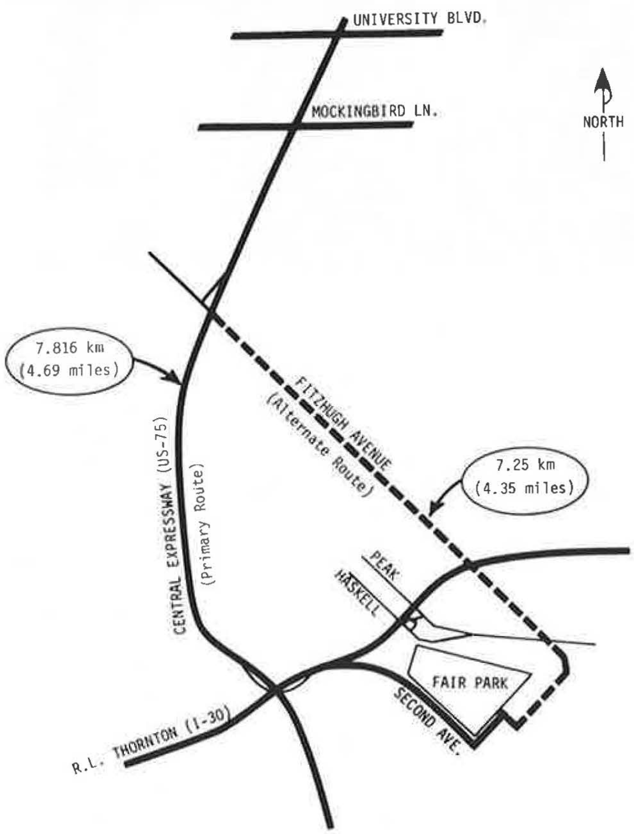


Figure 2. Primary and alternative routes.



marking the route to Fair Park are shown in Figure 4.

Field Study Procedure

The study procedure involved a comprehensive license-plate origin-destination technique identical with that used in the July 4th study and described in detail elsewhere (1).

Figure 3. Trailer-mounted matrix sign.

OCTOBER 9-10TH SINGLE-POINT-DIVERSION STUDY

The Texas-Oklahoma football game is played annually at the Cotton Bowl on the first weekend in October. The 1976 game was scheduled for October 9, which coincided with the opening day of the Texas State Fair. More than 277 000 people attended the opening day of the fair, of whom 72 000 attended the football game. More than 245 000 people attended the fair on Sunday, October 10, 1976. These weekend events were selected as the second series of events on which to conduct single-point-diversion studies.

Study Objectives

Historically, traffic congestion on the approaches to the fair park (including the adjacent freeways) has been severe on the opening day of the Texas State Fair. Also, at times when collegiate football-game attendance is near Cotton Bowl capacity, heavy traffic congestion is experienced, although for a shorter duration. The combination of the two events on one day was influential in selection of the matrix-sign message display. The laboratory studies indicated that it would be desirable to display both temporal information in terms of the expected delay on the primary route or the time savings on the alternative route and the reasons why the alternative route should be preferred. Because heavy traffic concentrations were highly probable on both dates, the messages were designed to display temporal information in several ways and also to supply downstream primary-route traffic information to assist drivers in making a diversion decision. The results of the July 4th questionnaire analysis (2) substantiated the laboratory study indication for the need for this information. These results also indicated that it would be desirable to display the fact that the alternative route was well marked. A further requisite was to investigate the use of only one changeable-message sign, rather than two, to produce an acceptable amount of route diversion.

The July 4th event was scheduled to start at a specific time. Similarly, the football game on Saturday afternoon was scheduled to start at a specific time, 3:00 p.m. Conversely, the state fair opened daily at 7:00 a.m. and attracted people throughout the entire day. Therefore, people arrived at the fair throughout the day, whereas those attending the football game arrived over the short time span immediately preceding the game. An additional objective of the study was to evaluate message displays in producing diversion under these two different arrival conditions; i.e., events having variable starting times and events having fixed starting times.

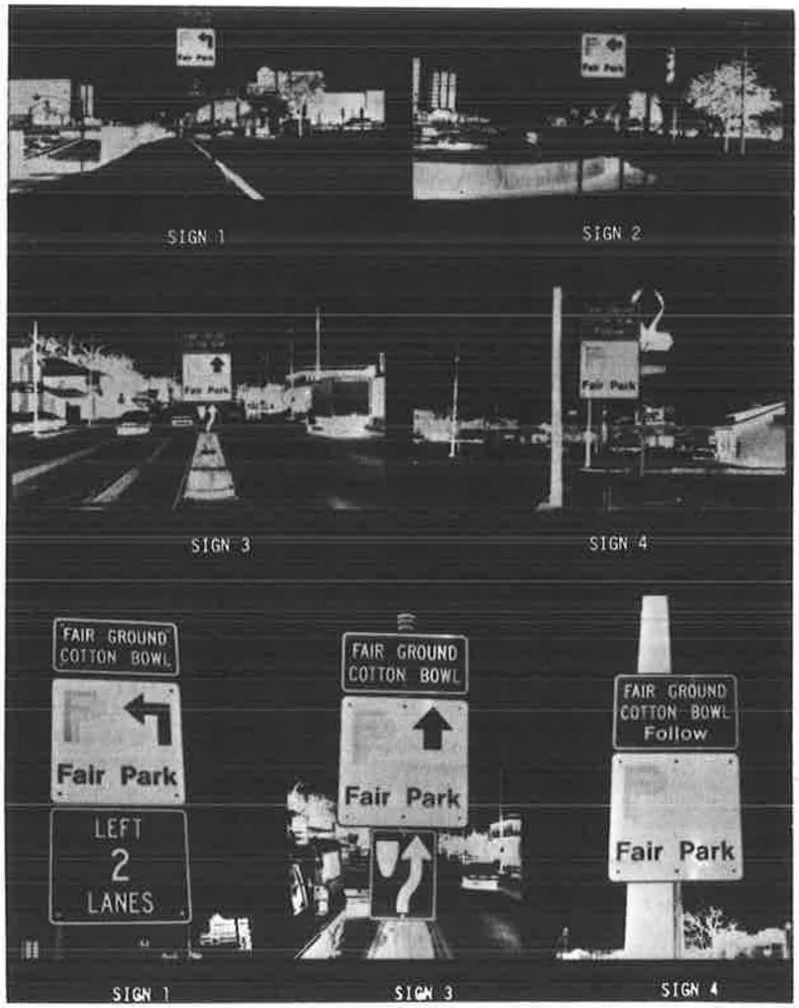


Figure 4. Trailblazer signs along Fitzhugh Avenue: October and January studies.

The objectives, therefore, are summarized:

1. To investigate the route diversion that could be produced by a display of downstream traffic-state descriptors (e.g., avoid traffic jam, congestion, and major delay),
2. To investigate the route diversion that could be produced by a display of temporal information (e.g., XX-min delay, save XX minutes),
3. To investigate the route diversion that could be produced by a display of information that adequate guide signing is provided along the alternative route,
4. To investigate the route diversion that could be produced by using only one matrix sign to further test the need for redundancy, and
5. To investigate the route diversion at events having variable starting times and at events having fixed starting times.

Matrix-Sign Messages

The message sets displayed are shown in Figure 5. Message set 1 was identical to the message displayed during the July 4th study to provide a common base condition. The downstream sign message was constant for all message sets in which two signs were used and was identical to the July 4th message. Message sets 2 through 6 displayed temporal information in terms of time saved and avoidance of delay by diverting to Fitzhugh Avenue. Message sets 7 and 8 displayed traffic-descriptor information pertaining to the primary route and advised drivers that, by diverting to the alternative route, they would avoid the particular traffic state. Message set 9 displayed the information that the alternative route was well guide marked (in response to the suggestions received in the July 4th questionnaires). Message sets 10 through 12 displayed the base message and two types of temporal information and used the downstream sign to investigate the need for redundant message display.

The original intention was to display the sign messages for 15-min intervals separated by 10-min blank conditions to allow traffic to stabilize. The messages were selected to be realistic in terms of the traffic conditions, particularly with respect to a display of the higher value time-saved or delay messages. Continuous monitoring of the traffic situation throughout the study cordon by using radio communication provided real-time information about expected delays, queue lengths, and traffic movement so that messages could be selected that were realistic in terms of temporal information. This was thought to be necessary to achieve credibility in the sign displays.

Response to Matrix-Sign Messages

Table 1 gives the data for and Figure 6a illustrates the percentages of Fair Park-bound traffic that used the Fitzhugh Avenue alternative route throughout the October 9th study periods. This was the day of the football game, which started at 3:10 p.m. Similar data for Sunday, October 10, are given in Table 2 and illustrated in Fig-

Figure 5. Maxtrix-sign messages: October 9-10, 1976.

Message Set	LOVERS LANE SIGN	YALE BLVD. SIGN
Message Set 1	BEST ROUTE TO FAIR GROUNDS / USE FITZHUGH AVE	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 2	FAIR GROUNDS USE FITZHUGH / SAVE 10 MINUTES	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 3	FAIR GROUNDS USE FITZHUGH / AVOID 10 MIN. DELAY	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 4	FAIR GROUNDS USE FITZHUGH / AVOID MAJOR DELAY	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 5	FAIR GROUNDS USE FITZHUGH / SAVE 20 MINUTES	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 6	FAIR GROUNDS USE FITZHUGH / AVOID 20 MIN. DELAY	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 7	FAIR GROUNDS USE FITZHUGH / AVOID CONGESTION	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 8	FAIR GROUNDS USE FITZHUGH / AVOID TRAFFIC JAM	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 9	BEST ROUTE TO FAIR GROUNDS / USE FITZHUGH FOLLOW SIGNS	FAIR GROUNDS BEST ROUTE / FITZHUGH AVE 2 MILES
Message Set 10	Sign Not Used	BEST ROUTE TO FAIR GROUNDS / USE FITZHUGH AVE
Message Set 11	Sign Not Used	FAIR GROUNDS USE FITZHUGH / SAVE 10 MINUTES
Message Set 12	Sign Not Used	FAIR GROUNDS USE FITZHUGH / AVOID 10 MIN. DELAY

ure 6b. The time periods shown reflect the travel time between the Yale Boulevard sign and the Fitzhugh Avenue exit. These data are summarized below.

Time Period	No. of Drivers Using Fitzhugh Avenue to Fair Park	No. of Drivers Using US-75 and I-30 to Fair Park	Total Fair Park Demand	Percentage of Total Demand Using Fitzhugh Avenue
Saturday a.m.				
Sign on	159	61	220	72.3
Sign off	40	70	110	36.4
Difference				35.9
Saturday p.m.				
Sign on	269	46	315	85.4
Sign off	107	128	235	45.4
Difference				40.8
Sunday a.m.				
Sign on	162	65	227	71.4
Sign off	10	85	95	10.5
Difference				60.9
Sunday p.m.				
Sign on	255	103	358	71.2
Sign off	29	145	174	16.7
Difference				54.5

To ensure that all vehicles passing the matrix sign were included in the analysis time interval, a floating vehicle left the matrix-sign position at the instant a message was activated or terminated and traveled along the Central Expressway at traffic-stream speed. As the floating vehicle passed the Fitzhugh Avenue exit ramp and overcrossing, the time was recorded. By knowing the travel time of this vehicle, vehicles could then be positively identified as having passed the sign with respect to a specific display.

Traffic response to the matrix signs was not expected to be identical during the pregame time period (Saturday afternoon) and the Saturday morning period and Sunday throughout the day when the fair activities were the only attraction.

A notable feature of the Saturday study was that, in the morning and again in the afternoon just before the football game, Fitzhugh Avenue traffic was much heavier when the signs were blank compared with the July 4th and October 10th studies. An average of 36.4 and 45.4 percent of the Fair Park-bound traffic used this route without a sign message on Saturday morning and afternoon. This is attributed to several factors: (a) the static trailblazer signs had been in place for 3 months and probably had been seen by many local drivers, (b) local drivers are aware of the expected congestion on the freeway near Fair Park, (c) the Fitzhugh Avenue route had been publicized through brochures distributed by the Fair Park publicity department, and (d) local drivers are well aware that the entrances to the fairground parking lots are located near Fitzhugh Avenue.

A review of the table above and Figure 6a shows that, on Saturday afternoon before the football game, an average of 85.4 percent of the Fair Park-bound drivers took the alternative route along Fitzhugh when a message was displayed on the matrix sign. This is almost double the nonmessage condition.

An important finding is that the percentage diversion was about the same regardless of the message used.

Table 1. Route choice by Fair Park drivers: October 9th study.

Time Period	Message Displayed	No. of Drivers Using Fitzhugh Avenue to Fair Park	No. of Drivers Using US-75 and I-30 to Fair Park	Total Fair Park Demand	Percentage of Total Demand Using Fitzhugh Avenue	Percentage Diversion to Fitzhugh Avenue[a]
Saturday morning: Texas state fair						
10:00-10:15	Blank	10	25	35	28.6	
10:15-10:29	Message set 1	47	24	71	66.2	29.8
10:29-10:39	Blank	14	24	38	36.2	
10:39-10:54	Message set 2	62	22	84	73.8	37.4
10:54-11:04	Blank	16	21	37	43.2	
11:04-11:15	Message set 3	50	15	65	76.9	40.5
Saturday afternoon: football game						
12:30-12:45	Blank	19	27	46	41.3	
12:45-12:52	Message set 10	49	6	55	89.1	43.7
12:52-1:24	Blank	50	63	113	44.2	
1:24-1:39	Message set 11	90	12	102	88.2	42.8
1:39-1:48	Blank	12	20	32	37.5	
1:48-1:58	Message set 12	42	9	51	82.4	37.0
1:58-2:06	Blank	13	10	23	56.5	
2:06-2:17	Message set 4	44	11	55	80.0	34.6
2:17-2:25	Blank	9	2	11	81.8	
2:25-2:38	Message set 5	29	4	33	87.9	42.5
2:38-2:45	Blank	4	6	10	40.0	
2:45-2:58	Message set 6	15	4	19	78.9	33.5

[a] Percentage of total demand with matrix message sign on minus average percentage using Fitzhugh Avenue with sign off.

Figure 6. Effects of matrix-sign messages on alternative-route choice: (a) October 9th study and (b) October 10th study.

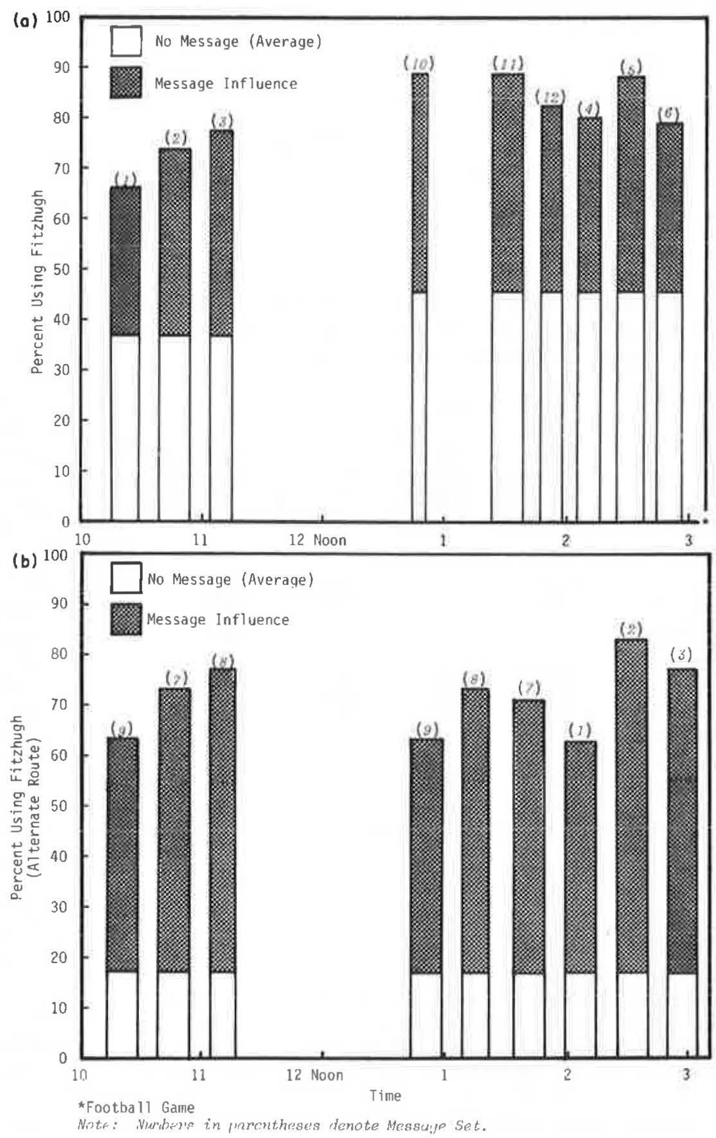


Table 2. Route choice by Fair Park drivers: October 10th study.

Time Period	Message Displayed	No. of Drivers Using Fitzhugh Avenue to Fair Park	No. of Drivers Using US-75 and I-30 to Fair Park	Total Fair Park Demand	Percentage of Total Demand Using Fitzhugh Avenue	Percentage Diversion to Fitzhugh Avenue[a]
Sunday morning:						
Texas state fair						
10:00-10:13	Blank	2	29	31	6.5	
10:13-10:28	Message set 9	45	26	71	63.4	52.5
10:28-10:38	Blank	3	32	35	8.6	
10:38-10:53	Message set 10	63	23	86	73.2	62.7
10:53-11:03	Blank	5	24	29	17.2	
11:03-11:16	Message set 8	54	16	70	77.1	66.6
Sunday afternoon:						
Texas state fair						
12:30-12:43	Blank	9	27	36	25.0	
12:43-12:58	Message set 9	41	24	65	63.1	46.4
12:58-1:08	Blank	2	26	28	7.1	
1:08-1:22	Message set 8	50	18	68	73.5	56.8
1:22-1:33	Blank	5	32	37	7.4	
1:33-1:48	Message set 7	55	23	78	70.5	53.8
1:48-1:58	Blank	5	33	38	13.2	
1:58-2:13	Message set 1	29	17	46	63.0	46.3
2:13-2:23	Blank	6	16	22	27.3	
2:23-2:38	Message set 2	35	7	42	83.3	66.6
2:38-2:48	Blank	2	11	13	15.4	
2:48-3:03	Message set 3	45	14	59	76.3	59.6

[a]Percentage of total demand with matrix message sign on minus average percentage using Fitzhugh Avenue with sign off.

During the afternoon study period, the best-route message and the general traffic-state descriptor and temporal information (XX-min delay, save XX minutes) message were displayed. No discernible difference in effectiveness was found between the messages. Thus, when the special event has a fixed starting time, any one of the messages tested would be effective, provided the information is credible to drivers.

Another important finding is that the diversion can be accomplished by using only one changeable-message sign. Message sets 10, 11, and 12 used only one sign and resulted in 82 to 89 percent use of the alternative route. From this, it may be inferred that one sign that displays a credible message of up to eight short words and is located properly will produce adequate diversion. This indicates that repetition of message elements on two signs is not always necessary if the messages are short and properly formulated.

On Sunday, October 10 (the table above and Figure 6b), an average of 10.5 and 16.7 percent of the Fair Park-bound traffic used the Fitzhugh route during the morning and afternoon study periods when no message was displayed. These increased to averages of 71.4 and 71.2 percent when messages were displayed.

All of the messages were effective. However, for the Texas State Fair study periods (Saturday morning and all day Sunday), for which the event did not start at a specific time, certain messages produced higher diversion than others.

Displays of general traffic-state descriptor information and temporal information resulted in higher diversion than the best-route message. This confirmed the results of the July 4th questionnarie study (2). No differences were noted between the traffic-state descriptor and temporal messages.

Informing drivers that adequate guide signing was provided along the alternative route was no better than the best-route message. Thus, it appears from this that drivers find traffic-state descriptor or temporal information more appropriate than merely the advice that the alternative route is well marked.

In summary, the results of the October studies, with respect to the stated objectives for special events that have variable starting times, are that

1. Although the base message, "Best Route to Fair Grounds—Use Fitzhugh Avenue," was effective, traffic-state descriptor messages (e.g., avoid traffic jam, congestion, or major delay) resulted in higher percentages of diversion;

2. The effectiveness of temporal messages (e.g., XX-min delay, save XX minutes) was about the same as the traffic-state messages—i.e., there does not appear to be any advantage in displaying specific temporal information;

3. The display of information that the alternative route is guide marked appears to about equal in effectiveness the base message, "Best Route to Fair Grounds—Use Fitzhugh Avenue"—i.e., both message sets produced lower diversion rates than did temporal or traffic-descriptor information displays for this type of events; and

4. Repetition of sign message elements does not appear to be critical to successful diversion for messages of up to eight words on properly located displays—i.e., either one or two signs produced equal diversion for both fixed and variable starting-time event conditions.

The results of the October studies, with respect to the stated objectives that have fixed starting times, are that

5. For the annual Texas-Oklahoma football game, more drivers anticipated congestion on the primary freeway route and apparently knew about the Fitzhugh Avenue alternative route; all messages studied were about equally effective in diverting traffic to the alternative arterial route.

JANUARY 1ST SINGLE-POINT-DIVERSION STUDY

The Cotton Bowl game is played annually on New Year's Day. The 1977 game attracted 58 500 spectators. This was selected as the third and final weekend event on which to conduct single-point-diversion studies.

Study Objectives

Historically, traffic congestion near Fair Park has been heavy for the 2-h period before the football game. Queues of traffic have extended more than 1.6 km (1 mile) from the Second Avenue exit on the US-75 to I-30

Figure 7. Matrix-sign messages: January 1, 1977.

LOVERS LANE SIGN		YALE BLVD. SIGN
Sign Not Used	Message Set 1	COTTON BOWL USE FITZHUGH AVOID MAJOR DELAY
Sign Not Used	Message Set 2	COTTON BOWL USE FITZHUGH AVOID 20 MIN. DELAY
Sign Not Used	Message Set 3	COTTON BOWL USE FITZHUGH AVOID 30 MIN. DELAY
Sign Not Used	Message Set 4	BEST ROUTE TO COTTON BOWL USE FITZHUGH AVE

ramp connection and for 3.2 km (2 miles) or more on I-30. It was expected that lengthy delays would occur before the game; therefore, this event was selected to evaluate different levels of delay information. Because only a small difference in diversion response was observed in the October studies between 10-min and 20-min delay messages, a 20-min and a 30-min delay message were used on January 1 to determine whether the higher delay value would produce significant diversion increases. Also, the human-factors laboratory studies indicated that the average driver considers a delay of about 20 min to be the maximum tolerable.

The October studies that used only one sign produced high diversion rates. To more freely evaluate the effects of using only one matrix sign, the January messages were programmed by using only the Yale Boulevard sign.

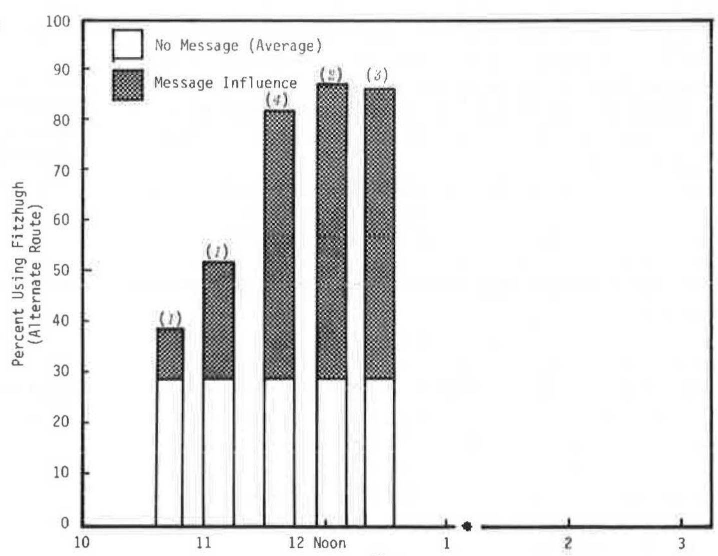
The July 4th study event descriptor on the matrix sign and the trailblazer signs had identified a specific event within the Fair Park complex—"Fireworks." The October study descriptor was the more general—"Fair Grounds." The specific destination descriptor, "Cotton Bowl," was selected for evaluation in the January matrix sign messages.

Table 3. Route choice by Fair Park drivers: January 1st study.

Time Period	Message Displayed	No. of Drivers Using Fitzhugh Avenue to Fair Park	No. of Drivers Using US-75 and I-30 to Fair Park	Total Fair Park Demand	Percentage of Total Demand Using Fitzhugh Avenue	Percentage Diversion to Fitzhugh Avenue[a]
10:15-10:37 a.m.	Blank	5	32	37	13.5	
10:37-10:49 a.m.	Message set 1	11	17	28	39.3	10.5
10:49-10:59 a.m.	Blank	2	14	16	12.5	
10:59-11:14 a.m.	Message set 1	27	25	52	51.9	23.1
11:14-11:29 a.m.	Blank	13	42	55	23.6	
11:29-11:44 a.m.	Message set 4	112	25	137	81.8	53.0
11:44-11:56 a.m.	Blank	33	42	75	44.0	
11:56 a.m.-12:11 p.m.	Message set 2	163	25	188	86.7	57.9
12:11-12:19 p.m.	Blank	14	30	44	31.8	
12:19-12:34 p.m.	Message set 3	137	22	159	86.2	57.4
12:34-12:43 p.m.	Blank	8	25	33	24.2	

[a]Percentage of total demand with matrix message sign on minus average percentage using Fitzhugh Avenue with sign off.

Figure 8. Effects of matrix-sign messages on alternative-route choice: January 1st study.



*Football Game

Note: Numbers in parentheses denote Message Set.

The study objectives are summarized below:

1. To more fully evaluate the use of only one matrix sign,
2. To further evaluate the general traffic-state descriptor message (i.e., avoid major delay),
3. To evaluate a "30-min delay" message in contrast to the "10-min delay" and "20-min delay" messages in the October studies, and
4. To further evaluate specific event descriptors in contrast to a general destination descriptor.

Matrix-Sign Messages

The message sets used are shown in Figure 7. All message sets display "Cotton Bowl" which, on January 1st in Dallas, is believed to convey only one meaning—the annual Cotton Bowl football game.

As was done in the October studies, the message selection was dictated by the traffic conditions. To achieve credibility, the "20-min delay" and "30-min delay" messages were displayed only when congestion in the Fair Park area indicated that delays of these magnitudes would actually occur.

Response to Matrix-Sign Messages

Table 3 gives the data on and Figure 8 illustrates the percentages of Fair Park-bound traffic that used both routes throughout the January 1st study periods. The time periods reflect travel times from the Yale Boulevard sign to the Fitzhugh Avenue exit as determined by the floating-vehicle method described above. The data are summarized below.

Sign	No. of Drivers Using Fitzhugh Avenue to Fair Park	No. of Drivers Using US-75 and I-30 to Fair Park	Total Fair Park Demand	Percentage of Total Demand Using Fitzhugh Avenue
On	450	114	564	79.8
Off	75	185	260	28.8
Difference				51.0

Table 3 and Figure 8 show that the traffic-state descriptor message "avoid major delay" (message set 1) produced lower diversion than did the specific temporal information (message sets 2 and 3) and the base message (message set 4). Message set 1 was displayed early in the study, i.e., from 2 to 2.75 h before the football game in the Cotton Bowl. It is believed that a considerable portion of the Fair Park demand at that time was for the midway activities. Drivers destined for the midway did not feel that the "Cotton Bowl" destination name on the matrix sign applied to them. In contrast, as game time approached, most drivers were destined for the Cotton Bowl. The results support the need for being very careful in selecting an appropriate destination name for use on the displays. The destination descriptor displayed is a critical factor in achieving the desired diversion for a particular event.

In the October studies (see Tables 1 and 2), there were almost no differences in diversion between the 10-min and 20-min delay messages that used two signs. Also, there were no significant differences between "10-min delay" messages displayed on either one or two signs. As shown in Table 3 and Figure 8, there was no appreciable difference in diversion between the "20-min delay" and "30-min delay" messages on one sign; however, these messages produced approximately 15 percent greater diversion in the January studies than did the "10-min delay" one-sign display in October.

As had been shown in the October studies, one-sign message displays produced a very acceptable level of diversion.

SUMMARY OF RESULTS

Data collection during the route-diversion studies included measurements of demand volumes (Fair Park-bound vehicles) on both southbound Central Expressway and Fitzhugh Avenue. Analyses of the data support the following findings:

1. All of the matrix-sign messages evaluated greatly influenced route diversion to the alternative route; the high degree of diversion obtained reflected the success of the signing approach.
2. Before events that have fixed starting times (e.g., football games), all of the candidate messages are very effective in diverting traffic, and there is no discernible difference in effectiveness among the messages.
3. Before events for which the arrival times are not set but spread over long periods of time (e.g., state fairs), traffic-state descriptor and specific temporal information messages produce the highest percentages of diversion. In addition, there appears to be no advantage in displaying specific temporal information over traffic-state descriptor information.
4. The display of information that the alternative route is guide marked produces no increases in diversion rates over the base message.
5. The repetition of sign-message elements by using two signs does not appear to be critical to successful diversion for messages of up to eight words on properly located displays. Only one changeable-message sign is necessary for effective diversion.
6. The results support the need to be extremely careful in selecting an appropriate destination name for use on the displays.
7. Because more drivers to the annual Texas-Oklahoma football game (which has a fixed starting time) anticipated congestion on the primary freeway route and apparently knew about the Fitzhugh Avenue alternative route, the diversion percentages for this event were the highest recorded.

ACKNOWLEDGMENTS

This study was conducted as part of a project sponsored by the Federal Highway Administration. The support and the assistance provided by the city of Dallas and the Texas State Department of Highways and Public Transportation are acknowledged. The contents of this report reflect our views; we are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

REFERENCES

1. G. D. Weaver, C. L. Dudek, D. R. Hatcher, and W. R. Stockton. Approach to Real-Time Diversion of Freeway Traffic for Special Events. TRB, Transportation Research Record 644, 1977, pp. 57-61.
2. G. D. Weaver, S. H. Richards, D. R. Hatcher, and C. L. Dudek. Human-Factors Requirements for Real-Time Motorist Information Displays: Vol. 14—Point Diversion for Special Events Field Studies. Texas Transportation Institute, Texas A&M Univ., College Station, Feb. 1978; Federal Highway Administration, Rept. FHWA-RD-78-18, Aug. 1978.

Publication of this paper sponsored by Committee on Freeway Operations.

Analysis of Driver Responses to Point Diversion for Special Events

Stephen H. Richards, William R. Stockton, and Conrad L. Dudek,
Texas Transportation Institute, Texas A&M University

In 1976, the Texas Transportation Institute conducted a series of field studies of point diversion of freeway traffic going to special events at the Fair Park complex in Dallas. These studies were specifically designed to evaluate primary candidate messages and displays that had been developed in the laboratory. In each of the studies, comprehensive origin-destination data were collected and analyzed to determine the effectiveness of the diversionary signing techniques being considered. In addition, questionnaires were mailed to both drivers who had diverted and those who had not to (a) obtain driver attitudes and reactions to the point-diversion signing systems, (b) explore trends in driver behavior related to the diversion decision, and (c) identify factors that influence a driver's decision to divert. This paper reports and assesses the significant findings of these questionnaires. The data obtained indicate that, although numerous factors appear to have some effect on diversion for special events, driver anticipation of conditions on the alternative route, driver familiarity with the routes, and the type of special event (fixed starting time versus variable starting time) have the most pronounced influence.

The Texas Transportation Institute has recently conducted a study of the human-factors requirements for real-time motorist information displays. The objectives of this project included the development of effective information displays for real-time incident management and route diversion. The initial project work involved extensive human-factors laboratory testing and concentrated on the development of candidate messages and displays. Then the emphasis was shifted to field evaluation of the laboratory results in an operational setting.

One special aspect of the project concerned the evaluation of single-point diversion of freeway traffic going to special events. The initial research in this area centered on the laboratory development of candidate diversion and trailblazer-sign messages and displays. The Texas Transportation Institute, with the assistance of the city of Dallas, then conducted a series of field studies at the Dallas Fair Park complex to evaluate the laboratory results. These studies were conducted in 1976 on July 4, October 9, and October 10 and coordinated with scheduled special events at Fair Park. Their objective was to investigate the use of lamp-matrix sign messages for diverting freeway traffic onto an arterial street to travel to Fair Park.

Origin-destination (O-D) data were collected during each of the field studies. These data provided input for determining the effectiveness of the signing techniques being evaluated. The findings of these O-D studies are reported elsewhere (1) and in a paper by Dudek and others in this Record. The field evaluation also included an extensive questionnaire study that was designed to (a) obtain driver attitudes and reactions to the route-diversion signing systems, (b) explore trends in driver behavior related to the diversion decision, and (c) identify factors that influence a driver's decision to divert. This paper reports the significant findings of this questionnaire study.

STUDY-SITE CHARACTERISTICS

The Fair Park complex is a 97.5-hm^2 (240-acre) area located in south central Dallas. It is the site of the Cotton Bowl Stadium, the Texas State Fairgrounds, and several public buildings including a music hall and museum. Special events are held periodically at Fair Park and attract large volumes of traffic to the area. For example, the annual fireworks display attracted more than 40 000 people to the Cotton Bowl on July 4, 1976, and the annual football game between the Universities of Texas and Oklahoma attracted more than 70 000 people on the afternoon of October 9, 1976. In addition, the Texas State Fair during its 1976 opening weekend (October 9th and 10th) attracted more than 520 000 people to Fair Park.

The location of Fair Park relative to the freeway and arterial street system is illustrated in Figure 1 in the paper by Dudek and others in this Record. For traffic originating in north Dallas, the primary access route to Fair Park and its associated parking facilities is US-75 (Central Expressway). Near Fair Park, US-75 connects with I-30 (R. L. Thornton Freeway). This traffic normally takes US-75 to I-30 and then exits from I-30 at either the Second Avenue or the Haskell Avenue exit. Within the study area, this freeway route to Fair Park is a six-lane controlled-access facility. A direct two-lane ramp joins US-75 and I-30.

However, as shown in this figure, there is also available an alternative arterial route, Fitzhugh Avenue, which crosses US-75 well in advance of the I-30 junction. Fitzhugh Avenue is a two-way arterial that becomes half of a pair of one-way streets (Bennett Avenue is the other) about 0.4 km (0.25 mile) east of US-75. The street narrows from three to two lanes approximately 1.6 km (1.0 mile) east of US-75. The availability of this route, coupled with the heavy congestion that occurs at the Second Avenue and Haskell Avenue exits before special events at Fair Park, prompted the selection of this site for the field evaluation studies.

During all of the studies, the traffic signals at the Fitzhugh-US-75 diamond interchange were retimed to accommodate the increased left-turn demand. The remaining signals along Fitzhugh Avenue were coordinated to provide progressive traffic movement toward Fair Park. Also, parking was prohibited along the alternative route to provide a minimum of two operating lanes at all points.

RESEARCH APPROACH

Study Design

Southbound motorists on US-75 were given route-diversion information in messages displayed on two lamp-matrix signs. These signs were mounted on trailers and placed directly over the US-75 southbound main lanes by positioning the sign trailers on overcrossing structures upstream of the diversion point.

The message sets evaluated in the studies are shown in Figure 1. Two messages were tested in the July 4th study; a total of 12 messages were evaluated in the October 9th and 10th studies.

Questionnaire Development

Separate questionnaires were developed and administered to drivers who had diverted and to those who had not.

Figure 1. Summary of candidate route-diversion messages.

July 4th Studies

UNIVERSITY BLVD. SIGN		MOCKINGBIRD LANE SIGN
BEST ROUTE TO FIREWORKS	Message Set 1	FIREWORKS BEST ROUTE
USE FITZHUGH AVE		FITZHUGH AVE 1 MILE
ROUTE TO FIREWORKS	Message Set 2	FIREWORKS BEST ROUTE
INFORMATION AHEAD		FITZHUGH AVE 1 MILE

October 9-10th Studies

LOVERS LANE SIGN		YALE BLVD. SIGN
BEST ROUTE TO FAIR GROUNDS	Message Set 1	FAIR GROUNDS BEST ROUTE
USE FITZHUGH AVE		FITZHUGH AVE 2 MILES
FAIR GROUNDS USE FITZHUGH	Message Set 2	FAIR GROUNDS BEST ROUTE
SAVE 10 MINUTES		FITZHUGH AVE 2 MILES
FAIR GROUNDS USE FITZHUGH	Message Set 3	FAIR GROUNDS BEST ROUTE
AVOID 10 MIN. DELAY		FITZHUGH AVE 2 MILES
FAIR GROUNDS USE FITZHUGH	Message Set 4	FAIR GROUNDS BEST ROUTE
AVOID MAJOR DELAY		FITZHUGH AVE 2 MILES
FAIR GROUNDS USE FITZHUGH	Message Set 5	FAIR GROUNDS BEST ROUTE
SAVE 20 MINUTES		FITZHUGH AVE 2 MILES

October 9-10th Studies (Continued)

FAIR GROUNDS USE FITZHUGH	Message Set 6	FAIR GROUNDS BEST ROUTE
AVOID 20 MIN. DELAY		FITZHUGH AVE 2 MILES
FAIR GROUNDS USE FITZHUGH	Message Set 7	FAIR GROUNDS BEST ROUTE
AVOID CONGESTION		FITZHUGH AVE 2 MILES
FAIR GROUNDS USE FITZHUGH	Message Set 8	FAIR GROUNDS BEST ROUTE
AVOID TRAFFIC JAM		FITZHUGH AVE 2 MILES
BEST ROUTE TO FAIR GROUNDS	Message Set 9	FAIR GROUNDS BEST ROUTE
USE FITZHUGH FOLLOW SIGNS		FITZHUGH AVE 2 MILES
Sign Not Used	Message Set 10	BEST ROUTE TO FAIR GROUNDS
		USE FITZHUGH AVE
Sign Not Used	Message Set 11	FAIR GROUNDS USE FITZHUGH
		SAVE 10 MINUTES
Sign Not Used	Message Set 12	FAIR GROUNDS USE FITZHUGH
		AVOID 10 MIN. DELAY

Samples of the questionnaires have been given by Weaver and others (3). Both questionnaires included four questions designed to group the drivers and determine whether their responses were valid to the study. All motorists were also asked to reproduce the displayed sign messages to the best of their ability. Space was provided at the end of every questionnaire for general comments.

The questionnaires sent to those drivers who had not diverted requested additional information related to their disregard of the matrix-sign information. These drivers were asked why they failed to use the alternative route in light of the information presented on the matrix signs and requested to list additional information that, if presented, would have altered their route selection.

Drivers who had diverted were asked whether the route-diversion displays influenced their selection of the Fitzhugh Avenue route for traveling to Fair Park. They were also queried about guidance along the alternative arterial route.

Data Collection

During all studies, comprehensive O-D data were collected by stationing observers at critical points to read the license plate numbers of passing vehicles into cassette recorders. This approach permitted the path tracing of almost every vehicle within the study boundaries. It also allowed the researchers to determine which drivers diverted to Fitzhugh Avenue and which ones did not divert but continued on the US-75 route.

Table 1. Summary of questionnaire sample sizes.

Study	No. of Drivers Observed	No. Who Received Questionnaire	Sampling Rate (%)	No. Who Returned Questionnaire	Return Rate (%)	Interview Rate (%)
Drivers who used Fitzhugh Avenue						
July 4	421	300	71.3	136	45.3	32.3
October 9	435	291	66.9	108	37.1	24.8
October 10	417	319	76.5	118	37.0	28.3
All	1273	910	71.5	362	39.8	28.4
Drivers who used US-75 and I-30						
July 4	251	197	78.5	74	37.6	29.5
October 9	118	69	58.5	14	20.3	11.9
October 10	168	125	74.4	30	24.0	17.9
All	537	391	72.8	118	30.2	22.0
Total	1810	1301	71.9	480	36.9	26.5

A mailing address was then obtained corresponding to each vehicle that had Texas tags observed traveling to Fair Park (both those that diverted and those that did not) by using vehicle registration information on computer files at the Texas State Department of Highways and Public Transportation. No address information was available for vehicles that had out-of-state tags.

After obtaining the necessary addresses, questionnaires and prepaid return envelopes were distributed to the drivers by first class mail. All questionnaires for an individual study were mailed at the same time, approximately 1 month after the field study. The 1-month delay resulted from the time required to reduce the field data, obtain mailing addresses, and address mailing envelopes. Based on the responses to the questionnaires, this delay was reasonable and affected only verbatim message recall.

Sampling rates for the questionnaire studies are given in Table 1, which shows that 71.9 percent of all drivers observed traveling to Fair Park were mailed questionnaires. Furthermore, the sampling rates for diverting and nondiverting drivers were about equal (71.5 and 72.8 percent, respectively). However, the sampling rates for individual studies did vary somewhat, because more out-of-state drivers were observed in some studies.

A summary of the questionnaire return rates is also given in Table 1. The return rate of completed questionnaires averaged 36.9 percent, indicating significant driver interest in the route-diversion studies. The interview rate, or effective rate of return, averaged 26.5 percent and was generally consistent; the only exceptions were the October 9th and 10th drivers who had stayed on the primary route. A possible explanation for these drivers' apparent reluctance to respond might be that a large portion of them were out-of-town visitors and not particularly interested in the Dallas study.

STUDY RESULTS

Driver Familiarity and Diversion

The importance of driver familiarity in relation to the diversion decision is indicated by the data below. In this table, the percentages of drivers who diverted to Fitzhugh Avenue are given according to the driver-familiarity categories used on the questionnaires.

Study	Percentage of Interviewed Drivers Who Diverted: Within the Last Month (N = 104)	Within the Last Year (N = 289)	Within the Last 5 Years (N = 43)	Never Before (N = 36)
July 4	49.0	67.9	86.7	57.1
October 9: a.m.	40.0	96.4	100.0	87.5
October 9: p.m.	75.0	92.7	90.0	87.5
October 10	66.7	84.9	81.8	84.6
All	57.7	79.2	88.4	80.6

These data suggest that drivers who are less familiar with the routes to their destination are more likely to divert than are more familiar drivers. This trend continues until a driver is totally unfamiliar with the routes to his or her destination. At this level of familiarity, an increasing percentage of drivers fail to heed diversion instructions. This decrease in diversion potential is probably due, in part, to those drivers' lack of confidence in their ability to navigate along an unfamiliar, unplanned arterial route.

Another interesting point to note from the table above is the large number (75.0 percent) of very familiar drivers who decided to use Fitzhugh Avenue on the afternoon of October 9th. Most of the drivers were destined for the Texas-Oklahoma football game and were probably aware of the unfavorable traffic conditions on the primary freeway route. They also were determined to arrive at their destination before the game started. These two factors combined to produce a higher diversion rate among these very familiar drivers during this time period.

Driver Destination and Diversion

The influence of driver destination on the diversion decision in the July 4th study is summarized below.

Driver Destination	Message Set 1: N	Message Set 1: % of N	Message Set 2: N	Message Set 2: % of N	All July Messages: N	All July Messages: % of N
Cotton Bowl fireworks show	125	77.6	54	44.4	179	67.6
Fireworks show and some other activity	17	64.7	5	0.0	22	50.0
Not fireworks show	5	20.0	3	66.6	8	37.5

In this table, the percentages of drivers who diverted to Fitzhugh Avenue are given according to the message set used. The messages used in the July 4th studies were directed at only those drivers en route to the fire-

Table 2. Relationship between driver destination and diversion: October studies.

Driver Destination	Interviewed Drivers Who Diverted Onto Fitzhugh Avenue: October 9th a.m. N	%	p.m. N	%	Total N	%	October 10th N	%	All October N	%
Texas State Fair	13	84.6	19	89.5	32	87.5	127	82.7	159	83.6
Texas-Oklahoma football game	3	100.0	28	92.9	31	93.5			31	95.3
Football game and some other Fair Park activity	30	93.3	21	85.7	51	90.2			51	90.2
Texas State Fair and some other Fair Park activity (excluding football game)	2	100.0	3	66.7	5	80.8	8	75.0	13	76.9
Fair Park activity other than Texas State Fair and football game	1	0.0	1	100.0	2	50.0	15	60.0	17	58.8

Table 3. Reasons cited by drivers who did not divert for not diverting.

Category: Reasons for Not Diverting	Percentage of Drivers Who Did Not Divert: July 4th (N = 74)	October 9-10 (N = 44)	All Studies (N = 118)
1: Driver anticipated unsatisfactory conditions on alternative route	62	61	62
Traffic conditions	48	56	51
Along arterial route	34	52	41
Time delay	14	16	14
Increased distance	1	4	3
Congestion and traffic jams	8	21	13
Intersection stops	11	11	11
Heavy traffic at exit ramp or freeway approach to Fitzhugh Avenue	3		2
Diversion of others made intended route more desirable	11	4	8
Road conditions	10		6
Surface of alternative route is rough; construction	3		2
Type of facility is not desirable because it is narrow and undivided and has uncontrolled access	7		4
Other	4	5	5
Undesirable area (bad neighborhood)	3		2
Uncertain about access to parking	1	5	3
2: Message did not reach driver	39	28	35
Driver failed to see messages or second sign of message set 2	31	14	25
Driver misread message	3	7	4
Driver could not read message	4	7	5
Driver was not certain of which fireworks display	1		1
3: Driver was not familiar with alternative route	32	27	30
General conditions	16	11	14
Driver was uncertain of adequate guidance along alternative route	7	2	5
Driver did not want to change normal route or habit	9	14	11
4: Driver lacked confidence in information	10	9	10
Driver uncertain of message audience (going to Fair Park but not Cotton Bowl)	3		2
Driver believed information not current	5	2	4
Driver believed message was an advertisement	1	2	1
Driver did not believe alternative route was best	1	5	3

Note: Columns do not add to 100 percent because some drivers cited more than one reason for not diverting.

works display; therefore, drivers visiting other Fair Park events were less inclined to divert. For example, 67.6 percent of the drivers en route to Fair Park to visit only the fireworks display diverted to Fitzhugh Avenue. In comparison, only 37.5 percent of the drivers going to Fair Park but not to the fireworks display, diverted. Also, 50.0 percent of the drivers going to more than one Fair Park event, one of which was the fireworks display, decided to divert.

All messages displayed during the October 9th and 10th studies referred only to the "Fair Grounds"; individual events (i.e., football games, rodeos, or horse shows) or specific destinations (i.e., the Cotton Bowl or the main gate) were not mentioned. Because of this inherent difference between the July and October messages, data from the two studies could not be combined.

Table 2 presents the percentages of interviewed October 9th and 10th drivers who diverted to Fitzhugh Avenue grouped according to their destination within the Fair Park complex. These data again indicate that a driver's decision to divert is affected by his or her specific destination. For example, 83.6 percent of the drivers destined only for the Texas State Fair diverted when given information concerning an alternative route to the "Fair Grounds." In contrast, only 58.8 percent of the 17 motorists destined for a Fair Park activity (e.g., rodeo, horse show, or music hall) other than the Fair or the Texas-Oklahoma football game diverted. Apparently, many of the drivers in this latter group did not associate the "Fair Grounds" message with their specific destination within the Fair Park complex.

Drivers destined for the football game, however, did not have this association problem. More than 93 percent of the drivers destined for the football game on October 9th elected to use the Fitzhugh Avenue route. Thus, most of these drivers knew that the stadium was part of the fairgrounds.

These significantly higher diversion rates just before the football game also support the finding that a higher percentage of drivers will use an alternative route for events that have a fixed starting time when heavy traffic congestion is anticipated on the primary freeway route [in comparison to events that have variable starting times (for which less severe congestion is anticipated)].

Table 4. Effect of driver familiarity on reasons cited for not diverting (all studies combined).

Category: Reasons for Not Diverting	Percentage of Drivers Who Did Not Divert — Familiarity Category[a] A (N = 41)	B (N = 63)	C (N = 6)	D (N = 7)
1. Driver anticipated unsatisfactory conditions on alternative route	63	63	17	57
Traffic conditions	51	53	17	57
Along arterial route	46	39	17	43
Time delay	17	14		14
Increased distance	2	3		
Congestion and traffic jams	12	11	17	29
Intersection stops	15	11		
Heavy traffic on exit ramp or freeway approach to Fitzhugh Avenue				
Diversion of others made intended route more desirable	5	6		14
Road conditions	6	6		
Surface of alternative route is rough; construction				
Type of facility is not desirable because it is narrow and undivided and has uncontrolled access	6	6		
Other	4	4		
Undesirable area (bad neighborhood)	2	2		
Uncertain about access to parking	2	2		
2: Message did not reach driver	39	37	33	14
Driver failed to see messages or second sign of message set 2	27	27	33	
Driver misread message	5	3		14
Driver could not read message	5	5		
Driver was not certain of which fireworks display	2	2		
3: Driver was not familiar with alternative route	24	33	33	14
General conditions	2	17	33	14
Driver was uncertain of adequate guidance along alternative route	5	6		
Driver did not want to change normal route or habit	17	10		
4: Driver lacked confidence in information	4	10	33	28
Driver uncertain of message audience (going to Fair Park but not Cotton Bowl)	2		17	
Driver believed information not current	2	3	16	14
Driver believed message was an advertisement		2		14
Driver did not believe alternative route was best		5		

[a] Familiarity categories are coded as follows: A = within the previous month; B = within the previous year; C = within the previous five years; and D = never before.

Attitudes and Requirements of Drivers Who Did Not Divert

Drivers who had not diverted were asked to specify the considerations that had prevented them from diverting. Table 3 presents a detailed breakdown of the responses to this question grouped into four major categories and several subcategories. The table includes the results of the July 4th study, the October 9th and 10th studies, and all studies combined.

Anticipated dissatisfaction with the alternative route was cited most often as a deterrent to diversion. When all studies were combined, 73 of the 118 drivers who did not divert (62 percent) cited dissatisfaction with the alternative route as a reason for not using Fitzhugh Avenue. Of particular interest in this group of responses is the subcategory related to congestion or traffic jams on the alternative route. In the July 4th study, only 6 of the 74 drivers who did not divert (8 percent) mentioned congestion on Fitzhugh Avenue as a deterrent to diversion. In the October 9th and 10th studies, however, the number rose to 9 of 44 drivers (21 percent). A possible reason for this increase was the noticeably higher volumes of traffic on the Fitzhugh exit ramp and frontage road intersection during the October studies.

When the major-category totals are compared, the response percentages for the July 4th study and October 9th and 10th studies are fairly consistent. The one notable exception occurs in the category entitled "Messages Did Not Reach Driver." In the July 4th study, 39 percent of the drivers cited this as a reason for not using Fitzhugh Avenue. Only 28 percent of the drivers mentioned this factor in the October 9th and 10th studies.

An analysis of the effect of driver familiarity on reasons for not diverting is given in Table 4 for all studies. For drivers of all familiarity levels, except those who had visited Fair Park in the previous 5 years but not in the previous year (familiarity category C), anticipated dissatisfaction with the alternative route was the most often stated reason for not diverting. In fact, this reason was cited by 26 of the 41 drivers who had not diverted (63 percent) who had visited Fair Park in the previous month (familiarity category A), 40 of the 63 drivers (63 percent) who had visited Fair Park in the previous year (familiarity category B), and 4 of the 7 (57 percent) who had never visited Fair Park (familiarity category D). The familiarity category C group that was the exception had a sample size of only 6 drivers and only 1 of them mentioned anticipated unsatisfactory conditions on Fitzhugh Avenue.

One additional point worth noting concerns the major category pertaining to driver confidence in the displayed information. The data in this category suggest that unfamiliar drivers lacked confidence in the information or were not certain about the meanings of the messages. However, the small sample sizes may be responsible for these results.

An additional analysis was conducted to determine whether there was a relationship between message content and reasons for not diverting. However, because the large majority of drivers diverted, the sample sizes of drivers who had not diverted associated with individual messages were quite small. Thus, no conclusions could be reached.

The drivers who had not diverted were also asked to state what information could have been given, but was not, that would have encouraged them to divert. The responses to this question are summarized below.

Additional Information Required to Divert	Percentage of Interviewed Drivers Who Did Not Divert — July Study (N = 74)	October Studies (N = 44)	Combined Studies (N = 118)
Intended route blocked or closed	16.2	15.9	16.1
Intended route congested or jammed	16.2	11.4	14.4
Specific traffic-state descriptors	9.5	11.4	10.2

Additional Information Required to Divert	Percentage of Interviewed Drivers Who Did Not Divert July Study (N = 74)	October Studies (N = 44)	Combined Studies (N = 118)
Assurance of alternative-route guidance	10.8	9.1	10.2
Advance publicity (i.e., radio, television, or newspapers)	9.1	–	5.1
Parking instructions	2.7	6.8	4.2
None or no response	25.7	50.0	34.7

In this table, the percentages of drivers who did not divert are given according to the survey date. Severely degraded operation of the primary freeway route (intended route blocked or closed) was most often cited as information that would have influenced diversion. Strangely enough, 7 of the 19 drivers who had not diverted and requested this information had viewed message sets 7 or 8 in the October studies (which informed them of congestion or jammed traffic on their intended route). However, many of these drivers either failed to see the signs or misinterpreted the messages to infer that the congestion was on the Fitzhugh Avenue route, rather than on their intended route.

Assurance of guidance along the alternative route was requested by 10.2 percent of the drivers. None of these drivers had viewed message set 9 (which offered this assurance). Also, the responses of these drivers were analyzed to determine whether driver familiarity was a factor, but there were no significant findings.

This table also shows that the "None or No Response" category increased from 25.7 percent of the drivers who did not divert in the July 4th study to 50.0 percent in the October studies. In the October studies, many drivers were given information related to primary route conditions and guidance assurance. Therefore, the drivers who failed to divert could not be further influenced by such information; it had already been given to them.

Attitudes and Requirements of Drivers Who Diverted

Drivers who had diverted were asked to critique the messages in terms of desirable information that was not displayed. The responses to this question are summarized below.

Additional Information Requested	Percentage of Interviewed Drivers Who Did Not Divert July Study (N = 136)	October Studies (N = 226)	Combined Studies (N = 362)
Parking instructions	12.5	15.0	14.1
Additional en route guidance	9.6	10.2	9.9
Guidance on return route	15.4	5.8	9.4
Distance to Fair Park	4.4	5.8	5.2
Traffic-state descriptors		2.7	1.7
Additional alternative routes	1.5	1.3	1.4

When the July and October studies are combined, parking information is first on the list; 14.1 percent of the drivers desired instructions for parking at Fair Park. Additional trailblazer signs were requested by 9.9 percent of the drivers. Return route guidance and the distance to Fair Park were requested by 9.4 and 5.2 percent of the drivers respectively.

The table below presents the percentages of drivers who had diverted who made various unfavorable comments related to the diversion attempts.

Unfavorable Comment	Percentage of Drivers Who Diverted July Study (N = 136)	October Studies (N = 226)
Parking facilities at Fair Park were inadequate	6.6	6.2
Diversion instructions would be ignored in the future	4.4	
Fitzhugh Avenue was too congested	3.7	0.4
I-75 and I-30 would have been a better route than Fitzhugh Avenue	2.2	2.7
Traffic was diverted through a bad neighborhood	2.9	
The return route was too congested	2.2	
Diverting traffic was a waste of money	0.7	0.4
Traffic was diverted to the wrong area within Fair Park		0.9
The pavement on Fitzhugh Avenue was in poor condition		0.9
The questionnaire study was an invasion of privacy	0.7	
The questionnaires arrived too late		0.4

The important point to note concerning the data in this table is that very few drivers made unfavorable comments. The one area that did receive a significant number of unfavorable comments in all studies concerned parking conditions at Fair Park. In the October studies, 6.6 percent of the drivers who diverted complained about parking; 6.2 percent of the July 4th drivers complained about parking conditions.

In the July 4th study, a few drivers (4.4 percent) remarked that they would not heed diversion instructions under similar conditions in the future because they were so dissatisfied with the Fitzhugh Avenue route. This prompted concern that, under congested conditions on the alternative route, many drivers might lose credibility in any route-diversion attempt. However, no drivers in the October 9th and 10th studies stated that they would not divert in the future under similar conditions, although traffic on the alternative route was moderately congested during these studies.

SUMMARY OF RESULTS

Similar questionnaires were used to interview drivers observed traveling to Fair Park during the July 4th, October 9th, and October 10th special-events point-diversion studies in Dallas. The questionnaires were administered to 497 drivers in the July 4th study and 804 drivers in the October 9th and 10th studies; the completed and returned questionnaires totaled 210 and 270 respectively, for an effective return rate of 42.3 and 33.6 percent respectively. The pertinent findings are summarized below:

1. In both the July and the October studies, drivers who were less familiar with the route were more likely to divert than were those who were more familiar with it. However, before the football game on the afternoon of October 9th, the percentage of drivers who were very familiar with the route who diverted was significantly greater than that at other time periods. This increase is attributed to the facts that the football game had a fixed starting time and severe congestion was anticipated before the game on the primary freeway route.

2. Anticipated dissatisfaction with the alternative route was the most frequently cited deterrent to diversion. Driver lack of confidence in the messages was the least frequently cited reason for not diverting. The reasons cited for not diverting were relatively consistent between both drivers who were familiar with the route and those who were not.

3. Advisories of severe traffic problems on the primary freeway route were cited most often in both the July and October studies as the information that would have influenced the drivers who did not divert to use the alterative route. This unexpected result for the October studies (traffic condition information was displayed) was primarily due to the fact that some drivers failed to read or correctly interpret the message(s).

4. More than 14 percent of the 362 drivers who diverted interviewed stated that additional information on parking in the Fair Park area was desirable. Also, 9.4 percent of the drivers requested guidance on the return route, and 9.1 percent requested additional en route guidance.

5. No drivers who had diverted during the October studies, despite the heavier volumes on the Fitzhugh Avenue route, stated that they would not follow diversion instructions in the future under similar circumstances.

ACKNOWLEDGMENTS

This study was conducted as part of a project sponsored by the Federal Highway Administration. The support and the assistance provided by the city of Dallas and the Texas State Department of Highways and Public Transportation are acknowledged. The contents of this report reflect our views; we are responsible for the facts and the accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration. This report does not constitute a standard, specification, or regulation.

REFERENCES

1. G. D. Weaver, C. L. Dudek, D. R. Hatcher, and W. R. Stockton. Approach to Real-Time Diversion of Freeway Traffic for Special Events. TRB, Transportation Research Record 644, 1977, pp. 57-61.
2. G. D. Weaver, S. H. Richards, D. R. Hatcher, and C. L. Dudek. Human-Factors Requirements for Real-Time Motorist Information Displays: Vol. 14—Point Diversion for Special Events Field Studies. Texas Transportation Institute, Texas A&M Univ., College Station, Feb. 1978; Federal Highway Administration, Rept. FHWA-AD-78-18, Aug. 1978.

Publication of this paper sponsored by Committee on Freeway Operations.

Incident Detection: A Bayesian Approach

Moshe Levin and Gerianne M. Krause, Bureau of Materials and Physical Research, Illinois Department of Transportation

A single-feature incident-detection algorithm based on Bayesian considerations is developed. The algorithm uses the ratio of the difference between the upstream and the downstream minute occupancies and the upstream occupancy as the traffic-flow feature followed and uses historical incident information. The historical incident data used are representative of the inner lane of the outbound Kennedy Expressway between the Chicago Loop and the Edens Expressway junction during the afternoon rush period. Mathematical expressions are developed for the distributions of the ratio from incident and incident-free data. The probabilities of incidents occurring on the outbound Kennedy are developed from available incident data. The optimal threshold to be used in the incident-detection process is determined mathematically by using Bayesian concepts. The efficiency of the algorithm is evaluated in terms of its detection rate, false-alarm rate, and mean time to detect and is compared with those of the California algorithm and two algorithms developed by Technology Services Corporation. The Bayesian algorithm compares favorably with the others with regard to detection rate (100 percent) and false-alarm rate (0.0 percent). However, its mean time to detect is greater than that of the other algorithms by almost 2.5 min. A preliminary on-line evaluation comparing the Bayesian algorithm and one of the others showed no significant differences in detection rate, false-alarm rate, and mean time to detect.

Freeway incident-management systems that offer various levels of service to the motoring public have been in operation for quite some time. In essence, these systems provide some or all of the following elements:

1. Detection of traffic-flow abnormalities,
2. Incident identification,
3. Traffic management strategies and tactics through communication and control systems, and
4. Early removal of incidents and the return to normal traffic-flow conditions.

The degree of comprehensiveness of the management system and the level of sophistication of its elements determine the operational efficiency of the system and its success in achieving its objectives.

The detection of traffic-flow abnormalities on a freeway is carried out by a surveillance system, usually through an electronic detector subsystem. The availability of such a subsystem allows continuous quantification of traffic-flow characteristics, the identification of incidents, and the application of appropriate control strategies.

The process by which traffic-flow abnormalities are identified as incidents is a key element in such a management system because a positive identification normally activates the control, driver-communication, and incident-handling subsystems. Obviously, missed incidents or false alarms will affect the efficiency of the management system and its credibility.

The incident-identification process uses an incident-detection algorithm that relates measured relationships between traffic characteristics to calibrated ones and yields a decision with regard to the occurrence of an incident. Throughout the years of freeway-control research, the basic approaches to the development of incident-detection algorithms have been

1. Pattern recognition: the comparison of current flow patterns with those expected based on historical data or theoretical traffic-flow considerations and the identification of consistent deviations as incidents (1, 2, 3) and
2. Statistical forecasting of traffic behavior: the comparison of current traffic characteristics with forecasted ones based on time-series analysis and the identification of calibrated deviations as incidents (4, 5, 6).

The efficiency of such algorithms can be determined by three related parameters:

1. Detection rate: the percentage of total capacity-reducing incidents occurring during a specified time period that are detected,
2. False-alarm rate: the percentage of total incident messages occurring during a specified time period that are false (on-line definition) [another definition (3) of the off-line situation is the percentage of incident messages ("1"s) out of all messages ("1"s and "0"s), where messages are produced at specific intervals (i.e., every 1 min) out of representative incident-free data], and
3. Mean time to detect: The mean delay between the apparent occurrence of the incident and its detection for all detected incidents occurring during a certain period of time.

The inherent positive correlation between the detection and false-alarm rates can be detrimental to the effectiveness of the incident-management system because the desired low false-alarm rate is coupled with a low detection rate. The implication of such a detrimental effect can best be illustrated by applying the detection characteristics of an existing algorithm (3), which is considered an efficient one, to the incident situation on Chicago expressways. It is estimated that the Chicago expressway system under electronic surveillance experiences approximately 20 capacity-reducing incidents/d during the afternoon rush period. By applying an algorithm optimally calibrated to have a 0.01 percent false-alarm rate coupled with a 34 percent detection rate, it can be shown that only 7 incidents will be detected but 20 false alarms will be reported. In reality, the false-alarm rate as viewed by the incident-management decision maker would be close to 75 percent. At such a high probability of making the wrong decision, no decision will be made without more information as to the reliability of the incident message.

High message reliability can be achieved by the use of sophisticated incident-identification and verification systems (such as closed-circuit television and CB radio), which will offset the weaknesses of the detection algorithm. However, for large freeway systems, this type of verification is probably not feasible. Another way to increase message reliability is to use an algorithm that considers the likelihood of the occurrence of a capacity-reducing incident at a particular location and time period. In addition, message reliability can be improved by a time trade-off in waiting for additional incident indication(s).

The incorporation of incident historical information into a probablistic model could be achieved by using Bayesian concepts. This approach will give the probability that an incident has actually occurred, given an incident signal or string of signals.

THEORETICAL CONSIDERATIONS

Consider a freeway section located between two detectors that have been placed on one of the section lanes. Let Z represent a certain traffic feature (characteristic) that can be measured at either the upstream or the downstream detector or at both. Let $f(Z/U_1)$ and $f(Z/U_0)$ represent the frequency distribution functions of feature Z during incident and incident-free situations respectively for certain environmental and traffic conditions.

The probability of an incident occurring on this section [$P(U_1)$] under certain environmental and traffic conditions can be derived based on its history of capacity-reducing incidents. The probability of not having any capacity-reducing incidents is then $P(U_0)$. Thus

$$P(U_0) = 1 - P(U_1) \tag{1}$$

and

$$P(U_0)\int_{a_0}^{b_0} f(Z/U_0)dZ + P(U_1)\int_{a_1}^{b_1} f(Z/U_1)dZ = 1 \tag{2}$$

where a_0, b_0, a_1, and b_1 are the upper and lower bounds of Z in the functions $f(Z/U_0)$ and $f(Z/U_1)$ respectively.

For any feature value Z (threshold), it can be shown that the probability of obtaining an incident signal [P(1)] can be expressed as follows:

$$P(1) = P(U_0)\int_{z_1}^{b_0} f(Z/U_0)dZ + P(U_1)\int_{z_1}^{b_1} f(Z/U_1)dZ \tag{3}$$

Similarly, the probability of obtaining a nonincident signal [P(0)] is

$$P(0) = P(U_0)\int_{a_0}^{z_1} f(Z/U_0)dZ + P(U_1)\int_{a_1}^{z_1} f(Z/U_1)dZ \tag{4}$$

By applying Bayesian considerations, one can develop an expression for the probability that an incident has actually occurred, given that an incident signal "1" was output. This expression is

$$P(\text{incident}/1) = P(U_1)\int_{z_1}^{b_1} f(Z/U_1)dZ \div \left[P(U_0)\int_{z_1}^{b_0} f(Z/U_0)dZ + P(U_1)\int_{z_1}^{b_1} f(Z/U_1)dZ\right] \tag{5}$$

The probability that no incident has occurred when a nonincident signal "0" is output can be expressed as

$$P(\text{no incident}/0) = P(U_0)\int_{a_0}^{z_1} f(Z/U_1)dZ \div \left[P(U_0)\int_{a_0}^{z_1} f(Z/U_0) + P(U_1)\int_{a_1}^{z_1} f(Z/U_1)dZ\right] \tag{6}$$

The optimal threshold (Z_1) can be obtained by maximizing the expression:

$$P(\text{incident}/1) + P(\text{no incident}/0) \tag{7}$$

Theoretically, this optimization procedure for Z_1 can be repeated for $f_i(Z/U_1)$ to give a set of optimal thresholds [Z_1] where i represents consecutive determined time intervals after the detection of an incident. However, by selecting a feature for which there are no statistically significant differences between $f_i(Z/U_1)$ and $f_{(i+1)}(Z/U_1)$, only one threshold value could be used in the detection process. The use of such a feature is important because of the delay between the occurrence of an incident and its detection. In such cases, the calculated values of Z_1, Z_2, and so on will not represent time intervals immediately after the occurrence of the incident. Thus, by selecting one appropriate threshold, the threshold-synchronization problem is eliminated.

The application of the Bayesian concepts can also be extended to the case of strings of signals. The evaluation of the probability that an incident has actually occurred, given an n-signal string (although this forces a wait of n time intervals before making an incident-management decision) can provide the decision maker with more reliable information. Obviously, for practical reasons, n could be limited to three; thus, the signal strings of interest would be 10, 11, 100, 101, 110, and 111. The nature of the probabilities of the occurrence of an incident, given any of the above signals, can be shown to be such that

$$P(\text{incident}/111) > P(\text{incident}/11) \quad (8)$$

and

$$P(\text{incident}/100) < P(\text{incident}/10) \quad (9)$$

These relationships are necessary conditions if the Bayesian approach is valid.

The probabilities of the actual occurrence of an incident, given the n-signal strings shown above, are given below.

Probability	Expression
P(incident/1)	P(1/incident) P(incident) ÷ [P(1/incident) P(incident) + P(1/no incident) P(no incident)]
P(incident/0)	P(0/incident) P(incident) ÷ [P(0/incident) P(incident) + P(0/no incident) P(no incident)]
P(no incident/1)	P(1/no incident) P(no incident) ÷ [P(1/incident) P(incident) + P(1/incident) P(no incident)]
P(no incident/0)	P(0/no incident) P(no incident) ÷ [P(0/incident) P(incident) + P(0/no incident) P(no incident)]
P(incident/10)	P(0/incident) P(incident 1) ÷ [P(0/incident) P(incident/1) + P(0/no incident) P(no incident/1)]
P(incident/11)	P(1/incident) P(incident/1) ÷ [P(1/incident) P(incident/1) + P(1/no incident) P(no incident/1)]
P(incident/100)	P(0/incident) P(incident/10) ÷ [P(0/incident) P(incident/10) + P(0/no incident) P(no incident/10)]
P(incident/101)	P(1/incident) P(incident/10) ÷ [P(1/incident) P(incident/10) + P(1/no incident) P(no incident/10)]
P(incident/110)	P(0/incident) P(incident/11) ÷ [P(0/incident) P(incident/11) + P(0/no incident) P(no incident/11)]
P(incident/111)	P(1/incident) P(incident/11) ÷ [P(1/incident) P(incident/11) + P(1/no incident) P(no incident/11)]

These probabilities can be computed for any particular freeway section and specific traffic and environmental conditions by using the history of capacity-reducing incidents.

For appropriate freeway sections and environmental and traffic conditions, the theoretical probabilities that an incident has actually occurred, given a certain signal string, can be correlated with actual string probabilities derived from on-line implementation of the Bayesian algorithm. Once the calibration is complete, a certain criterion value can be selected to use in making an incident-management decision.

ALGORITHM DEVELOPMENT

The process of quantifying these theoretical considerations requires the use of three data bases that were established within the framework of other incident-detection research. These are the

1. Incident data base,
2. Incident-free data base, and
3. Emergency patrol-vehicle-assists data base.

The first two data bases were used to develop $f(Z/U_1)$ and $f(Z/U_0)$ respectively, and the third one was used to develop historical probabilities of capacity-reducing incidents.

The incident data base contained 122 data segments for that number of incidents and represented various locations and environmental and traffic conditions. Each data segment consisted of main-line, 20-s based, occupancy and volume data for at least 15 min before the occurrence of the incident and at least 10 min after its termination. The same traffic characteristics were collected for similar environmental and traffic conditions on the same freeway locations. Data were collected through the Chicago expressway surveillance system, which features center-lane detectors spaced every 0.8 km (0.5 mile) and full detector coverage every 4.8 km (3 miles). These detectors are connected via leased telephone lines to a central computer that processes the raw pulses into occupancy, volume, and speed data.

The historical data base for the capacity-reducing incidents was developed by using the assistance-rendered reports submitted by emergency patrol vehicles for the years 1973, 1974, and 1975. These reports provide such information as the type of incident (e.g., accident or stalled vehicle), its location, estimated occurrence time, and environmental conditions.

Once the three data bases were available, it was possible to proceed with the selection of a study site for which the Bayesian model would be developed, the selection of the appropriate traffic-flow feature to be used in the model, and the development of historical probabilities of capacity-reducing incidents.

Study-Site Selection

Because the incident and incident-free data bases were prepared in the framework of other incident-detection research, the selection of the study site was confined to isooperational sections for which large enough samples of incident-free and incident data were available. This led to the selection of the outbound J. F. Kennedy Expressway between the Chicago Loop and the junction with the Edens Expressway. This section of the Kennedy Expressway is basically four lanes wide and has two reversible lanes that operate outbound in the afternoon rush period; it has an average daily traffic of approximately 115 000 vehicles. In 1975, the number of assists by emergency vehicles on weekdays (capacity-reducing incidents and others) was nearly 1500, and there was an average of approximately 1 capacity-reducing incident/afternoon rush period.

Feature Selection

Seven traffic features were considered in the feature-selection process. These features were taken from incident-detection algorithms developed by Payne and others (3) for Technology Science Corporation (TSC) and have been the subject of other research (7). The seven features considered are given below:

Name	Definition
OCC(t)	Minute-average occupancy measured at upstream detector at time t
DOCC(t)	Minute-average occupancy measured at downstream detector at time t
OCCDF(t)	OCC(t) - DOCC(t)
OCCRDF(t)	OCCDF(t)/OCC(t)
SPEED(t)	Minute-average speed measured at upstream detector at time t
DOCCTD(t)	[DOCC(t - 2) - DOCC(t)] /DOCC(t - 2)
SPDTDF(t)	[SPEED(t - 2) - SPEED(t)]/SPEED(t - 2)

The criteria for selecting a feature were that it have

1. Considerable difference between its values before and during the incident and
2. Stability of this difference during the incident.

High stability would allow the use of a single threshold throughout the detection process.

The feature OCCRDF(t) was selected for the Bayesian model. However, because theoretically $-\infty \leq \mathrm{OCCRDF}(t) \leq 1$, the feature $Z = 1 - \mathrm{OCCRDF}(t)$, where $0 \leq Z \leq +\infty$, was introduced for mathematical convenience.

The next step was to develop mathematical expressions for $f(Z/U_0)$ and $f(Z/U_1)$. These functions were developed for the data collected on the study site for the afternoon rush period of dry-weather weekdays. Statistical analysis by using the Kolmogorov-Smirnov test at the 5 percent level of significance confirmed the following truncated shifted gamma distributions.

$$f(Z/U_1) = 21.6[21.6(Z + 0.4)]^{16.1} \times \exp[-21.6(Z + 0.4)] \div 0.993\Gamma(17.1) \qquad (10)$$

$$f(Z/U_0) = 1.082[1.082(Z - 0.821)]^{-0.711} \times \exp[-1.082(Z - 0.821)] \div 0.991\Gamma(0.289) \qquad (11)$$

Probabilities of Capacity-Reducing Incidents

Once the study site was selected, the number of capacity-reducing incidents was determined through correlation of the assistance-rendered reports with the flow abnormalities indicated by available occupancy contour maps. The average number of incidents occurring on dry-weather weekdays during the afternoon rush period (2 - 6 p.m.) was found to be 1.04.

The probability of an incident occurring at a given detector at a specified minute in time is given by the ratio A:B*C, where A = average number of incidents occurring on the study section in the total time period, B = total number of detectors in the study section, and C = number of minutes in the time period. Hence

$$P(\text{incident}) = 1.04/(15 \times 240) = 0.000\,27 \qquad (12a)$$

and

$$P(\text{no incident}) = 0.999\,73 \qquad (12b)$$

Derivation of Optimal Threshold

Expressing $P(U_0)$, $P(U_1)$, $f(Z/U_0)$, and $f(Z/U_1)$ mathematically allowed the graphical derivation of the optimal threshold Z that maximizes Expression 7. The optimal threshold was found to be $Z_1 = 0.57$, which gives $\mathrm{OCCRDF}(t) = 1 - Z_1 = 0.43$.

ALGORITHM EVALUATION

The effectiveness of the algorithm was evaluated by determining its detection rate, false-alarm rate, and mean time to detect by running it through incident and incident-free data related to the study site. Also, these results were compared with those obtained by applying three other algorithms to the same data.

The threshold value of the traffic feature OCCRDF was compared with minute values of this feature for the incident and incident-free data. Any time the value of OCCRDF was greater than the threshold value, a "1" signal was output. A "0" signal was output when the threshold value was not exceeded. The first "1" signal was considered to be an indication of a tentative incident for both the incident and the incident-free data. A string of four consecutive "1"s was required to signal a confirmed incident. Detection time was defined as the difference between the time the fourth consecutive "1" signal was output and the apparent occurrence time of the incident.

A total of 17 incidents representing the afternoon rush period on dry-weather weekdays and 2 h of incident-free data taken at 15 subsystems were analyzed. The detection rate (percentage of incidents detected), false-alarm rate (percentage of "1111" signal strings in incident-free data), and mean time to detect found were as follows:

Parameter	Value
Detection rate, %	100
False-alarm rate, %	0.0
Mean time to detect, min	3.9

Note that the structure of the algorithm requires that the mean time to detect be at least 4 min. Thus, the mean time to detect actually achieved is as good as can be expected from this algorithm. (The slight discrepancy between 3.9 min and 4 min is due to inaccuracies in determining the apparent time of occurrence of the incident.)

The Bayesian algorithm was compared with the California algorithm (TSC-2) and the TSC algorithms 7 and 8 (3). The structure of these algorithms are shown in Figures 1, 2, and 3. The structure of the Bayesian algorithm is shown in Figure 4. By using the optimization routine developed by TSC for any chosen detection rate, a set of optimal thresholds for each of the TSC algorithms was established. This set yielded the minimum false-alarm rate for the particular detection rate. A comparison among the four algorithms by using the incident and incident-free data obtained at the study site is given below:

Algorithm	Detection Rate (%)	False-Alarm Rate (%)	Mean Time to Detect (min)
Bayesian	100	0.0	3.9
California	100	0.11	1.5
TSC 7	100	0.0	1.5
TSC 8	100	0.0	1.5

The Mann-Whitney U-test was conducted to determine the significance of the differences in mean time to detect between the Bayesian algorithm and each of the others. At the 5 percent level of significance, such difference was found.

From these results it can be seen that, for the 17 incidents on the outbound Kennedy Expressway, the Bayesian algorithm compared favorably with others tested. In some cases, a difference of 2-2.5 min in detecting an incident might not be particularly significant.

This could be the case if the variability in response time of the incident-handling subsystem is quite considerable or traffic messages are given by commercial radio frequently during the rush period. Also, a delay of such magnitude in implementing ramp-control strategies for incident situations should not be detrimental, especially if there is a dynamic control system that is responsive to flow changes.

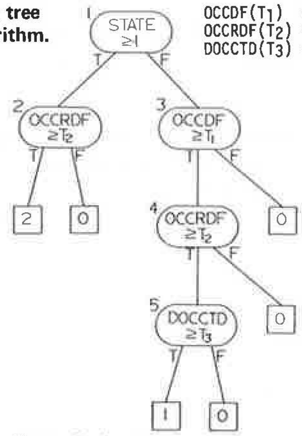
Probabilities of Incidents Given Various Signal Strings

As discussed above, theoretical probabilities for the actual occurrence of incidents given certain signal strings can be developed once $f(Z/U_0)$, $f(Z/U_1)$, $P(U_0)$, and $P(U_1)$ are defined quantitatively. Moreover, such values can be computed by considering the number of capacity-reducing incidents that occurred during the particular time slice just before the incident under consideration. These values are given below.

Signal String	Probability of Incident
P(incident/1)	0.003 05
P(incident/11)	0.033 51
P(incident/10)	0.001 04
P(incident/111)	0.282 09
P(incident/110)	0.011 66
P(incident/101)	0.011 66
P(incident/100)	0.003 54
P(incident/1111)	0.816 62

This table shows that, once a signal "1" is output,

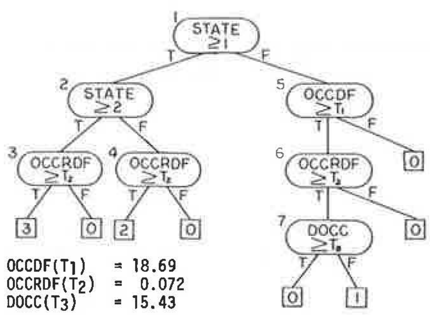
Figure 1. Decision tree for California algorithm.



OCCDF(T1) = 22.89
OCCRDF(T2) = 0.6418
DOCCTD(T3) = 1.4144

State	Designates
0	Incident-free conditions
1	Incident occurred
2	Incident continuing

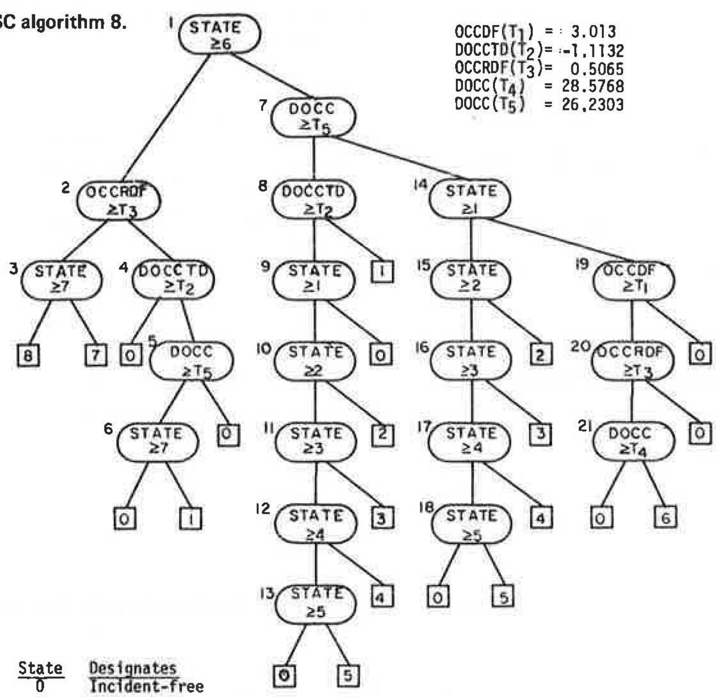
Figure 2. Decision tree for TSC algorithm 7.



OCCDF(T1) = 18.69
OCCRDF(T2) = 0.072
DOCC(T3) = 15.43

State	Designates
0	Incident-free conditions
1	Tentative incident
2	Incident occurred
3	Incident continuing

Figure 3. Decision tree for TSC algorithm 8.



OCCDF(T1) = 3.013
DOCCTD(T2) = -1.1132
OCCRDF(T3) = 0.5065
DOCC(T4) = 28.5768
DOCC(T5) = 26.2303

State	Designates
0	Incident-free
1	Compression wave downstream in this minute
2	Compression wave downstream 2 minutes ago
3	Compression wave downstream 3 minutes ago
4	Compression wave downstream 4 minutes ago
5	Compression wave downstream 5 minutes ago
6	Tentative incident
7	Incident confirmed
8	Incident continuing

the probability that an incident has occurred is 0.00305 and the probability of making a wrong incident-management decision is close to 100 percent. The probabilities of the actual occurrence of an incident do not improve significantly for two or three consecutive signals. For a string of four "1" signals, this probability increases to 0.817; obviously the longer the string of "1" signals, the higher the probability of actual occurrence of an incident. However, the decision about the appropriate string size to operate should come after an on-line evaluation.

Preliminary On-Line Evaluation

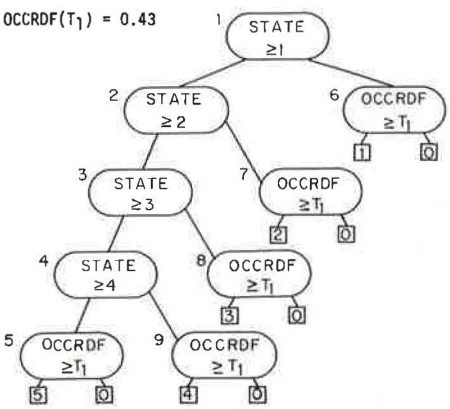
A preliminary on-line evaluation comparing the Bayesian algorithm with algorithm 7 was conducted on the Eisenhower Expressway (inbound and outbound) during the afternoon rush period on dry-weather weekdays.

In calibrating the Bayesian algorithm for use on the Eisenhower test section, P(incident) and P(no incident) were calculated by using incident data obtained in a previous study. Based on 25 days of data, it was found that, during the afternoon rush period on the Eisenhower Expressway P(incident) = 0.00054 and P(no incident) = 0.99946. By using the Bayesian criterion, the optimal threshold was determined to be 0.43.

In the on-line evaluation, algorithm 7 was run by using a pair of threshold sets—the first applied to typical freeway sections and the second applied to sections that had geometric anomalies such as bottlenecks. This fine tuning of algorithm 7 improved the false-alarm rate, which had been quite high at these geometric problem areas. This pair of thresholds was used for the entire 5-d evaluation period.

The Bayesian algorithm was run by using a threshold of 0.43 for the first 3 d of the evaluation. Inspection of the data indicated that increasing the threshold to 0.6 would substantially improve the false-alarm rate without affecting the detection rate. Thus, for the two remaining days of the study, the Bayesian algorithm was run by using a threshold of 0.60. Table 1 presents the comparison between the two algorithms. A t-test performed on the first 3 d of data showed no significant differences in detection rate, false-alarm rate, and frequency of false alarms. However, a significant difference in the number of false alarms exists at the 10 percent level of significance. No significant difference between the mean time to detect of incidents detected by both algorithms was found at the 5 percent level of significance. The increase in the Bayesian threshold in the next 2 d reduced the number of false alarms but also reduced the detection rate.

Figure 4. Decision tree for Bayesian algorithm.



State	Designates
0	Incident-Free Conditions
1	Tentative Incident
2	Tentative Incident
3	Tentative Incident
4	Incident Occurred
5	Incident Continuing

CONCLUSIONS

Based on the analyses presented in this paper, the single-feature Bayesian algorithm compared favorably with multifeature algorithms with respect to detection rate and false-alarm rate. There was an apparent difference of 2-2.25 min in mean time to detect between the TSC algorithm 8 and the Bayesian algorithm in the off-line evaluation. The preliminary on-line evaluation, however, showed no significant difference in mean time to detect between the Bayesian algorithm and the TSC algorithm 7.

The simplicity of the Bayesian algorithm allows one to estimate the likelihood that an incident signal will be a false alarm or an actual incident. Operationally, one is not limited to receiving only two possible messages—incident occurred or incident free. It would be easy to receive a printout of a string of feature values so that the operator could judge the severity of the situation. For example, if the string consists of low values close to 0.43, one may wish to wait for further evidence before responding to the message. On the other hand, if the string has extremely high values, it could be judged highly indicative of incident conditions. The Bayesian algorithm presents the operator with a set of incident signals to which he or she can apply other resources and so reduce the false-alarm rate.

Clearly, more complicated multifeatured algorithms do not permit the operator such flexibility.

The use of historical incident information in the Bayesian model allows the development of probability values that can be used in determining the occurrence of an incident at a particular location and time.

The Bayesian model could be used when the occurrence of nearly simultaneous incidents required the allocation of limited resources.

At this time, there should be a more comprehensive on-line evaluation of the Bayesian and other algorithms.

Table 1. Preliminary on-line comparison between TSC 7 and Bayesian algorithms.

		Detection Rate		False-Alarm Rate		No. of False Alarms	
Day	No. of Incidents	Algorithm 7	Bayesian Algorithm	Algorithm 7	Bayesian Algorithm	Algorithm 7	Bayesian Algorithm
1	1	1.0	1.0	0.75	0.84	3	5
2	2	0.5	0.5	0.66	0.84	2	5
3	5	0.2	0.6	0.80	0.75	4	9
4	2	0.5	0.0	0.84	1.0	5	4
5	5	0.6	0.2	0.40	0.66	2	2
Avg		0.46	0.40	0.69	0.80		

ACKNOWLEDGMENT

The research reported in this paper was conducted within the framework of the Illinois Department of Transportation Highway Planning and Research Program. We wish to acknowledge the support of the staff of the Chicago Area Expressway Surveillance Project in the data-collection phase. The opinions, findings, and conclusions expressed in this paper are ours and not necessarily those of the sponsoring agencies.

REFERENCES

1. K. G. Courage and M. Levin. A Freeway Corridor Surveillance, Information, and Control System. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 488-8, Dec. 1968.
2. H. J. Payne, D. N. Goodwin, and M. D. Teener. Evaluation of Existing Incident-Detection Algorithms. Federal Highway Administration, Rept. FHWA-RD-75-39, Feb. 1975; NTIS, Springfield, VA, PB 241 883.
3. H. J. Payne, E. D. Helfenbein, and H. C. Knobel. Development and Testing of Incident-Detection Algorithms: Vol. 2—Research Methodology and Detailed Results. Federal Highway Administration, Rept. FHWA-RD-76-20, April 1976; NTIS, Springfield, VA, PB 259 237.
4. C. L. Dudek and C. J. Messer. Incident Detection on Urban Freeways. TRB, Transportation Research Record 495, 1974, pp. 12-24.
5. M. Sakasita and A. D. May. Development and Evaluation of Incident-Detection Algorithms for Electronic Detector Systems on Freeways. U.S. Department of Transportation, Rept. DOT-TST-75-94, Aug. 1974.
6. A. R. Cook. The Detection of Capacity-Reducing Incidents on an Urban Freeway Using Traffic-Stream Measurements. University Microfilms, Ann Arbor, MI, 1972.
7. M. Levin and G. M. Krause. Off-Line Evaluation of Improved Incident-Detection Algorithms. Illinois Department of Transportation, Sept. 1977.

Publication of this paper sponsored by Committee on Freeway Operations.

Distance Requirements for Frontage-Road Ramps to Cross Streets: Urban Freeway Design

J. Michael Turner and Carroll J. Messer, Texas Transportation Institute, Texas A&M University

Heavy ramp volumes on urban freeway frontage roads can cause operating problems at adjacent diamond interchanges if the spacing between the cross street and the ramps is not adequate. Spacing requirements from diamond interchanges were studied for both exit ramps and entrance ramps. Exit ramps should provide sufficient storage for the vehicles queued at the cross street as well as an adequate distance for the weaving and braking maneuvers performed by exiting vehicles. Entrance ramps should provide sufficient spacing to prevent queueing through the cross-street intersection. Studies were made on several Texas freeways to provide data to be used to determine spacing requirements. Volume and queue counts were conducted to determine realistic volumes. These studies were used to determine design criteria and to further justify the analytic exit- and entrance-ramp models presented.

The modern urban freeway is conceived and constructed to move large numbers of persons and goods safely and efficiently over considerable distances. The basic design objective is to provide a high level of service in an economical manner. One of the consequences of this design objective is that the spacings of ramps and interchanges are relatively long because this minimizes the effects of weaving on the freeway flow. Apparently, little attention has been given to the resulting negative effects on the connecting ramps and frontage roads. Neither AASHO (1) nor Leisch (2) discuss this question. The traffic engineer currently has a minimum of design criteria or procedures available to use in selecting ramp spacings. Some designers have used rule-of-thumb procedures—e.g., 152 m (500 ft) for exit ramps and 229 m (750 ft) for entrance ramps.

Significant operating problems have been observed on urban freeway ramps and frontage roads near diamond interchanges, especially those where ramp-metering systems are in operation. In most cases, these problems on connecting exit and entrance ramps are directly related to insufficient ramp spacings. These problems are of three different types:

1. Signal queues at interchanges that block merge areas of exit ramps and the frontage road (see Figure 1),
2. Signal queues at interchanges that back into freeway main lanes, and
3. Ramp-metering queues that back into cross-street intersections (see Figure 2).

Freeway exit ramps should be long enough to store enough vehicles to prevent the spillback of the vehicles onto the freeway. The dangerous condition of spillback should not be tolerated as a recurring event and should occur only as a result of unusual circumstances. Freeway entrance ramps should be long enough to minimize queue spillback into the adjacent cross-street intersection because of ramp metering. The installation of ramp metering is a frequently used solution to the problem of congestion on urban freeways, even newly constructed ones.

The objective of this study was the investigation of the locations of entrance and exit ramps with respect to effects on the operation along the frontage-road approaches to an interchange and the development of criteria for the determination of the appropriate distances between ramps and interchanges. This paper will identify the different components that determine ramp-distance requirements and incorporate them into analytic models for exit and entrance ramps. The paper also provides information on various field studies conducted on freeways throughout Texas. Finally, design criteria and recommendations based on the data evaluated will be presented.

Figure 1. Signal queue at interchange that blocks merge area of exit ramp.

Figure 2. Ramp-metering queue that backs into cross-street intersection.

EXIT-RAMP SPACING

Exit-Ramp Model

The approach used in the determination of the storage length needed to prevent blockage of the ramp merge area considers three storage-length components. The geometric configuration shown in Figure 3 gives the three traffic-operating components used to compute the exit ramp-to-interchange spacing. These three components—weaving, braking, and queueing distances—are discussed below; perhaps it would be best to develop the length from the exit ramp to the signalized intersection of the interchange.

The distance needed to perform the weaving maneuver is the first distance to be provided. The basic weaving model presented in the Highway Capacity Manual (3) is used. Table 1 presents the required weaving distances for three levels of weaving quality of flow based on urban and suburban arterial-operating criteria. Total weaving volumes must be estimated from exit-ramp and frontage-road volumes and their respective turning movements at the interchange. No braking should be required during the weaving movement; the motorist should not be required to perform more than one basic driving task at a time. Therefore, the weaving movement should be completed before the motorist begins to brake to a stop.

The next distance required is the safe stopping distance. This length can be readily found by using Equation 1:

$$SSD = 2.78V \times T + (V^2/253f) \qquad (1)$$

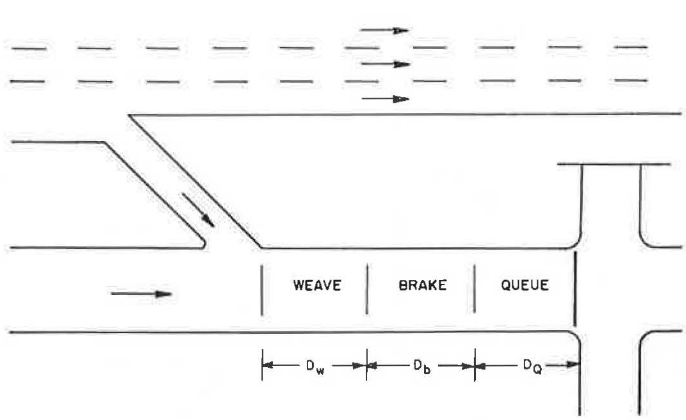
where

SSD = safe stopping distance (m),
V = frontage-road speed (km/h),
T = perception-reaction time (s), and
f = coefficient of friction.

The driver must safely stop before he or she reaches the end of the queue of vehicles stopped at the interchange traffic signal. Solutions to the SSD equation are shown in Figure 4 for perception-reaction times of 1.0 and 2.5 s. Deceleration rates and the resulting coefficients of friction vary with approach speed. Values used in this paper are those given by AASHO (1). A 1.0-s perception-reaction time provides only a minimum-condition reaction time; a 2.5-s time is more desirable.

The queue length at the interchange is the final component in the exit-ramp model. The design queue length can be obtained from Figure 5 (4).

Figure 3. Exit-ramp spacing required.



Exit-Ramp Studies

To develop and test the model, several types of studies were conducted on several different freeway locations. US-75 (North Central Expressway) in Dallas, US-59 (Southwest Freeway) and I-10 (Katy Freeway) in Houston, and some studies in Corpus Christi were chosen. These freeways vary with respect to geometrics and volumes experienced.

Three types of studies on exit ramps—volume counts, queue counts, and spacings between ramp and interchange—were conducted. Exit-ramp volumes were taken on I-10 in Houston and US-75 in Dallas. The cumulative exit-ramp volumes used to classify the volume levels are shown in Figure 6. Eighteen exit ramps are included in the morning and afternoon I-10 studies. Twenty-two exit ramps are included in the US-75 counts. A particularly troublesome high-volume exit ramp in Corpus Christi has a peak-hour volume of 1025 vehicles/h. These volumes provide a base for calculating weaving distances.

Figure 7 presents the measured exit-ramp distances versus the cumulative percentage of ramps for 20 ramps on US-75 in Dallas and 10 ramps on US-59 in Houston.

Table 1. Weaving lengths for variable weaving volumes and design levels.

Total Weaving Volume[a] (equivalent passenger automobiles/h)	Quality of Flow (level of service) 3-4 (A-B)[b]	4 (C-D)[b]	5 (E)[b]
100	15	15	15
200	15	15	15
300	15	15	15
400	30	15	15
500	30	15	15
600	30	15	15
700	60	30	15
800	75	30	15
900	90	45	15
1000	110	60	15
1100	120	60	15
1200	140	75	15
1300	150	90	15
1400	170	90	15

[a]Total weaving volume is assumed to be 63 percent of total frontage-road approach volume.
[b]Level of service is based on urban and suburban arterial criteria (3, pp. 173-5) and assumes that number of lanes is adequate for weaving.

Figure 4. Graphical solution of Equation 1 at perception-reaction times of 1.0 and 2.5 s.

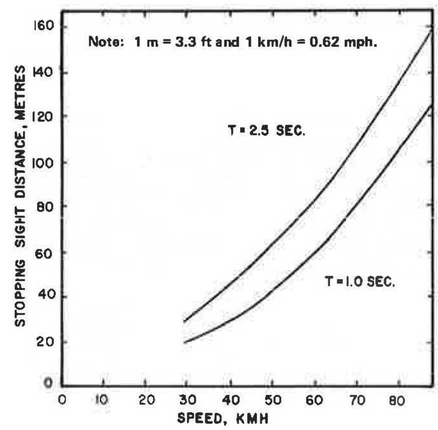


Figure 5. Storage length versus turning volume (modified Poisson distribution).

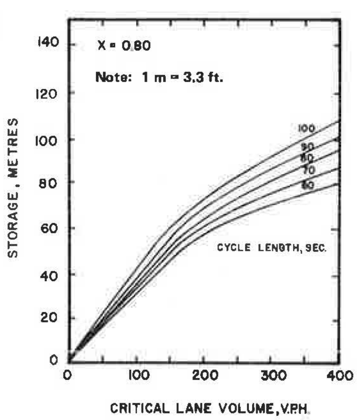


Figure 6. Cumulative percentages of exit-ramp volumes: US-75 in Dallas and I-10 in Houston.

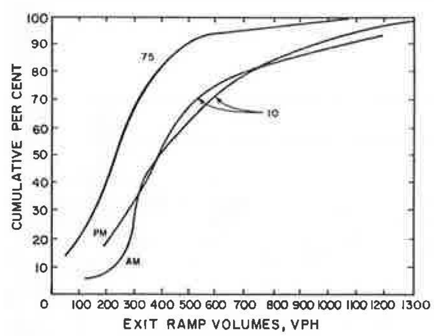


Figure 7. Cumulative percentages of exit-ramp distances: US-75 in Dallas and US-59 in Houston.

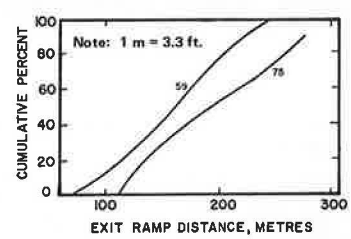


Figure 8. Heavy-volume traffic conditions.

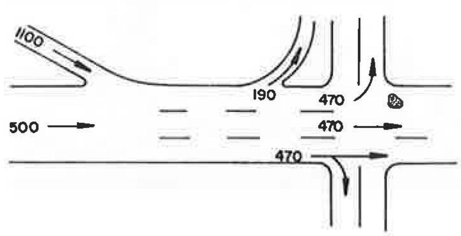


Table 2. Exit-ramp design criteria for three design levels.

Design Criterion	Design Level 1: Desirable	2: Usual Minimum	3: Absolute Minimum
Operating speed, km/h	50-58	45-50	30
Weaving quality	3-4	4	5
Weaving volume, %	63	63	63
Perception-reaction time, s	2.50	1.75	1.00
Stopping distance, m	25.91	16.76	7.01
Cycle length, s	90	80	70
Signal saturation, X	0.80	0.80	0.80
Maximum lane volume[a]	$F \times V_T$	$F \times V_T$	$F \times V_T$

Note: 1 km/h = 0.6 mph and 1 m = 3.28 ft.
[a]$F = 0.4 + 0.6 \exp(-0.13\ V_T C/3600)$ where V_T = total frontage-road approach volume.

The median (50th percentile) distances are about 152.4 m on US-59 and 182.9 m on US-75. Distances were measured from the physical nose to the stop line of the intersection. These studies can be used to compare the results of the model with existing ramp spacings.

Exit-Ramp Design Criteria

The exit-ramp spacing model has been formulated in terms of the three component lengths discussed above—weaving distance, braking distance, and queue length. The distance required for weaving is primarily related to the exit-ramp volumes and the total weaving volume. The exit-ramp-volume data indicate that the 95th percentile exit-ramp volumes of the two study sites shown in Figure 6 are approximately 690 and 1100 vehicles/h. These are termed moderate and high-volume conditions and are the two basic design-volume conditions defined in this paper. The assumed volume distributions for the 1100 vehicle/h exit-ramp flow is shown in Figure 8; frontage road, U-turn, and lane distributions were selected as being representative of the high-volume conditions. Any other exit-ramp volume is assumed to have the same percentage distribution of traffic movements as does the 1100-vehicle volume level; other volumes can be scaled to lower or higher levels than that shown in Figure 8 depending on how they compare to 1100. These volumes can then be used to determine total weaving volumes and resulting required weaving distances.

To determine the trade-off options between freeway level of service and frontage-road operating conditions, exit-ramp design levels of performance were defined as (a) desirable, (b) usual minimum, and (c) absolute minimum (corresponding to qualities of flow of 3-4, 4, and 5 respectively). Although these design levels are not defined specifically in terms of equivalent levels of service, they represent approximately levels of service C, D, and E respectively. Design criteria selected for the model variables are given in Table 2. These variables include quality of weaving, safe approach speed for stopping, perception-reaction time, and signalized-intersection cycle length. The values selected define reasonable and desirable conditions for operations but are certainly not ideal conditions. Design level 3 is an absolute minimum or capacity level and its use is not recommended.

Exit-Ramp Spacings

The exit-ramp spacings calculated by the model for total frontage-road volumes ranging from 200 to 2000 vehicles/h are given in Table 3 for the design levels defined above. Distances are from the exit-ramp centerline point of merge with the frontage road to the stop line at the signalized intersection. This distance may be 30-61 m (100-200 ft) less than the actual distance from the physical nose of the exit ramp to the intersection because of the exit ramp entry angle (4°) and any pedestrian crosswalks at the intersection.

Exit-Ramp Summary

Experience has shown that exit ramps may experience operating blockages at their point of merge with frontage roads because of queue spillback from the adjacent signalized intersection. It is highly probable that many exit ramps have this type of problem, and these ramps should be redesigned to provide greater frontage-road spacing. A desirable level of design should be used where possible. Trade-off analyses should be made between the freeway and frontage-road operations if providing a more desirable exit-ramp spacing would also lower the level of service on the freeway.

Careful consideration should be given before designing the spacing of exit ramps to less than that required for a frontage-road volume level of 600 vehicles/h. Planning data and projected volumes are based on numerous assumptions and estimations of future events and, consequently, exit-ramp volume projections may be in considerable error. Likewise, exit ramps that are expected to feed adjacent major arterials or traffic generators probably should not be designed for a volume level of less than 1600 vehicles/h.

ENTRANCE-RAMP SPACING

Entrance-Ramp Model

Ramp-metering systems are becoming an accepted practice on urban freeways. This has caused concern about the spacing provided between diamond interchanges and the point of merge from the entrance ramp to the freeway main lanes. Queues form at these metered ramps and sometimes back into the cross-street intersections, as shown in Figure 2. The number of vehicles stored behind the ramp signal over a period of time depends on the ramp demand volume and the operating capacity of the ramp-metering signal.

By assuming Poisson arrivals to the ramp and Poisson departures from the ramp-metering signal, Morse (5) has shown that the probability that a ramp of a known queue storage (N) will overflow is given by

$$\text{Probability of overflow} = (\text{volume}/\text{capacity})^{N+1} \qquad (2)$$

Results of this model are shown in Figure 9 for volume-to-capacity ratios (V:C) of the ramp-metering signal of 0.80, 0.90, and 0.95. The higher the V:C ratio, the longer the ramp storage required for a given probability of queue overflow. For a given ramp demand-volume level, the required ramp storage increases as the freeway level of service decreases.

From a theoretical viewpoint, the queue-length distribution over time and space can be determined if the demand volume and metering capacity of the ramp are known. Few studies have been published relating ramp-metering capacity to freeway-lane-one (outside-lane) volumes. Brewer and others (6), however, have developed a theoretical model of merging-control operations that was later validated in Houston. An approximation of this model is

$$C_r = 1620 - 0.81V_1 \qquad (3)$$

where C_r = capacity of metered entrance ramp (vehicles/h) and V_1 = volume of lane one (outside) of freeway (vehicles/h). The normal ramp capacity theoretically cannot be less than the minimum acceptable cycle length if 1 automobile/cycle is metered onto the freeway. Cycle lengths should not be less than 4.0 to 4.5 s (900 to 800 vehicles/h). Ramp volumes greater than 800 vehicles/h usually have high violation rates and multiple vehicle entries during the green.

The volume on the outside lane of the freeway (lane one) varies with the total freeway flow. An estimate of the volume in lane one can be determined from Figure 10. By using level-of-service criteria given in the Texas design manual (7), a range of level-of-service D lane-one volumes and the resulting ramp-metering capacities were developed. From these results, required ramp vehicle storages were calculated at a 5 percent probability of overflow by using Morse's equation; these results are given in Table 4. Level of ser-

Table 3. Distance requirements for separation between exit ramp and cross streets for different design levels.

Total Frontage-Road Volume[a] (vehicles/h)	Approximate Exit-Ramp Volume[b] (vehicles/h)	Design Level		
		1: Desirable	2: Usual Minimum	3: Absolute Minimum
200	140	152	116	79
400	275	170	140	110
600	410	192	152	122
800	550	210	165	131
1000	690	232	180	137
1200	830	265	195	146
1400	960	296	210	152
1600	1100	326	235	162
1800	1240	360	262	168
2000	1380	396	296	177

[a] Exit-ramp volume plus existing frontage-road volume.
[b] Exit-ramp volume assumed to be 69 percent of total volume.

Table 4. Required ramp vehicle storage for given ramp volumes.

Volume (automobiles/h)	Freeway Level of Service		
	Near C	Mid-D	D Near E
Lane no. 1	1000	1200	1400
Ramp metering capacity	810	648	486
Ramp demand			
300	2	3	6
400	4	6	15
500	5	11	-[a]
600	7	38	-[a]
700	11	-[a]	-[a]
800	25	-[a]	-[a]

[a] Large queue: cannot be calculated by theory.

Figure 9. Relationship between volume:capacity ratio of ramp metering and probability of ramp storage overflow.

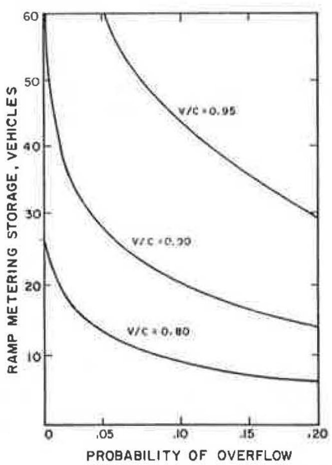


Figure 10. Approximate volume of through traffic in lane no. 1 in vicinity of ramp gores.

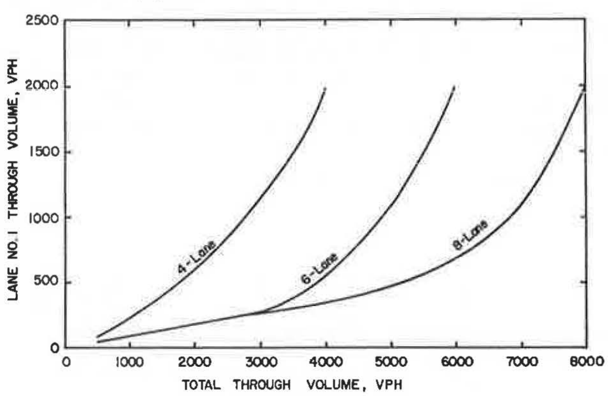
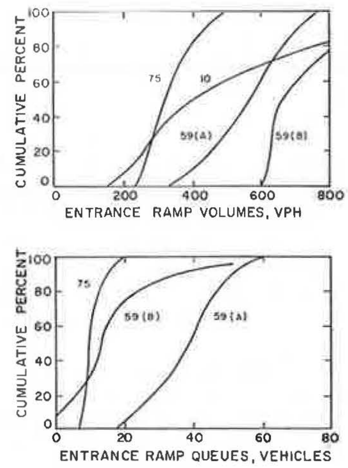


Figure 11. Relationship between entrance-ramp volumes and entrance-ramp queues.



vice D was selected as the base because peak-hour metering frequently operates within this high-volume stable-flow region.

Entrance-Ramp Studies

Several studies of geometric and operating characteristics were conducted to establish the model parameters, to confirm the realism of the model results, and to provide field data with which to objectively evaluate existing freeway geometrics.

Entrance-ramp volume data were collected at 21 metered entrance ramps on US-75 (North Central Expressway) in Dallas during the peak hour for 2 d during April 1976. A cumulative frequency plot (converted to percent) is given in the top illustration of Figure 11 as curve 75. Most metered ramp volumes ranged from 250 to 400 vehicles/h; the maximum was 510. No connecting roadways from interchanges to the freeway were included in this sample. All of these ramps were on continuous frontage-road sections.

Metered entrance-ramp volume studies were also conducted on US-59 (Southwest Freeway) in Houston during the spring of 1976. Four high-volume ramps—Bellaire, Westpark, Chimney Rock, and Enloe—were observed to study high volume and delay conditions. These ramps are in an area of southwest Houston beyond loop I-610 and near a large shopping center complex, and most do not have frontage roads or other convenient alternative freeway routes. Cumulative percentage plots of ramp-volume data taken during the peak hour are given in the top illustration of Figure 11 as curve 59(A). Most ramp volumes were between 450 and 650 vehicles/h; two-thirds were greater than 500 vehicles/h.

Another set of entrance ramps on US-59, these inside of loop I-610, were also studied. Four ramps near the Summit arena—inbound and outbound Buffalo Speedway and outbound Shepherd and Kirby drives—were counted. These ramps are on a continous frontage-road section of an eight-lane freeway section of US-59. Curve 59(B) in Figure 11 shows that high volumes occur on these ramps; most of the queues observed on these ramps ranged between 10 and 25 vehicles.

A fourth set of ramp-volume data was obtained from a study of I-10 (Katy Freeway) in Houston between the outer loop (I-610) and West Belt Drive. A total of 16 entrance ramps were counted during the peak hour in the inbound or outbound direction. These average

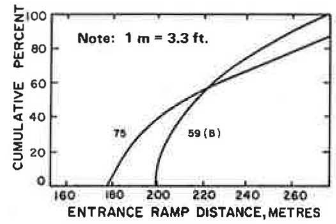


Figure 12. Relationship between recommended design spacings and existing ramp spacings.

volume results are shown in Figure 11 as curve 10. This curve is much flatter and has a wider range of volumes; about 20 percent exceed 800 vehicles/h. These ramps are not metered. The capacity of a metered ramp seldom exceeds 800 vehicles/h. Metered ramp volumes tend to balance the traffic demand over a section of freeway, thereby reducing hot spots or local pockets of high-density flow.

Queue-count studies at metered entrance ramps were also conducted on US-75 (North Central Expressway) in Dallas and US-59 (Southwest Freeway) in Houston. These peak-hour results are shown in the lower illustration of Figure 11. Most queues observed on US-75 in Dallas ranged from 5 to 15 vehicles, although none of the 21 ramps in the continuous frontage-road sections studied operate at what might be considered high volumes. Some of the lower volume ramp-queue counts were not included in this data set; the ramp metering combined with the continuous frontage roads and progressive operations tended to balance out the ramp loadings. Motorists were frequently observed to divert from joining a ramp queue if it exceeded 8 to 10 vehicles. By referring to the table above and Figure 11, one can see that ramp volumes in the 300-400 vehicle/h range would be expected to generate queue lengths in the 5-15 vehicle range during the peak hour (as did the US-75 ramps shown in Figure 11). Ramp volumes in the 400-600 range normally would be expected to generate moderately long queues (6-38 vehicles at midlevel of service D). This rather wide queueing range is illustrated by the higher volume results of curve 75 in Figure 11 and the lower volume queues of curve 59(A). Ramp volumes greater than 500 vehicles/h may generate large ramp queues, as did US-59 outside of I-610, where some queues frequently exceeded 50 vehicles when the freeway was operating near or at level of service E.

On the other hand, ramp volumes in the 600-800 range may generate only moderate queues, as did US-59 inside of US-10, if the freeway is operating at a level of service near C at the metered ramp.

Based on the results of this study, it appears that the number of vehicles in a ramp queue depends primarily on the operating level of service on the freeway, the ramp demand volume, and whether continuous frontage roads (and possibly frontage-road progression) are available.

Entrance-Ramp Design Criteria

It is recommended that the design of entrance ramps in urban areas provide adequate spacing between the diamond interchange and the point of entrance-ramp merge to the freeway such that ramp metering can be installed and operated and not generate queues that overflow into the adjacent interchange.

There are basically only two parts required in determining spacing requirements—the metering section and the queue storage. The first part of the ramp design must provide an adequate distance between the ramp signal and the point of merge to allow an automobile to accelerate to a reasonable merge speed and select a gap if available. Everall (8), indicates that 61-76 m (200-250 ft) is required to provide adequate time to merge. Ramp-metering studies in Dallas and Houston support these guidelines. However, 61 m should be considered the minimum distance to the merge point and 76 m the desirable.

Recent research in Texas has shown that vehicles store at about 7.6-m (25-ft) intervals behind traffic signals (4). Thus, the queue storage needs discussed above must be multiplied by 7.6 m/automobile and this added to the 61-76 m required from the ramp signal to the merge point to determine the required ramp spacings.

Entrance-Ramp Spacings

The recommended entrance-ramp-spacing design requirements (measured from the curb line of the adjacent diamond interchange to the point of merge of the entrance ramp and the freeway) are given below (1 m = 3.28 ft).

Ramp Demand Volume (vehicles/h)	Spacing (m) Desirable	Minimum
<300	229	137
400	305	175
500	381	213
600	457	251
700	533	290
>800	610	328

Desirable and minimum design spacings were selected based on the previous study results and the considered judgment of the researchers.

A comparison can be made between the recommended design spacings given above and the existing entrance ramp spacings determined for two of the freeways described above. These spacings for US-75 in Dallas and US-59 near the Summit arena are shown in Figure 12. In general, the minimum ramp spacings are being provided by the two current freeway designs although some ramps may be deficient. An analysis of the individual ramp volumes and spacings on US-75 showed that 55 percent of the entrance ramps did not meet the desirable spacing criteria. All the US-59 ramps inside I-610 studied fail to meet the desirable spacing criteria because of the high volume levels experienced. In the US-75 data set, there was no correlation between ramp volumes and the ramp-spacing provided. That is, a low-volume ramp was just as likely to have a long spacing as a higher volume ramp.

Entrance-Ramp Summary

The design of entrance ramps in urban areas should consider the possibility that ramp metering will be installed. Adequate ramp spacings are required between the adjacent diamond interchanges and the point of merge of the ramp and the freeway to ensure smooth ramp-metering operations and little queue overflow into the interchange. Minimum and desirable spacing design spacings have been presented that should be considered in future design work. An investigation should be conducted of the current adequacy of all entrance ramps in urban centers in Texas to evaluate the potential need to redesign those ramps that have deficient spacings.

ACKNOWLEDGMENTS

This paper discusses one phase of a research project on the development of a frontage-road level-of service evaluation program that was conducted by the Texas Transportation Institute and sponsored by the Texas State Department of Highways and Public Transportation in cooperation with the U.S. Department of Transportation, Federal Highway Administration. We wish to thank Harold D. Cooner, Herman E. Haenel, and Elmer A. Koeppe of the Texas State Department of Highways and Public Transportation for their technical inputs and constructive suggestions throughout the duration of this project. The assistance provided by Murray A. Crutcher and the staffs in Houston and Dallas of the Texas Transportation Institute and of district 12 of the Texas State Department of Highways and Public Transportation is gratefully acknowledged. The contents of this paper reflect our views; we are responsible for the facts and accuracy of the data presented herein. The contents do not necessarily reflect the official views or policies of the Federal Highway Administration. This paper does not constitute a standard, specification, or regulation.

REFERENCES

1. A Policy on Geometric Design of Rural Highways. AASHO, 1965.
2. J. E. Leisch. Spacings of Interchanges on Freeways in Urban Areas. Trans., ASCE, Vol. 126, Pt. 4, 1961.
3. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
4. C. J. Messer, D. B. Fambro, and D. A. Andersen. A Study of the Effects of Design and Operational Performance of Signal Systems. Texas Transportation Institute, Texas A&M Univ., College Station, Res. Rept. 203-2F, Aug. 1975.
5. P. M. Morse. Queues, Inventories, and Maintenance. Operations Research Society of America, Operations Research, No. 1, 1967.
6. K. A. Brewer, J. H. Buhr, D. R. Drew, and C. J. Messer. Ramp Capacity and Service Volume as Related to Freeway Control. HRB, Highway Research Record 279, 1969, pp. 70-86.
7. Highway Design Division Operations and Procedures Manual. Texas Highway Department, Pt. 4, 1970.
8. P. F. Everall. Urban Freeway Surveillance and Control: The State of the Art. Federal Highway Administration, 1972.

Publication of this paper sponsored by Committee on Freeway Operations.

Abridgment

Stimulus-Response Lane-Changing Model at Freeway Lane Drops

William A. Stock,* SRI International, Menlo Park, California
Jin J. Wang, JHK and Associates, San Francisco

Lane-changing behavior brought about by a freeway lane drop should differ considerably from the other types of lane-changing behavior considered in past studies because drivers in the dropped lane must merge into the adjacent through lane by or just downstream of the end of the lane-drop taper. Every lane-drop site has warning signs or pavement markings or both to alert drivers of the impending drop.

The stimulus-response lane-changing model was developed as a part of the freeway-lane-drop study (1) to relate driver lane-changing responses to various types of stimuli that forewarn of the lane drop. The model seeks to bridge the gap between macroscopic empirical observations of lane-changing behavior at freeway lane drops and microscopic human-behavior models that attempt to represent individual decision-making processes.

The lane-drop study collected approximately 2.5-3 h of traffic data at 18 mainline lane-drop sites by using a time-lapse photographic technique. Each site was divided into five 122-m (400-ft) subsections for data-reduction purposes; the boundary between the fourth and fifth subsections was always located at the end of the lane-drop taper. Traffic measures were reduced on a subsection or an entire-site basis. Lane-changing data reduced in this manner formed the basis for calibration of the model.

MODEL STRUCTURE

The basic behavioral assumption of the stimulus-response lane-changing model is that each sign or pavement marking or the view of the lane drop itself can be considered as a stimulus that will induce a certain proportion of drivers to change lanes. Thus, the model is most effective in representing the behavior of drivers who rely on warning signs and the visibility of the lane drop in their lane-changing decision-making processes. Drivers, such as commuters, who frequently traverse the lane-drop section probably respond more to their preexisting knowledge. The discussion below applies only to those drivers who are unfamiliar with the lane-drop site.

In the simplest form, suppose that each stimulus causes a certain proportion (p) of drivers to change lanes. Then, the probability (P_i) that a randomly chosen driver entering the lane-drop area within the dropped lane will change into the adjacent through lane in response to the ith stimulus obeys a geometric distribution with parameter p.

However, the geometric distribution is only an approximation, because it requires an infinite number of stimuli to cause all drivers to leave the drop lane. In actuality, only a finite number of stimuli, say n, are present. An improved representation can be obtained by assuming that the beginning of the lane-drop taper constitutes a final (n+ 1st) stimulus that will cause all

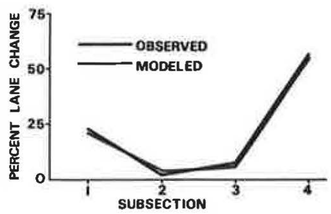


Figure 1. Observed and modeled lane-changing behavior at typical good-fit site.

drivers still in the drop lane to commence their lane changes. The complete response distribution with this modification is

$$P_i = \begin{cases} (1-p)^{i-1}p & (i \leq n) \\ (1-p)^n & (i = n+1) \\ 0 & (\text{otherwise}) \end{cases} \quad (1)$$

The actual lane-change-completion distribution over the length of the entire lane-drop area is of more interest than the response distribution. The drivers reacting to any one stimulus complete their lane changes over distance according to some probability density function [g (x)], where x represents distance relative to the location of that stimulus. This density function is referred to as the lane-change-completion density function. The density function of the gamma distribution, with parameters k and λ, was selected to represent lane-change completion in the model because this distribution has the flexibility to conform its shape to a wide range of possible lane-changing distributions. Furthermore, the gamma density function has a zero probability associated with x values less than zero (where it is anticipated that no lane changing in response to the stimulus should take place).

If the ith stimulus is located at coordinate d_i relative to a common origin, then the density function for lane changing in response to that stimulus relative to this common origin is $g(x - d_i)$. However, it is not reasonable to assume that the first lane-change completion arising from a stimulus will occur precisely at the location of that stimulus. Therefore, a location parameter (R) is included so that the density function for lane-change completion in response to stimulus i becomes $g(x - R - d_i)$.

Lane-change completion in response to a given stimulus is a random variable that is the sum of (a) distance traveled while perceiving the stimulus, (b) distance traveled while reacting to the stimulus, and (c) distance traveled while changing lanes in response to the stimulus. In this paper, the first two components are treated as constants, which allows handling their effects as part of R. It is realized that this treatment of the first two components is only approximate and so offers an area for further model improvement.

The overall lane-change-completion density function [f(x)] for the site is the weighted sum of the separate densities $[g(x - R - d_i)]$; the weights are the P_i given by Equation 1:

$$f(x) = \sum_{i=i}^{n} [(1-p)^{i-1} p \times g(x - R - d_i)] + [(1-p)^n \times g(x - R - d_{n+1})] \quad (2)$$

Finally, as presented here, the model has been generalized by assuming that signs and pavement markings are distinct types of stimuli and that each will affect drivers differently. Consequently, one impact value (P_s) will be used for signs and another impact value (p_m) will be used for pavement markings. An expression for the density function for such a generalization has been developed elsewhere (1).

PARAMETER ESTIMATION

The parameters k and λ were estimated from the mean and standard deviations of the lane-changing times given elsewhere (2). These estimated values are (1 m = 3.3 ft)

Side of Drop	k	$\lambda (m^{-1})$
Right	2.83	0.0305
Left	2.22	0.0174

The other parameters (R, p_s, and p_m) (at those sites that have pavement markings) were estimated individually from the observed data for each site by using a least-squares estimation procedure (1).

EVALUATION OF MODEL RESULTS

A total of 14 mainline lane-drop sites were selected for analysis with the model. The other four mainline sites were excluded from the analyses because lane-drop changing might be influenced by nearby same-side ramps not accounted for in the model.

Good model fits that had reasonable estimated parameter values were achieved at rural sites deemed to have few commuters, but poor fits occurred at urban sites that had heavy commuter traffic. This was not surprising; commuters can be expected to have limited reaction to signs and pavement markings because of their preexisting knowledge of the drop. These drivers will merge out of the drop lane farther upstream of the drop.

Observed and modeled behavior at a typical good-fit site are shown in Figure 1. The bimodal lane-changing pattern exhibited at this location can be explained very effectively by the model.

Based on the results of the seven sites for which the model fits the data well, the average values for the three parameters fitted are R = 67 m (220 ft); $p_s = 0.28$; and $p_m = 0.14$. For these seven sites, the impact of signing appears to be greater for left-side drop sites than for right-side drop sites. A second interesting finding is that pavement markings tend to have less impact than signs.

FUTURE WORK

A need exists for a more realistic portrayal of the perception and reaction distance components of the lane-change-completion distribution. Specifically, these two components should be treated as random variables. A critical need exists for better empirical data. Especially useful would be the locations of individual lane-change completion, which would allow the use of maximum-likelihood estimation techniques. Lastly, the model should be generalized to include the situation of driver familiarity with the lane drop. A possible model modification to handle this behavior is given elsewhere (1).

ACKNOWLEDGMENT

The model described in this paper was developed as a part of a study of freeway lane drops that was sponsored by the U.S. Department of Transportation, Federal Highway Administration.

REFERENCES

1. J. J. Wang, W. A. Stock, B. C. Johnson, and I. J. Fullerton. Improving the Safety and Operation of Freeway Lane Drops. JHK and Associates, San Francisco, Sept. 1977.

2. R. D. Worrall and A. G. R. Bullen. Lane Changing on Multilane Highways. Department of Civil Engineering, Northwestern Univ., Evanston, IL, June 1969.

Publication of this paper sponsored by Committee on Freeway Operations.

**Mr. Stock was at JHK and Associates when this paper was prepared.*

Abridgment

Derivation of Freeway Speed Profiles From Point Surveillance Data

Philip J. Tarnoff and Wayne K. Kittelson, Alan M. Voorhees and Associates, McLean, Virginia

In the past, detailed analyses of the traffic flow on freeways have been hampered by the cost of acquiring large samples of vehicle trajectories along the freeway sections of interest. The analyst has been forced to spend hours laboriously studying aerial photographs or performing numerous floating-automobile runs to obtain the required level of detail.

The use of freeway surveillance systems has been restricted by the inability of real-time surveillance systems to measure traffic flow over a continuous segment of roadway. The use of point sensors (usually implemented as loop detectors) in these systems has prevented them from being effectively applied to surveillance applications such as incident detection.

The purpose of this paper is to describe a technique that permits the derivation of individual vehicle trajectories from sensor data. This technique was developed by using data that were collected on a freeway surveillance and control system in the Los Angeles area and supporting data that were collected on the John Lodge Freeway in Detroit.

THEORETICAL BACKGROUND

This research is based on the application of sampling theory that has long been used in the communications field. A mathematical derivation of this theory is given elsewhere (1, 2).

Spectral Characteristics

One of the basic theorems of engineering analysis states that any arbitrary function of time [f(t)] can be described in terms of an infinite summation of sinusoidal components. If f(t) is periodic, the frequency of its sinusoidal components will be required. The description of a waveform in terms of its frequency spectrum rather than of its variation in time is known as the Fourier transform. It is the Fourier transform analysis that forms the basis for the development presented below.

Sampling Theory

It is a basic theorem of communications theory that (1)

If a signal whose highest frequency is W cycles has been sampled at a rate of 2W samples/s, the samples are in the form of impulses whose area is proportional to the magnitude of that instant, the sampled signal may be reconstructed by passing the impulse train through an ideal low-pass filter whose cutoff frequency is W cycles.

The process described by this theorem is shown graphically in Figure 1 (1, p. 24). This figure shows an arbitrary function of time [f(t)] being sampled by a commutator switch and then passed through an ideal low-pass filter where x(t) is reconstructed and matches f(t). Thus, the low-pass filter is a device that will pass only those frequency components of a time-varying waveform that are less than W cycles/s.

Black (2) has further extended this theory by a theorem that states that, if derivatives of f(t) are available, the sampling rate (Δt) must be

$$\Delta t = (R + 1)/2fc \qquad (1)$$

where R = number of derivatives available and fc = cutoff frequency (the highest component frequency) of input. Thus, the frequency with which a function must be sampled in time is inversely proportional to the number of derivatives used in the sampling process and directly proportional to its cutoff frequency.

Translation to Sampling in Space

To be useful for this application, it is necessary that the sampling theorem be modified to take into account the fact that the sampling is occurring in space rather than in time. Thus, by substituting space (s) for time (t) in the original theory, we obtain the overall version of the sampling equation as

$$f(s) = \sum_{k=-\infty}^{\infty} \{f(k\Delta s) + (s - k\Delta s)f'(k\Delta s) + [(s - k\Delta s)^2/2!]\, f''(K\Delta s) + \ldots + [(s - k\Delta s)^R/R!]\, f^{(R)}(k\Delta s)\} \times \{\sin[(\pi/\Delta s)(s - k\Delta s)]/(\pi/\Delta s)(s - k\Delta s)\} \qquad (2)$$

In equation 2, s is the location on the freeway at which the speed is being computed, $k\Delta s$ refers to the detector location (assuming equal detector spacing), and $(s - k\Delta s)$ is the distance between the location at which the speed [$f(k\Delta s)$] is being estimated and the detector from which the data is being used at a given instant of time. These relations are shown in Figure 2. Equation 2 shows that the further a detector is from s, the lower its influence will be on the overall speed computation. Thus, although the summation indicates that an infinite number of detectors must be sampled to reconstruct the speed at all points along the freeway, it will be possible to use a more limited number of detectors in the vicinity of s.

The required detector spacing (Δs) can similarly be derived by translating Equation 1 from the time domain to the space domain to give

$$\Delta s = (R + 1)/2fc \qquad (3)$$

where fc = cutoff frequency of the traffic-stream speed variations in space. The actual values of fc and Δs are discussed below.

Computing Derivatives

Examination of the terms in Equation 1 for the reconstruction of speeds based on speeds measured by individual sensors shows that a potential problem is created because it is necessary to compute derivatives of the speed [$f'(k\Delta s)$] at each of the sensors. This problem arises because the derivative is a function of space rather than of time; i.e.,

$$f'(k\Delta s) = df(k\Delta s)/ds \qquad (4a)$$

and

$$f''(k\Delta s) = d^2 f(k\Delta s)/ds^2 \qquad (4b)$$

Obviously, a point sensor such as a loop detector is not capable of measuring speed changes in space because, by definition, this equation is capable only of measuring data at a particular point along the roadway. Yet the analysis of cutoff frequency discussed below indicates that, to arrive at feasible sampling rates, it is necessary to compute at least the first and second derivatives of the speed change in space.

This problem is solved elsewhere (3); the relationship between the derivatives in time (t) and space (s) is

$$\partial u/\partial s = (-1/u_w)(\partial u/\partial t) \qquad (5a)$$

and

$$\partial^2 u/\partial s^2 = (1/u_w)^2 (\partial^2 u/\partial t^2) \qquad (5b)$$

where u = speed and u_w = speed-propagation rate.

u_w is related to the slope of the freeway flow (q) versus density (k) curve and can be expressed as

$$u_w = dq/dk \qquad (6)$$

Thus, for a particular section of freeway, it will be necessary to calibrate a flow-density curve as a part of the implementation process. Then, a value of u_w can be estimated by using the vehicle flows measured by the detectors. By knowing u_w and computing $\partial u/\partial t$ directly from the vehicle speed data, we can calculate the derivatives of speed with respect to space.

The use of this approximation permits the substitution of the derivative of speed with respect to time for the derivative of speed with respect to space. This derivative with respect to time is readily computed from the detector data.

FEASIBILITY OF APPLYING PROCESS

From the above discussion, it is obvious that the sampling theorem can be applied to the case of freeway-traffic-speed estimation only if the cutoff frequency of the speed profiles is low enough to permit a reasonable detector spacing as computed by Equation 2. To determine the minimum acceptable detector spacing, it was necessary to determine the frequency spectrum for actual freeway speeds as recorded during floating-automobile runs. Two sources of data were used for this determination—some collected on the John Lodge Freeway in Detroit at or near the Calvert, Chicago, Hamilton, and Gladstone overpasses and some were collected for this project on the 67-km (42-mile) loop in Los Angeles.

By using the worst-case frequency obtained from these sources, we can then determine the detector spacing (assuming that the first and second derivatives are computed) necessary to perfectly reconstruct vehicle speed in the following way: $s = (R + 1)/fc = 3/0.002 =$

Figure 1. Sampler and low-pass filter for reconstruction of signal (delay of low-pass filter neglected).

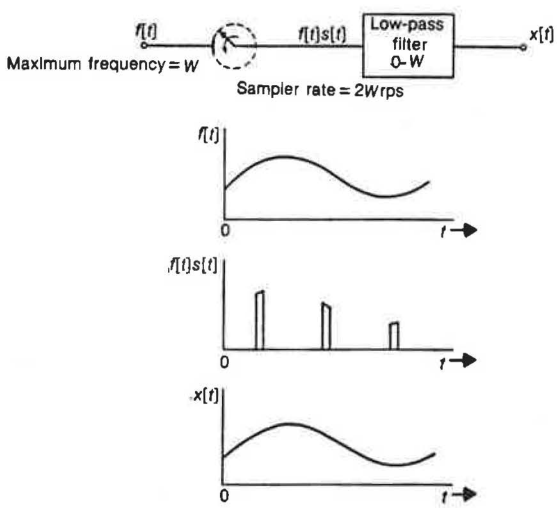
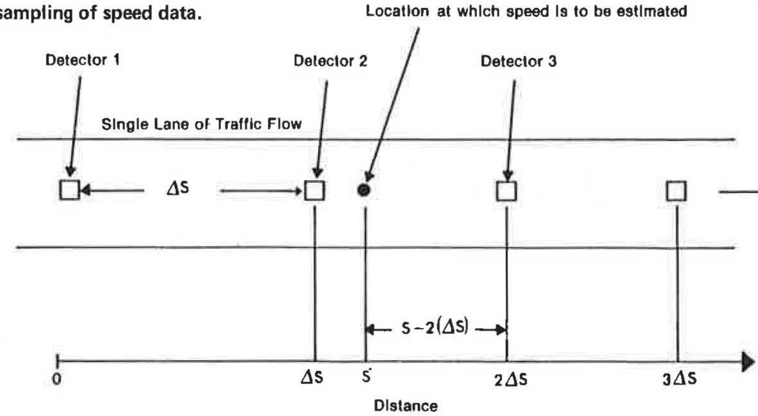


Figure 2. Relationships used in sampling of speed data.



1500 ft (457 m). This spacing would be required to provide accurate estimates of vehicle speeds. It should be used when there is a desire for reliable incident detection or other types of speed estimation. Thus, the spectral analysis represented can be used as a tool for the estimation of required detector spacing.

Raw detector data obtained from the California Department of Transportation were used in an attempt to demonstrate the feasibility of processing detector data to develop estimates of the equivalent of floating-automobile speeds. Unfortunately, the detector spacing exceeded that computed above and, as a result, a reliable validation of the process discussed was not possible.

CONCLUSIONS

The discussion above has described an approach to the derivation of vehicle speed profiles from point surveillance data. Unfortunately, the limitation of the available data precluded a comprehensive evaluation of the procedure.

All indications are that this process can be successfully used for freeway applications. It is recommended, therefore, that further validation of the process be pursued. This validation is especially important because of the possible errors introduced by variations in vehicle lengths, difficulties in estimating speed propagation rates, and lane-changing maneuvers between the detectors. Currently, variations in vehicle length are compensated for by smoothing the input data to reduce the impact on speed processing, but this also introduces errors in the computation process.

REFERENCES

1. J. C. Hancock. An Introduction to the Principles of Communications Theory. McGraw-Hill, New York, 1961.
2. H. S. Black. Modulation Theory. Van Nostrand, Princeton, NJ, 1953, chap. 4.
3. P. Brach, W. Kittelson, J. Mehra, and P. Tarnoff. Vehicle Operation, Fuel Consumption, and Emissions as Related to Highway Design and Operation. Alan M. Voorhees and Associates, McLean, VA, 2nd Interim Progress Rept., Aug. 1977; Federal Highway Administration, Aug. 1977.

Publication of this paper sponsored by Committee on Freeway Operations.

Development of Efficient Procedure for Recreating and Analyzing Traffic Flow Patterns on Urban Freeways

Wayne K. Kittelson and Philip J. Tarnoff, Alan M. Voorhees and Associates, McLean, Virginia

The purpose of this paper is to describe an analytical procedure for the synthesis and analysis of the speed profiles of typical vehicles traveling on urban freeways. Much past research has been conducted in this area in an attempt to define and describe traffic operating characteristics on urban freeways; this paper represents an extension of the past research and as such develops new analytical tools and relies on an additional data-collection effort. The paper includes a detailed discussion of the floating-automobile experiments and the data-collection procedures that were used and a description of the results that were obtained. This description includes a discussion of the variables found to be important in classifying freeway segments and a summary of the results of the data-collection effort. The paper concludes with a description of the methodology developed for using the results of the analysis to generate (in either a manual or an automatic mode) typical speed profiles for vehicles traveling on any section of urban freeway.

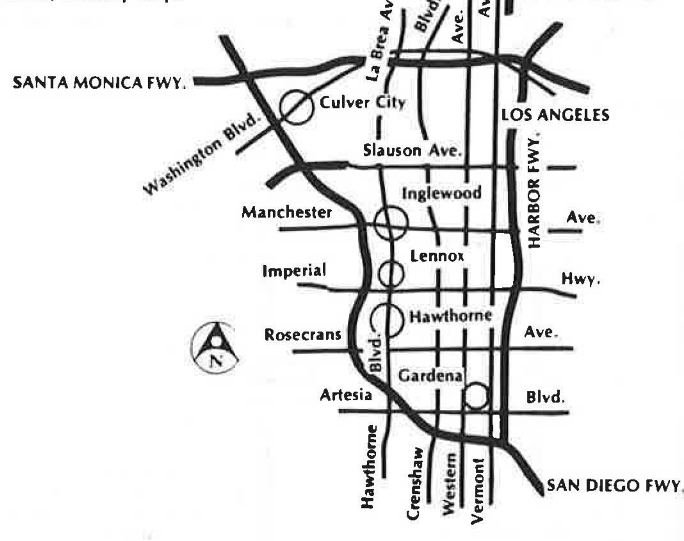
The urban freeway is a major component of the total urban transportation network. From a service point of view, urban freeways can offer relatively high average travel speeds to a very large volume of traffic; today's urban freeways carry an estimated 13 percent of the annual urban travel although they represent only 2 percent of the total urban street length available. Because of the important part that freeways play in the overall transportation network, a great deal of effort has been expended in an attempt to understand and describe freeway traffic flow patterns.

The purpose of this paper is to describe an analytical procedure for the synthesis of the speed profiles of typical vehicles traveling on an urban freeway. Much research has been conducted in this area in an attempt to define and describe traffic operating characteristics on urban freeways (1); this paper represents an extension of this past research and as such develops new analytical tools and relies on an additional data-collection effort. The paper begins with a rather detailed discussion of the floating-automobile experiments and data-collection procedures and continues with a description of the results that were obtained. This description includes a discussion of the variables found to be important in classifying freeway segments and a summary of the results of the data-collection effort. The paper concludes with a description of the methodology that was developed to use the results of the analysis to generate (in either a manual or an automatic mode) typical speed profiles of vehicles traveling on an urban freeway.

FLOATING-AUTOMOBILE EXPERIMENTS AND DATA-COLLECTION PROCEDURES

The purpose of this section is to provide a detailed description of the floating-automobile experiments used in collecting the data for the freeway analysis. Currently, only a rather limited set of urban freeway data

Figure 1. Schematic representation of 67-km (42-mile) freeway loop.



exists that contains detailed speed-profile data collected via instrumented automobile measurements. Thus, there is obviously a need for this kind of data collection, especially when it is associated with the particular geometric and environmental characteristics present on the freeway during the data-collection period. A potential alternative to the floating-automobile data-collection technique is to use data obtained from a fixed-point freeway surveillance system; however, although large amounts of fixed-point data can provide a measure of the variation of general freeway traffic conditions (such as average speed or average percentage occupancy), it is not yet known how well the fixed-point data can be manipulated to synthesize detailed speed profiles of individual vehicle trips along the freeway. It was decided, therefore, to concentrate the data-extension effort on the collection of additional detailed speed-profile data by using an instrumented automobile.

The data-extension effort was conducted on a 67-km (42-mile) instrumented freeway loop in the Los Angeles metropolitan area. This loop, shown in Figure 1 (2), consists of segments of the San Diego, Santa Monica, and Harbor freeways. These three freeways are all heavily traveled and subject to both recurring peak-period congestion and randomly occurring incident- or accident-caused congestion. Consequently, a vehicle traversing this system should encounter substantial variations in speed, making this a good test site for collecting detailed speed-profile data. Also, this freeway loop was immediately accessible to the laboratories that would be conducting the actual floating-automobile experiments, which meant that more of the available resources could be spent in the actual collection of data.

To obtain the required speed-profile data, an instrumented vehicle was used that had been equipped for the California Air Resources Board. This vehicle contained a digital data-acquisition system capable of recording on nine-track magnetic tape the instantaneous values of a number of automatically and manually entered variables at a sample rate of 1 complete scan/s. The important variables recorded during this investigation included the following:

Variable	Recording Mode
Direction of travel	Manual
Time of day (h-min-s)	Automatic
Speed (mph x 10)	Automatic
Location reference marker	Manual
General weather conditions	Manual
General traffic conditions	Manual

Before beginning the data-collection effort, a number of roadside reference markers were established for use in keeping track of the precise location of the floating automobile at any particular time. In addition, these markers were used to identify the beginning and ending points of critical sections of roadway for which data either were or were not to be included in the analysis. For example, all three transitions from one freeway to another were bounded by reference markers so that any data obtained during the transition could be deleted from the subsequent analyses. Each time that a reference marker was passed, the data-collection team would manually enter this information onto the magnetic tape by setting a thumb-wheel switch to the appropriate value.

The floating-automobile data were collected for 18 weekdays, beginning on January 20, 1977, and ending on February 14, 1977. The schedule followed is given below (c = clockwise and cc = counterclockwise).

Day No.	Direction of Travel	Starting Time	Ending Time
1	cc	6:00 a.m.	12:00 n.
2	c	6:00 a.m.	12:00 n.
3	cc	6:20 a.m.	12:20 p.m.
4	c	6:20 a.m.	12:20 p.m.
5	cc	6:40 a.m.	12:40 p.m.
6	c	6:40 a.m.	12:40 p.m.
7	cc	7:00 a.m.	1:00 p.m.
8	c	7:00 a.m.	1:00 p.m.
9	cc	7:20 a.m.	1:20 p.m.
10	c	12:00 n.	6:00 p.m.
11	cc	12:00 n.	6:00 p.m.
12	c	12:20 p.m.	6:20 p.m.
13	cc	12:20 p.m.	6:20 p.m.
14	c	12:40 p.m.	6:40 p.m.
15	cc	12:40 p.m.	6:40 p.m.
16	c	1:00 p.m.	7:00 p.m.
17	cc	1:00 p.m.	7:00 p.m.
18	c	1:20 p.m.	7:20 p.m.

As can be seen, data were collected for 6 h each day, all in one direction of travel. Every day, the direction of travel was reversed so that a total of 9 d of data were obtained for each direction. The starting and ending times for each day were also varied, so that data were

Figure 2. Mode-selection logic.

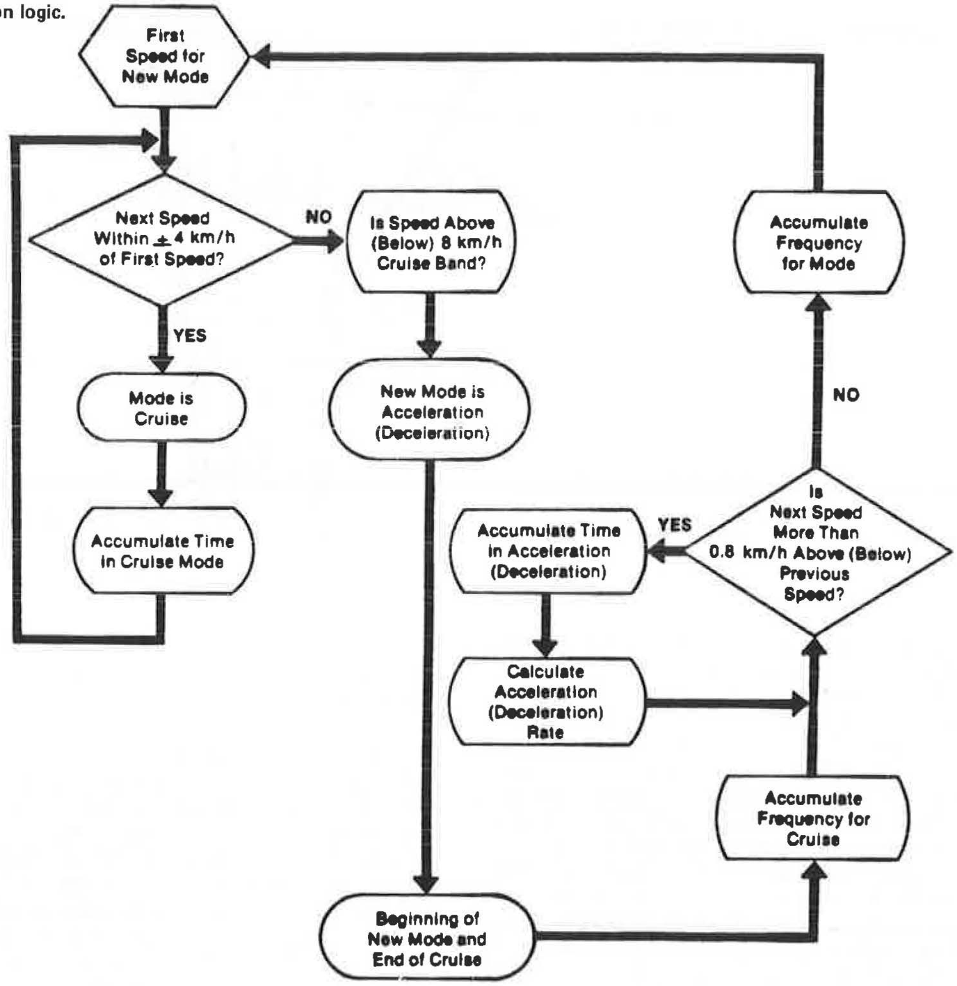
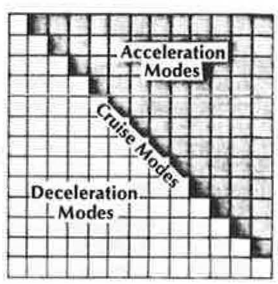


Figure 3. Matrix of vehicle modes of operation.

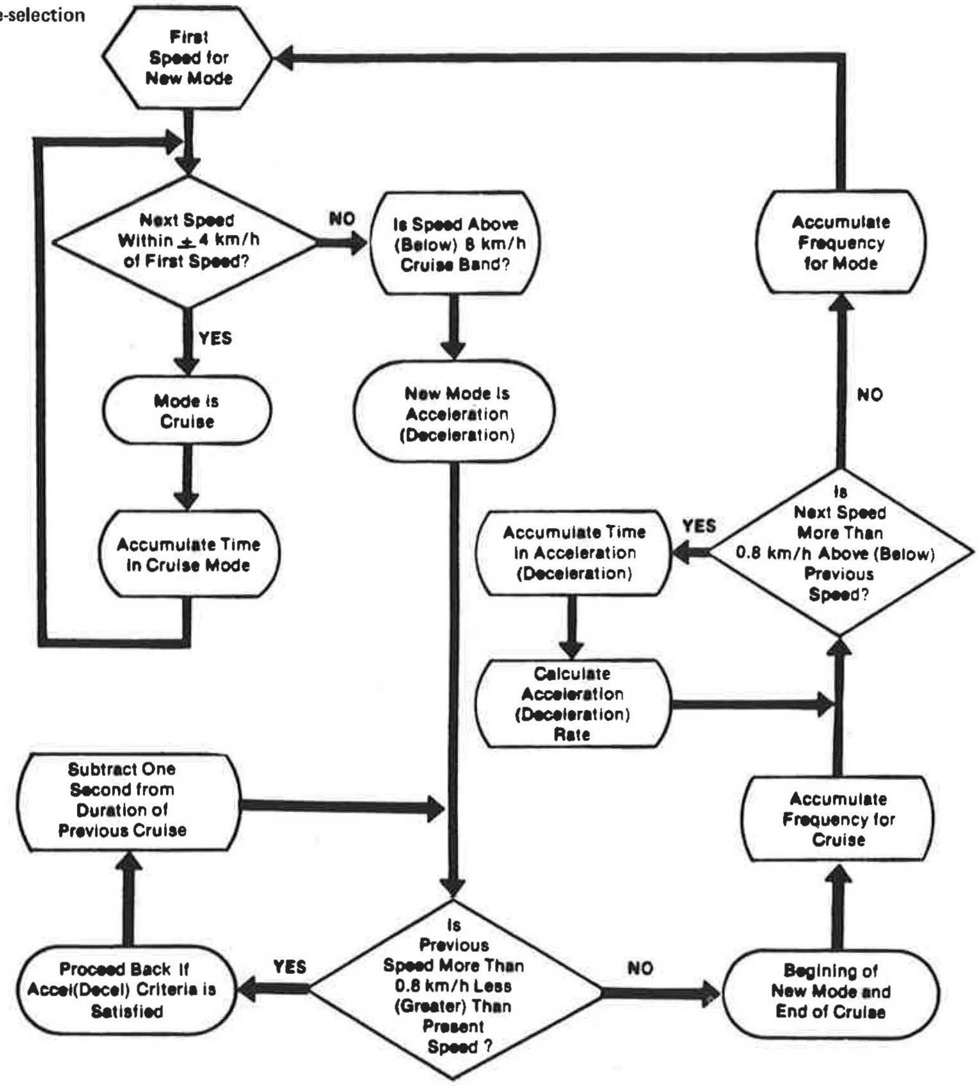


collected over a period of more than 13 h of the day. Because it took approximately 1 h to make one circuit on the freeway loop, between 5 and 7 loops were completed each day of data collection. In all, 54 loops were made in a clockwise direction and 49 in a counterclockwise direction; i.e., a little more than 100 h of data were collected during the 18-d period.

Several weeks before the floating-automobile runs were to be made, personnel of the California Department of Transportation (CalTrans) were notified of the upcoming data-collection effort and were requested to save specific outputs from the surveillance system during the days of interest (most notably, volume data on a 5-min basis). The actual data-collection effort went quite smoothly; only a few minor problems (which are to be expected) were encountered. CalTrans personnel were able to supply almost all of the data requested, and the floating-automobile runs were made without any significant problems. On completion of the data-collection effort, the second-by-second speed profiles obtained were reduced to a series of straight-line segments that could be summarized as a sequential series of cruises, accelerations, and decelerations. This summarization of the data was accomplished by using the logic described below.

Assume that the vehicle whose operation is being observed lies within a particular 8-km/h (5-mph) speed band. Successive speeds of this vehicle are checked until one is observed to be above or below that band. The last speed inside the band marks the end of the initial cruise mode and the beginning of the next mode. A cruise mode was also terminated in the original logic when successive speeds differed by more than 0.8 km/h (0.5 mph). The mode duration is then obtained from the difference in beginning and ending times for the mode. The mode following the cruise is declared to be an acceleration (or deceleration) if the cruise ended when a speed went outside the 8-km/h boundaries and differed from the previous speed by more than 0.8 km/h. For cases where the 0.8-km/h difference criterion is not met, two successive cruise modes are thus obtained (called step cruises). Once an acceleration (or deceleration) mode is established, successive speeds are checked to see whether they are greater (or less) than the preceding speed by more than 0.8 km/h. Failure to meet this criterion ends the acceleration (or deceleration)

Figure 4. Refined mode-selection logic.



mode. Figure 2 (3) shows the logical flow diagram for this methodology. The exercising of this logic produces a 13 × 13 matrix of initial speed versus final speed (in 8-km/h increments), in which both frequency of occurrence and total time duration of each mode (i.e., cruise speed, acceleration, and deceleration) are given [see Figure 3 (3)].

Analysis of the above logic showed that the initial 1 or 2 s of many acceleration and deceleration modes were included in the previous cruise mode because the data points still lay within the initial 8-km/h band. This deficiency was overcome by the use of the refined logic shown in Figure 4 (3). As an acceleration or deceleration is detected, the logic regresses backward, checking previous speeds, to find the actual beginning differences in successive speeds that satisfy the 0.8-km/h acceleration or deceleration criterion. The time originally thought to be spent in the preceding cruise mode is subtracted from this mode, and the same amount of time is added to the acceleration or deceleration mode being processed. The overall effect is to yield slightly shorter cruise modes and slightly longer acceleration and deceleration modes.

By using the mode-frequency and mode-duration matrices, two other useful matrices can also be computed. An average-time-in-mode matrix is computed by dividing each entry in the mode-duration matrix by the corresponding entry in the mode-frequency matrix. Also, an initial speed versus final speed transition-probability matrix can be computed by dividing each row entry (initial speed) of the mode-frequency matrix by the sum of all frequencies for that row. Note that each row of the transition-probability matrix sums to unity, because all different final speeds are possible for each initial speed. A sequence of modes can thus be generated by applying random-number sampling techniques to the transition-probability matrix.

DESCRIPTION OF RESULTS

The purpose of this section is to describe the results of the freeway analysis and the development of the mode-frequency and mode-duration matrices for use in estimating vehicle speed profiles along any section of freeway possessing a given set of geometric and volume characteristics. For ease of presentation, this section is divided into two subsections. The first subsection describes the freeway classification scheme that was used in preparing the experimental design matrix. The second subsection details the procedures that were used in preparing the mode matrices, presents examples of the mode-frequency and mode-duration matrices that were developed for the experimental design cells and shows how other types of information can be obtained from the mode matrices.

Freeway Classification Scheme

This subsection describes the generalized highway-design-features classification scheme that was used in the analysis of the relationships between freeway design and vehicle operating characteristics. The variables that were incorporated into the experimental design matrix were chosen on the basis of their apparent importance as reflected in a literature review conducted earlier in the study. The two principal variables of classification that were used are the interchange spacing and volume: capacity (v:c) ratio. The v:c ratio is an obviously important variable that has been shown through numerous studies to have a very important effect on the quality of flow along an urban freeway (4, 5). Interchange spacing is introduced into the classification scheme because of the major effect that freeway interchange maneuvers are believed to have on vehicle operating characteristics (4, 5, 6, 7, 8).

The v:c ratio is broken into a series of four intervals (<0.4, 0.4-0.6, 0.6-0.8, >0.8), each of which represents a distinctly different set of roadway operating conditions. At the outset of this investigation, an attempt was made to introduce the concept of a composite v:c ratio. This composite ratio was to be determined by computing v:c ratios for individual homogeneous subsections of freeway by using a weighted (by subsection length) average computed for the series of subsections under consideration. However, this technique proved to be too sophisticated given the amount of information available and the relatively large intervals that were being used for the v:c ratio variable. Consequently, a v:c ratio was calculated for each homogeneous subsection of the freeway under investigation. This technique proved to be much easier to apply and seemed to have no adverse impact on the results obtained.

With regard to the actual capacity computations, it was originally thought that this information would be readily available from existing files maintained by CalTrans personnel. However, subsequent conversations showed that subsection capacities are not part of the information normally acquired and maintained by CalTrans for this particular freeway system. However, CalTrans normally uses 1800 vehicles/h/lane as the effective capacity of any section of the loop. This procedure can be justified by the observations that there are no significant grades anywhere within the loop (the most severe is a 3 percent grade on the Harbor Freeway that runs for less than 0.8 km) and that truck traffic along these freeways remains fairly constant over all locations. More detailed capacity calculations were carried out for the Santa Monica Freeway and indicated that 1800 vehicles/h/lane is, in fact, a good approximation to the effective freeway capacity. Therefore, this investigation assumed the effective capacity of any freeway subsection located in the loop to be 1800 vehicles/h/lane.

The interchange-spacing variable is stratified into three intervals [<1.2 km (0.75 mile), 1.2-2.4 km (0.75-1.5 miles), and >2.4 km (1.5 miles)]. (A fourth stratification is discussed below.) This stratification reflects a general categorization of interchange spacing that previous studies have found to represent the primary dividing points for the impact that interchange spacing has on vehicle operation. The actual intervals identified in the experimental design matrix are somewhat lower than those identified in the existing literature because of the need in this investigation for a slightly different definition of interchange spacing. For this project, interchange spacing was defined as the distance between any two consecutive entrance or exit ramps not located in a major interchange area. Thus, an interchange spacing of, say, 1.6 km (1.0 mile) can be bounded by two exit ramps, two entrance ramps, or an exit and an entrance ramp. Furthermore, these exit-entrance ramps need not be located near a major interchange but can be anywhere between. It was felt that this definition of interchange spacing was more realistic than any other from the driver's point of view, because access and egress points can have a disturbing influence on local traffic flow. However, because this definition resulted in generally smaller interchange spacings, it was necessary to adjust the intervals with which this project was concerned to reflect the results of previous studies.

A significant amount of travel also occurs within interchange areas, and so a fourth category—within interchanges—has been included in the interchange-spacing variable. By definition, this category includes all travel on an urban freeway that is not included in any of the other three categories of the interchange-spacing variable. Because this category includes only travel within a major interchange, it is much more likely that vehicles would be encountering traffic-flow disturbances caused by entering or exiting vehicles. If this hypothesis is correct, then the results of the study should indicate lower speeds and more accelerations and decelerations for this category of the interchange-spacing variable than for others.

Freeway Mode Matrices

After completion of the data-collection effort, the following data items were available:

1. A second-by-second speed history of the instrumented floating automobile during all 103 circuits around the 67-km freeway loop,
2. A summary of the second-by-second speed history of the floating automobile by mode type (cruise, acceleration, and deceleration) and by mode duration (in seconds) for each of the 103 circuits, and
3. A summary (every 5 min) of the traffic volume sensed at each detector station in the surveillance system throughout the data-collection period.

In addition, schematic diagrams were obtained from CalTrans personnel identifying the milepost location of every on-ramp, off-ramp, and bridge along the freeway system. With these data, it was possible to stratify the collected data according to the experimental design matrix described above.

To begin, the schematic diagrams of the freeway system were used to identify the location of every on-ramp and off-ramp passed by the floating automobile. By using the second-by-second speed histories of the floating automobile, it was possible to accumulate the distance traveled for each run so as to always know where in space it was at any point in time. Therefore, the data obtained from the floating automobile could be stratified according to the interchange-spacing variable by associating its space and time location with the milepost location of all ramps that it passed.

In a similar manner, the flow rate (and hence the v:c ratio) in the immediate vicinity of the floating automobile was determined by associating the vehicle's space and time locations with the appropriate 5-min volume data obtained from the surveillance system.

After the guidelines for stratifying the floating-automobile data into the appropriate cells of the experimental design matrix were established, a computerized procedure was developed for performing this stratification and accumulating the data. The results of this pro-

Figure 5. Matrices for between-interchange data: (a) mode frequency and (b) mode duration in seconds.

(a)

Initial Speed (MPH)	Final Speed (MPH)												
	0	5	10	15	20	25	30	35	40	45	50	55	60
0	26	15	6	5	1	1	0	0	0	0	0	0	0
5	12	37	8	5	1	2	0	0	0	0	0	0	0
10	8	11	34	7	3	2	1	1	0	0	0	0	0
15	4	6	6	33	9	5	0	0	0	0	0	0	0
20	0	3	2	8	27	11	0	1	0	0	0	0	0
25	1	0	4	3	7	32	11	2	0	0	0	0	0
30	2	0	1	1	0	3	33	11	6	1	0	0	0
35	0	0	0	1	3	2	11	41	14	4	0	0	0
40	0	0	0	0	1	1	4	14	47	12	4	0	0
45	0	0	0	0	1	1	3	1	13	77	19	10	0
50	0	0	0	0	0	2	2	1	9	25	246	50	5
55	0	0	0	0	0	0	0	1	1	12	71	304	11
60	0	0	0	0	0	0	0	0	0	1	4	14	61

(b)

Initial Speed (MPH)	Final Speed (MPH)												
	0	5	10	15	20	25	30	35	40	45	50	55	60
0	223	68	37	47	12	0	0	0	0	0	0	0	0
5	51	315	37	38	16	22	0	0	0	0	0	0	0
10	58	39	257	38	23	15	14	11	0	0	0	0	0
15	31	37	23	248	39	38	0	0	0	0	0	0	0
20	0	17	8	30	211	48	0	9	0	0	0	0	0
25	15	0	32	21	21	216	42	12	0	0	0	0	0
30	29	0	6	9	0	12	267	41	33	0	0	0	0
35	0	0	0	9	19	11	30	406	58	28	0	0	0
40	0	0	0	0	13	9	23	73	413	71	16	0	0
45	0	0	0	0	12	11	27	7	80	1002	95	86	0
50	0	0	0	0	0	20	20	10	74	122	5191	241	46
55	0	0	0	0	0	0	0	9	13	77	343	8807	44
60	0	0	0	0	0	0	0	0	0	11	31	53	1306

Notes: 1 mph = 1.6 km/h.
Interchange spacing = 1.21-2.41 km (0.75-1.5 miles) and v:c ratio = 0.6-0.8.

cess were output as mode-frequency and mode-duration matrices, one pair for each cell of the experimental design matrix (see Figure 5). These matrices show several interesting facts. First of all, there were no data available for interchange spacings greater than 2.4 km (1.5 miles); to fill these matrices, therefore, it would be necessary to obtain data from other freeway systems on which interchange spacing is, at least occasionally, greater than 2.4 km.

The available data were spread among the remaining cells of the experimental design matrix, but the amount of data going to any particular cell appeared to be highly dependent on the value of the v:c ratio that helps to define that cell. In particular, the distribution of the data according to the value of the v:c ratio was as follows:

	Amount of Available Data Collected	
v:c Ratio	Hours	Percentage of Total
<0.4	0.99	2
0.4-0.6	8.84	13
0.6-0.8	21.69	33
>0.8	34.45	52
Total	65.97	100

This distribution of the data reflects an attempt to emphasize vehicle operating characteristics at relatively high-volume conditions. Above a v:c ratio of 0.6, vehicle interactions play an increasingly significant role in defining speed profiles; as a result, there are proportionally more accelerations and decelerations than occur at lower volumes. This was clearly indicated by the distribution of data within each of the individual mode matrices: At the lower v:c ratios, most of the data were clustered in the lower right-hand corner of the matrix, indicating relatively high travel speeds, and most of the data were contained within the diagonal cells of the matrix, indicating a high percentage of cruises. Conversely, at v:c ratios greater than 0.8, a substantial part of the data was found in the upper left-hand corner of the matrix, indicating slow speeds, and a higher percentage of the data was located outside of the diagonal cells, suggesting more accelerations and decelerations. Thus, a cursory review of the data reflects the same relationships between the variables as have been observed in past studies.

The 66 h of data included in these mode matrices represent a significant reduction over the 110 h of data originally collected. This discrepancy was brought about by the need to discard a large amount of the original data because: (a) some of the data was collected while the floating automobile was in transition from one freeway to another and (b) human errors in identifying reference-marker locations meant that the precise location of the floating automobile could not always be determined. Thus, approximately 40 percent of the original data base was deleted from the analysis.

Much of the data (nearly 31 percent) was collected within interchange areas. At v:c ratios lower than 0.8, there appeared to be a significant difference between the data collected within interchanges and that collected outside of interchanges: The within-interchange data reflected more speed fluctuations and lower overall speeds. This observation is intuitively appealing and suggests that it is appropriate to maintain this stratification of the data at these v:c ratios. However, at the higher v:c levels, there seemed to be very little difference between the data collected within interchanges and those collected outside of interchanges. Both data sets reflected a relatively low average travel speed and a large number of accelerations and decelerations. This observation implies that, at high v:c levels, intense vehicle interactions are taking place even outside of the interchange area and these interactions dominate in determining the speed profile of a vehicle. Therefore, it may be appropriate to combine data collected outside of interchanges with those collected within interchanges whenever the v:c ratio is greater than 0.8.

The stratification of the data according to interchange spacing has also yielded some interesting results. In particular, the mode matrices developed from data collected at interchange spacing of less than 1.2 km appear to be very similar to those developed from data collected within interchanges. This observation suggests that the effect of an interchange area on traffic flow patterns extends for distances of at least 1.2 km upstream and downstream of the physical location of the interchanges. It also suggests that it would be appropriate to combine all data collected within 1.2 km of any interchange into the within-interchange category. For interchange spacings greater than 1.2 km, substantial differences appear when the mode matrices are compared with those obtained from within-interchange data. (These differences are characterized by fewer accelerations and decelerations and a higher average travel speed.) Therefore, this particular stratification of the data seems to be necessary.

The representation of the speed-profile data as mode matrices allows detailed analyses of traffic-flow patterns to be made relatively quickly. Explicitly defined within these matrices is information such as the following:

1. Number of stops per vehicle;
2. Average duration of stops;
3. Percentage distribution of time spent (a) cruising at constant speed, (b) accelerating, (c) decelerating, and (d) idling;
4. Frequency and duration of nonstop speed-change cycles of various magnitudes; and
5. Attempted cruise speeds.

In addition, the information contained in these matrices is potentially applicable in a number of other areas. For example, by analyzing the distribution of cruise speeds in the mode-frequency matrix, it may be possible to develop a procedure whereby data collected by using a radar speed meter can be associated with a particular set of mode matrices; such a procedure would result in significant reductions in data-collection costs and allow much more detailed traffic-flow analyses. Data from mode matrices can also be used to estimate fuel consumption and exhaust emissions along any section of an urban freeway because they explicitly define not only the percentage of total time spent in each travel mode but also the amount of time that the average vehicle spends in the mode (from the average-time-in-mode matrix). Therefore, the use of mode matrices represents a significant advancement in the art of manipulating existing data sources so as to more easily reveal the true operating characteristics of the freeway.

METHODOLOGY FOR USING RESULTS

Although it was not possible to fill every cell of the experimental design matrix by using the data base collected on the 67-km freeway loop in Los Angeles, those cells that were filled represent a significant advancement in the state of the art of estimating vehicular speed profiles along sections of homogeneous urban freeways. The purpose of this section is to describe the methodology that has been developed for using these mode matrices to generate typical speed profiles along any segment of an existing or a planned urban freeway. The procedure will be described as a manual process, but it can easily be adapted to automatic data-processing techniques. The results would, of course, be a much more efficient use of resources when examining the impacts of alternative freeway configurations on vehicle speed profiles.

The foundation of the methodology is the mode matrices. As a result of the analyses described above, the potential user will initially have in hand a set of mode-frequency and mode-duration matrices, each pair of which can be associated with a particular cell in the experimental design matrix. At the outset, then, the user must segment the freeway section of interest according to the two variables that define the experimental design matrix (i.e., the v:c ratio and the interchange spacing). When this task is accomplished, the freeway section will have been converted into a set of segments, each of which is homogeneous with respect to these two variables. Furthermore, each of these segments can then be associated with the particular pair of mode-frequency and mode-duration matrices generated by using floating-automobile data collected on freeway sections possessing characteristics similar to those of the segment(s) being analyzed.

Having identified the appropriate mode-frequency and mode-duration matrices for each segment of the freeway under investigation, the user can now prepare an initial-speed versus final-speed transition-probability matrix by dividing each row entry (initial speed) of the mode-frequency matrix by the sum of all frequencies for that row. Similarly, an average-time-in-mode matrix can be computed by dividing each entry in the mode-duration matrix by the corresponding entries in the mode-frequency matrix. Together, these two matrices will allow the user to generate a typical speed profile for the freeway segment that is being investigated. This task can be accomplished through the following two steps.

First of all, a sequence of modes can be generated by applying random-number sampling techniques to the transition-probability matrix. For example, on entering the transition-probability matrix at any given initital speed, a final speed for the (as yet undefined) mode can be stochastically determined by applying a Monte Carlo sampling technique to the universe of potential final speeds. Sometimes called the crude Monte Carlo technique, this approach begins by generating some uniformly distributed, random decimal numbers, then sets each of them equal to the cumulative distribution function, and finally solves the equations individually to obtain random observations from the probability distribution. More specifically, for the purposes of this paper, the probability of choosing any particular final speed is defined as the value in the transition-probability matrix of the cell that is associated with that final speed and the original initial speed. For example, consider the transition-probability matrix shown below (1 km/h = 0.6 mph) and assume that the beginning initial speed is 72 km/h (45 mph).

Initial Speed (km/h)	Final Speed (km/h) 48	56	64	72	80	89
48	0.40	0.30	0.20	0.10	0.00	0.00
56	0.20	0.35	0.20	0.15	0.05	0.05
64	0.05	0.10	0.45	0.15	0.15	0.10
72	0.05	0.10	0.15	0.42	0.18	0.10
80	0.00	0.05	0.10	0.16	0.54	0.15
89	0.05	0.05	0.10	0.12	0.20	0.48

Therefore, from this transition-probability matrix, the universe of potential final speeds and the probability of achieving a final speed of X mph [$P_f(X)$] associated with achieving these final speeds can be defined as follows:

X (km/h)	P_f (X)
48	0.05
56	0.10
64	0.15
72	0.42
80	0.18
89	0.10
Total	1.00

After the initial and final speeds for the first mode are defined, the second step in the methodology is to determine the duration of the mode. This can be accomplished simply by looking up the appropriate value in the average-time-in-mode matrix (defined by the initial and final speeds obtained in the first step). At this point in the analysis, the user will know not only the initial mode for the hypothetical vehicle (cruise, acceleration, or deceleration), but also how long the hypothetical vehicle maintained that mode. By integrating the instantaneous speed of the vehicle over this time period, the user can also estimate the distance traveled during this first mode. At this point in the analysis, then, the user will know exactly where the hypothetical vehicle is on the road segment being analyzed, how long it took to get to that point, and what mode of travel was used. The final speed for the mode just analyzed can then be treated as the initial speed for the next (as yet undefined) travel mode, and the entire process can be repeated. This procedure will continue until the user's travel-distance calculations indicate that the hypothetical vehicle has passed the boundary of the segment being analyzed. If this freeway segment is bounded by another freeway segment possessing different v:c or interchange-spacing characteristics, then the entire process can be continued by using a new set of mode-frequency and mode-duration matrices taken from the appropriate cell of the experimental design matrix. If, however, this is the last freeway segment to be analyzed, then the procedure is complete and the user should have in hand a comprehensive speed history for the hypothetical vehicle as it traveled through the freeway segments that were analyzed.

In order that the final results be reasonable, it is desirable that the procedure described above be carried out for a number of hypothetical vehicles and the results be averaged into one typical speed profile. This is necessary because of the possibility of generating an atypical speed profile for any single vehicle as a result of using a table of random numbers to create the profile. By generating a number of profiles and filtering out the obviously bad ones, a more realistic speed profile will be ensured.

The methodology described above is one that is simple to use and, although it is presented here as a manual exercise, it is a procedure that could easily be automated if so desired. Such automation would certainly improve the efficiency with which the user could investigate the impact of alternative freeway configurations on vehicle speed profiles. In any case, the primary limitation of the methodology lies in the number of mode matrices that are currently available for use in the analysis, together with the quantity of data that was used to fill these matrices. As additional work is completed and the current voids in the data base are filled, the power and usefulness of this methodology will increase. Therefore, high priority should be given to the task of extending the existing data base and maintaining its relevance to current traffic flow characteristics. Any such effort would certainly have a high payoff in usefulness to existing and potential users of the methodology.

REFERENCES

1. P. Brach, W. Kittelson, J. Mehra, and P. Tarnoff. Vehicle Operation, Fuel Consumption, and Emissions as Related to Highway Design and Operation. Alan M. Voorhees and Associates, McLean, VA, 2nd Interim Progress Rept.; Federal Highway Administration, Aug. 1977.
2. Los Angeles Freeway Surveillance and Control Project. California Business and Transportation Agency, Annual Rept., Feb. 1973.
3. J. Harkins and others. Vehicle Operations Survey. Scott Research Laboratories, San Bernardino, CA; Coordinating Research Council, New York; Office of Air Programs, Environmental Protection Agency, Vol. 1, Dec. 17, 1971; NTIS, Springfield, VA, PB 210 640.
4. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
5. D. L. Gerlough and M. J. Huber. Traffic Flow Theory: A Monograph. TRB, Special Rept. 165, 1975.
6. J. E. Baerwald, ed. Transportation and Traffic Engineering Handbook. ITE, 1976.
7. A Policy of Urban Highways and Arterial Streets. AASHO, 1973.
8. D. O. Covault and R. C. Kirk. Influence of Off-Ramp Spacing on Traffic Flow Characteristics of Atlanta Freeway and Arterial System. HRB Highway Research Record 59, 1964, pp. 39-52.

Publication of this paper sponsored by Committee on Freeway Operations.

Application of Freeway-Corridor Assignment and Control Model

Brian L. Allen and Said M. Easa, Department of Civil Engineering, McMaster University
E. R. Case, Systems Research and Development Branch, Ontario Ministry of Transportation and Communications

An improved traffic assignment and control model—CORCON—has been applied to a real freeway corridor. This model is capable of predicting traffic behavior in a freeway corridor by assigning time-varying origin-destination demand to the freeway mainline and the surrounding network streets. The impact of queueing behavior on the selection of minimum-cost travel paths is incorporated by using a flexible traffic-diversion model. Because all or a portion of the traffic may be diverted from a particular queueing path, the effects of freeway-entrance-ramp control on adjacent roadway systems can be investigated. The model was calibrated and validated by using 15-min volume, travel time, and queueing data collected during the 2-h morning peak period on a section of the Queen Elizabeth Way freeway corridor for typical before-control and after-control periods. The calibration resulted in an overall correspondence of 3 to 5 percent, and the validation process predicted average traffic behavior within 5 to 10 percent of actual behavior. The testing was considered very successful because the model was shown to be capable of predicting with reasonable accuracy flows, queues, and travel times on a freeway and adjacent street network.

Over the past 10 to 15 years, freeway surveillance and control has evolved from being considered as a traffic engineer's toy to being recognized as a useful urban traffic management tool.

The inexorable growth of urban traffic demand forced investigators, as long as 10 years ago, to acknowledge that attention must be given to both the freeway and the adjacent street networks if the maximum utilization of existing facilities is to be attained for the entire freeway corridor. However, despite the efficacy of that acknowledgment, the analytic techniques available at that time were not capable of predicting and assessing traffic operations on the freeway and the street system simultaneously. Therefore, the impact of freeway control strategies on adjacent street networks (and vice versa) could not be easily forecast.

In the past several years, there has been considerable work in the development of the tools necessary to correct that problem. The methods proposed have tended to fall in three major categories:

1. Methods that deal with corridor operations and characteristics in a gross manner by ignoring detail in favor of analytic simplicity and efficiency,
2. More complex methods that treat the freeway and the street network in greater detail and thereby (presumably) achieve greater accuracy, and
3. Traditional transportation planning models capable of assigning traffic flows to network links.

Because the models representative of the third category are well known and have had little practical application in the freeway operations field, only the first two categories will be dealt with here. A method typical of the first category was presented by Allen and Newell (1, 2, 3) who suggested that operations on the entire freeway corridor network could be represented (for preliminary planning purposes) by only two routes—one route for the freeway and one route for all other streets. That work was further developed and applied to a real corridor by Liew and Allen (4, 5), who showed that the very simple model could be useful in preliminary development and planning of control strategies.

More extensive work on methods in the second category has been regularly reported by May and his coworkers (6, 7, 8). That work has focused mainly on developing an optimization package that treats the freeway and street network in considerable detail, in fact, in sufficient detail to permit reasonably comprehensive assessments of optimal design and control improvements to selected portions of the freeway corridor. The results of this work have received considerable attention and have been adopted for regular use by many operating agencies.

Parallel to the above methods, Yagar (9, 10, 11) has developed and tested a procedure that conceptually lies somewhere between them. This model uses a relatively conventional link-node configuration for network representation but includes considerably more corridor detail than models in the first category. The basic structure of this model has been improved by Easa and Allen (12, 13, 14, 15), who felt that it had great potential for evaluating the effect of control on freeway corridor operations and recommended that it be fully tested by application to a real corridor.

This evaluation was undertaken (16) on the recently established Queen Elizabeth Way (QEW) freeway surveillance and control system demonstration project study site (which is described by Case and Williams in the following paper in this Record). Data describing traffic volumes, link travel times, and queue lengths were collected for periods before and after implementation of freeway entrance-ramp control along a 9-km (5.5-mile) section of the QEW corridor. These data were used as the basis for a before-and-after test of the predictive capability of the model.

The results of the calibration and validation testing are presented in this paper. The basic model structure (including a brief description of its more novel features) is contained in the following section. The third section contains an introduction to the study site, examples of the data collected, and a brief discussion of the application procedures. Details of the results obtained from the calibration and validation process are reported in the fourth section. The final section contains a brief accounting of the major conclusions and recommendations stemming from the project.

MODEL

The freeway CORridor assignment and CONtrol model—CORCON—divides the peak period into equal lengths of homogeneous demand called time slices. The demand in each time slice and the queued demand of the previous time slice are assigned to the network by using the principle of minimizing individual travel cost (time). The flow versus travel time relationship for each link is an increasing function and is approximated by three linear components. Network features are represented by a simplified link-node method of representation. The model incorporates a procedure for turn prohibitions and overlapping minimum paths to avoid illogical paths within the corridor, a traffic diversion procedure, and a method for calculating turning volumes without the need to provide turning links.

During the development phase, particular attention was given to establishing a simple and efficient method of network representation, avoiding the occurrence of illogical paths in the minimum path assignment algorithm, minimizing the input data requirements, and considering traffic diversion and queueing characteristics. After a careful review of a first generation of the model (13), several improvements were made; a detailed description of the final version of operating characteristics and input-output formats is given elsewhere (14, 15). Therefore, only a brief summary of the key elements is presented here.

Any operational model of this type includes three principal steps:

1. The link-node representation,
2. The determination of the minimum paths, and
3. The traffic-demand assignment to those minimum paths subject to the corridor controls implemented.

These steps, as performed in CORCON, are outlined below.

Link-Node Representation

An essential task in the modeling process is to adequately represent the corridor network in terms of a set of links and nodes. It is obviously desirable to use the most efficient representation possible, either to reduce network coding effort or to permit investigation of larger networks. CORCON incorporates a new method of network representation that allows more than one directional roadway link to have common upstream and downstream nodes. This feature is particularly advantageous in simplifying the representation of complex merging, weaving, intersection, and interchange network sections (14, 15).

Determination of Minimum Paths

After representing the corridor by a complete set of links and nodes, the minimum paths from each origin in the corridor to all destinations are obtained. However, existing minimum-path algorithms are not capable of use with the simplified link-node representation adopted. Consequently, a completely new minimum-path algorithm that has provision for turn prohibitions was developed. The new algorithm allows the use of turn prohibitions without additional coding and simplifies the link-node representation by allowing more than one link in the same direction to have common upstream and downstream nodes. Unlike existing algorithms, the new algorithm also allows up to four costs to be reserved at each turn-prohibition node.

The new algorithm does not require direct input of network turn prohibitions because these are identified automatically by CORCON by using information regarding available upstream feeder links.

Traffic-Demand Assignment

Once the minimum paths have been determined, the traffic demand is assigned to those paths by using the principle of minimizing individual travel cost. A traffic-diversion procedure is incorporated in the assignment algorithm. This procedure first assigns demand to the minimum path. If this path contains any queueing links, a certain proportion of that demand is diverted to a non-queueing alternative minimum path, if such exists. The proportion (percentage) of traffic diverted is calculated according to the travel characteristics of those paths as follows:

$$P = 100/[1 + (\Delta T/Q)]^r \qquad (1)$$

where

- P = percentage of traffic diverted,
- ΔP = difference in travel cost (time) between non-queueing and queueing minimum paths respectively,
- Q = queueing cost (time) along queueing minimum path, and
- r = diversion parameter.

The diversion parameter (r) is used as a calibration control for the diversion characteristics. When $r = 0$, total diversion will occur. When $r = \infty$, no diversion will occur. A respective proportion of the traffic will divert when $0 < r < \infty$.

APPLICATION

The CORCON model was applied to an existing freeway corridor on which entrance-ramp metering was implemented in the summer of 1975. Specifically, CORCON was used to predict traffic operating characteristics in the corridor network for peak-period traffic demand before (1975) and after (1976) the implementation of the freeway-access control strategy.

Study Corridor

The study corridor is located southwest of Toronto in the city of Mississauga as shown in Figure 1. The more detailed view in Figure 2 shows that the major east-west routes in the corridor are the six-lane QEW and two arterial highways, Dundas Street (Highway 5) and Lakeshore Road (Highway 2). Traffic operations in the eastbound direction only were considered for these roadways. Major north-south crossing streets include Southdown Road-Erin Mills Parkway, Mississauga Road, Hurontario Street (Highway 10), and Cawthra Road. Traffic operations on these roadways were considered for both directions with the exception of Mississauga Road north (the northbound direction was not considered necessary because it is not used by the eastbound traffic under consideration). The corridor also included all freeway service roads within the study area.

The three-lane eastbound portion of QEW considered in this study extends approximately 9 km (5.5 miles) from the mainline origin west of Southdown Road to the mainline destination east of Cawthra Road and contains five entrance ramps and three exit ramps. This portion currently experiences congestion and queues during the morning peak period (7-9 a.m.) because of a bottleneck section immediately east of Highway 10. In addition, the corridor includes arterial and other street sections that have existing or potential operating problems because of the entrance-ramp control strategies implemented at the five QEW entrance ramps.

After the corridor configuration had been investigated, it was represented in traditional terms of links and nodes to form the necessary model input as shown in Figure 3. Each merging section on QEW was represented by two dummy links, e.g., 28-29, that had common upstream and downstream nodes. These links were used to regulate the merging capacities of both freeway and entrance-ramp approaches. The two links 35-40 represent exit ramps at the Highway 10 interchange. The two entrance ramps at this interchange required special treatment. As shown in Figure 4, the entrance ramps from northbound and southbound Highway 10 combine together to form one merging section with the freeway and each ramp has its own metering control system. Consequently, the two links 50-36 are used as dummy links to set the metering rates of these ramps and link 36-37 (lowest) is used to represent the merging section of the combined ramp traffic.

The corridor has 93 turn prohibitions; this includes U-turns located at intersections, interchanges, merging sections, and origin-destination nodes. (There are 17 turn prohibitions at the Highway 10 interchange alone.)

Data Collection and Reduction

After the boundaries and configuration of the study corridor had been established, the following data for model calibration and validation were collected:

1. Freeway-user origin-destination (O-D) demands (15-min basis),
2. Link volumes and queues (15-min basis), and
3. Network characteristics (physical and control).

To establish the demand characteristics of the entire corridor, the O-D demand at freeway entrance and exit ramps was obtained by conducting a license plate survey at each entrance and exit ramp in the corridor. For each 15-min time slice in the 2-h peak period, the license plate data were analyzed by first determining the origin location within the corridor for each vehicle surveyed and then establishing the number of vehicles from each origin area. The origin locations were ascertained from addresses in motor-vehicle registration files, each address was located within the study area, and the number of trips from each origin was computed to reflect the observed volume at each survey station. The destination locations of vehicles recorded at each entrance station and of those vehicles originating on the mainline were determined for each time slice by using a conventional license plate matching routine. Finally, the 15-min demand volumes emanating from a particular entrance station were assigned to the origin locations of that station in direct proportion to the total demand at each exit station. In addition, the observed link volumes and queue lengths were used to construct the O-D demand of freeway nonusers.

Data describing major corridor characteristics include principally the flow versus travel time relationships, capacity information, and network turn prohibitions. Standard travel time and delay runs were conducted for all corridor sections simultaneous with the volume counts. In this way, the correspondence between volumes and link travel times was obtained and the flow versus travel time relationship was constructed. Figure 5 shows a typical sample of the derived relationship for a particular corridor link. For use in CORCON, the relationship is approximated by three linear components of costs and capacities (15).

Capacities of corridor links and intersection approaches were calculated by using the Highway Capacity Manual procedure (17). The total merging capacity for each merging section was estimated from traffic volumes measured by the loop-detector surveillance system located along the freeway. Capacities of the controlled entrance ramps (i.e., metering rates) and queue lengths on both the freeway and the ramps were obtained from the Ontario Ministry of Transportation and Communications. Information on all turn prohibitions in the corridor network was also collected. Although much of the necessary data was collected for both the before-control (1975) and the after-control (1976) periods, the

Figure 1. Location of study corridor.

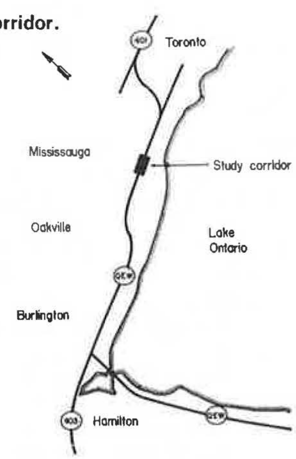
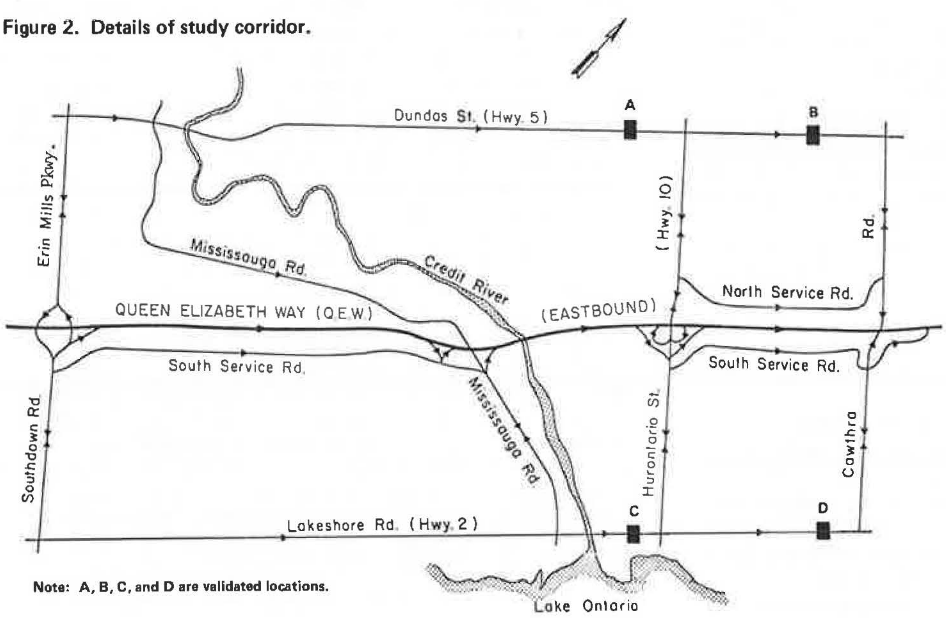


Figure 2. Details of study corridor.



Note: A, B, C, and D are validated locations.

after-control data were considerably more comprehensive. As a result, the CORCON model was first applied to the corridor and calibrated for the 1976 conditions.

CALIBRATION PROCEDURE AND VALIDATION RESULTS

The diversion parameter and the O-D demand patterns were calibrated for the after-control period, and the CORCON procedure was validated by comparing the predicted and the observed conditions for the before-control

Figure 3. Link-node representation of study corridor.

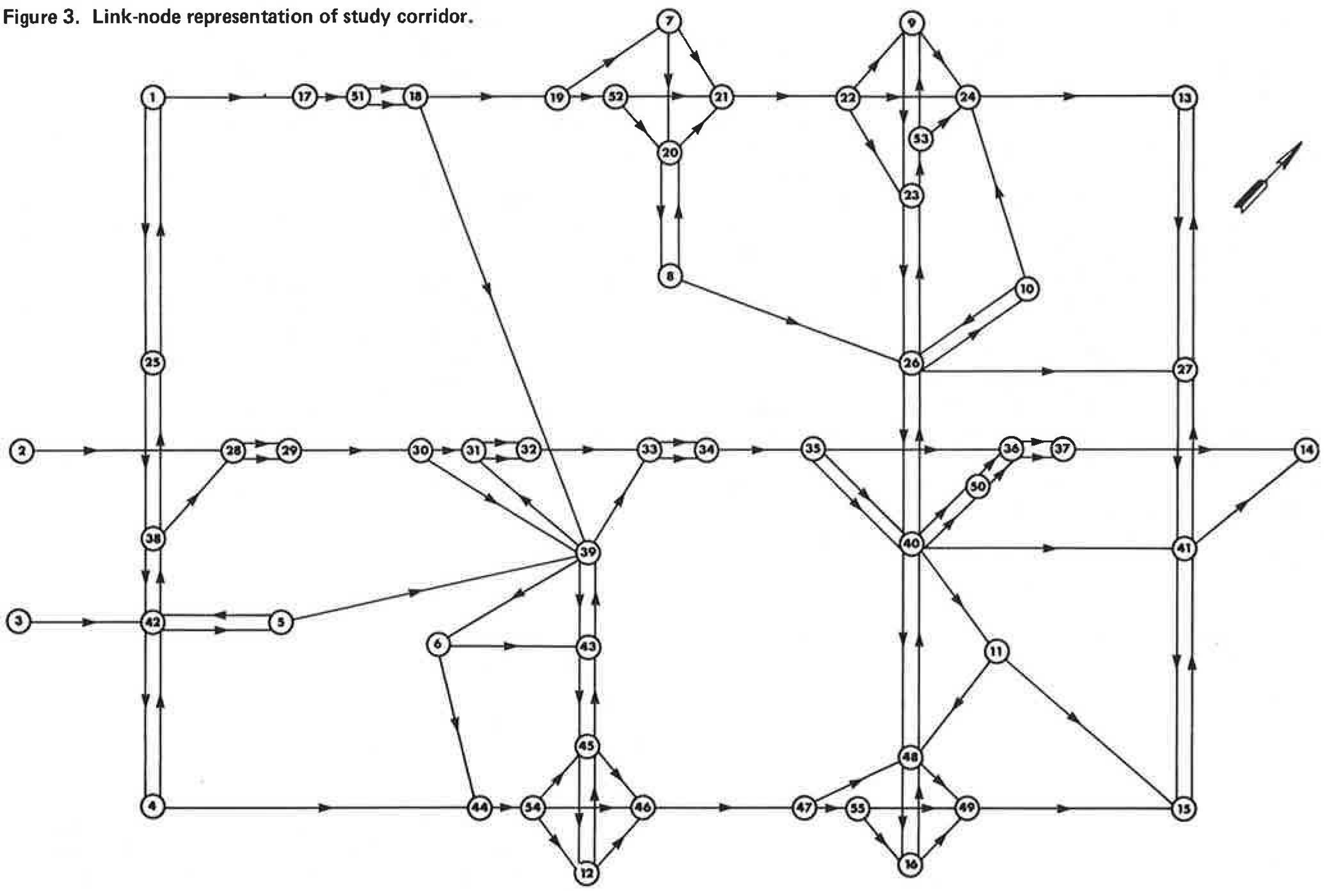
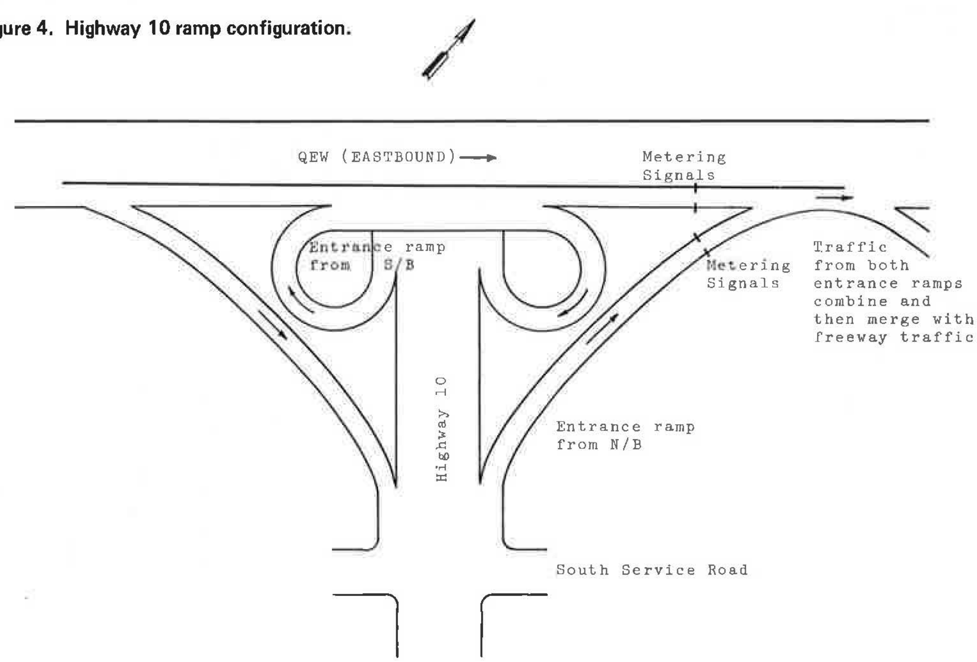


Figure 4. Highway 10 ramp configuration.



period. The calibration therefore included the preparation of a trip table that combined both freeway users and freeway nonusers and the determination of the best diversion-parameter value. It was assumed that identical O-D distribution patterns would be experienced for both study periods and the validation was evaluated on that basis.

Model Calibration

To fabricate the freeway nonusers O-D demand, a trial-and-error procedure was adopted. First, a preliminary O-D demand of freeway nonusers was established. The minimum paths that the freeway users would choose were hypothesized, and the freeway nonusers on each link were calculated (observed link volume minus freeway user volume). Although freeway nonusers have different origins and destinations, their route-selection processes are not likely to be particularly sensitive to varying control strategies. Knowledge of the actual origins and destinations was therefore not considered essential and the nonuser O-D demands were selected such that resultant link volumes matched observed values. Subsequently, demand volumes were established for a preliminary corridor O-D trip table.

To establish the value of the diversion parameter, it was assumed that diversion of freeway users from controlled ramps depends principally on the alternative-travel cost characteristics in the corridor. In CORCON, those characteristics are approximated by linear components that provide a range of flow within which the link cost remains constant (Figure 5). The exact nonuser link volume is therefore not necessary, and the preliminary nonuser O-D demand can be used in establishing the diversion-parameter value. Consequently, to estimate the value of the diversion parameter for the entire corridor, a representative link was selected on an alternative route for diverted traffic (the south service road east of Cawthra Road, shown as link 41-14 in Figure 3). The model was used to predict traffic volumes on that link for different values of the diversion parameter. These volumes were compared with observed volumes, and the best value of the diversion parameter was selected as that which minimized the discrepancy between the calibrated and the actual volumes.

This value was found to be 4. (In other words, if the queueing time on a minimum path is equal to the difference in travel time between the nonqueueing and the queueing minimum paths, approximately 6 percent of the demand will divert to the nonqueueing alternative route.) That value was then used in CORCON, and the preliminary O-D table was adjusted by trial-and-error to minimize the difference between predicted and actual link volumes for the entire corridor by adjusting the freeway-nonuser demand patterns until an acceptable difference average was achieved. In Table 1, an example of the resultant O-D demand for an arbitrarily chosen time slice is shown.

A summary of the differences for each time slice on major arterial and freeway links is given below.

Time Slice (a.m.)	Average Difference (%) Link Volume	Unit Travel Time
7:00-7:15	3.0	4.9
7:15-7:30	2.7	5.6
7:30-7:45	3.0	5.9
7:45-8:00	2.1	4.1
8:00-8:15	2.9	5.3
8:15-8:30	4.0	4.5
8:30-8:45	1.0	4.5
8:45-9:00	2.3	4.9

The average differences for all time slices are approximately 3 percent for volumes and 5 percent for unit travel times. Results for a typical time slice are illustrated in Figures 6 and 7, which show respectively the calibrated and actual volumes, unit travel times, and queues.

Model Validation

After the corridor O-D demand and the best diversion-parameter value for after-control conditions were determined, the model was used to predict corridor operations for the before-control conditions. The validation

Figure 5. Relationship between flow and travel time and its linear components.

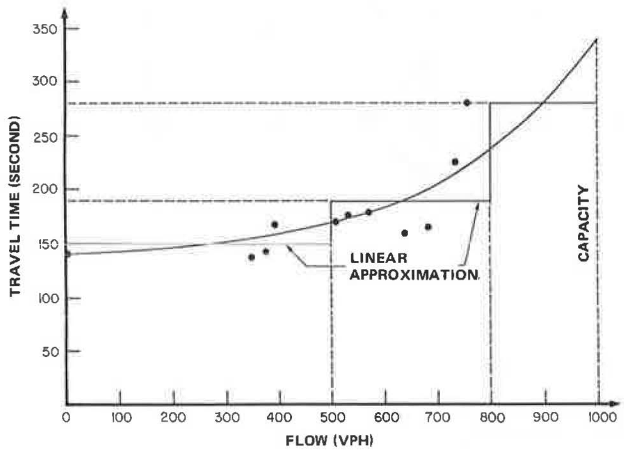


Table 1. Origin-destination demand table (8:00-8:15 a.m.).

Origin No.	Destination No. 1	4	6	7	9	10	11	12	13	14	15	16	25
1	0	171	0	176	180	0	0	20	0	0	0	0	67
2	0	0	8	0	36	0	8	0	0	738	0	0	0
3	118	0	0	0	0	0	0	0	0	32	0	0	0
4	63	0	0	0	0	0	0	36	0	50	94	0	0
5	118	0	0	0	0	16	0	0	0	61	0	0	0
6	0	0	0	0	0	16	0	0	0	213	214	0	0
7	0	0	0	0	0	0	0	0	98	50	0	0	0
8	0	0	0	0	21	0	0	0	60	50	0	31	0
9	0	0	0	0	0	27	0	0	60	45	0	180	0
10	0	0	0	0	127	0	0	0	203	44	0	0	0
11	0	0	0	0	20	40	0	0	71	60	0	0	0
12	0	0	0	0	0	0	0	0	0	50	9	15	0
13	0	0	0	0	0	0	0	0	0	141	43	0	0
16	0	0	0	0	20	24	0	0	0	36	46	0	0
18	0	0	0	0	0	0	0	0	0	51	0	0	0
25	0	75	0	86	0	0	0	0	0	40	0	0	0

process involved comparing predicted operating conditions with observed conditions at several locations on east-west arterials in the study corridor. (Only volume was used as a criterion for comparison because other criteria, such as travel time and queue length, were not available in sufficient detail for the before-control situation.) The four most important locations on Highway 5 and Highway 2 that were used for the final validation comparison are shown as blocks A, B, C, and D in Figure 2, and the measured and predicted volumes

Figure 6. Calibrated volumes, travel times, and queues (8:00-8:15 a.m.).

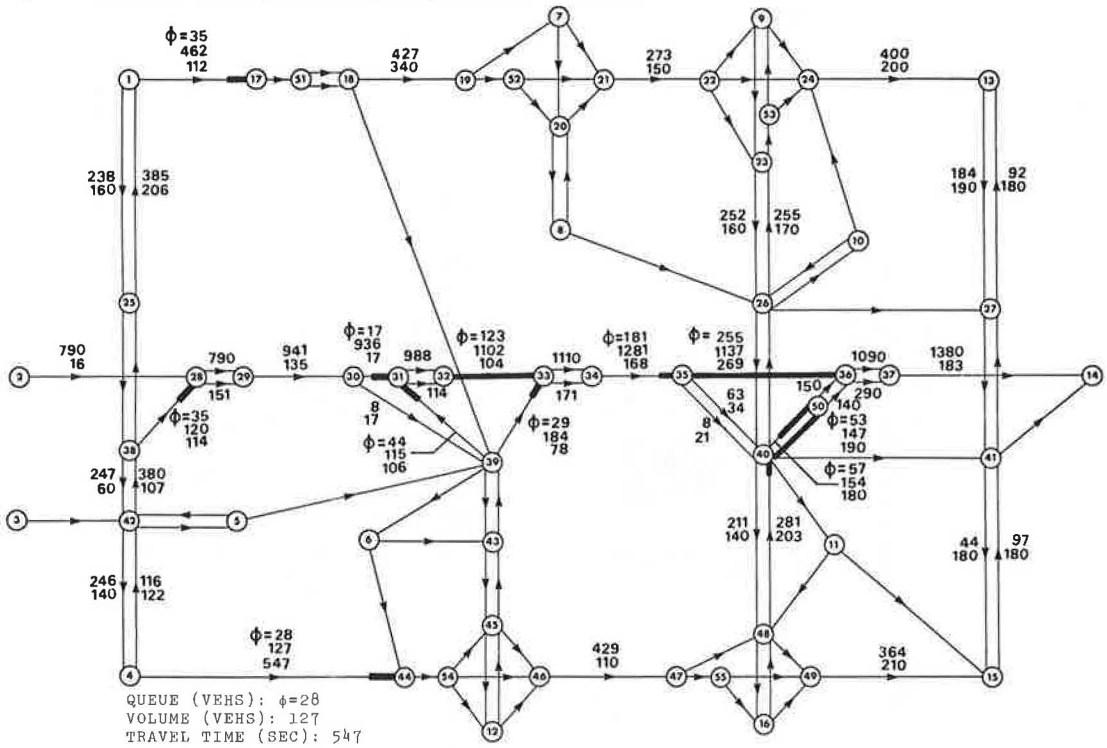


Figure 7. Measured volumes, travel times, and queues (8:00-8:15 a.m.).

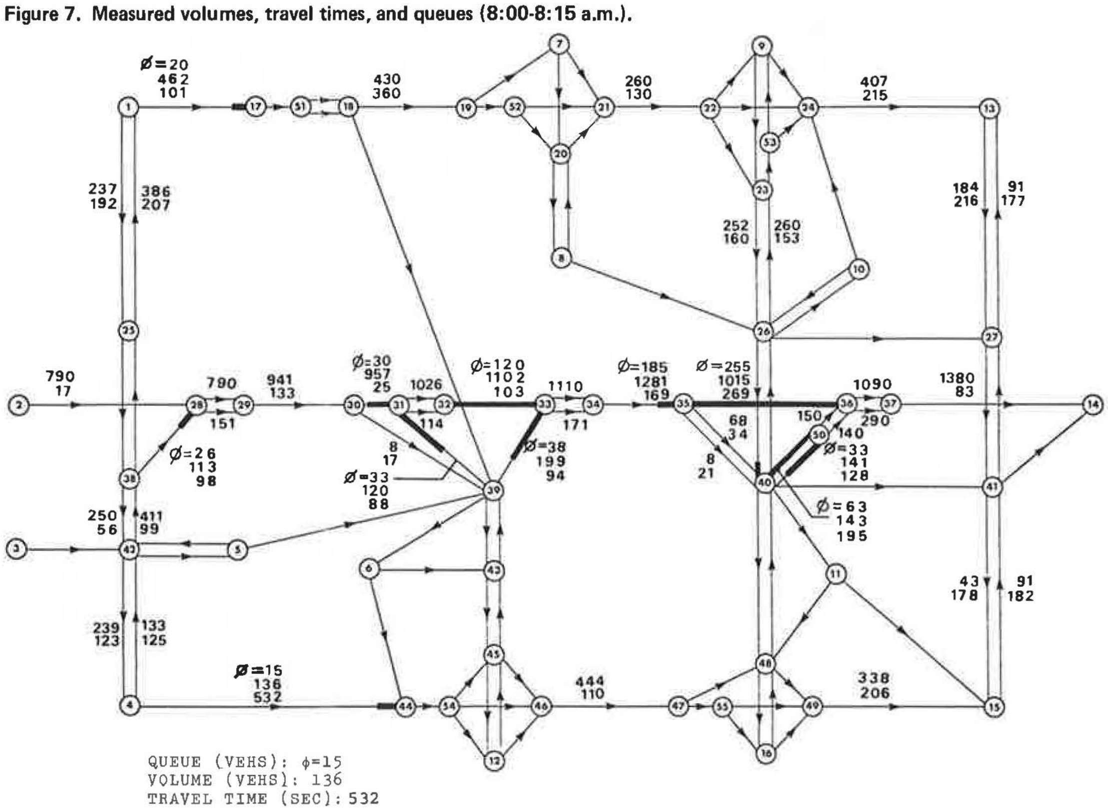
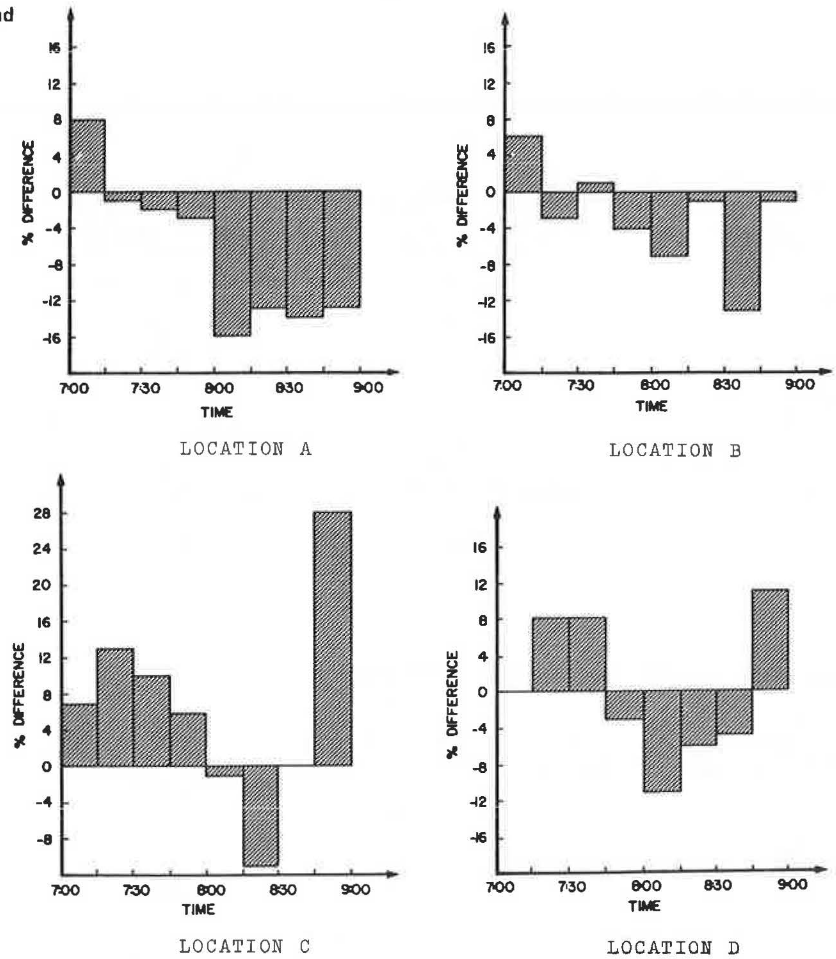


Table 2. Comparison between measured and predicted volumes.

	Location A			Location B			Location C			Location D		
Time Slice (a.m.)	Predicted No. of Vehicles	Actual No. of Vehicles	Percent Difference	Predicted No. of Vehicles	Actual No. of Vehicles	Percent Difference	Predicted No. of Vehicles	Actual No. of Vehicles	Percent Difference	Predicted No. of Vehicles	Actual No. of Vehicles	Percent Difference
7:00-7:15	212	196	+8	282	267	+6	138	129	+7	180	180	0
7:15-7:30	226	229	1	320	330	-3	270	239	+13	254	236	+8
7:30-7:45	256	260	-2	364	360	+1	357	324	+10	321	298	+8
7:45-8:00	243	251	-3	351	364	-4	381	358	+6	347	358	-3
8:00-8:15	268	319	-16	386	416	-7	404	410	-1	347	391	-11
8:15-8:30	277	317	-13	345	350	-1	370	415	-11	334	354	-6
8:30-8:45	239	278	-14	251	287	-13	301	300	0	250	263	-5
8:45-9:00	239	275	-13	240	242	-1	262	205	+28	184	166	+11

Figure 8. Differences between measured and predicted volumes.



for those locations are compared in Table 2.

The differences between predicted and actual volumes at all locations ranged from -16 to +13 percent, with the exception of one extreme value. The absolute difference was less than or equal to 13 percent in 91 percent of all results. The mean value of the absolute difference was 9, 5, 10, and 7 percent at the four validation locations respectively. A graphic view of the differences in each time slice for each of the four locations is shown in Figure 8. Results indicate that predicted volumes at locations A and B on Highway 5 tend to be underestimated and those at locations C and D on Highway 2 tend to be slightly overestimated.

Discussion of Results

Although the differences between measured and predicted volumes at the principal validation locations were as high as 16 percent and the mean difference was 10 percent, the correspondence was amazingly close when one considers how the O-D demands were established. Average growth rates in traffic demand for the entire corridor area were assumed for comparative purposes. In fact, however, considerably greater growth occurred in the northeast portion, contributing significantly to the underestimation of volumes on Highway 5, but the Highway 2 (southeast) area was overestimated. Closer correspondence would have obviously occurred had this change of growth pattern been reflected in the development of the O-D trip tables.

CONCLUDING REMARKS

As a consequence of the study results and experience gained during the conduct of this project, several observations are considered worthy of note.

It was shown that the new CORCON model was capable of predicting traffic volumes, travel times, and queueing characteristics on a freeway corridor; the maximum overall average difference between predicted and observed characteristics was 10 percent. Such a correspondence of predicted and actual traffic operating conditions is certainly sufficient to recommend CORCON for regular use as a planning tool for assessing alternative freeway-corridor control plans. (This is particularly apparent because fine tuning of the O-D demand table was not even attempted in this study; gross approximations are obviously sufficient.)

Furthermore, the new features of turn prohibitions (avoidance of illogical paths) and incorporation of a queue-diversion procedure operated extremely well and proved to be valuable additions to the corridor-assignment algorithm. By implication, the simplification of network link-node representation, the ability to simulate and introduce necessary turn-prohibition controls at critical network locations, and the ability to account for traffic diversion from entrance ramps where queues have formed are all useful and necessary elements of an acceptable model.

It should also be noted that CORCON, like many other models of this type, requires a complete matrix of O-D demand volumes for each time partition in the study period under investigation. The effort required to collect that information, either by postcard survey or by a comprehensive license plate survey, is not insubstantial. If the CORCON model is to be used regularly and relatively efficiently, O-D data could perhaps be best fabricated from more readily available data. (It is understood that a major effort to develop an O-D manufacturing process from volume counts is currently under way in New York. The results of that work could prove to be extremely useful to the operation of CORCON and other freeway-corridor assignment models. It is also important to note that the level of accuracy required is relatively low and that approximations of this type give perfectly acceptable results.)

Similar minor problems might be encountered in attempts to establish the value of the diversion parameter for use in CORCON. Although one would normally assume that diversion characteristics are similar at almost all entrance ramps, substantial amounts of data will be necessary to firmly establish the magnitude of the parameter. Should this prove difficult, one could calibrate to the best value by using representative alternative-route link volumes as the comparative baseline in a way similar to the procedure adopted in this study.

Because of dimensioning constraints in the currently available computer program, it may be difficult to apply CORCON to very large corridors [e.g., 8 × 40 km (5 × 25 miles)]. This can be accommodated by either increasing the dimension sizes, analyzing two or three smaller subsections of the corridors separately, reducing the amount of coded network detail, or some combination of the preceding. In any case, it should be possible to accommodate corridors of reasonable size without appreciable difficulty.

Despite these minor difficulties, the advantages of the model far outweigh its potential disadvantages. Consequently, we suggest that CORCON is the best available analytic procedure for reasonably predicting traffic-operating conditions on a complex urban freeway corridor. To this end, further work is currently under way using CORCON as the primary assessment tool for the investigation of the impact of proposed freeway-corridor traffic system management strategies on the Highway 401 bypass route in the metropolitan Toronto area. Preliminary indications are that the model is quite capable of handling the complex core-collector system of the freeway and the rather extensive network sections [which extend approximately 10 × 75 km (6 × 45 miles)].

The CORCON model can also be used very effectively for evaluating operating conditions in urban areas that may or may not have freeways. Furthermore, the model can be used not only to evaluate the effects of freeway ramp metering but also to evaluate the impact of a wide range of transportation system management strategies. A more detailed discussion of the use of CORCON for evaluating strategies in the fields of traffic demand management, traffic regulation, and traffic operations is given elsewhere (18).

ACKNOWLEDGMENT

This research was sponsored in part by the Ontario Ministry of Transportation and Communications and in part by the National Research Council of Canada. We express our appreciation to the many persons who contributed to this work over the past two or three years. In particular, we thank K. Williams of the ministry for organizing the extensive data collection efforts. His willingness to arise in the early hours of the morning to collect travel-time data exemplified the cooperation we received from all personnel in the traffic engineering office. Special thanks are also extended to A. Ugge who participated fully in the later validation process and whose thoughtful critiques were gratefully received. The opinions, findings, and conclusions expressed here are ours and not necessarily those of the Ontario Ministry of Transportation and Communications.

REFERENCES

1. B. L. Allen. Traffic Assignment and Control in a Transportation Corridor. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, PhD thesis and Special Rept., 1972.
2. B. L. Allen and G. F. Newell. Some Issues Relating to Metering or Closing of Freeway Ramps: I—Control of a Single Ramp. Transportation Science, Vol. 10, 1976, pp. 223-238.
3. B. L. Allen and G. F. Newell. Some Issues Relating to Metering or Closing of Freeway Ramps: II—Translationally Symmetric Corridor. Transportation Science, Vol. 10, 1976, pp. 243-268.
4. C. J. Liew. Entrance Ramp Closure as a Means for Controlling Traffic in a Transportation Corridor: Theory and Application. Department of Civil Engineering, McMaster Univ., Hamilton, Ont., MEng thesis, 1974.
5. B. L. Allen and C. J. Liew. Simplified Approach to Modeling Freeway Operations and Control. TRB, Transportation Research Record 533, 1975, pp. 88-96.
6. A. D. May and M. Orthlieb. Urban Freeway Corridor Control Model. TRB, Transportation Research Record 601, 1976, pp. 8-18.
7. J. F. Aidoo, R. W. Goedhart, and A. D. May. Bay Area Freeway Operations Study: Final Report—Analytic Techniques for Evaluating Freeway Improvements. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, 1971.
8. M. Orthlieb and A. D. May. Freeway Operations Study (phase 4): Freeway Corridor Control Strategies. Institute of Transportation and Traffic

Engineering, Univ. of California, Berkeley, Final Rept., 1975.
9. S. Yagar. Analysis of the Peak-Period Travel in a Freeway-Arterial Corridor. Institute of Transportation and Traffic Engineering, Univ. of California, Berkeley, PhD thesis and Special Rept., 1970.
10. S. Yagar. Dynamic Assignment of Time-Varying Demands to Time-Dependent Networks. Transportation Development Agency, Ministry of Transport, Canada, 1974.
11. S. Yagar. CORQ: Model for Predicting Flows and Queues in a Road Corridor. TRB, Transportation Research Record 533, 1975, pp. 77-87.
12. S. M. Easa. A Model for Investigating Traffic Assignment and Control in a Freeway Corridor. Department of Civil Engineering, McMaster Univ., Hamilton, Ont., MEng thesis, March 1976.
13. S. M. Easa and B. L. Allen. Analysis of Urban Freeway Corridors: Toward the Development of a Model for Assessing Freeway Control Plans. Department of Civil Engineering, McMaster Univ., Hamilton, Ont., Interim Rept. 3311-76/1, July 1976.
14. S. M. Easa and B. L. Allen. CORCON: Freeway CORridor Assignment and CONtrol Model. Paper presented at the 57th Annual Meeting, TRB, 1978.
15. B. L. Allen and S. M. Easa. Analysis of Urban Freeway Corridors: A Corridor Model for Assessing Freeway Control Plans. Department of Civil Engineering, McMaster Univ., Hamilton, Ont., Final Rept. 3311-77/2: Vol. 1, March 1977.
16. B. L. Allen and S. M. Easa. Analysis of Urban Freeway Corridors: Field Validation of CORCON—A Corridor Model for Assessing Freeway Control Plans. Department of Civil Engineering, McMaster Univ., Hamilton, Ont., Final Rept. 3311/77/3, Vol. 2, March 1977.
17. Highway Capacity Manual. HRB, Special Rept. 87, Washington, DC, 1966.
18. S. M. Easa and B. L. Allen. A Technique for Evaluating TSM Strategies in Urban Areas. Paper presented at the Annual Conference, Canadian Institute of Traffic Engineers, Calgary, Alta, 1978; Department of Civil Engineering, McMaster Univ., Hamilton, Ont., March 1978.

Publication of this paper sponsored by Committee on Freeway Operations.

Queen Elizabeth Way Freeway Surveillance and Control System Demonstration Project

E. Ryerson Case and Kenneth M. Williams, Systems Research and Development Branch, Ontario Ministry of Transportation and Communications

The design, operation, and effects of a freeway surveillance and control system that became operational near Toronto, Ontario, in July 1975 are described. The system consists of a low-light-level closed-circuit television system, microprocessor-based ramp-metering controls, ramp and mainline loop-detector installations, a central traffic-control computer, and a cathode ray tube graphic display. A single broadband coaxial cable is used for both television and two-way data transmission. The system provides traffic-responsive control that is based on both mainline and ramp conditions, incident detection, hardware-status monitoring, and a performance evaluation and reporting capability. The control center is located in a local Ontario Provincial Police facility near the freeway. Operating experience is discussed in terms of the effect of adverse public reaction to ramp metering, driver behavior, and system reliability. Substantial improvements in travel time and freeway speeds have been achieved, even under poor operating conditions, and the accident rate appears to have decreased. The closed-circuit television system has proved to be a valuable tool for traffic and incident management, particularly because of the close interaction with the police. The incident-detection system has been operating satisfactorily but requires verification by using the television system to eliminate false alarms. Overall, the project is considered to be successful.

Freeway surveillance and control is used in many cities. Even so, its introduction in a new area can still be a noteworthy event. And, because of changes in technology and public attitudes, the design, operation, and results of a new system can still add to the general pool of knowledge about the subject.

The Queen Elizabeth Way freeway surveillance and control system demonstration project represents the first venture by the Ontario Ministry of Transportation and Communications into the field of freeway surveillance and control. The project was considered for the following reasons:

1. The continuous increase in traffic on the freeway system,
2. The appearance of congestion on the freeway system,
3. The high cost of constructing or reconstructing freeways,
4. Public aversion to more or bigger urban freeways, and
5. Favorable results from similar projects in the United States.

Two broad goals were established for the overall freeway surveillance and control program:

1. To operate the freeway system at a high volume rate and a reasonable level of service while maintaining the best quality of service possible on nearby arterial roads and
2. To minimize collisions on the freeway system by

(a) recognizing conditions likely to cause collisions and providing adequate warning and (b) rapidly recognizing and responding to collisions (thus reducing the risk of secondary collisions).

Within the framework of these goals, specific objectives were established for the QEW demonstration project:

1. To provide relief to a congested component of the freeway system,
2. To introduce to the public the freeway surveillance and control concept and to demonstrate its effectiveness,
3. To provide local data and experience in system operation to project personnel, and
4. To obtain and develop expertise in the technology of freeway surveillance and control.

A portion of the Queen Elizabeth Way (QEW) just west of Toronto was selected for the demonstration project. There was localized congestion and, because it is close to a metropolitan area, the area was felt to be consistent with the above objectives. There was an additional advantage in that space for a control center was available in a nearby Ontario Provincial Police detachment office.

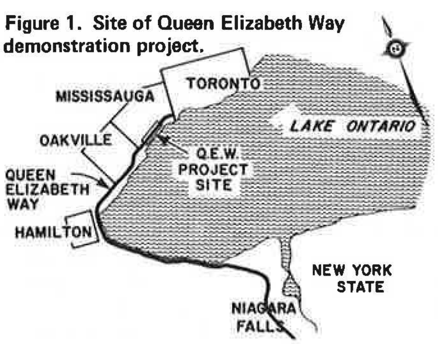
This paper describes the QEW freeway surveillance and control system (FSCS) and outlines the operational experience and the results achieved since the system was implemented in July 1975.

CORRIDOR DESCRIPTION AND SYSTEM REQUIREMENTS

Physical Description

The QEW connects the Niagara peninsula and Buffalo with Toronto (see Figure 1). Near Toronto, it also serves as a major commuter access route from the nearby cities of Burlington, Oakville, and Mississauga. The QEW FSCS controls the heavy flow toward Toronto in the morning. The actual project site is in Mississauga between Southdown Road and Highway 10, a distance of 6.2 km (3.9 miles), as shown in Figure 2. This is the area of greatest congestion.

Figure 1. Site of Queen Elizabeth Way demonstration project.



The Credit River, which crosses the QEW just east of Mississauga Road is a barrier to the movement of east-west traffic. The nearest alternative routes that cross the river are located about 2 km (1.75 miles) away. Both of these, Highway 2 and Highway 5, are four-lane urban arterials that have traffic signals and are not attractive as alternative routes. There is also another east-west freeway (Highway 401) about 10 km (6 miles) to the north.

In the project area, the QEW has three lanes in each direction; there are four or more lanes closer to Toronto. There are right-side paved shoulders except on the 224-m (800-ft) Credit River Bridge. Southdown Road and Highway 10 are four-lane arterials, and Mississauga Road is a two-lane residential collector. The control center is at the Highway 10 interchange.

Traffic Conditions Before Metering

Traffic data collection was begun in 1973 for project justification and to provide a data base for before-and-after studies.

As expected, traffic volumes in the morning peak period increased from west to east; 15-min volumes reached a peak of 5900 vehicles/h between 7:15 a.m. and 7:30 a.m. and then dropped off rapidly even though there were still waiting vehicles. Seven percent of the traffic is trucks; therefore, the through lane capacity is 5580 vehicles/h (1). Congestion, as represented by low speeds, extended from west of Mississauga Road to Highway 10, and its extent increased by 1.5 km (0.9 mile) in 1 year. At the height of the peak period, there were extensive queues of vehicles on the Highway 10 ramps and shorter ones at Mississauga Road but none at Southdown Road. Concentrations of collisions were located at the Credit River and at Highway 10. Residential development was increasing to the west in the QEW corridor.

System Requirements

In compliance with the goals and objectives described above, the following broad functional requirements and specifications were identified:

1. Ramp metering that included provision for several metering rates and control modes was to be installed at each of the five entrance ramps.

Figure 2. Queen Elizabeth Way freeway corridor and locations of metered ramps.

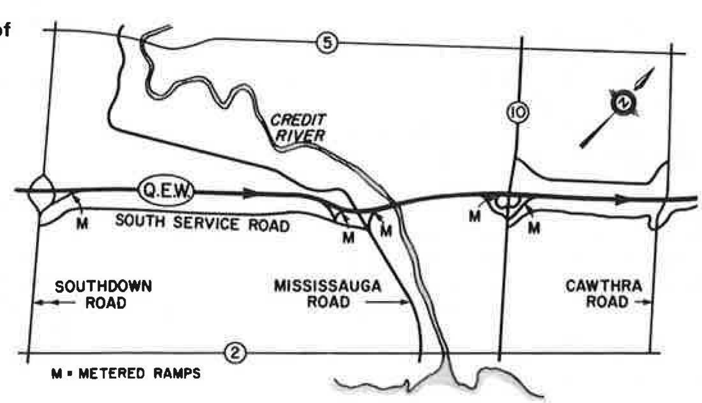
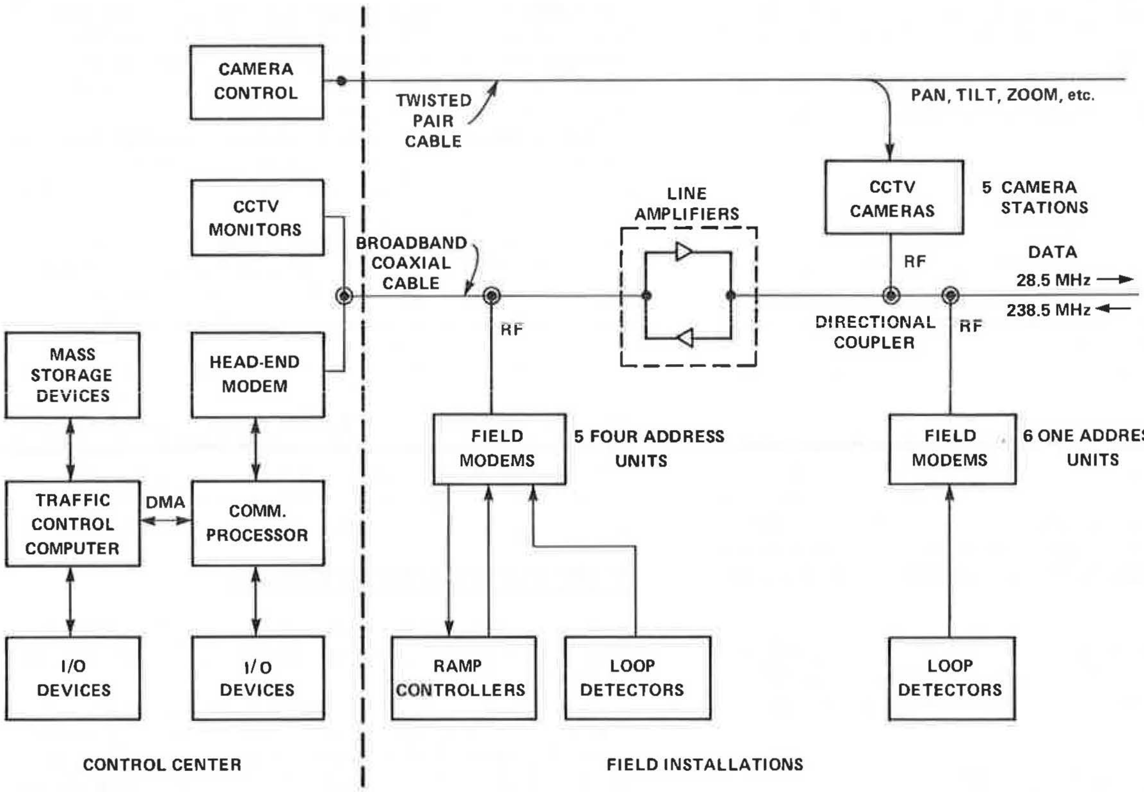


Figure 3. Functional block diagram of freeway surveillance and control system.



2. Closed-circuit television (CCTV) was to be installed to provide visual surveillance of the complete project area.

3. A central traffic control computer was to be installed at the control center to provide ramp-metering control, data collection, recording and analysis, incident detection, and hardware monitoring.

4. The feasibility of using a coaxial cable for broadband data communications was to be examined.

5. The project was to be carried out in-house to the greatest extent possible to give ministry staff in-depth expertise in the technology of freeway surveillance and control.

6. Sufficient instrumentation was to be provided to allow a comprehensive evaluation of system effectiveness.

SYSTEM DESCRIPTION

The six basic systems that compose the QEW FSCS (see Figure 3) are as follows:

1. Broadband coaxial cable system,
2. CCTV system,
3. Broadband data-communication system (BDCS),
4. Field installations,
5. Traffic-control computer (TCC) system, and
6. Application software for the TCC system.

The final configuration of the FSCS for the demonstration project was determined to some extent by the chronological order in which the various parts of the system were implemented. For example, the CCTV and coaxial cable systems, as well as the local ramp-metering controllers, were specified well in advance of the data-communication and TCC systems. This led to the installation of a 12-pair cable along with the coaxial cable to provide for voice-band data communications in the event that broadband data communications could not be used. As it turned out, the twisted-pair cable was used for CCTV camera control and central manual control of the ramp-metering controllers before the central traffic-control system was operational.

Broadband Coaxial Cable System

The broadband coaxial cable system is the main communication link between the control center and the remote field installations for both data and video transmission. It is essentially a high-quality coaxial television trunk line and uses a 1.91-cm (0.75-in) diameter aluminum-jacketed cable and bidirectional repeater amplifiers spaced about every 0.75 km (2200 ft) along its 7-km (4-mile) length. As shown in Figure 4, both the cable and the amplifiers are strung overhead on a pole-mounted, steel messenger cable. The system uses a subsplit configuration in which the central-to-remote bandwidth is 5 to 30 MHz and the remote-to-central bandwidth is 50 to 300 MHz. It can accommodate up to 30 television channels and several duplex data channels.

Closed-Circuit Television System

Channels 4, 6, 10, 12, and 13 are used for the five television cameras, which are spaced approximately 1.6 km (1 mile) apart along the freeway. Each camera location uses a standard audiovisual modulator and is connected to the main trunk line by a directional coupler; the audio carrier of each channel is currently unmodulated. The cameras, which use a low-lag, low-bloom type of silicon-intensified-target tube for both nighttime and daytime operation, are enclosed in a weatherproof housing that has air filters, exhaust fans, and windshield wipers. They are mounted on

Figure 4. Coaxial cable and repeater-amplifier installation.

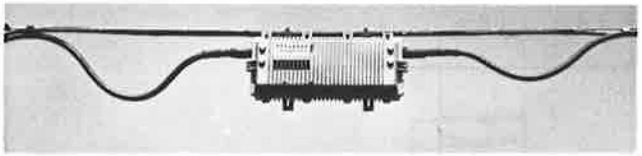
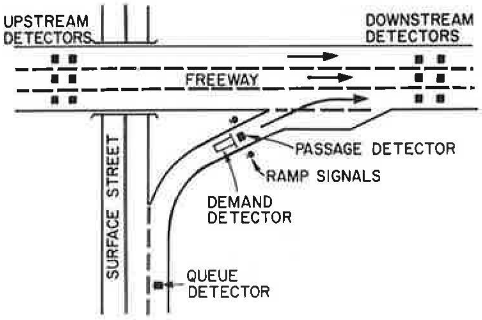


Figure 5. Basic ramp-metering installation.



15-m (50-ft) stressed concrete poles (which are used to minimize pole whip in high winds). All the usual remote-control functions are provided, such as pan, tilt, zoom, focus, iris opening, and windshield-wiper actuation. Five standard black-and-white television sets are used to separately monitor each camera. These are mounted in a console, which also includes a camera control panel, in the traffic-control center. The camera-control signals are carried over the twisted-pair cable shown in Figure 4.

Broadband Data-Communication System

Broadband data communications was a natural choice for the QEW FSCS because of the system requirement for CCTV. This type of system is cost-competitive with a voice-band system only when part of the cable cost can be apportioned to a CCTV system, but it has a number of advantages that make it an attractive choice from the technical point of view. These include high immunity to electric interference, large data-handling capacity, and ease of expansion.

The QEW BDCS (see Figure 3) interconnects the field loop detectors and the microprocessor ramp controllers with the central TCC by means of 11 field-data modulator-demodulator units (modems) located in 11 field cabinets. These field-data modems communicate with a head-end master-data modem through a duplex data channel by using a central-to-remote carrier frequency of 28.5 MHz and a remote-to-central carrier frequency of 238.5 MHz. The head-end master-data modem is connected to a communications processor (a 16K minicomputer) that serves as a front-end preprocessor to control the time sequencing of incoming and outgoing traffic data and to communicate with the central TCC. The communications processor has a teletype and a high-speed paper-tape reader for loading the communications-processor system software. Central-to-remote data communications is accomplished on a single time-share data channel that has a terminal addressing scheme. Remote-to-central data communications is accomplished by a time division multiplexing (TDM) system.

The maximum system data rate is 48 000 bits/s; this corresponds to a total of 64 terminal addresses that can be handled by the system, assuming a polling or sampling rate of 50/s. This sampling rate was chosen to limit the maximum absolute error of individual vehicle-speed-trap estimates to 10 percent at 100 km/h (60 mph), which corresponds to a 1-sigma sampling error of about 2.7 km/h (1.7 mph).

Field Installations

The field installations include five ramp-metering locations and 10 mainline detector stations. The specific configurations used (Figures 5 and 6) followed the recommendations given by Everall (2) wherever possible. Items worthy of comment are as follows:

1. Two signal heads are used on each ramp.
2. The right-side signal head is mounted 3.0 m (10 ft) above the roadway (Figure 7) to comply with legal requirements for traffic signals in Ontario.
3. An amber signal is included in each cycle, again to meet legal requirements. The amber time is 0.5 s. A 5.0-s amber is used during system turn-on.
4. The signal controller is a microprocessor-based device designed for ramp metering. The five cycle lengths can be selected on the basis of a fixed program, a remote device, mainline occupancy, or queue length.
5. The detector dimensions shown in Figure 6 will be adjusted in future installations because some motorists stop just before reaching the arrival detector and others stop just on the leading edge of the passage detector. In either case, the result is no green signal. The distance from the signal head to the leading edge of the passage detector will be increased from 2.4 m (8 ft) to 3.0 m (10 ft), and the length of the arrival detector will be increased to 6.1 m (20 feet) by moving the leading edge further away from the stop line.
6. The mainline detector stations are spaced approximately 800 m (0.5 mile) apart. Occupancy data are supplied to the local controllers, and occupancy, volume, and speed data are supplied to the central computer.

Traffic-Control Computer System

The TCC system consists of a 32 K-word minicomputer, a 1.25 million-word disk unit, a nine-track tape unit, a teletype terminal, and a cathode ray tube graphic display unit. The computer has a real-time clock, power-failure monitoring, and automatic-restart capability. The computer software runs under a real-time disk-operating system, which provides a real-time multitask environment and system-overlay facility. As shown in Figure 3, the TCC communicates with the communications processor through a duplex direct-memory-access (DMA) data channel. In addition to its normal communications task, the communications processor preprocesses freeway detector data before it is presented to the TCC. The graphic display unit uses a 48.3-cm (19-in) tube to plot graphic data on a 1024 × 781-dot matrix under computer control. An accompanying cathode ray tube hard-copy unit is also available.

Applications Software

The various software programs for real-time traffic-control-system applications are shown as a simplified functional block diagram in Figure 8. They are written in FORTRAN and assembly language and are under control of the master task program.

The purpose of the communications programs is to poll the communications-systems processor at 1-s intervals for the most recent traffic and field hardware-

Figure 6. Ramp detector configuration.

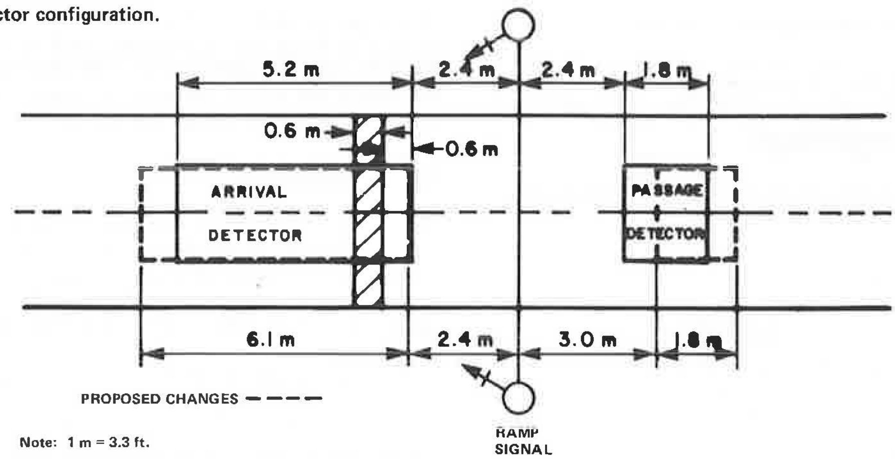


Figure 7. Typical ramp-metering installations.

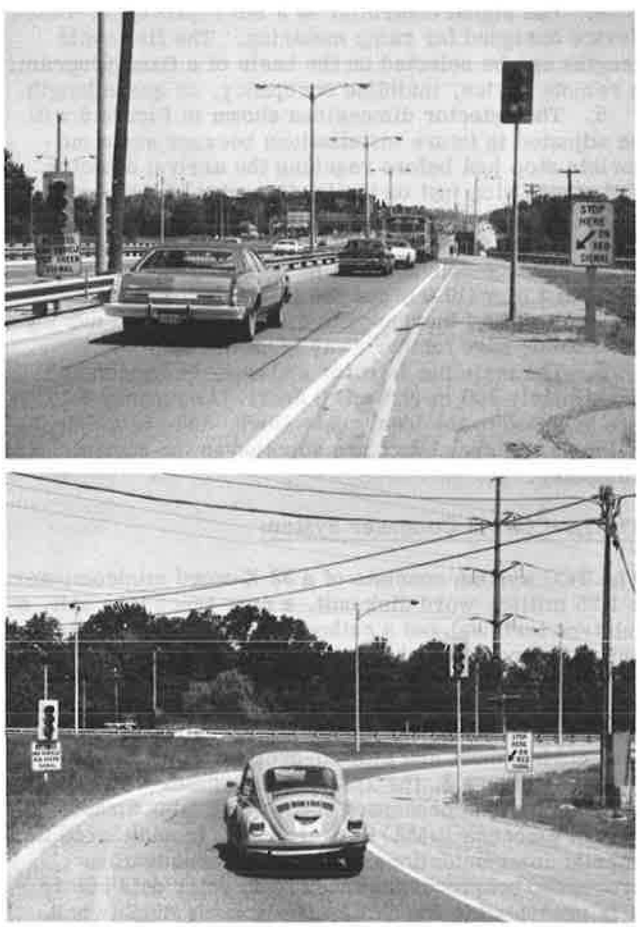
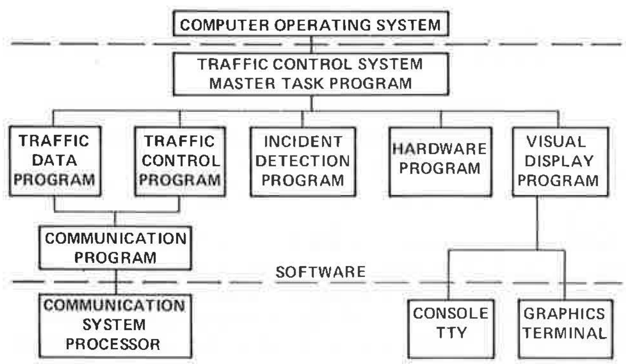


Figure 8. Software for real-time traffic-control system.



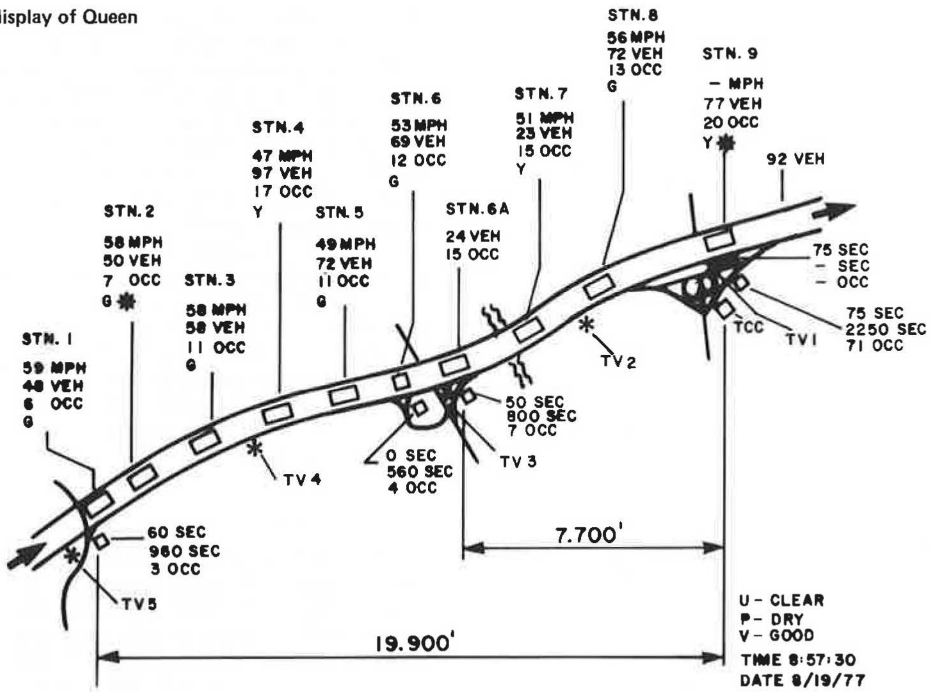
status data and to transmit the appropriate ramp-metering cycle-length commands to the communications-systems processor through the duplex DMA data channel. Data communication is accomplished through the execution of different protocols that are designed to accommodate different data transmission needs.

The traffic-data program accumulates and computes the real-time traffic data and saves the data files on the system data-bank disk. The traffic data base created and managed by this program contains both lane and station values of volume, occupancy, speed, and vehicle-length distribution, each updated at 30-s, 1-min, and 5-min intervals. The continuity of the data files allows enquiry and reporting capabilities of data ranging from the most recent to historical data up to 1-d old. This traffic data base is available to the other programs.

The traffic-control program is based on a traffic-responsive plan-selection strategy in which ramp-signal cycle lengths are selected on the basis of mainline and on-ramp conditions. Mainline conditions are determined at mainline detector stations located upstream, downstream, and adjacent to each ramp. Ramp-queue-detector occupancy data provides a measure of ramp congestion and is used to activate a queue-override function when it exceeds a preset threshold value, at which point the metering rate is changed to the minimum cycle length until the queue-detector occupancy drops below a second preset threshold value. The difference between these two preset occupancy values is a hysteresis band that must be large enough to prevent hunting between the minimum cycle length and the cycle-length corresponding to mainline traffic conditions. This program also uses the ramp-queue-detector occupancy data to estimate ramp queue lengths empirically for each ramp. Waiting time, calculated by multiplying cycle length by ramp queue length, provides an optional queue-override control mode.

The incident-detection program is based on the automatic freeway incident-detection concepts developed by the state of California. The method of incident detection and clearance uses a series of tests based on comparing the current and the previous 30-s mainline-detector-station occupancy data. This program also provides the information for the visual display

Figure 9. Graphic map display of Queen Elizabeth Way.



program that displays incident detection on the freeway map display and sounds an audible alarm signal.

The primary function of the visual display program is to display real-time traffic data on a freeway map displayed on the cathode ray tube graphics terminal. The program provides the capability of displaying the whole freeway map as shown in Figure 9 or of zooming in on individual interchanges for more detailed information. (Because the operation of these programs required that the data be given in U.S. customary units, SI units are not given in Figure 9.) The freeway graphic displays show real-time traffic data, including occupancy, volume, speed, ramp-queue waiting time and ramp-metering cycle lengths, all updated every minute. The location of an incident is indicated by an "X" midway between the upstream and downstream detector stations on the map display (Figure 9). This type of display was chosen rather than the conventional wall map because of its inherent flexibility and because it is representative of the next generation of wall-map displays that use color-projection TV.

The hardware program is designed to identify and record malfunctions in the system hardware. A malfunction is indicated by an audible alarm generated on the teletypewriter, and a record is kept on disk for maintenance-analysis purposes. The data modems are checked every second for the proper response; a particular modem is declared to be in an error state if there is no response after five successive interrogations. The microprocessor ramp controllers and the ramp-metering signals are checked every 30 s for possible malfunction, and the loop detectors are checked every 5 min.

A number of off-line programs have been written for purposes of system-performance evaluation. The results obtained are reasonably precise because the high level of detectorization used for the demonstration system provides data on total system input and output as well as mainline speeds and volumes throughout the entire system. Typical of the plots produced by these programs is the travel-time plot shown in Figure 10, which illustrates the large difference in travel time between good and poor conditions. Other plots include ramp waiting times (Figure 11) and freeway-speed contour diagrams. These plots, as well as printouts of total freeway travel time and travel distance, are all provided on a daily basis. Also, to provide access for research and other purposes to the large amount of traffic data generated by the QEW FSCS, an historical traffic data base has been established in the ministry's large-scale time-sharing computer system.

OPERATING EXPERIENCES

The QEW FSCS is now fully operational and has met the requirements laid down at the beginning of the project. However, it has gone through several evolutionary stages (Figure 12) since it was first implemented and it has experienced numerous problems along the way. Public reaction to ramp metering was the most serious problem.

Operational History

The first operational stage began on July 3, 1975, when isolated ramp metering that used local percentage-occupancy control at the four ramps at Mississauga Road and Highway 10 was implemented. To meet the project objectives, it was planned to increase the freeway speed to about 56 km/h (35 mph), the condition for maximum throughput according to the traditional relationship between volume and speed. Although some queuing had been expected, the metering rates chosen proved to be much too restrictive, which resulted in very long queues, excessive waiting times, and considerable congestion on the ramps and intersecting streets. At this time, there was no interconnecting cable or CCTV system to give an overview of system operation. Adjustments to the signal timings made during the succeeding few days were not sufficient to reduce the large amount of adverse public reaction. As a result, the system was shut down for a week.

It took several days for traffic to return to the congested premetered state, and it then appeared to deteriorate further and become even worse than before the metering. Clearly, the first few days of metering

Figure 10. Typical travel-time plots.

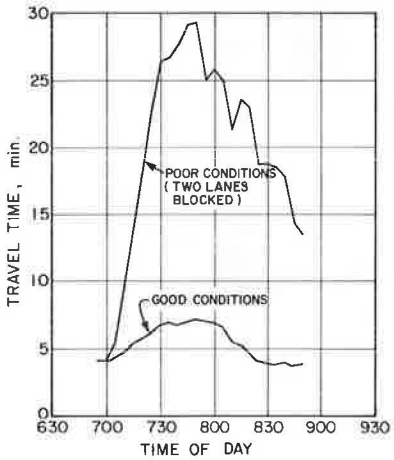


Figure 11. Typical plot of ramp waiting time.

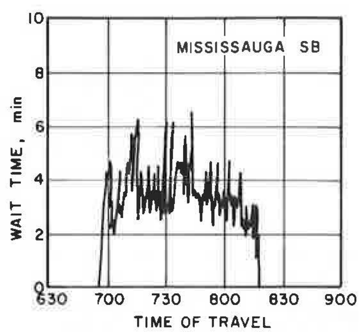
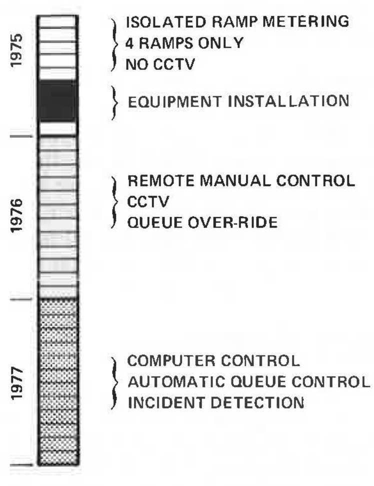


Figure 12. Operating history of freeway surveillance and control system.



had produced a radical change in the demand patterns on the entrance ramps and this had carried over into the nonmetered period.

During the shutdown period, it was decided that the only way the project could continue was to introduce a queue-override strategy. Maximum acceptable values of ramp queue length and waiting time were defined. When these values were exceeded, metering rates were increased even if this adversely affected mainline traffic. Combined with lower summer volumes, the queue-control strategy worked well until the end of August. But when volumes increased in September, adjustments could not be made quickly enough to prevent excessively long queues from forming and, after a few days, the system was shut down again—this time for 3 months. As before, conditions on the freeway deteriorated to premetering levels, but the merging interference with the mainline traffic returned more gradually.

During this second shutdown period, the communications cables and the CCTV system were installed and brought into operation. A remote, manual cycle-length selection system was also added by using spare wires in the 12-pair cable. When the system was started up again at the end of November, visual monitoring and manual adjustment of metering rates to maintain mainline flow while avoiding excessive ramp queues could be accomplished from the control center. This type of manual control constituted the second stage of ramp metering. There were no further shutdowns.

In January 1977, the third and final stage of ramp metering began when the data-communications system and TCC became operational. When necessary, the local controllers can still operate independently as a backup to the computer.

Public Relations

Public relations efforts included displays in local schools and a shopping center, a telephone survey, and distribution of leaflets explaining the project. The 3-day shopping center display was held a month before the start of ramp metering and included a ramp-metering signal, video tapes of traffic conditions on the QEW, charts and maps showing existing volumes and speeds, and handout leaflets. Many people visited the exhibit, and the staff were available to answer questions. The leaflets were also passed out on the ramps the day before ramp metering started.

During the 3-month shutdown, a similar display was set up for an evening in each of two local schools. A large and vocal turnout of people was expected in view of the adverse public reaction that had been generated. In fact, however, very few people came, although most of those who did attend were vehemently against the project. Their main argument was that it was inequitable to favor mainline users over Mississauga (ramp) users.

A telephone survey of Mississauga and Oakville residents was carried out by an independent consultant in the spring of 1976. At that time, the number of letters and telephone calls complaining about the project was diminishing and it was becoming more difficult to assess whether or not the public was still hostile. Of Mississauga residents who used the QEW in the morning peak period, 33 percent expressed some degree of satisfaction, 45 percent showed dissatisfaction, and 22 percent had no opinion. In Oakville, where a more favorable opinion would be expected because Oakville residents would not have experienced any ramp delays, the results were 70 percent satisfied, 16 percent dissatisfied, and 14 percent no opinion. About 60 percent of all motorists reported no change in travel time and of those who did report a change, 93 percent of Oakville residents reported a decrease and 65 percent of Mississauga residents reported an increase. Despite the publicity surrounding the project, only 30 percent of the users had noticed the television cameras. Only 6 percent had any objection to television surveillance. Two general conclusions obtained from the survey were that (a) there was not as much objection to the project as had been indicated earlier and (b) positive aspects of the project, such as motorist aid and incident management, should be publicized more fully.

Driver Behavior

It took only a few days for motorists approaching a ramp signal to adjust their driving habits. Stop-and-go movement followed by rapid acceleration when the green signal was obtained soon was replaced by a gradual approach and less extreme starting and stopping. A few motorists continue to stop in the wrong place, but the revised detector placements discussed above should reduce that problem. At Southdown Road, the ramp tapers from two lanes to one at the signal location. At times of day that the ramp is metered, motorists form two lines but, during the rest of the day, only one lane operates at the signal. No problems have occurred, and the extra metering storage and 2 vehicles/green throughput have been beneficial on this high-volume, high-speed ramp. Red-signal violations, which are highest toward the end of the metering period when congestion is easing up, have generally been quite low and have not been a cause for concern. A few consistent violators have been cautioned by the police, but no convictions have been registered.

Ramp motorists appear to be less aggressive than they used to be when merging with mainline traffic. During times when queue control permits motorists to enter at a faster rate than they can merge, there does not appear to be as much interference with the mainline flow as previously.

System Reliability and Maintenance

The QEW FSCS failure statistics for the period between April 1976 and July 1977 are shown below (listed according to the major subsystems identified in Figure 3).

System	No. of Failures/No. of Units	Failures of Total System
Coaxial cable	0/1	0
CCTV		0
Cameras	Many/5	
Controls	Many/5	
Data-communication system		0
Processor	0/1	
Modems	8/13	
Field installation		0
Ramp controls	6/6	
Detector installations	36/77	
Computer system		1
Computer	0/1	
Disk drive	1/1	
Tape drive	1/1	
Input-output	0/1	

The coaxial cable system had no failures in its first year of operation and has required little maintenance except a periodic routine monitoring of signal levels. The life of a high-quality coaxial television trunk line of this type is estimated to be in excess of 15 years under normal conditions.

The original CCTV system was unreliable almost from the day of installation, and its performance deteriorated as time progressed. The problems appeared to be in both the cameras (electronics) and the control (particularly the zoom drive). Both cameras and lenses were completely replaced after the first year and have since operated satisfactorily. The picture reception during both daylight and darkness has been excellent.

In general, the operation of the BDCS has been satisfactory. Proper heating, ventilation, and insulation are necessary to maintain the cabinet temperatures within the desired temperature range of operation of the data modems [0° C to 40° C(32°F to 105°F)]. The radio-frequency cards in the remote data modems required tuning every 6 months, and maintaining the integrity of the automatic gain-control system and the repeater amplifiers is an absolute necessity.

The only failures in the TCC system have been in the disk-drive and the magnetic-tape-drive units, both of which have been due to poor maintenance. Maintenenace is now carried out by ministry staff, and all repairs are performed by the manufacturer on an as-required basis.

In general, the modular design of the system has resulted in considerable ease in locating component failures and, as a result, there has been a minimum of disruption of system operation. It has been found that shipping and customs delays are very significant in determining the number of spares that should be kept on hand.

SYSTEM EFFECTIVENESS

An assessment of the effectiveness of the QEW FSCS during its first 2 years of operation has been carried out with reference to the original requirements for ramp metering, incident detection, and television surveillance.

Ramp Metering

As indicated above, ramp metering was implemented in three stages. The first, which used only local percentage-occupancy control, was short lived because of the unacceptably long queues generated. One can only speculate as to its ultimate effectiveness if it had been allowed to continue operating; perhaps the residual improvement shown in the immediate postmetering period gives some indication, but it is unlikely that the project could have continued if the ministry had tried to maintain mainline flow without sufficient regard for downstream ramp users.

The second stage of ramp metering involved centralized manual control in which cycle lengths were selected strictly on the basis of visual observation by CCTV of the mainline and ramp traffic conditions from 7:00 a.m. to 9:00 a.m. each weekday. The performance of the system during this period is summarized in Table 1. The operating statistics for the before condition were obtained from floating-automobile runs and extensive direct observations. The after statistics were derived almost entirely from data taken daily by direct television observations. Floating-automobile runs were made on occasion, but travel times and speeds were determined mainly by tracking an easily identified vehicle through the system. The other measures of effectiveness are similarly the result of direct television observations.

During this period, 49 collisions occurred in the eastbound lanes of the QEW on weekdays between the hours of 7:00 a.m. and 9:00 a.m. This compares with the 66 collisions that occurred under the same conditions a year earlier. The improvement is quite impressive, but the fact that the speed limit on freeways in Ontario was reduced from 112 to 100 km/h (70 to 60 mph) early in 1976 may be partially responsible.

The first-year operating statistics shown in Table 1 have been divided into two categories—good and poor—of conditions because of the different operating characteristics associated with the two categories. Poor conditions are defined as those that occur because of rain, snow, disabled vehicles, or collisions; good conditions are those that occur when poor conditions are not present. On good days, the system is operated to produce a downstream mainline volume of 5700 to 5800 vehicles/h, which is slightly over capacity. Under these conditions, there is obviously no reserve for unusual events that reduce the mainline capacity. On poor days, there are longer mainline queues, slower

Table 1. Operating statistics: December 1975 to December 1976.

Measure of Effectiveness	Good Conditions Before	After	Percentage Gain	Poor Conditions Before	After	Percentage Gain
Minimum speed, km/h	26.3	37.6	43	15.3	22.4	47
Avg speed, km/h	34.4	49.7	45	21.5	34.2	59
Congested period, min	105	85	20	131	113	14
Mainline queue length, km	4.7	3.1	34	6.3	4.9	24
Total travel time, vehicles/h	1308	1090	21	2020	1540	24

Note: 1 km = 0.62 mile.

speeds, longer or excessive ramp queues, and decreased ability to influence the overall operation of the system. Congestion is defined as the condition where the mainline density is greater than that which occurs at a speed of 64 to 72 km/h (40 to 45 mph), and mainline queue length is defined as the maximum extent of the congestion measured westerly from Highway 10. System travel time includes the waiting time on the ramps but does not include any allowance for vehicles that have sought alternative routes. Averages are taken over the 7:00 a.m. to 9:00 a.m. peak period.

Freeway traffic demand did not remain steady during this period, as can be seen from the average daily entering volumes before and after ramp metering was started given below.

Entrance	Before (1973-1974)	After 1975-1976	May-June 1977
Highway 10 ramps			
Northbound	710	790	930
Southbound	1 335	1 260	1 120
Mississauga Road ramps			
Northbound	820	1 370	1 105
Southbound	1 055	1 000	965
Southdown Road ramp	1 820	1 880	1 270
Mainline	4 890	5 700	6 280
Total	10 630	12 000	11 670

Approximately 600 vehicles leave by the ramps at Mississauga Road and Highway 10; the remainder represents the downstream mainline volume. Thus, the entering volume was increasing steadily during the time that the project was being implemented and during its first year of operation.

From these results, it is apparent that a significant improvement in overall performance has been achieved by ramp metering, even under poor conditions. These results are even more impressive in view of the increase in demand. The system will operate at the capacity of the mainline on good days. The number of collisions has been reduced, and there have been substantial reductions in travel time, congestion time, and mainline queue length. There is no question that mainline traffic benefits from the system. Many people, however, have questioned whether the ramp users receive any benefits.

An early analysis of ramp delays compared with mainline benefits showed the following:

1. At Highway 10, which is at the downstream end of the system, it was assumed that there were no mainline benefits for the ramp users. However, it was found that the ramp delays had decreased after metering, so that there was an overall benefit.
2. At Mississauga Road, there was some benefit from improved mainline operation but there was also increased delay on the ramps. The result was a net loss.
3. At Southdown Road, there was considerable mainline benefit but, of course, some ramp delay where there had been none before. The result was a net gain.

Since then, attempts have been made to adjust the metering rates to reduce the loss to the Mississauga Road ramp users; the results have yet to be evaluated.

The third stage of ramp metering started in January 1977, when the data communications and TCC systems became operational. Totally automatic operation under computer control did not start, however, until the end of March 1977. Preliminary evaluations have been carried out since that time by using the extensive performance-evaluation software that was developed for this purpose. The results are summarized in Figure 13, which shows a comparison of average travel times through the system with manual control (stage 2) and with computer control (stage 3). It is based on two samples of 36 good days chosen from travel-time data taken during the springs of 1976 and 1977 respectively. The improvement achieved by using computer control is apparent from the shape of the histograms; average travel times when computer control is used are less than those when manual control is used 72 percent of the time, and the mean value has decreased by almost 1 min. Some of this improvement may be due to a net decrease in total volume during the first half of 1977. The decrease occurred on some of the ramps, although the mainline volume beyond Southdown Road has continued to increase. Completion of the widening of Highway 401 a few kilometers to the north and the extension of a new east-west arterial has probably attracted some motorists away from the QEW ramps, which would account for some of the observed shift in demand.

Incident Detection

The QEW incident-detection system has been evaluated for the period from the time it became operational in early January 1977 to the end of May, a period corresponding to about 100 d of operation. CCTV was the main tool used to assess the accuracy of the system.

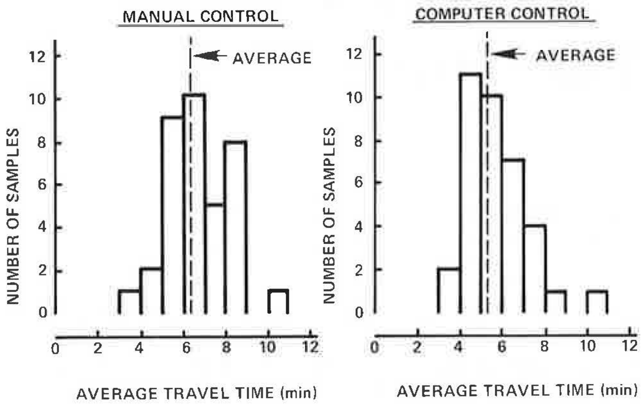
During this period, 45 incidents were recorded; 38 of these were detected by the system, which corresponds to a detection rate of about 84 percent. The total number of false alarms recorded was 142 but, of these, 50 were directly attributable to hardware failures; the false alarm rate was therefore 0.9/d.

It is interesting that, during this period, 49 percent of the false alarms had an alarm period of less than 2-min duration and 87 percent of the true incidents had an alarm period greater than 2-min duration. This would suggest that a persistence test should be added to the incident-detection algorithm; this feature is being added to the software and will be evaluated in the near future.

Television Surveillance

The CCTV system has been a useful tool for both the project staff and the Ontario Provincial Police. In con-

Figure 13. Comparison of manual and computer control.



junction with the computerized incident-detection system, it has been used to verify the occurrence of incidents and to locate them precisely between the mainline detectors, which is important for guiding police and emergency vehicles to the proper entrance ramp when the incident is located near an interchange. Incidents have also been detected independently by the CCTV by noting congestion in one or more lanes or an unusually low-volume condition. Even incidents in opposing lanes have been detected on occasion by noting slowdowns due to gawking motorists.

Proper camera placement at the interchange allows the use of the cameras to observe traffic on the ramps and the intersecting roadways. This becomes quite significant when queue control is used because the ramp queue detector provides information about the area at the detector but cannot indicate the extent of the queue in exceptional circumstances. In the event of detector failure, data can be gathered from the CCTV by direct observations. Mainline queues, undetected congestion, and the effects of increasing or decreasing ramp metering rates can likewise easily be observed by using the CCTV. In fact, the expression "a picture is worth a thousand words" is very apt for traffic management using CCTV.

The fact that the provincial police radio staff and the television monitors are in the same room has made the television even more effective in incident management. The incident can be precisely located, and its seriousness can be evaluated. The police will know how urgently they are needed at the scene. The television cameras are also used to verify and locate incidents reported by others. Major incidents can be quickly assessed by the police staff, who do not have to go out on the road and risk becoming delayed in the congestion.

The television cameras, together with very restrictive ramp metering, have been used to locate and clear a path for an ambulance on an emergency run. They are also used in inclement weather to give up-to-date traffic conditions to local radio stations. During complex lane-closing procedures, the cameras have been used to coordinate the procedure and keep vehicle delays to the minimum. The cameras have also been used to detect and send help to motorists in distress, possibly saving them from having to walk along the freeway. In one instance, help was quickly sent to a motorist whose car was dragging a mattress that appeared to have caught fire.

Although the other parts of the demonstration project operate only during the morning peak period, the television cameras operate 24 h/d and 7 d/week. The police radio staff operates and monitors the television system when the project staff are absent.

FUTURE PLANS

As its name implies, the QEW FSCS demonstration project was a limited-duration project to evaluate the effectiveness of freeway surveillance and control system technology in the Ontario context. In view of the favorable results obtained, it has been decided to add additional equipment and proceed with a local expansion program. In addition, an active research and development program has been undertaken to find ways to improve system operation and to explore the possibilities of applying this technology elsewhere in Ontario.

A changeable-message sign that uses two lines of 22 characters, each 45.8 cm (18 in) high, will be installed west of Southdown Road, facing the eastbound traffic. The sign will use reflective rotating magnetic disks to display up-to-date information obtained from the CCTV, incident-detection, and police radio facilities. Economy in operating costs is achieved because the sign requires power only for a message change. The coaxial cable will be extended to the sign site, and communication will be via a data modem.

Three new interchanges are to be constructed adjacent to the present control area, one downstream and two upstream. The QEW FSCS will be expanded to these areas, which will provide the opportunity for improved overall coordinated control. The television camera control signals will be sent via the coaxial cable, thus removing the need for the 12-pair data cable.

The research and development work under way includes the development of improved ramp-metering and incident-detection strategies, a study of the effect of the level of detectorization on the implementation of such strategies, the development of more refined performance-evaluation programs, and the development of a practical freeway-corridor model for use in systems analysis.

ACKNOWLEDGMENT

We wish to acknowledge the many valuable contributions made by other members of the Ontario Ministry of Transportation and Communications staff during the design and implementation phases of this project. In particular, we would like to mention Joe Tsai, who was responsible for the development of the traffic-control-system software and for the specification and procurement of the broadband data-communication system. Terry Blanks was responsible for the design, procurement, and installation of the closed-circuit television system.

REFERENCES

1. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
2. P. F. Everall. Urban Freeway Surveillance and Control: The State of the Art. Federal Highway Administration, June 1973.

Publication of this paper sponsored by Committee on Freeway Operations.

Evaluation of Reduction in Minimum Occupancy for Car Pools That Use a Priority Freeway Lane

Kenneth G. Courage, Thomas H. Culpepper,* Charles E. Wallace, and Joseph A. Wattleworth, Transportation Research Center, University of Florida

One of the primary control parameters associated with the use of preferential treatments for high-occupancy vehicles, specifically car pools, is the minimum occupancy of automobiles qualified to use the treatment. This parameter varies among the many projects around the nation but is usually between two and four persons per vehicle. The significance of the value used is in terms of the person-moving performance compared with the degree of priority given the high-occupancy vehicles. A priority lane was provided on I-95 in Miami for buses and car pools of three or more persons but, when the person-moving performance failed to meet the desired goal, the minimum car-pool occupancy was reduced to two persons. This unique action had predictable but significant results. The priority advantage previously afforded the high-occupancy vehicles was reduced to a large degree, but the total system effectiveness of the freeway, as well as its safety, were significantly improved. As a result, enforcement and operating problems were reduced and public acceptance was increased.

In recent years, there has been a proliferation of priority-treatment projects for high-occupancy vehicles (HOVs) in urban areas. These programs are a result of the need for low-capital-cost transportation improvements to enhance the person-moving capacity of existing highway facilities and to conserve energy (and improve the environment) by promoting the use of transit and other HOVs, notably car pools.

The techniques used to provide preferential treatment vary from situation to situation and have been reported extensively in the literature (1, 2, 3, 4). One of the more common ones is to reserve a lane on an urban freeway for use by buses and car pools during the peak hours.

A demonstration projection of this type was implemented in Dade County, Florida, in December 1975 on the I-95-NW 7th Avenue corridor; the associated express bus service was named the Orange Streaker. The complete project consisted of a 4-year program to implement and evaluate HOV priority-treatment techniques on both an arterial street [NW 7th Avenue (US-441)] and an urban freeway (I-95). The results of the project are detailed elsewhere (5, 6, 7).

The portion of the project in which the express buses and car pools used I-95 (phase 2 of the project) had a unique operational feature that warrants special attention. On January 10, 1977, the minimum occupancy of car pools authorized to use the priority lanes was reduced from three to two persons. This paper discusses the system in general, describes some theoretical considerations of the car-pool definition, and reports on the actual results of the change.

PROJECT DESCRIPTION

I-95 is the primary highway facility in the northern corridor of Dade County, connecting major residential areas in northern Dade and southern Broward counties with major employment areas in the greater Miami area. The freeway corridor is about 16 km (10 miles) long, extending between the Golden Glades interchange and the Miami central business district. The project area and its principal geographic and transportation facilities are shown in Figure 1 of another paper by Courage and others in this Record. I-95 is an 8- to 10-lane divided, fully access-controlled Interstate highway. The HOV priority lanes were constructed in the median between the Golden Glades and 36th Street interchanges, a length of about 11.7 km (7.3 miles). Thus, the capacity of the general lanes was not reduced by providing the priority lanes. Typical views of the freeway are shown in Figure 1. All lanes are the standard 3.7 m (12 ft) wide, and the right shoulders are 2.4 m (8 ft) wide. The freeway is divided by a concrete barrier wall that is separated from the traveled way by only 0.6 m (2 ft) in each direction.

The project also included a temporary 967-space park-and-ride lot located at the Golden Glades interchange. In March 1977, a new, permanent, 1320-space lot and a direct flyover ramp were added, but the effects of these improvements are not included in this report. Also, 30 specially equipped 47-passenger buses were purchased for the Orange Streaker service.

The Orange Streaker provided express bus service between the residential area north of the park-and-ride lot and three major employment areas: downtown Miami, the Miami Civic Center, and Miami International Airport (with limited continuing service to downtown Coral Gables).

Transit users and car poolers who formed their car pool at the park-and-ride lot used several modes of access to the priority system: park-and-ride, kiss-and-ride, Orange Streaker feeder bus, or local bus. Walking was not a feasible mode because of the isolated location of the lot.

Traffic control of the priority lanes was provided through the use of overhead fixed-message signs such as the one shown in Figure 1a and was reinforced by the standard restricted-lane diamond symbol painted at 76.2-m (250-ft) intervals on the pavement of the priority lanes. Operating hours were originally 6:00-10:00 a.m. (southbound only) and 3:00-7:00 p.m. (northbound only), but when the required car-pool occupancy was reduced, these times were changed to 7:00-9:00 a.m. and 4:00-6:00 p.m. respectively.

THEORETICAL CONSIDERATIONS

The original decision to require three or more persons to qualify for the priority treatment was based on two factors. First, the observed number of vehicles carrying two or more passengers was high enough that it was feared that no travel-time benefit would result if all of these vehicles actually used the lanes. Second, the prevailing tendency on a national basis is to specify three or more persons per vehicle (PPV) as a car-pool occupancy requirement. But after nearly a year of operation of express buses and three-PPV car pools, it was found that the anticipated degree of car-pool attraction did not materialize de-

Figure 1. Views of I-95 priority-lane system.

spite a travel-time advantage of about 3 min (during average morning and afternoon peak periods). Indeed the priority lanes were carrying fewer persons than the average general lane.

The underuse of the lane and the high violation rate generated substantial public pressure for relaxation of the priority-lane regulations. Two approaches were used to determine the probable effects of such a change. A theoretical approach was used to determine the optimum car-pool definition. A car-pool definition model was developed (7) that uses a two-stage traffic assignment technique. The model considers a system that consists of a freeway section that has both priority and general lanes and vehicular demand that is stratified by level of occupancy and origin-destination patterns. Demand is then assigned to the facility such that preferential treatment is given to HOVs, and the overall or passenger hours of travel. Both violation and nonuse rates are assigned to reflect actual conditions, and the model is iterated until equilibrium is established between the estimated average vehicular (or person) flow and the assigned vehicular demand. Assignments are constrained by internal capacities.

It was found that, in practically all freeway subsections, the preferred minimum car-pool requirement was between two and three PPV, as is shown in Figure 2. Additionally, these analyses indicate that a car-pool definition of two PPV would result in both minimum vehicle hours and minimum passenger hours of travel. However, it was also found that the 2-PPV requirement would fail to provide a significant level of preferential treatment for priority vehicles, as shown in Figure 3 (degree of priority is defined as the ratio of the general-lane demand-to-capacity and priority-lane demand-to-capacity ratios and should be greater than unity if preferential treatment exists). In fact, it was evident that, at this lower requirement, the priority lane could be expected to effectively operate as a general-use freeway lane that has developed user equilibrium. [Further investigation showed that the degree of priority could be improved by providing access-egress restrictions on the lane by using discrete entry-exit strategies (7)].

These findings tended to justify the use of three PPV over two PPV for preferred priority treatment, but, as seen in Figure 2, there was some indication that a surplus person-moving capacity was available for a car-pool definition of fewer than three PPV.

The Florida Department of Transportation (FDOT) conducted a further analysis to test the feasibility of reducing the car-pool definition (8). The significant findings were as follows:

1. When operating with the three-PPV requirement, the priority lane was carrying only about 44 percent as many persons as the average general lane (although in only 5 percent of all vehicles using the freeway).
2. The violation rate was about 63 percent of the priority-lane traffic and steadily deteriorated because of the lack of effective enforcement (and ultimately reached 78 percent in the afternoon peak hour).
3. Only about 32 percent of all vehicles eligible to use the reserved lanes actually did so.

It was then estimated that, by assuming a 50 percent increase in two-PPV car pools the probability of demand versus capacity approaching equilibrium with the general lanes would not exceed 40 percent in any 0.5-h period in the most critical section of the freeway. Overall, the probability of breakdown was estimated at less than 25 percent in the critical half hour.

Based on these estimates and accepting the risk of some deterioration in priority operations, the Florida DOT decided to reduce the occupancy level (and simultaneously, the periods of operation).

EFFECT OF CAR-POOL REDEFINITION

The effects of the change in car-pool occupancy were measured with respect to the following measures of efficiency:

Variable	Source of Data
Peak-period volumes and vehicle occupancies	Volume and occupancy studies
Bus travel times	Travel-time observations
Bus schedule adherence	Observations at Golden Glades terminal (provided by Metropolitan Transit Agency)
Automobile travel times and comfort measures	Instrumented moving-vehicle studies
Exclusive-lane-occupancy violators	Moving-vehicle observations
Weaving difficulties	Instrumented moving-vehicle studies
Transit ridership	Metropolitan Transit Agency records
Accident history	Dade County accident records

Effect on Transit Operations

The effects of the two-PPV car-pool definition on bus speeds and travel times in the reserved lane section of I-95 are summarized below (1 km/h = 0.62 mph).

	Car-Pool Definition		
Item	Three PPV	Two PPV	Change (%)
Morning peak			
Travel time, min	8.22	8.22	0
Speed, km/h	78.4	78.4	0
Afternoon peak			
Travel time, min	7.77	9.38	+21
Speed, km/h	83.0	68.7	-21

Figure 2. Optimum car-pool definitions for minimizing passenger hours: 3:30 to 6:30 p.m.

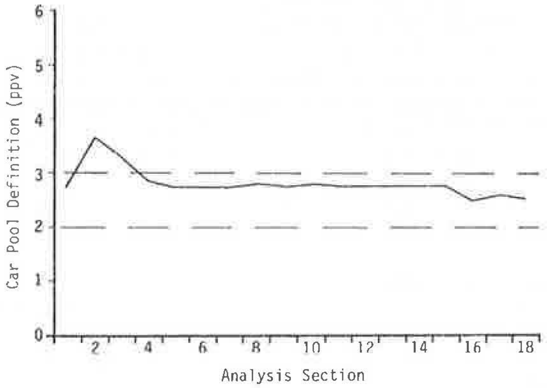
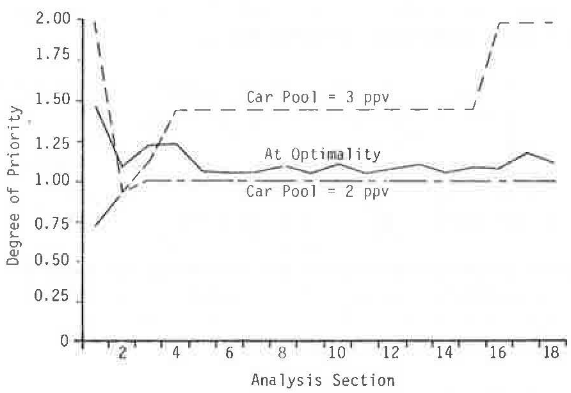


Figure 3. Degree of priority for minimizing passenger hours: 3:30 to 6:30 p.m.



The travel times are for the portion of the bus trip on I-95 between NW 36th Street and NW 151st Street, which represents most of the exclusive-lane section. These results indicate that the change in regulations had no measurable effect on the average bus travel time in the morning but that travel times were increased by approximately 21 percent during the afternoon peak period.

A more detailed analysis of the bus travel times—the effect of time of day on the variation of travel times—is shown in Figure 4. For example, in the morning peak the travel times remained more or less constant during the entire peak period. It is interesting that, although the average travel times were not altered by the change in regulations, the variation in travel time (as indicated by the width of the confidence limits) was noticeably greater when two-person car pools were allowed in the exclusive lane.

In the afternoon peak, on the other hand, the two-person regulation resulted in increases in both the average travel time and the variability of travel time. Furthermore, a strong peaking trend is evident in the two-PPV case during the more congested portion of the afternoon period.

Bus schedule-adherence studies were conducted during the afternoon peak period. The primary measure of effectiveness used was the difference between the scheduled and the actual arrival times for buses at the Golden Glades terminal. This measure is termed the arrival-time discrepancy. The distributions of arrival-time discrepancies representing the three-PPV and two-PPV car-pool stages are shown in Figure 5. The dispersion

Figure 4. Variation in express bus travel times during peak periods.

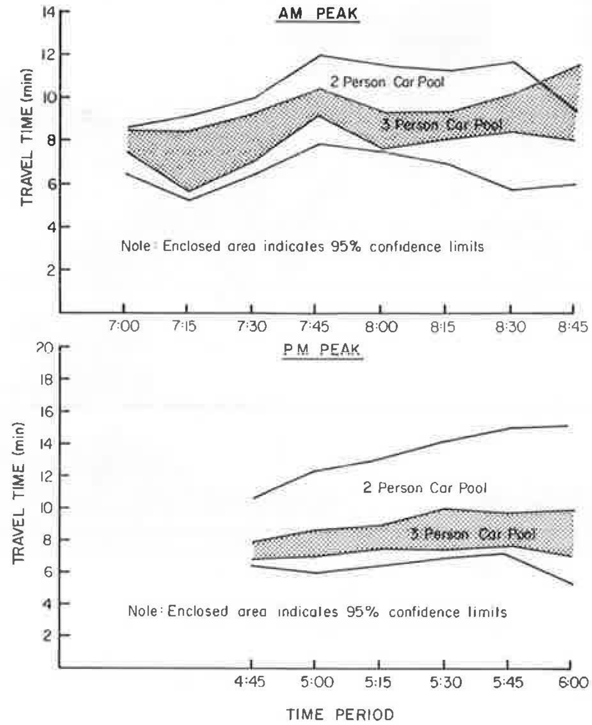
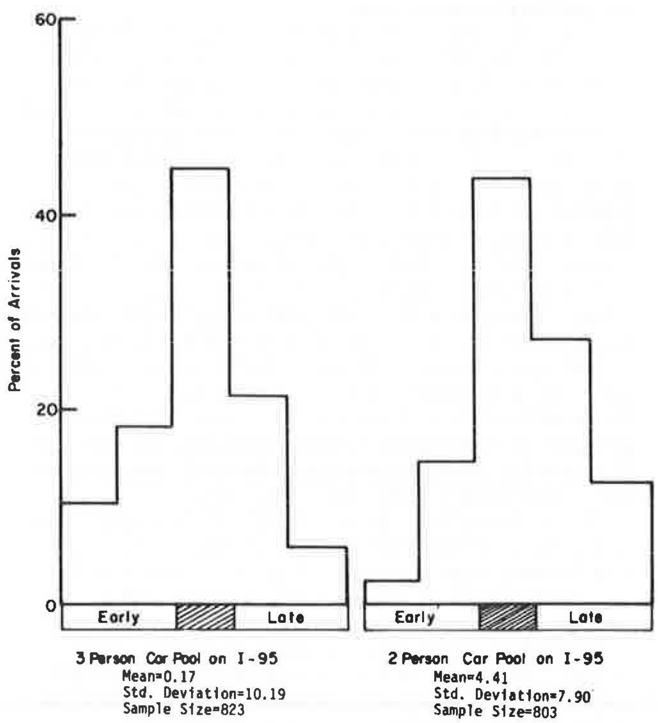


Figure 5. Distribution of differences between actual and scheduled arrival times of Orange Streaker buses at Golden Glades terminal.



of the distribution reflects the degree of schedule adherence (a more dispersed distribution represents a lower degree of adherence). Another measure of schedule adherence used was expressed in terms of the average lateness of buses. It is observed, for example, that the average bus arrived 4.4 min late at

Table 1. Effect of two-person car-pool definition on automobile travel times and speeds.

Item	Priority Lanes			General Lanes		
	Three PPV	Two PPV	Change (%)	Three PPV	Two PPV	Change (%)
Morning peak						
Travel time, min	7.51	7.50	0	10.52	10.73	+2
Speed, km/h	85.9	85.9	0	60.2	60.0	-2
Afternoon peak						
Travel time, min	8.02	7.94	Negligible	11.26	9.63	-14
Speed, km/h	80.5	81.3	Negligible	57.3	66.9	+14

Note: 1 km/h = 0.6 mph.

Table 2. Comparison of traffic volumes.

Time Period	Car-Pool Volume (vehicles/h)			Violator Volume (vehicles/h)			Total Volume (vehicles/h)		
	Three PPV	Two PPV	Change (%)	Three PPV	Two PPV	Change (%)	Three PPV	Two PPV	Change (%)
Morning peak	106	623	488	182	352	93	288	974	238
Afternoon peak	110	638	480	188	380	102	298	1017	241

the Golden Glades terminal with two-person car-pool operation and 0.2 min late with three-person car pools. This difference agrees generally with the difference in travel times observed on the freeway. The dispersion of arrival-time discrepancies between these two stages of operation dropped by approximately 20 percent. This indicates that, although travel times were longer, the predictability was improved, primarily because fewer buses arrived earlier than scheduled.

Effect on Automobile Operations

The effects of the two-person car-pool definition on automobile travel on I-95 are summarized in Table 1. The only comparison in the above summary that proved to be statistically significant (95 percent level of significance) was the improvement in travel time in the general lanes during the afternoon peak period. The automobile travel-time comparisons during the morning peak were consistent with the bus travel-time comparisons; i.e., no noticeable change occurred with the reduced car-pool requirement. It may therefore be concluded that morning peak operations were not substantially affected by the operational changes that were implemented.

During the afternoon peak, on the other hand, noticeable changes were observed in the bus travel times, which increased by 21 percent, and in the automobile travel times in the general lanes, which decreased by approximately 14 percent. Some increase in automobile travel times in the exclusive lane would be anticipated in view of the relatively large increase in bus travel times; however, no such increase was recorded in the field. The average speed for automobiles in the exclusive lane remained at approximately 80.5 km/h (50 mph) throughout both stages of the study. This is generally consistent with a level of service B operation. The corresponding travel time in the general lanes was 67.6 km/h (42 mph), which represents level of service C.

The difference between the bus travel times and the automobile travel times in the exclusive lane may be due to a number of factors, including the concentration of bus travel during the more heavily congested portion of the peak period and the difference in general maneuverability between the two classes of vehicles. It is interesting that, during the afternoon peak period, the average bus travel time with two-person car-pool operation was nearly the same as the average automobile travel time in the general lanes. This suggests that, with the reduced car-pool requirement, the system fell into user equilibrium during the congested portion of the peak period, and therefore the potential benefits of the exclusive lane did not materialize during that period.

Trip comfort was measured in this study as speed noise. This measure, defined as the coefficient of variation of individual vehicle speeds, provides an indication of the variability of speed as the vehicle proceeds along the route. A trip that is made at constant speed will experience no speed noise. A value that exceeds 1.0 generally reflects a noticeably stop-and-go type of operation.

Speed-noise measurements were carried out for automobiles using both the general and the reserved lanes on I-95 during both peak periods. The results followed the same pattern as the travel-time studies; i.e., no statistically significant differences were observed, except in the general lanes during the afternoon peak period when speed noise was reduced by 35 percent with the two-person car-pool regulations. This indicates that a generally more comfortable trip was experienced under this condition.

Bus travel-time measurements were made by direct observations of departure and arrival times. It was not therefore possible to provide a quantitative speed-noise comparison. Some deterioration in transit-passenger trip comfort would, however, be anticipated during the afternoon peak period because of the increased travel times experienced by buses when the car-pool regulations were relaxed.

The reduction in passenger-occupancy requirements for the exclusive lane changed the definition of a violator substantially. A reduction in violation rates would, therefore, be anticipated.

A comparison of car-pool volumes, violator volumes, and total traffic volumes violation rates is given in Table 2, and a comparison of noncompliance ratios for both peak periods is given below.

Time Period	Violation Rate (%)		Noncompliance Rate (%)	
	Three PPV	Two PPV	Three PPV	Two PPV
Morning peak	63	36	2.8	5.8
Afternoon peak	63	37	3.1	7.3

The violation rate is the percentage of vehicles using the reserved lane that are ineligible to do so. The noncompliance rate is the percentage of all ineligible vehicles that use the reserved lanes. Thus, these measures are two methods of measuring violations.

Table 3. Comparison of time and distance required for entry to and exit from exclusive lane.

Item	Weaving Time (s)			Weaving Distance (m)		
	Three PPV	Two PPV	Change (%)	Three PPV	Two PPV	Change (%)
Morning peak						
Entry	46	36	+22[a]	732	671	+8
Exit	46	53	-15	792	762	+4
Afternoon peak						
Entry	62	45	+27[b]	1006	701	+30[b]
Exit	47	29	+38[b]	945	457	+52[b]

Note: 1 m = 3.3 ft.
[a]Comparison between stages significant at 95% level.
[b]Comparison between stages significant at 99% level.

Table 4. Summary of system performance measures on I-95.

Performance Measure	Morning Peak		Afternoon Peak	
	Three PPV	Two PPV	Three PPV	Two PPV
Vehicle demand, km	171 219	194 230	199 702	237 035
Vehicle travel time, h	2 630	2 960	3 210	3 250
Avg vehicle speed, km/h	65.2	65.7	62.3	72.9
Passenger demand, km	227 380	255 380	279 840	344 047
Passenger travel time, h	3 450	3 810	4 440	4 720
Avg passenger speed, km/h	66.0	66.9	62.9	72.9
PMI, km/vehicle-h	86.6	86.6	87.2	105.7
HOV priority index	1.011	1.020	1.012	1.000

Note: 1 km = 0.6 mile.

The same trends were evident during both peak periods; specifically,

1. Car-pool volumes more than quintupled,
2. Violator volumes increased by an average of 97 percent, and
3. Violation rates decreased by an average of 41 percent.

The net result was an appreciable increase in total exclusive-lane volumes (approximately 24 percent on the average), indicating substantially greater use of the exclusive lane. All of the comparisons given above were statistically significant at the 99 percent level.

One of the potential problems of the HOV priority-lane concept is the difficulty of crossing several congested lanes of traffic to gain access to the priority lane. Studies were carried out to assess the degree of difficulty of the weaving maneuver under both conditions of car-pool definition. The measures of effectiveness, obtained by moving-vehicle studies that used an instrumented vehicle, were the time and distance required to complete the weaving maneuver. The entry movements were studied downstream of an entrance ramp in the most congested area of the freeway during each peak period, and the exit movements were studied upstream of the last exit ramp in the system (these are the locations where the majority of weaving activities are concentrated).

The results, as summarized in Table 3, indicate that reducing a car-pool requirement from three to two PPV significantly decreased both the time and distance required for executing the lane-changing maneuver during the afternoon peak. The morning peak showed a slight reduction in the time necessary to complete the weaving maneuvers but not the distance.

There appears, therefore, to be a strong indication that, during the evening peak, a reduction in car-pool requirements from three to two PPV altered the lane distribution to the point that weaving maneuvers were significantly easier to perform. This conclusion is based on the significantly lower times and distances required to perform weaves from an entrance ramp to the exclusive lane or from the exclusive lane to an exit ramp.

The same phenomenon did not hold true for the morning peak when the times and distances associated with weaving across the freeway generally showed no statistical differences. The lack of differences during the morning period was due primarily to the fact that the change in car-pool definition had generally little effect on the morning peak operations.

Effect on System Operating Characteristics

The operating characteristics were compared for the two stages of the pool demonstration project. To develop these comparisons, field data were collected to determine

1. Average traffic volumes on I-95 during each of the peak periods,
2. Average passenger occupancy for exclusive-lane automobiles and automobiles traveling in the general lanes,
3. Travel times for each mode of travel, and
4. Bus passenger volumes.

From the field data, the following measures of effectiveness were calculated for each peak period:

1. Total vehicular demand on the freeway (vehicle kilometers),
2. Total passenger demand (passenger kilometers),
3. Total vehicular travel time on the freeway (vehicle hours),
4. Total passenger travel time on the freeway (passenger hours),
5. Average vehicle speed (vehicle kilometers divided by vehicle hours),
6. Average passenger speed (passenger kilometers divided by passenger hours),
7. Passenger movement index (PMI) (passenger kilometers divided by vehicle hours), and
8. HOV priority index (average passenger speed divided by average vehicle speed).

The vehicle and passenger speeds are relatively simple from a conceptual point of view. The PMI is defined

Table 5. Accident analysis data.

Item	Morning Peak Three PPV	Morning Peak Two PPV	Afternoon Peak Three PPV	Afternoon Peak Two PPV
No. of accidents	75	33	92	27
No. of days	211	101	211	101
No. of accidents per day	0.36	0.33	0.44	0.27
Vehicle demand, 10^6 km	34.1	18.8	30.3	17.4
No. of accidents per million vehicle kilometers	2.20	1.75	3.04	1.55
Passenger demand, 10^6 km	45.4	24.3	42.8	25.3
No. of accidents per million passenger kilometers	1.65	1.36	2.15	1.07

Note: 1 km = 0.6 mile.

for purposes of this study as the number of passenger kilometers of travel per vehicle hour of travel time. It is suggested that this measure provides the most meaningful relationship between the service provided by the facility, in terms of passenger throughput, and the cost of providing that service, in terms of traffic congestion.

Another derived measure of effectiveness is the HOV priority index. This measure is defined for purposes of this study as the ratio of average passenger speed to average vehicle speed. An HOV priority index of 1.0 would indicate that no travel-time advantage was experienced by high-occupancy vehicles. To achieve an index greater than 1.0, it would be necessary to move vehicles carrying larger number of passengers at higher speeds than vehicles carrying fewer occupants.

The results of each of the operational stages are summarized in Table 4. In general, the system performance measures were not changed substantially in the morning peak period. The HOV priority index increased by 1 percent, and the PMI was unchanged. The improvement in the HOV priority index resulted primarily from the ability of the system to accommodate the transfer of additional two-person vehicles to the priority lane during this period without adversely affecting speeds in the lane.

In the afternoon peak, the changes were more pronounced. A 21 percent improvement in the PMI for the two-person car-pool stage was observed. This improvement was, however, achieved at the expense of the degree of priority given to HOVs. Thus, the HOV priority index for the two-person car-pool stage was reduced to 1.0, indicating that the system was in user equilibrium. Some advantages were gained by car pools using the priority lane during the afternoon peak, but this advantage was offset by the operating difficulties apparently experienced by the buses, whose scheduled movements tended to concentrate in the more congested portion of the peak period.

The accident rates were examined during the peak periods (using the 2-h peaks for consistency) in terms of accidents per day, accidents per million vehicle kilometers (MVK) and accidents per million person kilometers (MPK). Data were obtained from computerized records of the Dade County Public Safety Department. The data are given in Table 5 (6), and the effects are summarized below (1 km = 0.6 mile).

Item	Change (%) Morning Peak	Change (%) Afternoon Peak
Accidents per day	-8	+39
Accidents per million vehicle kilometers	-20	-49
Accidents per million passenger kilometers	-18	-50

The results indicate a negligible change in the accident frequency during the morning peak, although the accident rates per MVK and MPK decreased by about 20 percent (which was not statistically significant). The more substantial improvement in safety occurred in the afternoon peak when accident frequency decreased by 39 percent and the accident rates decreased about 50 percent (both statistically significant). Overall, the accident frequency and both rates decreased significantly in the combined peak periods. This indicates that the improved quality of general traffic flow in the two-PPV car-pool operation was much safer than that in the three-PPV car-pool operation.

On the other hand, an examination of the accident severity rates indicated that the percentage of all accidents that involved injuries increased from 27 to 39 between the two stages in the morning peak period (there was no change in the afternoon). Although this change does not represent an increase in the injury-accident rate, it does suggest that there was a higher probability of more severe accidents in the two-PPV car-pool operation in the morning peak.

SUMMARY AND CONCLUSIONS

The operating changes on the I-95 bus and car-pool priority system were implemented largely in response to public concern over the apparent underuse of the facility. The initial minimum car-pool requirement of three PPV was based on analyses that demonstrated that no substantial priority for HOVs would materialize if the car-pool definition was set at a lower level (because the analyses assumed no violations and that all car pools would use the priority lane). The same analyses indicated, however, that the lower level would result in a higher passenger-carrying capability due to more effective use of the freeway capacity by lower occupancy vehicles.

Field studies that compared the two operating strategies indicated that the degree of use of the exclusive lane by qualified vehicles was somewhat lower than had been anticipated. This factor has maintained a consistent travel-time advantage in the exclusive lane throughout the morning peak period and through the noncongested portions of the afternoon peak, even with the reduced car-pool requirement. During the more heavily traveled portion of the afternoon period, however, the system falls into user equilibrium (i.e., the general lanes became equally attractive from the user's point of view). The express buses experience particular difficulty under these conditions because their maneuverability is more limited than that of automobiles. Travel times, delays, and overall trip comfort deteriorated during the afternoon peak for HOVs in general and for buses in particular after the car-pool redefinition.

On the other hand, some appreciable benefits have resulted from the reduction in the car-pool-occupancy requirement. Overall travel times and delays were reduced. The passenger throughput per vehicle hour of travel was improved by 21 percent in the afternoon peak period. Lane-changing problems were significantly reduced. The problems of enforcement were greatly alleviated by eliminating the two-person car pool as a violator of the traffic-control regulations, and the accident rates were improved appreciably. Although the two-person car-pool requirement has compromised, to some extent, the HOV priority advantages, it has also improved system operation and safety.

ACKNOWLEDGMENT

We wish to acknowledge the support of the Florida Department of Transportation and the Federal Highway Administration and Urban Mass Transportation Administration, U.S. Department of Transportation, who sponsored this research. All opinions, findings, and conclusions expressed in this publication are ours and not necessarily those of the sponsoring agencies.

REFERENCES

1. Priority Techniques for High-Occupancy Vehicles. Transportation Systems Center, U.S. Department of Transportation, Cambridge, MA, Nov. 1975.
2. R. H. Pratt and J. H. Dupree. A Study of Low-Cost Alternatives to Increase the Effectiveness of Existing Transportation Facilities: Vol. 2—Results of Case Studies and Analysis of Busway Applications in the United States. U.S. Department of Transportation, Jan. 1973; NTIS, Springfield, VA, PB 223 922.
3. H. S. Levinson, W. F. Hoey, D. B. Sanders, and F. H. Wynn. Bus Use of Highways: State of the Art. NCHRP, Rept. 143, 1973.
4. M. J. Rothenberg. Priority Treatment for High-Occupancy Vehicles. Federal Highway Administration, Rept. FHWA-RD-77-56, March 1977.
5. J. A. Wattleworth and others. Evaluation of the NW 7th Avenue Express Bus System and Bus-Priority Systems: I-95-NW 7th Avenue Bus-Car-Pool Systems Demonstration Project—Phase 1. Transportation Research Center, Univ. of Florida, Gainesville, Repts. 1-9, 1977.
6. J. A. Wattleworth and others. Evaluation of the I-95-NW 7th Avenue Bus-Car-Pool System: I-95-NW 7th Avenue Bus-Car-Pool Systems Demonstration Project—Phase 2. Transportation Research Center, Univ. of Florida, Gainesville, Repts. 1-4, 1978.
7. K. G. Courage and others. Traffic Control of Car Pools and Buses on Priority Lanes. Transportation Research Center, Univ. of Florida, Gainesville; Federal Highway Administration, Rept. FHWA-RD-77-148, 1977.
8. R. D. Deuser and M. Lopatin. I-95 Exclusive-Lanes Demonstration Project: Project Review—December 1, 1976. Florida Department of Transportation, Internal working paper, Dec. 1976.

Discussion

B. Tom Shin, McMaster University

It was a disappointment to see that the minimum occupancy that qualified an automobile as an HOV was lowered from three to two in the I-95 project, thus weakening its persuasive power to achieve the original objective of increasing transit patronage and car pooling. The disappointing move seems to be justified by Courage and others through an extensive evaluation study. However, the evaluation technique used does not seem to agree with the original project objective.

The most important point to be investigated in the evaluation process would be the impact on transit ridership of the move. Also, periodic counts of parked vehicles at the park-and-ride lot during the study period would be of interest. The parked vehicles are indeed the results of such a project. Drivers will leave their automobiles at the parking lot and take transit if the project measures are effective but take the reverse course if the measures are not convincing. The number of vehicles classified by the number of occupants before and after the change will be another significant statistic. Similar to the decision to take transit, occupants of low-occupancy vehicles may decide to form an HOV if such a tactic is perceived advantageous. A survey of the drivers of the parked vehicles and the occupants of HOVs might explain why and how they became project supporters, and the findings of the survey could suggest the future directions toward which the project effort should be aimed.

These are the points one would expect in an evaluation of this type. Unfortunately, none of them were discussed.

Courage and others showed that, by lowering the minimum requirement for an HOV from three occupants to two, more effective use of the freeway capacity had been achieved. Specifically, they say that "car-pool volumes more than tripled; violation rates decreased by 45 percent." These changes, according to them, were statistically significant at the 99 percent level. These seemingly striking statistics are in fact meaningless and can easily mislead readers. The changes occurred simply because vehicles carrying two persons on the HOV lane were counted as violators in one case and as legitimate users in the other. This does not at all mean that the average vehicle occupancy increased.

Generally, Courage and others were interested in the impact on operating conditions when different definitions of HOV were used rather than the impact of road-user behavior in choosing their mode. They investigated the numbers of vehicles and persons that passed the study location and the level of service that was provided. However, the fundamental objective of the I-95 project should be, rather than to provide the highest total vehicular capacity on the freeway section, to increase the average occupancy of automobiles and buses by discriminating against low-occupancy vehicles. This could be achieved through frustration of the users of low-occupancy vehicles, but such frustration will not necessarily occur simultaneously with the overall highest use of the road surface. One more frustrated automobile driver who decides to leave his or her vehicle at the park-and-ride lot and take transit is a greater sign of success of the project than is one more vehicle passed through during a peak period.

The other parameters used—the PMI and the HOV priority index—do not seem to be relevant for this evaluation task. The PMI is meaningful only when the minimum-occupancy requirement is unchanged. The use of this factor in evaluating the impact of the change of occupancy requirement is unfortunate. For example, if the project is completely abandoned and vehicles are not differentiated, the overall vehicle speed will be increased because of the wider use of the HOV lanes. This apparent failure will yield an increased value of PMI. At the same time, both vehicle and passenger speeds will be increased and the HOV priority index will not explain the impact clearly. Even when the minimum occupancy is held fixed, the conversion of low-occupancy vehicles to

HOVs will cause an increase of speed not only of the new HOVs but also of the remaining low-occupancy vehicles because there will be fewer vehicles on the general lanes. Therefore, this situation will not necessarily increase the HOV priority index.

Finally, the high number of violators on the HOV lane (78 percent in the afternoon peak periods) and the low rate of legitimate users on it (only 23 to 37 percent of the total qualified vehicles) make one think that stricter enforcement might have been more effective than lowering the minimum-occupancy requirement in making the I-95 project successful.

Authors' Closure

We appreciate the comments of Shin on our paper. Although some of his comments are well taken, many of them are addressed to points that are beyond the scope of our paper, which was expressly limited to the operating effects of the change in the minimum car-pool occupancy requirement on traffic stream characteristics. Many of these comments are addressed directly in the reports that were prepared as part of the complete evaluation of the demonstration project. In particular, the papers by Wattleworth and others (5, 6) and by Courage and others (7) will answer many of his comments.

Shin inquired about many important measures that were not included in the original paper; these will be presented in summary from here. He asked about the use of the park-and-ride lot and about transit ridership. The average number of vehicles using the lot and the average number of morning peak-period bus passengers are given below for each operational condition.

Condition	Average No. of Vehicles Using Parking Lot	Average No. of Express Bus Passengers
Base	418	726
Three PPV	464	816
Two PPV	525	870

It can be seen that both of these transit-use measures continued to improve when the car-pool definition was reduced from three to two PPV. Thus, the change in car-pool definition did not seem to have a serious effect on the use of the transit system.

Shin points out that the reduction in violation rates in the reserved lane from 63 percent when the three-PPV car-pool definition was used to 37 percent when the two-PPV car-pool definition was used represents, at least partially, the change in the base of vehicle types on which the definition of violator is based. This is certainly a valid observation and, in fact, Table 2 shows that the volume of violators actually increased when the car-pool definition was changed from three to two PPV. However, the violation rate is still an important measure from two points of view, enforcement and system operations. If the violation rate is high, it breeds disrespect for the system, which can lead to a further increase in violators and may foster disrespect for other traffic regulations.

Shin states that the objective of an HOV priority system should be to increase the average occupancy of automobiles. However, the average automobile occupancy is of less significance from a system operation point of view than is the passenger movement capability, and the low passenger movement capability when the three-PPV definition was used was the major reason for redefining the minimum car-pool requirement for the reserved lane. Table 6, Figure 6, and the percentage of total vehicles that are single-occupancy vehicles (SOVs) given below point this out quite dramatically.

Item	SOVs (percentage of total)	Item	SOVs (percentage of total)
Morning peak		Afternoon peak	
No. of vehicles		No. of vehicles	
Base	79.5	Base	76.1
Three PPV	76.6	Three PPV	70.7
Two PPV	79.0	Two PPV	68.2
No. of passengers		No. of passengers	
Base	62.0	Base	56.2
Three PPV	57.3	Three PPV	49.7
Two PPV	61.0	Two PPV	46.7

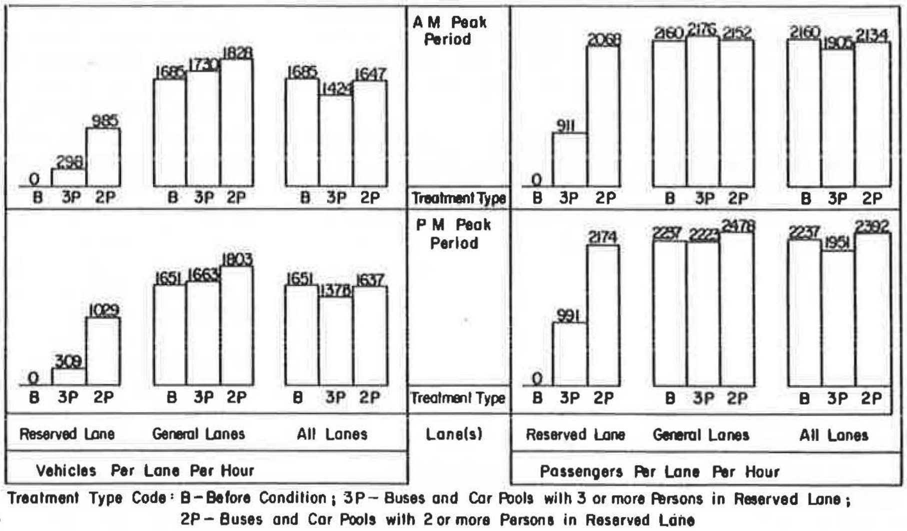
1. When three-PPV operation was used, the total number of automobiles that were eligible to use the reserved lane was extremely small (611 vehicles/2 h in

Table 6. Summary of vehicular and passenger movements.

		Morning Peak: I-95 Southbound						Afternoon Peak: I-95 Northbound					
		No. of Vehicles			No. of Passengers			No. of Vehicles			No. of Passengers		
Lanes	Type of Vehicle[a]	Base	Three PPV	Two PPV	Base	Three PPV	Two PPV	Base	Three PPV	Two PPV	Base	Three PPV	Two PPV
Reserved	SOV	0	225	703	0	225	703	0	241	759	0	241	759
	2OV	0	139	1 058	0	278	2 116	0	135	1 021	0	270	2 042
	3OV	0	211	187	0	691	613	0	219	254	0	734	849
	Buses	0	21	21	0	628	704	0	23	23	0	656	697
	Total	0	596	1 969	0	1 822	4 136	0	618	2 057	0	1 981	4 347
General	SOV	9 827	9 967	11 456	9 827	9 967	11 456	9 215	8 820	9 662	9 215	8 820	9 662
	2OV	2 185	2 325	1 656	4 370	4 670	3 312	2 230	2 846	2 789	4 460	5 692	5 778
	3OV	334	400	305	1 111	1 338	1 029	648	541	782	2 175	1 806	2 751
	Buses	21	0	0	548	0	0	24	0	0	572	0	0
	Total	12 367	12 705	13 417	15 856	15 975	15 797	12 117	12 207	13 233	16 422	16 320	18 191
All	SOV	9 827	10 192	12 159	9 827	10 192	12 159	9 215	9 061	10 421	9 215	9 061	10 421
	2OV	2 185	2 474	2 714	4 370	4 948	5 428	2 230	2 981	3 810	4 460	5 962	7 620
	3OV	334	611	492	1 111	2 029	1 642	648	760	1 036	2 175	2 542	3 600
	Buses	21	21	21	548	628	704	24	23	23	572	656	697
	Total	12 367	13 298	15 386	15 856	17 797	19 933	12 117	12 825	15 290	16 422	18 221	22 338
	Percent SOV	79.5	76.6	79.0	62.0	57.3	61.0	76.1	70.7	68.2	56.1	49.7	46.7

[a]SOV = single-occupant vehicle, 2OV = vehicle that has two occupants, and 3OV = vehicle that has three or more occupants.

Figure 6. Productivity of reserved lane, general lanes, and all lanes.



the morning and 760 vehicles/2 h in the afternoon). Most people would consider these numbers too small to justify a reserved lane.

2. When two-PPV operation was used, the number of eligible automobiles that actually used the reserved lane was 1245 vehicles/2 h in the morning and 1275 vehicles/2 h in the afternoon.

3. The number of persons moved in the reserved lane was about 1900 persons/2 h when the car-pool definition was three or more, and the number of persons moved increased to about 4250 persons/2 h when the minimum car-pool requirement was two persons.

4. The number of persons moved in all lanes of I-95 was about 18 000 persons/2 h when the car-pool definition was three and increased to about 21 000 persons/2 h when the car-pool requirement was two persons.

Thus, if one considers the person-moving capability of the freeway, it can be seen that the freeway (and the reserved lane) was able to move many more persons when the lower car-pool requirement was used. Because the primary function of a freeway (or other transportation system) is to move people, it is believed that this is a significant point.

Figure 6 presents some of this information on a basis of persons moved per lane per hour. Some of the significant points from this figure include the following.

1. When the three-PPV car-pool definition was used, the reserved lane carried only about 300 vehicles/h and about 950 persons/h, and the general lanes carried 1700 vehicles/lane/h and about 2200 persons/lane/h. Thus, the productivity of the reserved lane was much less than that of the general lanes.

2. When the car-pool definition was changed to two PPV, the reserved lane carried about 1000 vehicles/h and about 2125 persons/h. This represents a significant increase in productivity.

3. When the car-pool definition was three PPV, the general lanes averaged about 1700 vehicles/lane/h and about 2200 persons/lane/h. When the car-pool definition was reduced to two PPV, the general lanes averaged about 1800 vehicles/lane/h and about 2300 persons/lane/h.

Thus, the reduction in the car-pool definition in the reserved lane increased the productivity of both the reserved lane and the general lanes.

In view of these results, it is difficult to conclude other than that the reduction in the minimum car-pool requirement for reserved-lane use from three to two PPV produced a much more efficient operation from almost every practical viewpoint. Perhaps from an academic point of view, the change represented some philosophical sacrifices but from a practical, engineering point of view the changes were extremely beneficial.

Publication of this paper sponsored by Committee on Freeway Operations.

**Mr. Culpepper is currently with the Department of Civil Engineering, Auburn University.*

Abridgment

Methodology for Evaluation of Alternative Low-Cost Freeway Incident Management Techniques

Gary L. Urbanek, Peat, Marwick, Mitchell, and Company, Washington, D.C.
Samuel C. Tignor, Federal Highway Administration

In some cities, the problems caused by recurrent congestion have been attacked through the implementation of high-technology systems that use electronic loop detectors and closed-circuit television. The high costs of these solutions, however, make them feasible for only the largest freeway systems; in most places, less expensive approaches must be considered. This paper presents a methodology that may be useful for planning low-cost management actions to reduce nonrecurrent traffic delays and congestion caused by freeway incidents. The methodology components include estimating (a) total incident rate; (b) delay-causing incidents; (c) detection, response, and clearance times; and (d) total delay. The final step provides for the selection of appropriate low-cost freeway incident management (FIM) options and solutions.

ILLUSTRATION

To illustrate the magnitude of the FIM problem, consider a six-lane facility, 16 km (10 miles) long, that has shoulders, an interchange spacing of 2.4 km (1.5 miles), a rush-hour volume split of 5000 vehicles/h in the heavy direction and 2600 vehicles/h in the light direction, and two identical rush-hour periods. Our research has shown that the estimated annual incident rate for this facility is 10 000 incidents/year during the rush periods (1, 2, 3, 4, 5, 6).

If taken a step further, it can be shown that a subset of these incidents (the delay-causing incidents) will result in an estimated 550 000 vehicle-h of delay and the wasting of 1 362 711 L (360 000 gal) of gasoline.

INCIDENT-RATE DETERMINATION

As would be expected, systems that have better detection mechanisms identify more incidents and, thus, report greater occurrence rates. This suggests that, to determine an absolute rate, it is necessary to use the results of a continuous detector or of closed-circuit television surveillance. By using and interpreting data obtained on the John C. Lodge Expressway in Detroit (7, 8), it is possible to determine that approximately 124 incidents/million vehicle km (200 incidents/million vehicle miles) is a reasonable estimate of the occurrence rate. [A derivation of this can be found elsewhere (2)].

However, not all incidents cause delay. An incident is a delay-causing incident if it is (a) in-lane and requires a response or (b) an accident that either occurred on the shoulder or has been removed to the shoulder. By using this definition, it is possible to derive an incident distribution that can be used to determine delay-causing incidents, given the total incident rate.

Based on the incident distribution given elsewhere (2), the 10 000 incidents of the illustration are probably distributed as follows:

1. On the traveled way, 400 incidents will occur; of these, in 120 cases, the problem will be solved by the driver and 280 will require assistance in the form of a fire vehicle, an ambulance, or a tow truck.
2. There will be 60 accidents.
3. There will be 220 disablements (out of gasoline, blown tires, fires, out of water, transmission failures, and such).
4. There will be 9600 incidents that make their way to the shoulder.
5. There will be 400 accidents, of which 230 will require assistance from wreckers, ambulances, or fire services.
6. There will be 9200 disablements.

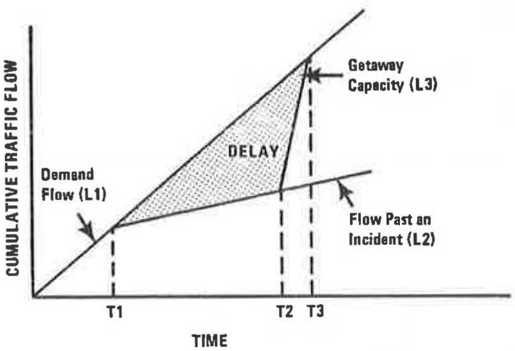
From this distribution and a few simple calculations, it can be shown that about 5 percent of the total number of incidents (i.e., the 280 incidents and the 230 accidents) cause delay. On the surface, this might give the impression that freeway incidents are only a small national problem; however, it has been estimated that the delay caused by urban delay-causing incidents in the United States is about 750 000 000 vehicle-h/year (9) or an amount equivalent to the gasoline produced from the oil flowing through the Alaskan pipeline for 10 d (10).

QUANTIFYING DELAY

To step from the number of delay-causing incidents to the vehicle hours of delay requires an understanding of what takes place when an incident occurs. Graphically, the FIM delay problem can be visualized as indicated in Figure 1, in which the ordinate is defined as the cumulative traffic volume and the abscissa is a time axis. L1, therefore, is a measure of vehicles per hour and can be labeled to represent the total number of vehicles wishing to use the freeway as the demand flow. When an incident occurs, the capacity of the freeway at the incident site is reduced until the incident is cleared, as is indicated by L2. L3 represents the getaway capacity as the maximum rate at which the vehicles behind the blockage can leave the backup. Graphically, T1 is the beginning time of the incident, T2 is the point in time at which the clearance of the incident was completed, and T3 is the point in time at which the last vehicle in the queue begins normal traffic-flow speed. Therefore, the shaded area is a measure of the vehicle hours of delay caused by incident; in terms of this graphic, the object of FIM is to reduce the shaded area.

The theoretical aspect of this model is intuitively appealing; however, its application to the FIM environment is difficult because of the lack of research that has been conducted relative to certain

Figure 1. Impact of incident on traffic flow.



values. In particular, values of the flow past the incident, the duration of the critical T1-to-T2 time period, and the getaway capacity are not widely known. Goolsby's work (4) has been particularly helpful in determining approximations of the various flow rates, although his information is based on a relatively small sample. Nonetheless, it is one of the few published research efforts in this respect and has been used to estimate the flow rates for the example.

Similarly, relatively few values have been published of typical incident durations and those that are available are for specific sites or freeways and vary widely (11, 12) because they are site-specific.

The fact that incident duration is site-specific dictates that a method of determining these values for the delay equation take into account local conditions. In addition, numerous interviews with police officials have indicated that major incidents (such as an overturned and burning tanker) are reported much more quickly than are minor incidents (such as a flat tire in the roadway). This fact and the fact that there are relatively few major incidents suggest also that the methodology should be capable of addressing minor and major incidents separately. Here, the method used to estimate major incidents is simply to project historical information. However, for the much larger minor-incident problem, each component of the duration of the incident (i.e., detection time, response time, and clearance) is calculated separately.

TIME COMPONENTS OF DELAY

As shown in Figure 1, the duration of the incident is T2 minus T1. This time period consists of an incident-detection time, a response time, and a site-clearance time.

1. Detection time is measured from the moment an incident occurs to the moment when it is detected by or reported to an official agency that has incident management responsibilities. Detection time includes related activities such as recognition and notification.
2. Response time is measured from the moment when the official agency becomes aware of the occurrence of the incident to the moment when all of the resources necessary to effect clearance have arrived at the incident site. There is great variability in the length of this time interval. It normally includes activities such as on-site evaluation of incident severity, communication, and the travel time of emergency vehicles and personnel to the incident site.
3. Clearance time is defined as the time required to remove an incident from the freeway and restore full freeway capacity. It begins as soon as the first response unit arrives at the site and ends when the last unit leaves. Clearance time varies considerably and is a function of type of incident and of the capability and availability of response resources.

It is apparent from Figure 1 that these times should be minimized to reduce total delay. To minimize delay requires that the decision maker have estimates of detection, response, and clearance times. If local estimates of these times are not available, then the techniques discussed elsewhere (2) can be used to estimate them.

LOW-COST FIM SOLUTIONS

After the total delay has been quantified and computed, the final step of the methodology involves examining the potential low-cost FIM options available and developing system solutions to reduce nonrecurrent type delay. About 30 options have been developed to solve various aspects of the FIM problem. These options have been classified into four categories, each of which is defined below with examples.

Surveillance Options

Surveillance options enhance the detection capability of the existing system and directly affect detection time. These options include actions such as increasing the frequency of police patrols and monitoring citizens band radio to capture incident information.

Administrative Options

Administrative options are intraagency activities intended to improve the ability of an agency to fulfill its present incident management role or to expand that role to include new responsibilities. Some of the options that have been developed are an emergency-light policy that reduces indiscriminate use of lights and lessens the natural tendency of an emergency vehicle to cause a general traffic slowdown and the use of accident-investigation sites to investigate accidents off the freeway.

Organizational Options

Organizational options may involve several agencies and seek to facilitate the cooperative effort that characterizes a successful incident management system. Typical options include the development of contracts, ordinances, or agreements to control the predominantly private towing services on urban freeways and the use of an FIM response team (13) (which is a multidisciplinary team of individuals who respond to large incidents to coordinate and direct multiagency activities and expedite the incident removal).

Preplanning Options

Preplanning options are designed to prepare an agency and other participating agencies and their personnel for the eventuality of an incident that requires their resources. Some of these include traffic-operations training and alternative route planning (which involves predetermining alternative routes for certain portions of a freeway system, identifying the routes on maps, and distributing the maps and information to FIM personnel).

PRESENT TESTING

At present, this methodology is in the process of being field tested by a cooperative effort between the Florida Department of Transportation and local agencies in the Tampa-St. Petersburg area. The site of the demonstration is the Howard Frankland Bridge between Tampa and St. Petersburg on I-275. The demonstration has progressed through all of the steps of the methodology and is currently in progress. In the near future, additional data will be collected so that the estimated delay savings can be compared with the actual delay being saved.

REFERENCES

1. G. L. Urbanek and R. W. Rogers. Alternative Surveillance Concepts and Methods of Freeway Incident Management: Vol. 1—Executive Summary. Federal Highway Administration, Rept. FHWA-RD-77-50, 1978.
2. J. R. Owen, J. M. Bruggeman, and G. L. Urbanek. Alternative Surveillance Concepts and Methods for Freeway Incident Management: Vol. 2—Planning and Trade-Off Analyses for Low-Cost Alternatives. Federal Highway Administration, Rept. FHWA-RD-77-59, 1978; NTIS, Springfield, VA, PB 279 497/AS.
3. G. L. Urbanek and J. M. Bruggeman. Alternative Surveillance Concepts and Methods of Freeway Incident Management: Vol. 3—Computational Example for Selecting Low-Cost Alternatives. Federal Highway Administration, Rept. FHWA-RD-77-60, 1978; NTIS, Springfield, VA, PB 282 497/AS.
4. G. L. Urbanek and J.R. Owen. Alternative Surveillance Concepts and Methods of Freeway-Incident Management: Vol. 4—Guidelines for Specific Low-Cost Alternatives. Federal Highway Administration, Rept. FHWA-RD-77-61, 1978; NTIS, Springfield, VA, PB 279 498/AS.
5. G. L. Urbanek and J. R. Owen. Alternative Surveillance Concepts and Methods of Freeway Incident Management: Vol.5—Training Guide for On-Site Incident Management. JHK and Associates, Alexandria, VA; Federal Highway Administration, Rept. FHWA-RD-77-62, 1977; NTIS, Springfield, VA, PB 279 499/AS.
6. J. M. Bruggeman and G. L. Urbanek. Alternative Surveillance Concepts and Methods of Freeway Incident Management: Vol.6—Delay Time and Queue Tables for Trade-Off Analyses. Federal Highway Administration, Rept. FHWA-RD-77-63, 1978; NTIS, Springfield, VA, PB 284 781/AS.
7. F. DeRose, Jr. An Analysis of Random Freeway Traffic Accidents and Vehicle Disabilities: Freeway Operations. HRB, Highway Research Record 59, 1964, pp. 53-65.
8. S. E. Bergsman and C. L. Shufflebarger, Jr. Shoulder Usage on an Urban Freeway. John C. Lodge Freeway Traffic Surveillance and Control Research Project, Detroit. Study 417, 1962.
9. H. Lunefeld and others. Postcrash Communications. U.S. Department of Transportation, National Highway Safety Board, Rept. DOT-HS-800-289, July 1970.
10. Now That the Pipeline's Almost Built, Who's Going to Take the Oil? National Journal, Washington, DC, Vol. 8, No. 50, Dec. 11, 1976, p. 1764.
11. M. E. Goolsby. Influence of Incidents on Freeway Quality of Service. HRB 349, 1971, pp. 41-46.
12. W. R. McCasland. Experience in Handling Freeway-Corridor Incidents in Houston. TRB, Special Rept. 153, 1975, pp. 145-155.
13. A. Gianturco. Incident Response Team Brings Order out of Chaos. Public Works, Vol. 108, No. 2, 1977, pp. 51-53.

Publication of this paper sponsored by Committee on Freeway Operations.

Abridgment

Traffic-Condition Grade: Evaluation of Concept

Robert J. Benke, Traffic Engineering Section, Minnesota Department of Transportation

The quality of service as perceived by drivers in a traffic stream is a function of how they perceive the various traffic-flow characteristics. Because the mental and physical attributes of drivers vary, different drivers will perceive the same condition as being of a higher or a lower relative quality on some idealistic, undefinable scale. Drivers differ in their degree of acceptance of slower travel times, heavy traffic volumes, and unpredictable events. They also differ in their attitudes toward the driving task itself.

Despite the variability of driver attitudes and characteristics, several assumptions can be made relative to how quality of flow is perceived by most drivers. The first of these is that, for each situation experienced, there will be a median perceived value of quality; half of the drivers will rate the instance lower and half of the drivers will rate it higher on the scale. A correlated requirement is that the surrogate scale must be understood and accepted by the highest proportion of drivers possible. In this situation, actual measures of the differences among drivers need not be known. Rather, we can assume that most drivers will be reasonably close to the median estimate of the quality of flow.

A second assumption (simplistic but essential) that can be made regarding perceived quality is that most drivers recognize smooth, fast flow as evidence of a

Figure 1. Traffic-condition grade sign.

good quality of service. Conversely, they perceive congestion conditions to be symptomatic of a poorer quality of service. From this assumption follows the ability to anchor the surrogate quality scale to actual measurements of various traffic-flow parameters. The problem then becomes one of calibrating the surrogate scale so that relative placement on it is correlated with what most drivers perceive the real-world quality of service to be. This calibration process is complicated by the fact that standard technical terms are not understood by many drivers. Terms and phrases such as capacity, lane-occupancy percentages, density, and volume-to-capacity ratio are usually meaningless to them. Likewise, average delay to be encountered or ratings of heavy congestion are too broad to be of more than general applicability.

The scholastic grading system of ABCDF is proposed as a quality scale that is simple and easily understood by most people and might be adaptable to this need. It remains then to determine the traffic flow that can be rated A (or excellent) and that which should be considered F (complete failure). Because most computerized traffic-control systems use the lane-occupancy percentages and volumes in their control algorithms, it is logical to base the various grade levels on these measures, just as scholastic grades are based on test scores and the quality of work done. If the grade limits are defined and validated by actual experience, then the traffic-condition grade can be determined quantitatively and serve as a surrogate measure of the qualitative level of acceptability of the traffic flow.

SIGN SYSTEM DESCRIPTION

Figure 1 illustrates the signs by which the traffic grade is presented to drivers approaching the two study-area ramps. In-place structures were used, and the TRAFFIC CONDITION legend was added to the route guide sign. A single 36-cm (14-in), seven-by-five-bulb matrix module is used to display the grade. The sign is controlled by a direct wire connection to the traffic management center. The estimated cost of the three signs was $9500 each, including modifications in the center operations room and the direct-wire communication links. The signs are operated only during the evening peak period (3 to 6 p.m.). The display grade is determined and updated on the basis of average lane-occupancy rates at detector stations on the 9.7-km (6-mile) link of I-35W between the Minneapolis central business district and Hennepin County Highway 62.

The average lane occupancy is compared with threshold values, and a grade is selected for display. The initial grade-change points were selected after comparing system-operator, field-observer, and computer ratings of several weeks' experience to provide a less conservative grading. The current thresholds are shown below.

Grade	Lane-Occupancy Range (%)	Volume (vehicles/min)
A	0-12.0	
B	12.1-16.0	
C	16.1-22.0	
D	22.1-40.0	15
F	22.1-40.0 or >40	15

The traffic grade being displayed is also used by the system operator when reporting traffic conditions to local radio stations, which provides earlier coverage for some motorists.

PREOPERATION PUBLICITY

Before the system was turned on, there was an extensive publicity campaign in the local media to teach drivers the grade meanings. All publicity was generated as news coverage; there were no paid advertisements. The initial coverage included taped interviews and video tape coverage on three of the four local commercial television channels and photographs and descriptions of the signs in areawide, downtown employee, and suburban newspapers.

EVALUATION SURVEY

A mail-back postcard survey was distributed to outbound evening drivers to measure their understanding of the traffic grades and their reactions to the concept. Survey forms were distributed at two entrance ramps on Tuesday, June 21, 1977. A total of 5890 forms were distributed among an estimated 6100 automobiles between 3 and 6 p.m., a 97 percent coverage rate. The forms were designed so that a driver was asked to define the meaning of only one letter, thus eliminating any learning bias. They were arranged sequentially in groups of five so that identical numbers of forms with each letter were distributed. The mix of letters over time and location was therefore constant; however, the letter received by a given driver was determined on a random-chance basis. A total of 2271 surveys were returned in time to be included in the subsequent analysis, a 38.6 percent return and 37.4 percent coverage of all entering drivers.

The returned questionaires were examined to determine whether they were representative of the traffic that had entered the freeway during the survey period. The tests showed that the return rates varied significantly by 30-min time increments ($\chi^2 = 313$); the peak rate occurred between 4:30 and 5:30 p.m. This finding shows that the sample contains, proportionally, too many peak-hour users. However, these drivers are those most frequently faced with congested flow and their response is of greatest concern. In this instance, oversampling during the peak hour does not create a problem. An analysis by letter grade, of the number of surveys returned, showed a uniform distribution, indicating a generally random response and no statistically significant variation.

SURVEY RESPONSES

A tabulation of the responses to the question about grade meaning showed that 1173 of 2257 people (52 percent) correctly identified the correct definition. The correct-response rate was variable, probably because of the word-choice options presented. In the case of the grades A, B, and C, a second response can be considered correct, depending on the person's perceived value scale. For example, some persons might consider a B to be a good grade relative to a D or F but others might consider it to be a fair grade relative to an A. These almost-correct answers illustrate the difficulty of agreeing on commonly defined descriptions of traffic conditions. In retrospect, further pretesting of the questionnaire was warranted.

The inclusion of almost-correct answers as correct responses results in a comparable percentage correct for grades A, B, and D but not grades C and F. The lower percentage correct is correlated with a higher degree of uncertainty and a higher selection rate for the alliterative word option (F = fair) for grade F. The grade C difference resulted from a 14.8 percent response to the poor option. A second tabulation (given below) of the responses of persons who use the freeway at least 1 time/week between 3 and 6 p.m. showed a slightly more accurate response.

Grade	N	Correct (%)	Almost Correct (%)	Total Correct (%)
A	383	45.7	33.2	78.9
B	379	39.3	31.4	70.7
C	392	55.0	7.9	62.9
D	377	77.5		77.5
F	400	60.3		60.3
Total	1931	54.2	14.2	68.4

The non-peak-hour users were less sure and less accurate in their responses.

Subsequent analysis of the accuracy of peak-hour users showed that the media coverage had played a significant role in educating them. More than 80 percent of those who indicated that they had heard or read about the grade signs were able to correctly define the meanings (see below).

Coverage	N	Total Correct (%)
Newspaper only	151	83.4
Radio only	288	73.3
Television only	222	84.2
Subtotal	661	79.3
Two of three media	279	80.3
All media modes	216	84.3
Total media	1156	80.4
None	793	50.8

Only half (51 percent) of those who had not read or heard about the signs gave correct responses. However, this level of response based only on exposure to the sign suggests that there is a latent ability to decipher the code or to guess the correct meaning. The high level of response accuracy, given the rather limited press coverage, suggests that, with additional education, understanding of the grade concept could be improved to include almost all peak-hour and most casual users of the system.

When asked about their typical use of the traffic grade signs, 51 percent of the peak-hour drivers who correctly defined the letter grade replied that the signs alerted them to problems ahead and 30 percent used the grades to make route choices (see below).

Sign Use	Commuters (%) (N = 1692)	Others (%) (N = 1077)
Route choice	29.8	14.1
Alerts to hazard	50.6	26.8
Of no use	17.8	22.7
Do not understand sign	1.8	36.4

Most of the drivers who did not use the signs responded that they could not because they did not understand them or found them of no use. Many of the commuters who said that the signs were not useful to them indicated by written comments that they were more or less forced to use the freeway for some reason so that the signs were irrelevant to them.

In retrospect, the high rate of response to the two use options (route choice and alert to hazard) may have been a result of the structure of the question and the options presented. Attempts to validate the route-choice rate have not been attempted thus far because of the infrequency of the use of the grade F. The grade D is not expected to divert a measurable number of drivers because most drivers will still prefer the freeway over alternative routes. The study structure was not designed to provide a validation check of the response to the alert-to-hazard use. It is reasonable to assume that, although the rates of use for route choice and alert are not high, there are still a significant number of drivers who understand and use the traffic-grade information. The problem that remains is to determine more accurately the true level of use and the level of use needed to make the sign system (or any driver information system) an effective concept.

The reactions of the commuter drivers who gave correct responses to the traffic-grade concept were very positive (see below).

Reaction	Commuters (%) (N = 1333)	Others (%) (N = 938)
Like it (as is)	48.6	24.3
Like it (needs work)	27.9	26.5
Not sure	8.9	14.8
Do not like it	6.5	6.8
Stop work on it	3.4	6.3
No answer	4.8	21.2

Only 9.9 percent did not like the concept and 76.5 percent did. The other drivers (non-peak-hour users or those who gave incorrect definitions) were less impressed but were still positive in their reaction.

CONCLUSIONS

Concept Validity

The survey findings demonstrate that a significant majority of drivers can correctly interpret the meanings of the traffic-condition grades A, B, C, D, and F. This ability could be improved by more effective media publicity. The survey also demonstrates that many drivers use the information communicated by the grade in their subsequent route-choice process and freeway travel. We can therefore conclude that the criteria of understanding and use are met and that the traffic-grade concept is valid.

Driver Reaction

The survey also demonstrates that drivers generally like the traffic-grade concept but that some shortcomings should be corrected. Specific areas of concern include the following:

1. The use of a conservative grade-selection algorithm creates some credibility problems among drivers.

2. The freeway zone represented by the sign is a composite of two areas that have different operating characteristics, which means that some drivers use the freeway without encountering the situation that determines the displayed grade and creates the impression that the signs are unreliable.

3. The locations of the three signs are close to the entry points to the freeway but still allow for diversion to surface streeets. Driver requests for locating the signs further away are reasonable. However, the abundance of in-place signs, communication problems, and the number of additional signs that would be needed to provide earlier warning on all routes mitigate against major system expansion. The use of a flashing mode for grades D and F has not been evaluated but does provide partial relief.

4. The responses of most drivers were positive and encouraging. The greatest shortcoming identified was the failure to reach all drivers with an explanation of the grade system. Additional efforts to reach both peak-hour and off-peak-hour users are required.

Publication of this paper sponsored by Committee on Freeway Operations.

Predicting Effectiveness of Transportation System Management Measures

Mark H. Scheibe, Puget Sound Council of Governments, Seattle
Neil J. Pedersen, JHK and Associates, Washington, D.C.

Transportation planning is moving into an era in which the object is to manage the existing system rather than to expand the present facilities. Most transportation planning tools, however, were designed for the latter task. Consequently, many planners are foregoing quantitative analyses when developing plans for management of the transportation system while awaiting the development of new forecasting and analysis tools. This paper reports on a recent study in the Washington, D.C., area that was oriented toward evaluating various measures for possible inclusion in a revised transportation control plan to achieve air-quality standards and used existing models to analyze a wide variety of transportation policy and low-cost improvement measures. The study did not demonstrate that no further model development is needed, but it did show that the quantitative information needed for the development of transportation system management plans can be obtained from existing models.

The planning of transportation services is becoming increasingly complex. On the one hand, many of the transportation improvements that were planned 10 to 20 years ago, such as major freeways and rapid transit facilities, are no longer feasible because of decreasing transportation revenues and increasing citizen opposition. At the same time, air-quality and energy concerns are focusing on the present transportation system as a significant cause of pollution and waste. As a result, many current transportation implementation actions involve low-cost modifications to the existing system and policy changes to encourage the use of high-occupancy modes and discourage the use of modes that are less energy efficient and cause more pollution per passenger.

Most transportation planning tools were not designed to be used in determining methods of managing the transportation system to achieve better efficiency, a major planning task today, but rather to be used in identifying the most effective use of high-cost capital investments in new facilities. As a consequence, transportation planners frequently feel compelled to forego quantitative analysis when developing plans for management of the transportation system while awaiting the development of a new breed of transportation forecasting and analysis tools. This decrease in quantitative analysis is evident in many of the transportation control plans that have been prepared in various metropolitan areas as a result of the U.S. Environmental Protection Agency (EPA) regulations and in many of the transportation system management plans prepared in response to the regulations issued by the Urban Mass Transportation Administration and the Federal Highway Administration. Many of these plans consist of tabulations of actions that are either already programmed or have appealed to the planners and do not have an accompanying assessment of either the overall effect of the entire package of measures or a comparison of the relative effectiveness of the individual measures. In some regions, this has given rise to confrontations between transportation and air-quality agencies in which one side claims that a certain set of measures will allow air-quality standards to be met and the other claims that a harsher set of measures is needed.

Obviously, all differences of opinion between transportation and air-quality professionals will not dissolve as a result of the use of predictive models. However, increased quantitative analysis of both transportation and air-quality phenomena will improve the information available to decision makers. Contrary to the feelings of some in the transportation planning field, the existing models are not totally inadequate for the task of evaluating the effectiveness of transportation policy decisions and low-cost transportation improvements. Disaggregately calibrated behavioral models are designed to replicate travelers' choices when faced with decisions involving time and cost. Most transportation system management measures, even those that have rarely been implemented, ultimately primarily impact the time and cost variables involved in transportation decisions.

An example of the use of existing models for the analysis of transportation management measures can be found in the national capital region transportation and air-quality study (1). The purpose of this study was to analyze in detail the transportation and air-quality implications of existing and alternative elements of the transportation control plan in the national capital interstate air-quality control region. This study was prompted by questions that had been raised regarding the implementation of certain of the transportation control measures included in the plan promulgated by EPA in December 1973 and by the desire to perform a detailed technical analysis of these and other proposed control measures.

The steering and policy body for the study was a management working group composed of a representative from each of the six agencies affected and a nonvoting member who represented the Washington Metropolitan Area Transit Authority.

The evaluation of a transportation control or transportation system management plan or plan modifications potentially involves forecasting base conditions and also changes in travel that could be expected to result from the application of candidate transportation control or system management measures. This study serves as an example of the way in which existing transportation planning methods can be used in forecasting the effectiveness of transportation system management plans.

STUDY CONTEXT

Pursuant to the Clean Air Act of 1970, EPA promulgated national ambient primary and secondary air-quality standards and also requirements that each state submit implementation plans for meeting the primary standards within 3 years of obtaining EPA approval. The regulations stipulated that the state plans set forth

1. Legally enforceable regulations and compliance schedules for implementation of the central strategy,
2. Contingency plans for preventing the occurrence of levels of pollution that would cause significant harm to the health of sensitive receptors,
3. Source surveillance procedures,
4. Procedures (permit systems) to ensure that construction or modification of stationary sources will not interfere with attainment or maintenance of the national standards,
5. Provisions for air-quality surveillance,
6. Descriptions of the resources needed to carry out the state plan, and
7. Provisions for intergovernmental cooperation.

As part of the Clean Air Act, Congress included a number of requirements relating to motor vehicles, the most important of which is a requirement that 1975 light-duty vehicles emit 90 percent less pollutants than 1970 light-duty vehicles. In addition, because it was felt that such measures might be insufficient to enable some cities to meet air-quality standards, Congress required the states to consider land-use and transportation controls in preparing their implementation plans.

In January 1972, the District of Columbia and the states of Maryland and Virginia (hereafter referred to as the three jurisdictions) submitted their plans. On reviewing the plans on a nationwide basis, EPA found them inadequate and allowed the states an additional year for study and making recommendations.

Pursuant to the Administrative Agreement Covering Air-Quality Planning for the national capital interstate air-quality control region, signed by the states of Maryland and Virginia, the District of Columbia, and the Metropolitan Washington Council of Governments, an air-quality planning committee (AQPC) was created and began to hold meetings in February 1972. This committee was given the power to adopt and recommend plans for interjurisdictional cooperation in the improvement of regional air quality, to adopt emergency-episode plans, and to engage in interstate air-quality planning.

The AQPC initially compiled a preliminary list of strategies that might help lead to regional compliance with the national ambient air-quality standards. The staff selected a package of favorable strategies, quantified their impacts on travel costs and related parameters, and conducted sensitivity analyses and macroanalyses to estimate approximate reductions in vehicle travel and emissions. Approximate estimates of fiscal impacts were made to allow assessment of direct costs versus results. On reviewing the tentative plan compiled by the staff, the AQPC made suggestions that they felt would make the plan more acceptable politically, economically, socially, legally, and technologically. The staff then quantified the impacts of the alternative approaches, and the process was repeated.

The transportation control plan as recommended by the AQPC was submitted to the three jurisdictions. They in turn used the AQPC report as a basis for their individual plans, which they individually submitted to EPA in April and May of 1973. In June 1973, the EPA administrator issued approval-disapproval notices, based on his interpretations of the Clean Air Act requirements. The jurisdictions then modified their plans to make them more acceptable to EPA and resubmitted the plans in June and July of 1973.

In August of 1973, EPA proposed a transportation control plan for each of the three portions of the region. The proposals were largely based on the submissions made by the three jurisdictions. Public hearings on these EPA proposals were held in September 1973. Certain changes were made in response to testimony given in the public hearings and, in December of the same year, the transportation control plans were officially issued by EPA in the form of approval of certain state-submitted measures and the promulgation of federal regulations for additional measures [38 Federal Register, 33702-33731 (1973)].

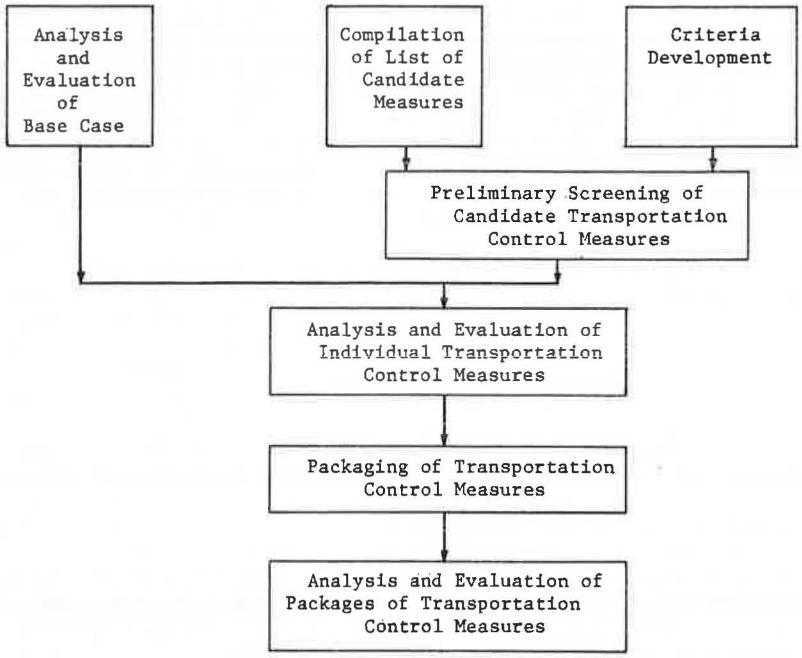
In the period following the EPA issuance of the transportation control plan for the region, a number of political, legal, economic, and technological issues were raised regarding the feasibility of implementing the plan. In addition, a number of the technical assumptions and procedures used in estimating the plan's effectiveness were questioned (2). These questions, combined with a desire on the part of local officials to assess the effectiveness of alternative means of achieving air-quality standards in the region, prompted the national capital region transportation and air-quality study.

ANALYSIS PROCEDURES

The major effort in the study was focused on the analysis and evaluation of a number of transportation control measures and packages of measures that might be candidates for inclusion in a revised transportation control plan for the region, including those measures already in the existing plan. The process used in the analysis and evaluation of the impacts of candidate control measures and packages of measures is outlined in Figure 1.

The first part of the process was divided into three major sections. Travel and air-quality conditions for a 1977 base condition were simulated by using travel-demand procedures and air-quality prediction techniques. A set of criteria was developed for use in the screening and evaluation of existing and proposed transportation

Figure 1. Analysis chronology.



control measures, and a list of candidate control measures was developed by the study management working group. The candidate measures were then screened by using the criteria and reduced to a list of 62 measures to be studied in detail. Most of the measures were analyzed (by using travel-demand modeling procedures) as to their impacts on travel patterns and traffic parameters, such as the number of vehicle trips and the amount of vehicle travel. The analysis gave comparisons among the measures and allowed inferences to be drawn concerning the relative abilities of the measures to reduce vehicle emissions. Other measures, such as inspection and maintenance and vehicular retrofits (which have no direct impact on traffic parameters), were analyzed by using air-pollutant-emission prediction techniques to determine their estimated impacts on vehicle emissions. In addition, the appropriateness of each of the 62 measures was evaluated qualitatively by using criteria addressing the cost, feasibility, and community impact of the measure and its effect on other measures.

After each control measure was individually analyzed and evaluated, the working group combined the measures into four packages of alternatives of varying degrees of harshness and effectiveness. Each package was analyzed in detail by using a process that forecast district-level and localized traffic parameters, which were then used to predict pollutant emissions and concentration. Regional oxidant-prediction methodology was not available at the time in the metropolitan Washington area. The impacts of the packages on oxidant precursors (hydrocarbons and nitrogen oxides) were forecast, however. Each package was also evaluated by using criteria that describe the cost, impact on regional travel accessibility, and effect on energy consumption.

EVALUATION CRITERIA

One of the initial steps in the analysis and evaluation process used in this study, as noted above, was the development of a set of screening and evaluation criteria that could be used in measuring the feasibility of implementation and the effectiveness of existing and proposed transportation control measures. The criteria adopted were then divided into three groups: screening criteria, quantitative criteria, and qualitative criteria. All candidate measures were subjected to a preliminary screening to qualitatively evaluate their potential for emissions reduction and their technological, administrative, and legal feasibility. Those measures that passed this preliminary screening were then carried forward to a detailed evaluation in which the additional qualitative and several of the quantitative criteria were used to gauge impacts. The individual transportation control measures were then combined to form packages of individual measures, and each package was evaluated in accordance with the full list of quantitative criteria.

TRAVEL-FORECASTING PROCEDURES

The travel-forecasting procedures used in this study consist of a chain of existing, calibrated models designed to replicate travel behavior in the Washington region (3). The basic models in the chain and the decisions they are designed to replicate are as follows:

1. Trip generation—the decision to make a specific type of trip (e.g., work or shopping),
2. Trip distribution—the decision to go to a particular destination for the trip,
3. Modal choice—the decision to use a particular travel mode for the trip,
4. Time-of-day choice—the decision to travel at a particular time of day, and
5. Trip assignment—the decision to use a particular route.

The trip-generation, trip-distribution, and trip-assignment models used were those contained in the TRIMS model chain (4). The trip-generation model is a district-level model based primarily on the socioeconomic characteristics of the potential trip maker and the characteristics of the potential destinations of the trip. It forecasts home-based, work-purpose person trips; home-based, shopping-purpose automobile-driver trips; home-based, other non-work-purpose automobile-driver trips; and non-home-based automobile-driver trips. Trip distribution is accomplished by using a gravity model. Trip assignment is done by using a minimum-time-path all-or-nothing assignment that also produces vehicle travel by speed range for each district.

For the work-purpose modal split, the Washington zonal-level modal-choice model (Washington model) (5) was applied to the home-based, work-purpose, person-trip estimate as output by the trip-distribution model. The basic zonal-level modal-choice model consists of two elements. The first splits trips between highway and transit as a function of the differences in time and cost for the two modes. The second splits highway person trips into automobile-driver and automobile-passenger trips as functions of the parking cost at the trip destination and of a measure of the density of highway travel between origin and destination. In addition, modifications were made to the model to forecast the impacts of car-pool incentive measures and the establishment of a van-pooling program.

The time-of-day distribution of travel was estimated by using factors derived from 1968 home-interview survey data that relate the percentage of trips occurring during specific time periods of the day to the daily totals. Separate factors were used for each major purpose of travel.

An important issue that must be addressed when attempting to interpret the results of the modeling of impacts of transportation control measures and packages is the reliability of the modeled results. Many of the candidate measures involve controls for which there is little past experience on which to base predictions. Others involve cost or time changes that are much larger than those that are normally modeled. Estimates must sometimes be derived from extrapolated portions of empirical relationships or be based on professional judgment.

Transportation models have primarily been used in the past for long-range transportation planning and, although a few tests have been made to gauge their accuracy in forecasting future travel, there have been several historical examples in which long-range forecasts of traffic volumes have proved to be badly off. Long-range travel forecasts must rely on highly uncertain projections of future land-use patterns and, thus, travel forecasting models should not necessarily be condemned for these poor forecasts. When the models are used for short-range forecasting, land-use changes are usually slight, so that the error due to what is normally considered the weakest link in the travel forecasting chain is minimized.

In the analysis of the impacts of transportation control measures and packages, trip-generation and trip-distribution models were not used, partly because it was felt that the short-term impacts on trip frequency and destination choice resulting from these measures would be minimal and partly because of their insensitivity to the travel parameters that would be changed as a result of the control measures. Although this modeling deficiency is probably not a serious one, it should be noted that any short-term changes in trip generation and distribution that might in fact occur have not been accounted for. Also, long-term trip-generation and trip-distribution characteristics could change significantly, particularly if a harsh transportation control plan were to be adopted.

The amount of confidence that can be placed in the results of the forecasts depends primarily on the reliability of the modal-choice and trip-assignment models and the assumptions made in applying them. When originally calibrated, the modal-choice model used in this study was shown to estimate transit and automobile-driver trips within 10 percent for all but the most sparsely populated areas of the region (5). It was particularly accurate in the most densely populated areas of the region where transportation controls are likely to have their greatest impact. Although tests have not been made to examine the accuracy of the model in predicting traveler response to specific system changes, when applied to 1965 Philadelphia transit network and travel data, it was found to estimate total Philadelphia area transit ridership to within 5 percent of the actual surveyed number of trips (6).

Generally, transportation models are more reliable when used for predicting modal shifts caused by changes in the transportation system, rather than when used for predicting absolute numbers of transit riders, automobile drivers, and automobile passengers for a given system. The models are based primarily on travelers' sensitivity to relative travel times and costs and thus are best for predicting traveler responses to small changes in these parameters. However, the Washington zonal-level modal-choice model was calibrated with 1968 travel data and, therefore, it is not based on actual observations of traveler responses to extremely high parking or gasoline costs. To model the impact of high parking costs on automobile occupancy, the calibrated automobile-occupancy curves that relate parking cost and trip density to automobile occupancy had to be extrapolated. In addition, professional judgment was used to conclude that increases in gasoline cost should have impacts on automobile occupancy similar to those of increases in parking cost.

As the long-term impacts of the gasoline price increases of the early 1970s and the imposition of a city tax on parking in the District of Columbia have become more apparent, it appears that the extrapolated portion of the automobile-occupancy model may be too sensitive to parking and gasoline cost increases. The automobile occupancies reported in this study may be high because of this oversensitivity and of the assumed gasoline price of 18.5 cents/L (70 cents/gal) (in 1975 dollars).

The modal-choice model is sensitive only to relative travel times and costs of alternative modes. Therefore, its applicability with respect to measures that are primarily information oriented, such as the implementation of a car-pool locator system, is limited. To model the impacts of these measures, models based primarily on professional judgment and experience in other cities were used.

The Washington modal-choice model as formulated is also insensitive to the impacts of measures that are designed to give car pools time or cost advantages over low-occupancy vehicles. To overcome this problem, the modal shifts that might be expected from these measures were modeled by using a multinomial logic formulation that examined trade-offs among three modes: transit passenger, drive alone, and shared ride. This model was recently developed for Denver (7) and, as a part of testing it, it was also calibrated by using the Washington home-interview-survey data file. However, the model has not been validated for the Washington area. In addition, the automobile-occupancy rate for shared-ride commuters had to be estimated by using data from other cities and professional judgment.

Another limitation of the Washington model is that it is not directly responsive to measures that restrict supply, in particular, measures that limit the availability of parking. To model these measures, it was necessary to assume that, if the parking supply was limited, market forces would increase the cost of parking to the point that enough commuters would switch to transit and car pooling and make supply and demand reach equilibrium. Although this is not a true reflection of what would happen if the supply was artifically restricted without concurrent price increases, the results modeled are probably similar to those that would occur if employers were forced to implement car-pool and transit requirements because of supply restrictions.

Because of budget and time restrictions, the impacts of control measures on nonwork and non-home-based trips were modeled by using a manual sensitivity approach in which sample interchanges were modeled. The results were used to obtain adjustment factors to apply to the nonwork and non-home-based automobile-driver trip estimates. This produced what is admittedly probably a very rough approximation of control measure and package impact on these trips. However, the control measures for the most part were designed to primarily impact home-based work trips and are expected to have only minimal impacts on nonwork and non-home-based trips.

The trip-assignment procedure used in this study to produce traffic-volume estimates was an all-or-nothing technique that loads all traffic on the shortest time paths and does not consider capacity constraints. Although this technique can give artificially high volumes for particular facilities, estimates of total vehicle travel should be fairly accurate provided the automobile-driver travel estimate from the modal-choice model is correct. Volumes on particular facilities are of interest primarily for carbon monoxide hot-spot analysis. For those hot-spot locations analyzed in this study, it was found that the localized volumes forecast by using the travel forecasting chain were usually within 10 to 15 percent of the volumes projected on a basis of expansion of observed volumes from existing counts. Comparable to the situation of the modal-choice model, percentage changes in traffic volumes (on links of the system that are not over capacity) and in total vehicle travel can be more accurately predicted than can absolute volumes or totals of vehicle travel.

Evaluation of Individual Transportation Control Measures

After examination of the current status of the existing transportation control plan, the working group developed a list of suggested transportation control measures to be considered for inclusion in the revised transportation control plan. These measures were evaluated by using a set of screening criteria, and the list was then narrowed to the following 62 measures to be analyzed in detail for their potential impacts on transportation and air quality.

1. Provide additional exclusive bus lanes,
2. Route buses into rail stations to facilitate intermodal transfer (included in 1977 base-case alternative),
3. Give car-pool vehicles priority at Metro stations,
4. Establish pedestrian malls in the core area,
5. Provide preferential bus treatments at Metro stations,
6. Allow car pools to use selected exclusive bus facilities,
7. Build additional bicycle lanes and install bicycle racks,
8. Increase express bus frequencies by scheduling extra buses displaced by Metro rail facilities,
9. Develop regional and localized car-pool matching services (included in 1977 base-case alternative),
10. Allow only car-pool vehicles in 50 percent of garages during morning peak periods,
11. Establish a reciprocal violations-enforcement system,
12. Expand parking at outlying Metro stations,
13. Encourage contract bus services,
14. Reserve convenient spaces in parking facilities for car-pool vehicles,
15. Institute flexible work times to stagger work hours and encourage car pooling,
16. Provide signal preemption for transit vehicles,
17. Encourage van pooling in the private sector,
18. Reduce peak-hour headways on commuter rail lines to 15 min,
19. Stagger work times primarily for federal employees,
20. Develop payroll deduction plans for transit passes,
21. Encourage employers to subsidize employee transit fares,
22. Include Metrobus information with car-pool locator information,
23. Provide car-pool locator information with Metrobus information services,
24. Provide freeway ramp metering where alternative routes exist and give preferences to buses, van pools, and car pools,
25. Reduce parking fees for car-pool vehicles in lots and garages,
26. Provide pedestrian priority signalization,
27. Reserve one lane for car pools on all facilities that have three or more lanes in the peak direction,
28. Institute commercial parking rates at all federal parking facilities,
29. Encourage commercial enterprises to offer free or discounted transit rides to off-peak customers,
30. Establish five park-and-ride lots for tourists near the Beltway,
31. Install progressive signalization so as to increase regional average speeds 1 percent,
32. Develop inspection and maintenance programs,
33. Establish a monthly transit pass at a cost of $20,
34. Develop seasonal controls such as smog days for federal workers,
35. Eliminate subsidized parking in private facilities for government employees,
36. Increase fuel taxes 2.5 cents/L (10 cents/gal),
37. Double registration fees for second and subsequent automobiles,
38. Restrict peak-hour truck deliveries in the District of Columbia,
39. Eliminate all on-street nonresidential parking of more than 2 h in all districts that have an accessibility level of four (most of the metropolitan core),
40. Establish a monthly transit pass at a cost of $15,
41. Impose a parking surcharge of $1 in any district that has an accessibility level of four (most of the metropolitan core),
42. Institute commercial hourly parking rates at all federal parking facilities,
43. Impose a parking surcharge of $2 in any district that has an accessibility level of four (most of the metropolitan core),
44. Eliminate all on-street nonresidential parking of more than 2 h in all districts that have an accessibility level of three or four (the metropolitan core and the immediately surrounding districts),
45. Impose a parking surcharge of $1 in any district that has an accessibility level of three or four (the metropolitan core and the immediately surrounding districts),
46. Impose a parking surcharge of $2 in any district that has an accessibility level of three or four (the metropolitan core and the immediately surrounding districts),
47. Eliminate daily and monthly parking rates and charge hourly rates throughout the day at all parking facilities,

48. Retrofit heavy-duty vehicles,
49. Reduce transit fares 10 percent,
50. Convert one lane to buses and one lane to car pools on all facilities that have three or more lanes in the peak direction,
51. Double the registration fee for all automobiles,
52. Increase fuel taxes 5 cents/L (20 cents/gal),
53. Allow no more than two lanes for low-occupancy automobiles on highway facilities inside the Beltway in the peak direction (limited-access facilities excluded),
54. Provide circumferential bus service on the Beltway and bus pull-offs at interchanges,
55. Eliminate all on-street nonresidential parking in the core area,
56. Establish a flat peak-period transit fare of 40 cents in the District of Columbia and 50 cents elsewhere,
57. Establish a flat peak-period transit fare of 40 cents,
58. Make ring 0 automobile free,
59. Reduce transit fares 25 percent,
60. Retrofit medium-duty air and fuel controls,
61. Eliminate 10 percent of all privately owned parking spaces in parking lots and garages in the core area, and
62. Increase fuel taxes 12.5 cents/L (50 cents/gal).

Because of time and budget constraints, most of the impact analyses of the individual control measures were not carried through in detail to the ultimate impact on air quality. Rather, the detailed air-quality analyses were reserved primarily for evaluation of the packages of control measures. Comparisons of the individual measures were primarily based on their impacts on transportation, social, and economic parameters.

The primary transportation characteristic used to compare the individual measures was the percentage reduction in work-purpose automobile-driver trips. According to the definition used in the Washington model chain, this reduction is equivalent to the reduction in work-purpose vehicles, which constitute the majority of vehicles on the road during the morning peak period. This time period is significant because there is a general correlation between morning peak-period vehicular traffic and peak oxidant concentrations, although many other factors are also important in the complex photochemical oxidant formation reaction. The effects of some of the individual measures are compared with the base case (given below) (8) in Table 1.

Base-Case Item	Value
Transit use, percentage of trips	20
Automobile-driver use, percentage of trips	59
Automobile occupancy, persons per vehicle	1.35
No. of transit trips	431 400
No. of automobile-driver trips	1 283 400

Evaluation of Transportation Control Measure Packages

After evaluating the 62 individual transportation control

Table 1. Changes to 1977 mode choice for home-based work trips due to individual candidate transportation control measures.

Measure	Change in Transit Use (%)	Change in Automobile-Driver Use (%)	Change in Automobile Occupancy (%)	Change in No. of Transit Trips	Change in No. of Automobile-Driver Trips
8 (increase express bus frequencies by scheduling extra buses displaced by metrorail facilities)	+1	0	0	+3 100	-2 000
9 (regional and localized car-pool matching service)[a]	-4	-1	+2	-17 600	-13 700
10 (allow only car-pool vehicles in 50 percent of garages during a.m. peak periods)	0	-4	+2	+2 100	-48 000
13 (encourage contract bus services)	+1	0	0	+4 100	-2 100
14 (reserve convenient spaces in parking facilities for car-pool vehicles)	0	-1	+1	-2 000	-14 000
17 (encourage van pooling in the private sector)		-1			-8 400
21 (encourage employers to subsidize employee's transit fares)	+2	0	0	+7 100	-4 300

[a] Included in the base case alternative.

Table 2. Changes in vehicle travel.

Condition	Automobile Vehicle Travel (km)	Percentage Change From Base	Truck Vehicle Travel[a] (km)	Percentage Change From Base	Total Vehicle Travel (km)	Percentage Change From Base
Morning peak (6-9 a.m.)						
Base	10 997 280		1 377 440		12 374 720	
Package 1	10 663 360	-3	1 377 440	0	12 040 800	-3
Package 2	9 097 280	-17	1 327 200	-4	10 424 480	-16
Package 3	8 633 600	-21	1 377 440	0	10 011 040	-19
Package 4	8 386 720	-24	1 384 960	+1	9 771 680	-21
Daily 8-h peak (2-10 p.m.)						
Base	32 166 080		2 863 360		35 029 440	
Package 1	31 902 880	-1	2 863 360	0	34 766 240	-1
Package 2	30 329 760	-6	2 863 360	0	33 193 120	-5
Package 3	29 792 160	-7	2 863 360	0	32 655 520	-7
Package 4	29 437 280	-8	2 870 880	+0	32 308 160	-8
24-h total						
Base	60 960 000		8 347 680		69 307 680	
Package 1	60 362 880	-1	8 347 680	0	68 710 560	-1
Package 2	57 195 200	-6	8 347 680	0	65 542 560	-5
Package 3	56 222 400	-8	8 347 680	0	64 700 080	-7
Package 4	55 546 240	-9	8 373 440	+0	63 919 680	-8

Note: 1 km = 0.62 mile.

[a] Commercial trucks only. Trucks used for personal use are classified as automobiles.

Table 3. Changes in vehicle trip origins.

Condition	No. of Automobile Trip Origins	Percentage Change From Base	No. of Truck Trip Origins	Percentage Change From Base	No. of Total Trip Origins	Percentage Change From Base
Morning peak (6-9 a.m.)						
Base	694 000		110 500		804 500	
Package 1	673 700	-3	110 500	0	784 200	-3
Package 2	591 500	-15	101 900	-8	693 400	-14
Package 3	555 900	-20	110 500	0	666 400	-17
Package 4	543 000	-22	110 500	0	653 500	-19
Daily 8-h peak (2-10 p.m.)						
Base	2 622 500		229 700		2 852 200	
Package 1	2 607 400	-1	229 700	0	2 837 100	-1
Package 2	2 520 700	-4	229 700	0	2 750 400	-4
Package 3	2 479 700	-5	229 700	0	2 709 400	-5
Package 4	2 456 100	-6	229 700	0	2 685 800	-6
24-h total						
Base	4 969 000		759 300		5 728 300	
Package 1	4 937 700	-1	759 300	0	5 697 000	-1
Package 2	4 758 700	-4	759 300	0	5 518 000	-4
Package 3	4 672 800	-6	759 300	0	5 432 100	-5
Package 4	4 625 900	-7	759 300	0	5 385 200	-6

Table 4. Changes to 1977 mode choice for home-based work trips due to candidate transportation control packages.

Package	Change in Transit Use (%)	Change in Automobile-Driver Use (%)	Change in Automobile Occupancy (%)	Change in No. of Transit Trips	Change in No. of Automobile-Driver Trips
1	+3	-2	+2	+12 200	-31 300
2	+10	-17	+17	+45 000	-214 300
3	+3	-24	+30	+13 600	-307 200
4	+12	-26	+32	+50 800	-338 000

measures, the working group combined a variety of them into four candidate packages of measures to be tested and evaluated in accordance with the adopted quantitative criteria. The compositions of the packages are shown below.

Package	Measures Included
1	5, 11, 15, 28, 31, 32, 39, 48
2	1, 3, 5, 8, 10, 11, 12, 13, 14, 15, 17, 18, 25, 28, 31, 32, 33, 38, 39, 43, 45, 48
3	1, 3, 5, 7, 8, 12, 13, 15, 17, 18, 25, 28, 31, 32, 44, 47, 53
4	1, 3, 5, 7, 8, 12, 13, 15, 17, 18, 25, 28, 31, 32, 44, 47, 48, 52, 53, 58, 59

When the packages were formed, the intention was to make each one progressively harsher and thus presumably more effective than the preceding one.

In the adopted evaluation process, the most important transportation variables were the vehicle travel and the vehicle trip origins. The regional effects of the different packages of measures on these variables are shown in Tables 2 and 3. The estimated changes in regional work-purpose modal choice due to the packages (compared with the base case given above) are shown in Table 4.

CONCLUSIONS

This study demonstrated that existing transportation planning structures and procedures can be used to develop useful information about the relative effectiveness of transportation system management measures of widely differing characteristics. This information can be used to assist in the rational development of regional transportation control plans and transportation system management plans.

The study did not show that no further model development is needed in response to changing demands on the transportation planning profession. The Washington models used in the study are particularly weak when examining car-pool incentives and measures that restrict the supply of transportation service, fuel, or parking. The study did show, however, that many existing models, when mixed with a little professional judgment, are sufficient for use until new models are available. In those urban areas where vehicle-use disincentives are not contemplated, those models currently available that are capable of examining high-occupancy modes (such as the newly developed modal choice for the Minneapolis-St. Paul area (9) may be sufficient for quite some time to come. Transportation system management is a new term in the transportation vocabulary. Its newness, however, should not cause the transportation planning professional to abandon the knowledge of travel behavior that has been learned over the past several decades and is embodied in our present analysis, evaluation, and design processes.

ACKNOWLEDGMENT

We wish to thank Paul Maxwell and Al Grant of the Transportation Planning Department of the Metropolitan Washington Council of Governments. We also wish to thank Richard Pratt of R. H. Pratt Associates and the other staff members who were involved in the study. The air-quality aspects of the study were handled by Environmental Research and Technology, especially William Olender. The air-quality side of the study was monitored by Skip Krause and Dennis Bates of the Health and Environmental Planning Department of the Metropolitan Washington Council of Governments. We also wish to thank the members of the study management working group, James Clark, Hal Kassoff, Richard Lockwood, V. Ramadass, Alvin Bowles, and James Alexander, who kept us searching for more data on which to base decisions. The national capital region transportation and air-quality study was performed in cooperation with the U.S. Department of Transportation, Federal Highway

Administration. It was administered by the Metropolitan Washington Council of Governments and sponsored and technically directed by the following six agencies: District of Columbia Department of Transportation, Maryland Department of Transportation, Virginia Department of Highways and Transportation, District of Columbia Department of Environmental Services, Maryland Department of Health and Mental Hygiene, and Virginia State Air Pollution Control Board.

REFERENCES

1. Transportation Controls for Air-Quality Improvement in the National Capital Region. R. H. Pratt Associates, Washington, DC; Environmental Research and Technology, Reston, VA, Oct. 1976.
2. Critique of Earlier Transportation Control Plan Development Methods. R. H. Pratt Associates, Washington, DC; Environmental Research and Technology, Reston, VA, Technical Memorandum 4, Jan. 1976.
3. Evaluation of Strategy Development Procedures. R. H. Pratt Associates, Washington, DC; Environmental Research and Technology, Reston, VA, Technical Memorandum 3, Sept. 1975.
4. W. Mann. TRIMS Model Users Manual. Metropolitan Washington Council of Governments, Washington, DC, Nov. 1975.
5. Development and Calibration of the Washington Mode-Choice Models. R. H. Pratt Associates, Washington, DC, June 1973.
6. T. Zakaria, W. T. Walker, and G. W. Schultz. Analysis of a Trip-End Versus a Trip-Interchange Modal-Split Model in the Delaware Valley Region. Delaware Valley Regional Planning Commission, Philadelphia, April 1973.
7. J. McLynn. Modal-Choice Models for Denver. Denver Regional Transportation District, Denver, Working paper, Dec. 1974.
8. Transportation Controls for Air-Quality Improvement in the National Capital Region: Technical Appendix 2—Assessment of the Impact of the Proposed Individual Transportation Control Measures. R. H. Pratt Associates, Washington, DC; Environmental Research and Technology, Reston, VA, Oct. 1976.
9. Development and Calibration of Mode-Choice Models for the Twin Cities Area. R. H. Pratt Associates, Washington, DC; DTM, Bethesda, MD; Minneapolis-St. Paul Metropolitan Council, MN, Nov. 1976.

Publication of this paper sponsored by Committee on Transportation and Air Quality.