SIMPLIFIED APPROACH TO MODELING FREEWAY OPERATIONS AND CONTROL

Brian L. Allen, Department of Civil Engineering, McMaster University; and
C. Jen Liew, De Leuw, Cather and Company of Canada

This paper develops a simplified model for simulating freeway operations influenced by entrance-ramp metering or closure. The model's application to a real freeway corridor is demonstrated. The model is based on the assumption that the entire corridor can be adequately represented by only 2 routes interconnected by equally spaced entrance and exit ramps. Optimal control is achieved by minimizing total corridor travel time. The effectiveness of 3 control strategies (entrance-ramp metering, entrance-ramp closure, and total interchange closure) is investigated. Traffic flow on a real freeway corridor was simulated with this model. The model compared favorably with observed conditions. When the effects of the 3 control strategies were investigated, freeway flow rates resulting from optimal control conditions were found to be nearly identical for each strategy. Identifying the optimal flow rate permitted accurate calibration of the model and reliable results. The model can be useful for initial planning evaluations. Data requirements for the model are minimal, and its application is straightforward.

• TRAVEL demand continues to increase, and with it, congestion on urban freeways spreads. This spread can be stopped or slowed by exercising some form of restrictive control. One common form is limiting access to freeways by either closing or metering entrance ramps.

Entrance-ramp metering has been widely accepted and successfully implemented in such freeway corridors as the Eisenhower Expressway in Chicago (1), the Gulf Freeway in Houston (2), the Van Wyck Expressway in New York City (3), and several freeways in the Los Angeles area (4). Vast amounts of monetary and human resources have been spent in metering research, development, and implementation (5, 6, 7).

Freeway-entrance-ramp closure has not been so widely accepted although it appears to be gaining in popularity as existing corridors become more congested. Several operating agencies have closed entrance ramps during peak travel periods, and usually they have had successful results (8, 9, 10). Lack of wider application seems to be because of the method's lack of political popularity, misunderstanding of its potential uses and benefits, and an absence of reasonable locations in which to implement it.

Detailed design of the method, and evaluation of its effectiveness for improving traffic operations, have proved to be a time-consuming and difficult task. To alleviate this burden, a significant proportion of development effort has been expended to provide sophisticated analytic models. These models simulate traffic flow on a freeway or in a corridor subject to a specified ramp-control strategy (11, 12, 13). Usually the models require extensive and accurate data input for successful operation, and, not surprisingly, such data are seldom readily available. For example, the FREQ model series (14, 15, 16) requires the user to supply complete details on freeway physical features, origin-destination patterns of traffic, and metering rates for all entrance ramps. By the time one considers, say, 30 different freeway subsections and 12 time intervals during the peak period, the magnitude of information required is formidable. Undeniably, that amount of detail is necessary if one is to place any degree of confidence in the final design of a control strategy. However, use of such techniques for preliminary analyses of freeway control seems impractical. There appears to be a need for a simplified technique that could be applied, for example, when an operating agency wished to ascertain the need for more detailed analyses on existing or future freeways for which comprehensive traffic data did not already exist.

In this paper, an analytic model is proposed that will fill the need dictated by such an application. The model requires a minimum amount of data for operation, gives reliable results, and serves as a useful first approximation of the detailed design of a freeway-control strategy. In addition, it permits direct comparison of the potential effectiveness of 3 control methods:

1. Entrance-ramp metering,
2. Entrance-ramp closure, and
3. Interchange (entrance- and exit-ramp) closure.

We suggest that this model can be applied directly to preliminary control and deficiency studies of existing freeway corridors and to similar studies for freeway corridors that are being planned or designed. Only a simple trip length distribution for the freeway corridor, speed-flow relationships for the freeway and surface streets, and freeway interchange spacing are required as data input. Numerical output can be used to suggest required metering rates, entrance ramps requiring closure, and optimal interchange spacing. The need for and the effectiveness of the 3 control methods can be directly ascertained.

MODEL DEVELOPMENT

A detailed description of model development is available elsewhere (17). With the goal of a simplified model in mind, we chose the freeway corridor representation shown in Figure 1. It consists only of 2 parallel routes, route 1 (freeway) and route 2 (city streets), interconnected by equally spaced access links (interchanges). All trips in the corridor are generated on route 2 and are destined for some point downstream that also is on route 2. They can enter route 1 on the entrance ramps and can exit by using the exit ramps. These entrances to and exits from route 1 may be selectively closed to permit investigation of the effects of entrance-ramp closure and total interchange closure strategies. $m - 1$ is the number of adjacent entrance ramps that will be closed; ms is the spacing between adjacent accessible entrance ramps.

To enable representation of entrance-ramp metering, one must impose a toll, δ ($\delta > 0$), at all accessible entrance ramps. This toll is considered to be in the form of a travel cost (time) penalty for each trip entering route 1. It represents the wait in queue behind a metering signal.

The segments of route 1 and route 2 between 2 adjacent access links are cells. The corridor comprises a series of individual cells, connected at common access links. Trips begin in an origin block containing the corridor segment between 2 adjacent accessible entrance ramps and terminate in a downstream destination block similarly defined. Each block contains $m + 1$ ramps and is ms long. The distance between corresponding ends of the origin and destination blocks is ns.

Within an origin block, x is the distance measured downstream between the first available entrance link and any specified origin within that block. Similarly, y_k is the distance measured upstream between the last available access link in destination cell $(n + k)$ or the kth cell in a destination block and any destination within that block where $k = 1, 2, \ldots, m$. X and Y are the respective distances to these origins and destinations measured from some arbitrary point upstream. The trip length, L, therefore is $Y - X$, the distance between the origin and destination of any trip.

The travel cost per unit of distance of travel on route 1, c_1, is an increasing function of the flow on route 1, f_1. The travel cost per unit of distance of travel on route 2, c_2, is assumed to be independent of the flow on route 2. As shown in Figure 2, $c_1 < c_2$ over rates of flow expected under control conditions.

It is assumed that travelers, because they are aware of the costs of using alternate routes, choose paths with the lowest cost. Travelers making short trips would find

Figure 1. Transportation corridor.

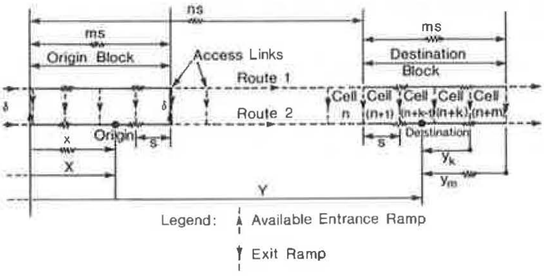


Figure 2. Travel cost per unit of distance of travel on routes 1 and 2.

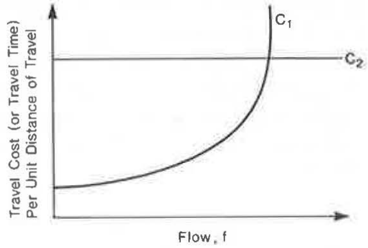


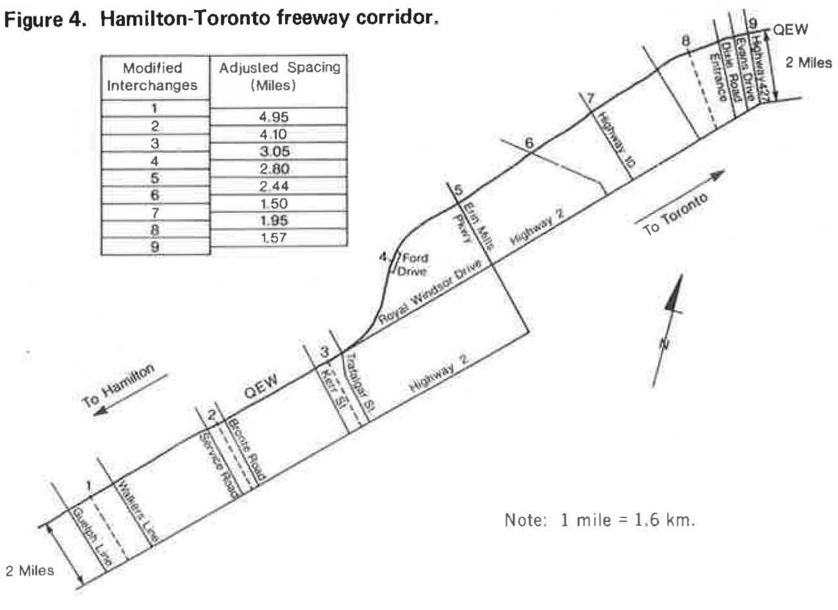
Figure 3. Average distance traveled and total average travel cost for short, intermediate, and long trips.

	Short Trips	Intermediate Trips	Long Trips
	$0 \le L \le ns-\mu\Delta$	$(n+k-1)s-(k-1+\mu)\Delta \le L_k \le (n+k)s-(n+\mu)\Delta$	$(n+m)s-(m+\mu)\Delta < L < \infty$
$f_1(L) =$	0	$\frac{1}{2m}\sum_{k=1}^{m}\{2(n+k-1)(L_k-ns+2\mu\Delta) - (s-2\Delta)(k^2-k) + 2\Delta(k-1)(n-\mu)\}$	$L - \frac{(m+1)}{2} s \frac{c_1}{c_2}$
$C(L) =$	c_2L	$\sum_{k=1}^{m} [c_2L_k - \frac{(\mu+k-1)}{m}(c_2-c_1)(L_k-ns+2\mu\Delta) + \frac{1}{2m}\{\mu(c_2-c_1)\mu\Delta + (k-1+\mu)(c_2-c_1)(k-1+\mu)\Delta + (k-1)(c_2-c_1)[(k-1)(s-2\Delta)+(s-\Delta)]\}]$	$c_1L + \delta + \frac{(m+1)}{2}(c_1+c_2)\Delta$

$$\Delta = \frac{s}{2}(1-\frac{c_1}{c_2}); \quad \mu = n-\frac{s}{s(c_2-c_1)}, \text{ where } 0 \le \mu \le 1$$

Figure 4. Hamilton-Toronto freeway corridor.

Modified Interchanges	Adjusted Spacing (Miles)
1	
2	4.95
3	4.10
4	3.05
5	2.80
6	2.44
7	1.50
8	1.95
9	1.57



Note: 1 mile = 1.6 km.

that route 2 cost the least time because they would avoid backtracking, queuing at entrance ramps, and extra travel to and from the freeway. Travelers making long trips would find that route 1 cost the least time even with these penalties. Unfortunately, travelers making trips of intermediate length cannot be assigned so easily. Depending on the location of the origin and destination within the blocks, these travelers might use either route 2 or route 1. All trips will be classified as being either short, long, or intermediate.

If all trips can be assigned to the corridor, route flows can be computed. If route flows can be computed, then c_1 and average total travel cost for trips of length L, C(L), can be determined.

To enumerate the number and pattern of trip origins and destinations, we defined trip density function as g(L). There are g(L) trips originating in the corridor segment (X, X + dL), destined for the segment (Y, Y + dL). Thus g(L)dL trips per unit of length are generated at any point along the corridor. By using the average travel cost computed for each of the 3 trip length ranges as shown in Figure 3, integrating $C_i(L)g(L)dL$ over all trip lengths in range i, and summing the 3 numbers, one can calculate total travel cost per unit of corridor length. Similarly, one can compute f_1 by integrating average travel distance, $f_1(L)$, shown in Figure 3, over the 3 trip length ranges.

When total corridor travel cost has been determined, optimization can start. Optimization involves choosing the appropriate metering rate, entrance-ramp closure configuration, or interchange spacing that minimizes cost.

MODEL APPLICATION

It is obvious that each expression in Figure 3 contains 3 unknowns, c_1, δ, and s. Even if both s and δ were fixed and known, c_1 and f_1 would be interrelated. Consequently, an iterative procedure must be used for solution. One must first compute values of f_1 by assuming various values of c_1. The known function $c_1(f_1)$ can be equated with those values and the intercept of the 2 functions will yield the correct c_1. Then all expressions can be solved. To aid in this tedious trial-and-error computation, an interactive computer program was developed, and data from a portion of the Hamilton-Toronto freeway corridor were used as input for a sample computation.

Study Area

The corridor shown in Figure 4 lies between Guelph Line and Highway 427, a distance of about 20 miles (32 km). Route 1 is the 3-lane eastbound portion of the Queen Elizabeth Way (QEW). Route 2 is Highway 2 and all parallel surface streets within 2 miles (3.2 km) of QEW. Perpendicular city streets connect these routes at 14 interchanges. All but the following interchanges have both entrance and exit ramps:

1. Guelph Line, Trafalgar Street, and Mississauga Road, which have additional entrance ramps, and
2. Royal Windsor Drive and Evans Drive, which have no entrance ramps.

The Dixie Road entrance is closer to the Evans Drive exit than it is to the Dixie Road exit. The Highway 427 entrance was outside the chosen study area.

The distance between exit and entrance ramps at any interchange was to be zero to conform with model assumptions. Closely spaced interchanges at Guelph and Walkers Lines, Service and Bronte Roads, Kerr and Trafalgar Streets, and the Evans Drive exit and Highway 427 were combined to form single representative interchanges because they serve a common area and could be considered as single interchanges. The locations of these modified interchanges are shown in Figure 4 by the dashed lines.

The operational characteristics of traffic in this corridor that are partially described by the speed-flow relationship for the QEW shown in Figure 5 were computed from an empirically derived, linear speed-density relationship supplied by the Ontario Ministry

Figure 5. Speed-flow relationship for Queen Elizabeth Way.

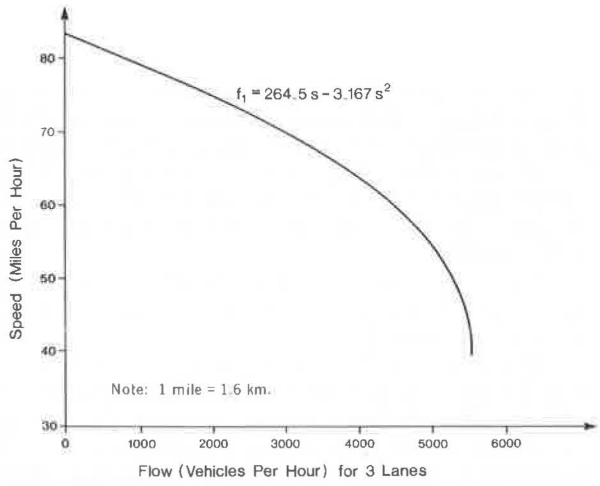


Figure 6. Travel cost-flow relationship for Queen Elizabeth Way.

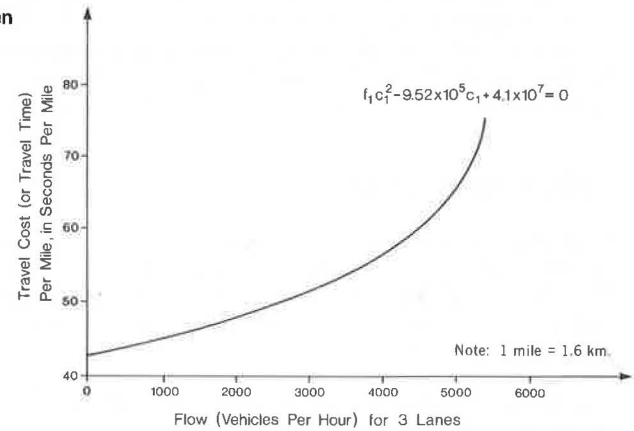


Figure 7. Trip density distribution for Queen Elizabeth Way corridor.

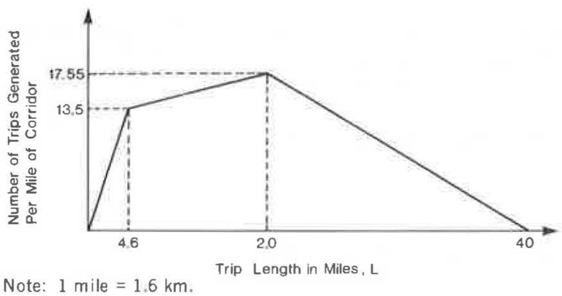
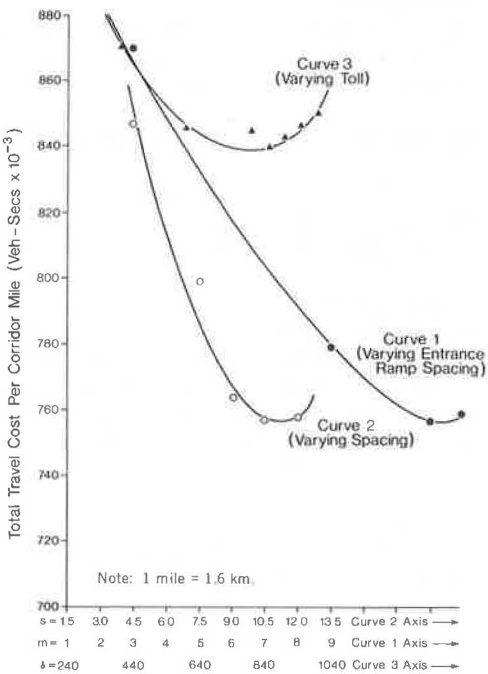


Figure 8. Total corridor cost for varying tolls, entrance-ramp spacings, and interchange spacings.



of Transportation and Communications (MTC). Travel time per unit of travel distance on the QEW as a function of flow was obtained from that speed-flow relationship and is shown in Figure 6. Average speed of travel on Highway 2 and parallel surface streets was assumed to be 30 mph (48 km/h). Average speed also was assumed to be essentially independent of flow variations on these streets. These assumptions were considered acceptable because field observations indicated low volume-to-capacity ratios and because any additional flow diverted from the freeway should not affect significantly the speed on these streets.

The magnitude and origin-destination patterns of traffic in this corridor were described by developing the trip density distribution shown in Figure 7. The distribution was computed by using data obtained from a study carried out in the Toronto area (18, p. 53).

People making trips generated in this corridor were assumed to have to travel an additional 2 miles (3.2 km) on lateral surface streets (those connecting routes 1 and 2) if they were assigned route 1 paths. Basing our calculations on an average speed of 30 mph (48 km/h), we assumed that this distance would add a 4-min penalty. Knowledge of existing corridor characteristics indicated that a spacing of 1.5 miles (2.4 km) between interchanges would be most representative of the critical section from Mississauga Road to Highway 427.

Results

All of the information on the chosen corridor was used as input into the computerized model. Results from this application are shown in Figure 8. Curve 1 is the total cost due to varying the spacing of available entrance ramps. It indicates that a spacing of 18 miles (29 km) between available entrance ramps (or closing 12 adjacent entrance ramps) would minimize the total cost of travel to all users in this corridor; the flow on QEW would be 5,100 vehicles per hour (vph).

Curve 2 is the total cost due to varying the spacing of entrance and exit ramps (interchanges). In this case, an interchange spacing of 10.5 miles (16.9 km) would minimize total user cost in this corridor; the flow on QEW would be 5,000 vph. Minimum total cost obtained by varying interchange spacing was not significantly different from the minimum total cost obtained by varying entrance-ramp spacing.

Curve 3 is the total cost obtained by varying the toll imposed on all users entering the QEW. A toll of 900 sec (or an additional penalty of 660 sec) would minimize total user cost in this corridor; the flow on QEW would be 5,100 vph.

Additional runs from this computerized model in which the trip density function and the penalty charged to all users assigned route 1 paths confirmed that a flow of 5,000 to 5,100 vph on the QEW would give the minimum total user cost regardless of the freeway-control strategy used. This flow range corresponds with MTC field observations of optimal travel conditions on the QEW through the critical section.

However, the recommendation for optimum spacing of entrance and exit ramps and entrance-ramp metering rates cannot be realistically applied to this corridor because the chosen 20-mile (32-km) corridor is relatively short. To accommodate this and to make the application more meaningful, trips with lengths greater than 20 miles (32 km) should be considered as external through trips that make up only a constant through flow on the QEW.

After thorough consideration of trip characteristics in this corridor, an external flow of 2,500 vph was computed. A new trip density distribution with a maximum trip length of 20 miles (32 km) was derived from the previous distribution by deleting the portion with trip lengths greater than 20 miles (32 km). The modified input was then fed into the computerized model. The results are as follows (1 mile = 1.6 km):

m	s (miles)	δ (sec per vehicle)	f_1 (vph)
1	1.5	240	5,200
1	3.0	240	5,000
2	1.5	240	5,100

The results indicate that an entrance-ramp spacing of 3 miles (4.8 km) (m = 2, s = 1.5) will reduce the flow to 5,100 vph, whereas an interchange spacing of 3 miles (4.8 km) (m = 1, s = 3.0) will reduce the flow further to 5,000 vph. Both of these flow rates are within the optimal range.

From this second application, one can recommend that some form of freeway-ramp-control strategy be implemented between Erin Mills Parkway and the Dixie Road interchange because of the shorter spacing. If ramp closure is preferred, then the entrance ramps (and exit ramps, if necessary) at Mississauga and Dixie Roads may be closed during the morning peak period to effect the desired optimal spacing.

COMMENTS AND CONCLUSIONS

The simple model of freeway corridor operations and control reported here most certainly will be subject to criticism. The simplifying assumptions used to decompose a complex system of interdependent variables into an extremely simple one are obviously suspect. For example, there never has been a corridor in which all traffic origin-destination patterns were identical along its entire length; neither will there ever be a corridor in which the physical characteristics of the roadways are invariant over length. The formulation of the speed-to-flow or travel-time-to-flow relationships also is open to question. Although no one can strenuously argue that the form used to represent travel on a freeway (route 1) is incorrect, the independence of travel time on flow on city streets is at least a dubious simplification. Oversaturation of critical signalized intersections in the street network could very quickly obviate any benefits realized on the freeway. Finally, the assumption of constant flow along the freeway, regardless of the number of available entrance ramps, is strictly incorrect. If, for example, every second entrance were closed and exits were open, flow would obviously decrease in the subsections immediately downstream of the exit ramps. The equations in Figure 3 that were used to compute travel times on the corridor are also strictly incorrect.

Despite these severe shortcomings, results from the example application seem to indicate, on a gross scale, a strong correlation between actual and simulated conditions. Most importantly, the method reliably predicts flows generated in the most critical sections of the freeway. Field observations also confirm that the model accurately predicts the optimal flow rate for critical freeway sections, that is, the maximum rate of flow that can be maintained without severe travel time increases. These results indicate that the method proposed here can be taken more seriously than we first thought. Because of this, one can also look seriously at several other interesting conclusions drawn from the example.

Perhaps the single most important observation concerns the optimal flow rates computed for critical sections. Results indicated a difference of only 2 percent between the optimum flows computed for each control strategy. This observation not only is intuitively appealing but also has important practical implications. When implementing a given control strategy, one should exercise control so as to obtain the prescribed optimal flow rate on the freeway. Although in practice there may be slight variations in that flow, it appears that sensitivity to minimizing total corridor travel cost would be minimal. Adherence to this procedure would reduce considerably the effort required to provide a final control design.

An application would consist of using the observed (or calculated) trip density function for a wide range of interchange spacings to yield the minimum total corridor travel time. The flow at that minimum would be chosen as the optimal flow. Because a proportion of the total trips on the freeway are likely to be through trips, the optimal spacing inferred by this first computation should be ignored. The trip density function then should be truncated to remove the cumulative influence of those trips and should be replaced as a constant nonadditive flow. The revised density function should then be used in the model to obtain the interchange spacing that yields the optimal flow obtained from the first computation. The spacing thus computed would be the recommended optimal spacing. Using this procedure, one can obtain general recommendations for control by total interchange closure, entrance-ramp closure, or entrance-ramp metering.

Although such a procedure may sound complicated and time-consuming, it is simple and easy to perform with a computer. In addition, the results are extremely easy to interpret. Identification of critical or potentially critical sections simply requires that one compare optimum spacing to existing spacing. If optimum spacing is greater than existing spacing, one should design improvements accordingly.

Data requirements for using this procedure are minimal. Trip length distributions are usually available from operating or planning agencies for almost every major urban corridor, and an indication of the proportion of through trips is obtained easily from a license plate survey or simple truncation of the trip length function. Together with the addition of travel-time functions, these are the only data required to obtain an indication of the degree of control required on the corridor.

Admittedly, this procedure could not be used for detailed design of a control scheme. Although rates established for entrance-ramp metering are unlikely to be equal for all ramps within critical freeway sections, the model results would indicate required rates. The spacings recommended for either accessible entrance ramps or interchanges could not be obtained precisely on a real corridor, but close approximations are usually possible. Finally, no detail concerning operations on the adjacent street network is used in the model so that final consideration of storage lane needs, revised signal timing, and intersection signing could not be established. However, the procedure could be used as a workable first approximation of control requirements. Perhaps it could be used in conjunction with standard deficiency studies, or it could be associated with planning and design procedures in which the availability of detailed data is limited. In any of these cases, final control specifications are not required, so use of more complicated models would not be warranted.

In addition to specifying approximate freeway control needs, output from the model would also be useful for comparing the relative effects of the 3 control strategies. Although such comparisons would be qualitative, they would be useful when one is considering trade-offs between strategies or contemplating a combination of control modes. Current activities should be expanded and continued so that a better understanding of various control modes will result.

We suggest that the model reported here offers considerable advantages over currently available methods for examining freeway corridor operations and control. Although it is not comprehensive in nature, it provides reliable indications of the extent and degree of control required and demands very little in the way of data preparation and output interpretation. It provides an essential link between awareness of problems, understanding the applications of various control modes, and final implementation of control.

ACKNOWLEDGMENTS

We thank J. Thomas and R. Walshaw, formerly of the Ontario Ministry of Transportation and Communications, for their assistance in providing data used in the example applications.

This work was made possible by sponsorship of the National Research Council of Canada.

REFERENCES

1. Chicago Area Expressway Surveillance and Control. Illinois Division of Highways, July 1971.
2. M. E. Goolsby, E. Merrell, and W. R. McCasland. Freeway Operations on the Gulf Freeway Ramp Control System. Texas Transportation Institute, Texas A&M Univ., Rept. 24-25, 1969.
3. L. Pignataro and K. Crowley. The Van Wyck Expressway Experiment. Traffic Engineering, Vol. 41, No. 4, Jan. 1971.
4. Los Angeles Area Freeway Surveillance and Control Project. California Division of Highways, 3rd Annual Report, Feb. 1973.
5. L. Newman, A. Dunnet, and G. Meis. Freeway Ramp Control—What It Can and Cannot Do. Traffic Engineering, Vol. 39, No. 9, June 1969.
6. W. Wainwright and R. Jain. Techniques of Ramp Control. Univ. of Connecticut, C.E. 72-47, Feb. 1972.
7. P. Everall. Urban Freeway Surveillance and Control: The State of the Art. Federal Highway Administration, Nov. 1972.
8. R. Loutzenheiser. The Effects of Ramp Closure on Freeway Operation During Morning Peak Periods. Texas Transportation Institute, Texas A&M Univ., Rept. 139-9, Aug. 1971.
9. L. Newman, A. M. Dunnet, and G. J. Meis. An Evaluation of Ramp Control on the Harbor Freeway in Los Angeles. Highway Research Record 303, 1970, pp. 44-55.
10. E. Gervais and W. Roth. An Evaluation of Freeway Ramp Closure. Michigan State Highway Department, April 1965.
11. H. Payne. Aggregate Variables Models of Freeway Traffic. Univ. of Southern California, Aug. 1971.
12. Y. Makigami et al. The Freeway Model. In Bay Area Freeway Operations Study—Final Report, Part 1, Univ. of California, Berkeley, Aug. 1970.
13. M. Orthlieb et al. Progress Toward a Freeway Corridor Model. In Freeway Operations Study—Phase III, Univ. of California, Berkeley, June 1973.
14. W. Stock et al. The FREQ3 Freeway Model. In Freeway Operations Study—Phase III, Univ. of California, Berkeley, June 1973.
15. M. Eldor and A. May. Cost-Effectiveness Evaluation of Freeway Design Alternatives. In Freeway Operations Study—Phase III, Univ. of California, Berkeley, June 1973.
16. J. Wang and A. May. Analytic Techniques for Ramp Control. In Freeway Operations Study—Phase III, Univ. of California, Berkeley, Interim Rept. 2, June 1972.
17. C. J. Liew. Entrance Ramp Closure as a Means for Controlling Traffic in a Transportation Corridor: Theory and Application. McMaster Univ., master's thesis, May 1974.
18. Metropolitan Toronto and Region Transportation Study. April 1966.