TRANSPORTATION RESEARCH RECORD

531

Traffic Control and Motorist Information

6 reports prepared for the 54th Annual Meeting of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
Washington, D. C., 1975

Transportation Research Record 531
Price $3.60
Edited for TRB by Marianne Cox Wilburn

subject areas
51 highway safety
52 road user characteristics

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

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

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

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

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

LIBRARY OF CONGRESS CATALOGING IN PUBLICATION DATA

National Research Council. Transportation Research Board.
Traffic control and motorist information: 6 reports prepared for the 54th annual meeting of the Transportation Research Board.

(Transportation research record ; 531)
1. Traffic engineering—Congresses. 2. Electronic traffic controls—Congresses. I. Title. II. Series.
TE7.H5 no. 531 [HE332] 380.5'08s [388.3'1] 75-15549
ISBN 0-309-02378-5

CONTENTS

FOREWORD

The 5 papers and 1 discussion contained in this RECORD have defined some of the operational problems that traffic, safety, and freeway surveillance specialists must face while fulfilling their responsibilities. The methods of study and the specific findings will be of value in solving other similar but unrelated problems.

In the opening paper, Tarnoff identifies the presence of data errors inherent in the operation of a computerized traffic-control system and suggests measures to minimize the effects to have an operation that is better than a pretimed system. Ross, in a discussion of the report, lists additional factors that must be investigated regarding the sensitivity of real-time control to data errors and stresses the need for additional research on this subject.

Hanscom presents 2 data collection techniques that were employed to evaluate various combinations of signs to warn motorists of potentially icy bridges. It was found that activated signing at the bridge elicited the highest response, particularly during periods of greater hazard.

In research completed in an HPR study, Roberts, Reilly, and Jagannath installed a series of diagrammatic signs and judged their effectiveness by evaluating the unusual maneuvers of exiting motorists. They report that diagrammatic signs can reduce stopping and backing maneuver rates under certain conditions but may increase unusual exit gore maneuver rates at the same location.

Hall and Dickinson developed 2 questionnaires to determine motorists' preferences for the type of real-time information desired for route-diversion signing. The responses indicated that motorists want to know the length and cause of congestion and to be provided with alternate route information.

Gartner and Little developed a systematic procedure for the determination of signal settings in a network. Offsets, green splits, and cycle time also were considered. The preliminary results from the traffic-flow model, which uses the generalized combination method as its basic building block, indicate that a potential for significant improvements in performance of traffic-signal systems exists.

Leuthardt discusses the theory and fundamentals of progressively timed signal systems whose objective is the achievement of maximum possible bandwidth. The relationship of the bandwidth to speed, cycle length, splits, and distances is defined. Based on these relationships, a model is developed.

—J. R. Doughty

DATA ERRORS IN URBAN TRAFFIC-CONTROL SYSTEMS

Philip J. Tarnoff, Federal Highway Administration

The presence of data errors in a traffic-control system is unavoidable. These errors result from the inadequacies of the surveillance system, inherent characteristics of the vehicle traffic, and inaccuracies in modeling the traffic system. If these errors are not controlled during system design and implementation, they can cause degradation of system operation to the point where it is less effective than that of a pretimed system. One of the measures that can be taken to prevent this is the design of a surveillance system that introduces errors that are no greater than the errors introduced by the other elements in the system. A second measure is the collection of data before system design that will permit identification of the parameters that must be varied on a time-of-day and link-specific basis in the prediction and optimization algorithms. This paper emphasizes the errors associated with the processing of vehicle volumes because the effectiveness of the control strategy depends most on the accuracy of this variable. Consideration is also given to the limitations inherent in the prediction process and the effect of system errors on vehicle delay at controlled intersections.

•THE URBAN Traffic Control System (UTCS) is a computer-controlled traffic signal system that has been installed by the Federal Highway Administration for developing advanced traffic-signal control strategies. The system development began in 1968 and continues at the present. A fully operational traffic-control system of 114 intersections has been installed in Washington, D.C. The system has been implemented to serve as a research facility to support the development of advanced control strategies that respond automatically to changes in traffic demand. To support these strategies, the design has included expanded detectorization, display, and data processing equipment beyond that that would be found in an operational system.

The surveillance system consists of approximately 500 loop detectors that have been installed to measure vehicle presence. From the detector outputs, the data processing system derives:

1. Volume—number of vehicles per lane per unit of time;
2. Occupancy—percentage of time of vehicle presence that is measured by the detectors;
3. Speed—average rate at which vehicles cross the detectors (this variable is proportional to occupancy divided by volume);
4. Queue length—number of vehicles waiting at the intersection approach at the end of the red phase;
5. Stops—number of vehicles on an approach that are required to wait for the red (this variable differs from queue length in that it represents the cumulative numbers of vehicles stopped over a 15-min period); and
6. Delay—estimated cumulative time that stopped vehicles are required to wait for the red.

Publication of this paper sponsored by Committee on Traffic Control Devices.

Recent experiments related to the development of traffic-control strategies for UTCS have shown that these strategies are very sensitive to errors in input data and that large errors exist in these data. This has led to a comprehensive analysis that includes

1. Identification of the sources of data errors in the surveillance and prediction elements of the traffic-control system,
2. Quantification of the characteristics of individual errors,
3. Evaluation of the sensitivity of the control strategies to data errors,
4. Development of surveillance and prediction techniques for minimizing the effects of these errors, and
5. Evaluation of the magnitude of the fluctuations in traffic volumes to which the control strategies are to respond.

Hopefully this analysis will be successful. It is obvious that an error exceeding the variations in the quantity being controlled will reduce the control system to total ineffectiveness. In fact, based on observations to date, this might be the cause of the lack of success that past researchers have had in developing control strategies (1, 2). Many of these strategies have been developed without prediction techniques or effectiveness evaluation. This paper focuses on 1 aspect of this question—the errors in input data and their effect on control-system design. It also presents a brief summary of the control strategies being developed for the UTCS project. These control strategies are discussed in greater detail in other reports (3, 4, 5).

FIRST GENERATION CONTROL STRATEGY

The UTCS control-strategy development consists of the implementation of 3 generations of control (Table 1). The first generation of control is based on the use of signal timing patterns generated off-line and stored in a peripheral storage device. The system is capable of using 3 possible modes of pattern selection.

1. The operator select mode is one in which the system operator determines the operation pattern and makes a selection through the control panel. This selection can be made at any time during system operation.
2. The time-of-day mode is one in which the computer selects timing patterns every 15 min according to a predetermined schedule.
3. The traffic-responsive mode is one in which the computer attempts to select the pattern that is best suited for current traffic conditions every 15 min.

The first generation software also is capable of making adjustments in the timing patterns at selected intersections in response to fluctuations in traffic demand at each signal cycle [critical intersection control (CIC)]. The adjustment is accomplished by measuring vehicle volumes and modifying the signal split in such a way that the percentage of green time given to the competing demands is approximately proportional to the approach volumes.

SECOND GENERATION CONTROL STRATEGY

The principal difference between the first and second generation control strategies is that the second generation strategy computes the traffic signal timing on-line at a fixed rate of 4 to 7 min. (The exact rate has not yet been determined.) The optimization technique used for this computation is based on the SIGOP optimization, which computes and implements signal timing directly and does not require operator intervention. Obviously, under these circumstances, the traffic engineer loses the capability to make adjustments to the computed pattern that he or she would typically have when operating with the first generation system.

Table 1. Urban Traffic Control System strategies.

Strategy	Update Interval (min)	Prediction	Pattern Generation (selection)	Critical Intersection Control
First generation	15	None	Off-line pattern timing Time of day Traffic responsive Operator select	Comparison of A-phase and B-phase demand to determine A-phase yield point
Second generation	4 to 7 (precise value not yet determined)	Historically based	Employs modified version of SIGOP offset optimization	Both split and offset computation based on previous phase demand
Third generation	3 to 5 (variable)	Statistical predictor (form not completely determined)	Cycle-free optimizations for moderate flow and congested flow	Not applicable

Table 2. Control-strategy data requirements.

Algorithm	Measurement Interval	Critical Lane Variable	Range of Error[a]
First Generation			
Traffic-responsive pattern selection[b]	15 min (total)	Volume	Must be consistent indicator of traffic conditions
		Occupancy	Must be consistent indicator of traffic conditions
Critical intersection control[c]	Each phase	Volume	1 to 3 vehicles per cycle
		Queue	1 to 2 vehicles
		Speed	5 to 10 percent
Second Generation			
Network optimization	4 to 7 min	Primary volume	1 to 3 vehicles per cycle
		Queue	1 to 3 vehicles
		Speed	5 to 10 percent
Critical intersection control[d]	Start of each signal phase	Primary volume	1 to 3 vehicles per cycle
		Queue	1 to 3 vehicles
		Speed	5 to 10 percent
Third Generation			
Undersaturated control	3 to 5 min	Primary volume	1 to 3 vehicles per cycle
		Secondary volume	1 to 3 vehicles per cycle
		Speed	5 to 10 percent
Saturated intersection control[e]	Continuously monitored	Link content	1 to 3 vehicles

[a]Computed on the basis of keeping timing errors below 2 to 5 sec.
[b]Measured at locations that provide data representative of need for timing pattern selection.
[c]Volume and queue updated continuously on minor phase.
[d]For saturated intersections, only total approach volume is required.
[e]Data needed at all saturated intersection control intersections.

The second generation strategy also has CIC. In this case, the CIC adjusts both split and offset for every signal cycle. As with first generation, split is adjusted in response to relative approach volumes. Offset is adjusted to accommodate queues that have built up because of secondary flow and variations in vehicle speeds.

THIRD GENERATION CONTROL STRATEGY

The most complex of the control strategies being developed for UTCS is the third generation of control. This strategy consists of 2 levels of control selected on the basis of traffic demand.

1. Medium flow control computes timing patterns at intervals of approximately 5 min. This control mode permits cycle length to vary at adjacent intersections. Cycle length also can vary at a given intersection from one cycle to the next. Under these conditions, split and offset also will vary constantly in both time and space. Thus, it is no longer convenient to treat signal timing in terms of the variables cycle, offset, and split. In both this mode and the congested mode, optimization computes signal timing as green-on and green-off times for each approach.
2. Congested flow control operates when traffic at either a single intersection or a group of intersections builds up to the point at which the intersection can no longer accommodate all of the vehicles arriving during a signal cycle. In this mode of operation, congested intersections are identified and cycle lengths are increased to maximize their throughput. In addition, traffic from upstream intersections is gated into the congested intersection in a manner that will prevent spillback across the upstream intersection. The gating is designed to prevent buildup of congestion around a closed loop of streets; in effect, traffic backs up around the block. Signal timing is computed continuously in this mode of operation to determine green switching times. Because of its cycle-free characteristics, third generation strategy does not require a critical intersection control capability.

CONTROL-STRATEGY DATA REQUIREMENTS

It is evident that each of these control strategies will have differing data requirements (Table 2). In the table, queue length can be replaced by secondary flow because either of these variables can be used to determine the number of vehicles that must be discharged before the main group for the upcoming cycle.

Both first and second generation control strategies have areawide control as well as single-intersection control, which is intended to fine-tune the areawide control settings. This implies 2 distinctly different levels of data requirements existing within the same control strategy. As a result, the costly deployment of large numbers of detectors can be limited to those intersections requiring critical intersection control. Error range is provided in the table as an indication of the level of accuracy that can be anticipated from the surveillance system rather than from a reflection of the actual requirements of the control strategies. Other data requirements not included in this table fall into the following 2 categories:

1. Threshold values used to determine the mode of operation of the control strategies and
2. Parameters used to model the traffic system in the optimization process of the control strategies.

In the first case, threshold values are most often used to identify the existence of saturation. For example, the first generation of control defines the existence of saturation as the buildup of the standing queue past the furthest upstream detector at any time during the red signal state for that link. This threshold is necessary because the CIC algorithm requires the measurement of B-phase demand during A-phase green.

Figure 1. Start-up delay frequency distribution.

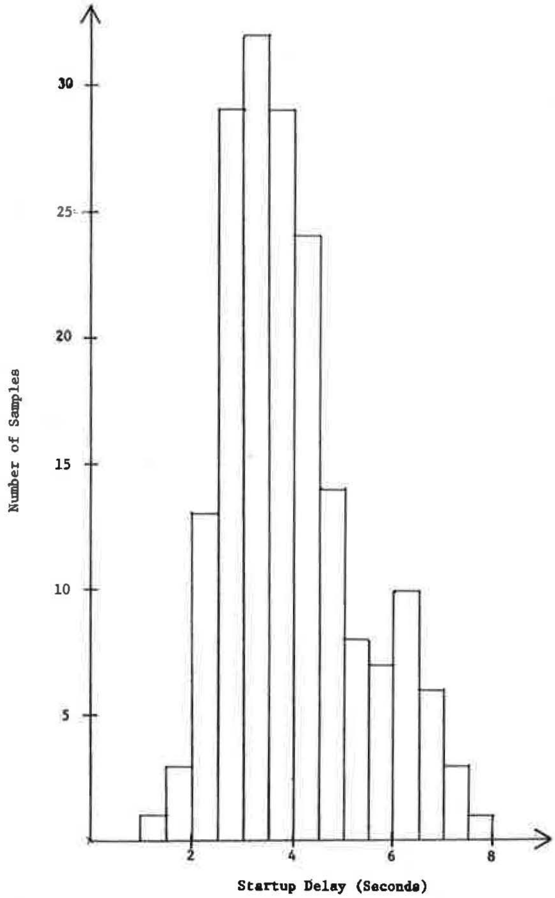
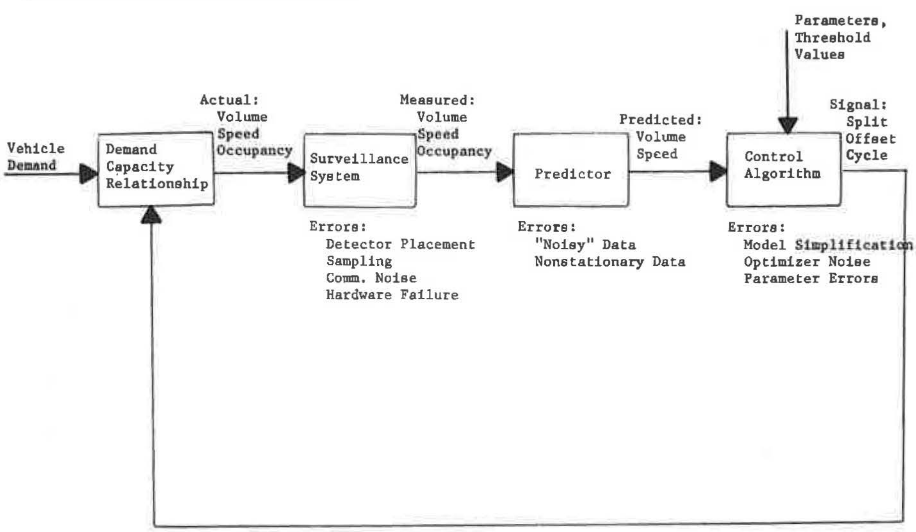


Figure 2. Traffic-control system and sources of error.



This approach is used to determine A-phase duration in approximately the same manner as the computation of a yield point in a semiactuated controller. Street surveys have determined that queue buildup past the upstream detector on a 2-detector link [with a detector at the stop line and a second detector 210 ft (64 m) upstream in the same lane] can be identified reliably by a value of occupancy of 35 percent. In this case, occupancy is computed as the percentage of time that vehicle presence is measured by the upstream detector. These data are smoothed by using first-order smoothing as follows:

$$\overline{O}_i = \overline{O}_{i-1} + k\,(O_i - \overline{O}_{i-1})$$

where

k = smoothing constant (a value of 0.5 currently is being used),
O_i = value of occupancy measured during signal cycle i, and
$\overline{O}_i$ = smoothed value of occupancy at signal cycle i.

Second generation software uses a similar technique to control the operation of its critical intersection control algorithm; third generation requires this type of threshold to change from moderate-flow to congested-flow modes of operation.

In the second case, data requirements for control strategies are often overlooked in the design of the strategies. These requirements are the parameters used in the optimization process. The following are examples of these parameters:

1. Start-up delay,
2. Discharge headways,
3. Number of lanes,
4. Link lengths (intersection spacing), and
5. Group dispersion.

Because a surveillance system rarely is designed to measure these parameters in real time, the developer of the control strategy must treat these input parameters as systemwide, link-specific, or time-of-day constants. A link-specific constant is rarely selected because the cost of detailed link-by-link data collection for all times of day is extremely high. Yet to treat these parameters as systemwide constants can result in serious errors in the optimization process. For example, Figure 1 is a histogram of the start-up delay measured at 14 locations in Washington, D.C. This figure indicates that start-up delays of between 2 and 7 sec are common. The variance in this parameter can result in offset errors that will cause increases in stops within the network because inadequate queue discharge times will be used to account for the larger start-up delays. It will not have as great an effect on delay unless use of incorrect start-up delay causes inadequate green time to be assigned to a phase resulting in saturation. This study was undertaken as an attempt to determine the cause of numerous incorrect values of offset arising from the TRANSYT optimization of the first generation signal timing. Obviously, this type of problem, which arises in an off-line optimization, is equally likely to occur in an on-line control strategy.

SOURCES OF ERROR

Some of the potential sources of error in an on-line traffic control system have been discussed. They occur throughout the control process and can be controlled only by more complex surveillance, off-line data gathering, and sophisticated processing techniques. All of these measures will result in increased system cost, which must be balanced against the potential benefits of a responsive control system. None of these measures will completely eliminate error in the control process.

Figure 2 shows a summary of the various sources of error in a traffic-control system. Each of these errors results in a control computation that is suboptimal for the "actual" conditions on the street and will result in a degradation of system effectiveness. Such a degradation can easily result in a responsive system whose operation is less effective than that of a first generation system that could be implemented at a much lower cost.

It is convenient at this point to select vehicle volume as the variable that will be emphasized in the remaining discussion because in most cases it will have the most significant effect on degrading the quality of the control. Furthermore, it is clearly beyond the scope of this paper to analyze the effects of each of the many other variables in the traffic-control-system operation.

It can be seen from Figure 2 that the sources of error related to vehicle volume are detector placement in the surveillance system and the characteristics of the data processed by the surveillance system (noisy data and nonstationary data). It is difficult to present the generalized statistics of errors that can be expected from the limitations of detector placement because these errors are closely related to the characteristics of the street on which the detectors are installed. Assuming that the correct critical lane (flow lane with maximum volume) has been instrumented, volume errors will result from lane changing, channelization, midblock sources and sinks, and queue buildup. Many of these factors will cause errors that are not zero mean errors (because these errors are correlated serially with the volume). This is significant because a zero mean process is often assumed when the effect of volume errors on operation is analyzed.

An indication of the magnitude of surveillance system errors can be seen from a limited study of 4 locations in the UTCS network; the results showed hourly volume errors with mean error values of 38.5 vehicles per hour and standard deviations of 65 vehicles per hour (9). This type of error should not necessarily be considered typical because it was measured at some of the worst locations in the UTCS network and instrumentation changes are currently under way to reduce their effects. They are presented as an indication of the potential magnitude of the problem and to point out that surveillance errors can have large mean values. Surveillance system errors can often be controlled by increasing the number of detectors in a network. Prediction errors are a result of the characteristics of the data being processed. These data can be described in the following terms:

1. Volume data contain both a time-varying mean and variance (nonstationary); and
2. Spatial and temporal correlations of volume data are low and might also be time varying.

Typical spatial and temporal correlations are shown in Figure 3 (5, 10). All data in this figure refer to the L Street approach to the intersection of L and 15th Streets. This is the reason for the correlation of 1.00 in this link. The correlations of 4 cycles indicate the value of upstream data for predictor operation. Obviously, the poor correlation shown in Figure 3 implies that there are inadequate data on which to base the prediction. For this reason, the most successful predictors developed to date have relied heavily on historic data, that is, data derived from previous days with similar characteristics. Complete reliance on historical data would eliminate the need for a traffic-responsive system because the use of the same data from 1 day to the next would result in the same signal-timing patterns each day. This would be, in effect, a fixed-time operation.

There have been approximately 9 different predictors developed for the UTCS project (4, 5). Each of these predictors has been developed on a different basis, yet most have resulted in error distributions of the type given as follows for 100 links, 46 predictions per link (4):

Figure 3. Spatial and temporal correlations of link volumes.

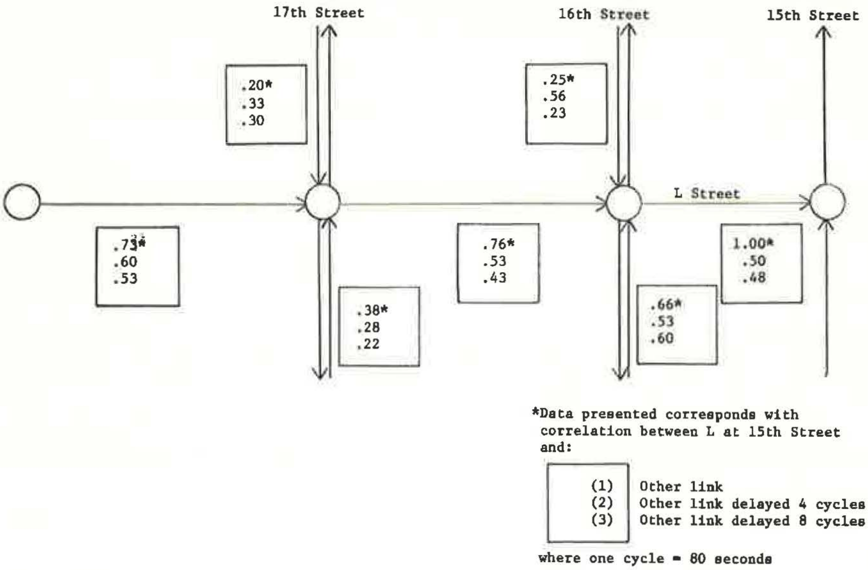
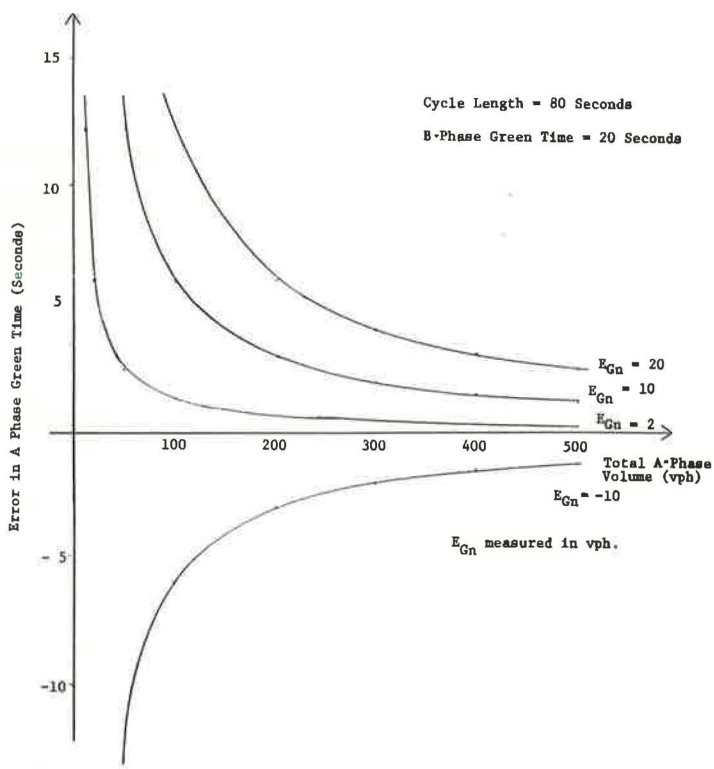


Figure 4. Errors in green time resulting from volume errors.



Volume Predictor	Error
$\overline{AM1}$	0.123
P05	0.719
P10	0.466
P20	0.175

where

$\overline{AM1}$ = mean AM1, which is $\left|\frac{f(t) - p(t)}{f(t)}\right|$,
P05 = Pr(AM1 $\geq$ 0.05),
P10 = Pr(AM1 $\geq$ 0.10),
P20 = Pr(AM1 $\geq$ 0.20),
f(t) = actual volume, and
p(t) = predicted volume.

Standard deviations of predictor error range between 7 and 18 vehicles per hour depending on the variability of the volume data on which the predictions are based. This value of error for the predictor operation, which is not controllable, provides the system designer of a traffic-responsive system with an indication of the acceptable level of surveillance errors. Thus it can be concluded that a good surveillance system design is one that results in volume measurement errors with a standard deviation that is less than 7 vehicles per hour when it is used in the modes given in Table 2. This information would be applied to detector location at an intersection approach in the following manner:

1. Identify through lane carrying the largest volume;
2. Select detector location as far back from intersection as possible but downstream from any major sources or sinks such as parking garages;
3. Measure lane volume at intersection and compare it with volume passing over selected detector location; and
4. Compute standard deviation of difference between measurement of volume entering intersection and volume at selected detector location at each signal cycle.

Perform test for 30 signal cycles during both peak periods and midday. Standard deviation should be less than 7 vehicles per hour (vph). Use of 30 samples is recommended based on past experience with similar measurements.

SENSITIVITY OF SPLIT COMPUTATION TO VOLUME ERRORS

An example of the effect of volume errors on control-strategy operation is the relationship between these errors and split computation. If the split for each intersection is computed by using green demand, which is assumed to be equivalent to total approach volume, the time for the nth phase (t_n) can be written

$$t_n = \frac{G_{Dn} \cdot C}{G_{Dt}}$$

where

C = cycle length,
G_{Dn} = green demand on phase n, and
G_{Dt} = total green demand on all phases.

For the purpose of this discussion, green demand is flow-lane volume in vehicles per cycle arriving at the approach to the intersection serviced during phase n. An error in volume measurement is equivalent to an error in green demand on phase n (E_{G_n}) that causes error in t_n (E_{tn}); the following expression results:

$$t_n + E_{tn} = \frac{(G_{Dn} + E_{Gn}) \cdot C}{G_{Dt} + E_{Gn}}$$

Assuming that E_{G_n} is much less than G_{Dt}, we can solve this expression for E_{tn} by substituting the expression for t_n that yields the result

$$E_{tn} = \frac{E_{Gn} \cdot (C - t_n)}{G_{Dt}}$$

If this equation is applied to the major phase of a 2-phase intersection, C - t_n is equivalent to the minor-phase time at the intersection. The results are plotted in Figure 4.

One approach to relating the effect of split errors to network performance is to compute the increase in intersection delay resulting from the incorrect computation of green time. If the split error does not cause the intersection to become oversaturated, and the offset is not affected, only the random vehicle delay at the intersection will be changed. Random delay is defined as the correction added to computation of vehicle delay to allow for cycle-by-cycle variations from average behavior. The random delay correction is modeled in the TRANSYT signal timing program as follows (6):

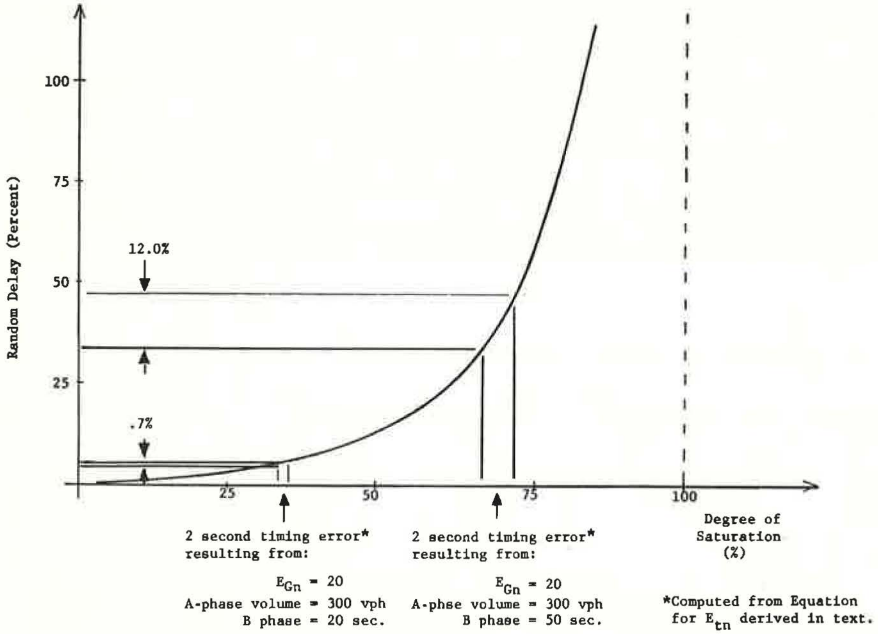
$$D_R = \frac{1}{4}\frac{x^2}{1 - x}$$

where x = the degree of saturation or, in other words, the fraction of green time during which vehicles are discharged through an intersection. From this relationship, which has been considered by other investigators (5, 7), it can be seen that a split computation error resulting in a reduction of available green time and an increase in degree of saturation will have a greatly increased effect on the delay experienced by motorists at that intersection. These effects are shown graphically in Figure 5, which indicates that the sensitivity of delay to errors in green time depends on the degree of saturation existing at the intersection. This is not a surprising result because it is equivalent to the statement that incorrect signal timing at an intersection will have a more serious effect on the intersection's operation under heavy traffic. What is surprising about this result is that a split error of only 2 sec for a degree of saturation of 75 percent can produce an increase in delay of 12 percent. This is equivalent to the level of improvement anticipated from a traffic-responsive system (Fig. 5). The 2-sec error was produced from an error in estimated green demand of 20 vehicles per hour, which is a value that is probably less than the standard deviation of that total error resulting from combined surveillance and prediction errors.

CONCLUSIONS

This paper has attempted to present some recent results of research resulting from the UTCS project. The research has demonstrated the existence of rather large data errors within a traffic-control system that have the potential for significantly degrading the operation of that system. These errors can be minimized only through careful surveillance of system design and creation of a large and detailed data base to serve the control-strategy operation.

Figure 5. Random delay at undersaturated intersections.



If the system designer is not willing to undertake these measures in the implementation of a real-time-responsive control system, the resulting system operation could be less effective than that of a pretimed system.

REFERENCES

1. J. Holroyd and J. A. Hillier. The Glasgow Experiment: Plident and After. Transport and Road Research Laboratory, England, RRL Rept. LR384, 1971.
2. Improved Control Logic for Use With Computer-Controlled Traffic. Final Rept., NCHRP Project 3-18(1), Nov. 1973.
3. Urban Traffic Control and Bus Priority System—Operator's Manual, Volume 2. FHWA, NTIS No. PB 214 641, Sept. 1972.
4. Urban Traffic Control System and Bus Priority System Traffic Adaptive Network Signal Timing Program—Software Description. FHWA, Aug. 1973.
5. Variable Cycle Signal Timing Program—Volume 3, CYRANO: Cycle-Free Responsive Algorithms for Network Optimization: Moderate, Congested, and Light Flow Regimes. FHWA.
6. D. I. Robertson. TRANSYT Method for Area Traffic Control. Traffic Engineering and Control, Oct. 1969, pp. 275-280.
7. R. Wormleighton. Queues at a Fixed Time Traffic Signal With Periodic Random Input. Canadian Operations Research Society Journal, Vol. 3, No. 3, Nov. 1965, pp. 129-141.
8. K. Keitt and H. Lum. Study of the Characteristics of Startup Delay at Urban Intersections. FHWA, Internal Rept., Sept. 1974.
9. G. Smith, H. Lum, and J. MacGowan. Traffic Volume on UTCS Links. FHWA, Internal Memorandum, Feb. 1974.
10. N. R. Draper and H. Smith. Applied Regression Analysis. John Wiley and Sons, Inc., New York, 1966, p. 33.

DISCUSSION

Dale W. Ross, DARO Associates, Inc.

Tarnoff has made a number of important points on how various errors can degrade the performance of a computerized traffic surveillance and control system. It is indeed time that some serious research was devoted to this subject. The paper does not, however, substantiate the implied hypothesis that data errors can cause a real-time, traffic-responsive control system to be less effective than a pretimed system. The paper tends to exaggerate these effects beyond what may be the actual situation. Some other points should be considered to see that the effects of data errors may have been exaggerated.

RANDOM DELAY FORMULA IS UPPER BOUNDING ONLY

The analysis and example in the paper assume that the formula

$$D_R = \frac{1}{4}\frac{x^2}{1 - x}$$

is an accurate or representative model of random delay at a signalized intersection. Careful reading of the field experimental results from which this formula was developed shows that this relation is not actually a model for random delay itself (11), but rather that the formula represents an upper bound or envelope for the field data taken on random delay. Robertson points out that there is considerable scatter in field observations of random delay (11). This is particularly the case at higher degrees of saturation. The effect on the paper is that the use of this formula must be considered a worst-case analysis, and that the actual sensitivity of random delay to data errors can be an order of magnitude less than that derived by using the above formula. The use of this formula in TRANSYT was based not so much on its being an accurate model of random delay as it was on its being a means of forcing the TRANSYT model to select phase durations and cycle lengths that led to saturation levels of less than 90 percent (6, 10). The deliberate exaggeration of the random delay by this formula at large degrees of saturation thus served as a built-in means of ensuring that the TRANSYT model would not select unreasonable phase durations. The paper misconstrues the use and meaning of the formula.

RANDOM DELAY IS LESS THAN TOTAL DELAY

The paper considers only the random delay component of total intersection approach delay. The other primary component is the deterministic delay due to offset and phase durations. Robertson shows that even when degree of saturation is as large as 90 percent and the offset is the best possible, it is typical to find that the random delay is no more than half of the total delay (11, Fig. 9). Consequently, the sensitivity of total delay due to timing errors is less than the paper indicates.

ASSUMED VOLUME ERRORS ARE LARGE

The text table on distribution of prediction errors indicates that the UTCS has had a volume prediction mean error of about 12 percent. This seems to be about twice as large as results that are being obtained in other current experimental work (12). In fact, according to J. Lam and D. Kaufman of the Corporation of Metropolitan Toronto, experimental results with a predictor similar to that used in the ASCOT system, which has been described elsewhere (2), have shown prediction to be on the order of 4 to 5 percent. Thus the inferences made in the paper may be overstated because of the assumption of fairly large prediction errors.

EXAMPLE USES CYCLE LENGTH THAT EXAGGERATES ERRORS

If one carefully examines the example given in the paper pertaining to Figure 5, one finds that the example assumptions (an A-phase degree of saturation of 0.75, an A-phase volume measurement error of 20 vehicles per hour, a B-phase duration of 50 sec, an A-phase timing error of 2 sec, and an A-phase volume of 300 vehicles per hour) are consistent among themselves only if the sum of the A-phase and B-phase volumes is 500 vehicles per hour, and if the cycle length is 125 sec. The assumption of a 125-sec cycle length for the 2-phase signal is somewhat unrealistic. If one had used Webster's method (13) to select a near-optimum cycle length for the intersection, the cycle length would have been chosen to satisfy the relation

$$C_o = \frac{1.5\ L + 5}{1 - Y}$$

where

C_o = optimum cycle length,
L = total lost time for the intersection, and
Y = sum of the volume-to-saturation flow ratios for the phases.

With the parameter values used in the paper, one can verify that this formula would have yielded an optimum cycle length of 125 sec only if L had been approximately 17.5 sec. This is an inordinate amount of lost time for a 2-phase signal. If a more reasonable lost time of 6 to 8 sec per cycle were used, one would find that approximately 60 sec would be the optimum cycle length. Thus a more reasonable cycle length for the parameter values given in the paper would have been 60 sec instead of 125 sec. If the 60-sec cycle had then been used in the analysis of the paper, it would have been seen that the effect of the 20-vph volume measurement would then have been only a 0.96-sec timing error instead of the 2.0-sec error in the paper. Thus the cycle length assumed in the paper perhaps overstates the magnitude of the error by a factor of about 2.

These 4 points indicate that the paper probably exaggerates the effects of data errors. The paper makes a good point that these errors need further study, but one should not make hasty conclusions regarding the effect of such errors. In particular, the analysis in the paper should not be misconstrued as meaning that real-time, traffic-responsive control systems are likely to be less effective than pretimed systems. One needs to be careful to draw such conclusions only from well-founded research results.

RECOMMENDATIONS

The introduction of the paper identifies 5 factors that UTCS researchers are investigating regarding the sensitivity of real-time control to data errors. It is suggested that the list of factors be expanded to 7; the 2 additional factors in the analysis would be

1. Investigation of the surveillance and control algorithms that provide the best compromise between good signal timings and insensitivity to data and parameter errors, which would require cross-testing of the UTCS work, British work (1), Canadian work (12), and ASCOT work (2), and
2. Investigation of programming and programming-induced errors such as

round-off and truncation errors in computations.

The last factor is one that should not be overlooked. Real-time software systems for traffic control are fairly intricate, and even the most brilliant programmers and engineers can make several subtle errors in the programming that do not evidence themselves in an obvious, consistent manner.

The issue of software or programming errors was, in fact, a significant factor in the ASCOT field results (2), and because I was principal investigator for ASCOT development and have continued to apply ASCOT techniques, I rebut Tarnoff's comment that the ASCOT development and tests met with a "lack of success." First, "success" can be measured in different ways. In some respects the ASCOT development was quite successful. It demonstrated that it is possible to achieve highly flexible methods of traffic control by using limited computational resources and that such control was well within the capabilities of most minicomputers. Also, the city of San Jose, California, continues to use the ASCOT system on a day-to-day basis, and has even expanded its use of ASCOT from a 12-hour control day to an 18-hour control day. Furthermore, techniques and methods used in ASCOT are finding application in other cities, including Chicago (14) and Toronto (12).

Tarnoff is correct in stating that in the San Jose tests of ASCOT, results were inconclusive regarding the effectiveness of ASCOT versus the effectiveness of pretimed operation. Some of the lack of improvement has been attributed to deficiencies in the offset optimization logic of ASCOT, and these are reported elsewhere (2). It is now known that several major software errors have been discovered in ASCOT, and these are major reasons for the lack of improvement.

The field tests of ASCOT were conducted in the summer of 1973, and at that time every possible effort was made to ensure that the software had been carefully screened and tested for programming errors. In the spring of 1974, a study was begun to develop documentation of ASCOT for San Jose's operating and engineering personnel, to develop additional programs for the evaluation of the system by using surveillance data, and to conduct a review of the software system to identify possible improvements and errors. The work revealed several software errors that had not been known at the time of the field tests and later (2). Here are some of the major errors that were found.

1. The ASCOT logic for computing offsets depends on the TRANSYT traffic-flow model for modeling the platooning of traffic and choosing offsets tailored to the platooning. It was found that in programming this model, link IN-patterns were incorrectly computed from the sum of upstream OUT-patterns. If one refers to the equations given by Robertson (11, p. 18), the correct equation is

$$q'(i + t) = F \cdot q(i) + (1 - F) \cdot q'(i + t - 1)$$

Instead of that equation, ASCOT had been programmed with the equation

$$q'(i + t) = F \cdot q(i) + (1 - F) \cdot q(i + t - 1)$$

(Primed variables are IN-pattern variables; unprimed variables are upstream OUT-pattern values.) The consequence of this error was that platooning was not correctly represented, and offsets could not be selected properly.

2. TRANSYT GO-patterns were organized in disk memory in groups of 10 links. One indexing error prevented any GO-patterns that had been computed for the last group of 10 links in any intersection group (subset) from ever being written to disk. Consequently, GO-pattern data for such links were missing, which led to erroneous TRANSYT platoon modeling.

3. The ASCOT CIC method depended on a subroutine to add the variable controller intervals on the current phase of each CIC-controlled intersection to determine the optimum time to switch from 1 phase to the next. This summing subroutine was programmed incorrectly, which led to incorrect estimates on the best time to switch phases.

Further investigation of ASCOT beyond the initial field tests has revealed that programming errors existed that had major consequence. The programming errors discovered were subtle errors and others may still exist. The point of all this is that data errors are only a part of the picture and that software errors also should be recognized as important. It often takes years to completely "iron out" a new software system, and further research should be devoted to improving means of reducing such errors.

REFERENCES

11. D. I. Robertson. TRANSYT: A Traffic Network Study Tool. Transport and Road Research Laboratory, Harmondsworth, England, RRL Rept. LR253, 1969.
12. Improved Operation of Urban Transportation Systems. Ministry of Transport, Canada, Ministry of Transportation and Communications, Ontario, and Corporation of Metropolitan Toronto, study in progress.
13. F. V. Webster. Traffic Signal Settings. Transport and Road Research Laboratory, Harmondsworth, England, Technical Paper 39, 1958.
14. Specifications and Contract Document 80.65-73-1. Dept. of Purchases, Contracts, and Supplies, city of Chicago, Sept. 1973.

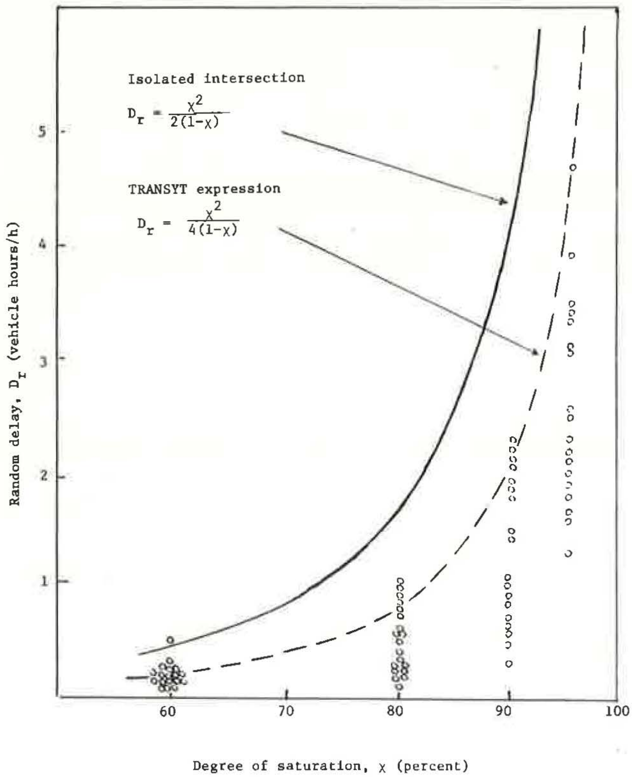
AUTHOR'S CLOSURE

Ross interpreted this paper as being a negative viewpoint on traffic-responsive strategies in general and the ASCOT program in particular. The conservative outlook expressed in this paper relative to potential improvements in traffic flow that might be possible with traffic-responsive strategies was not intended to reflect adversely on his work, which is recognized to have been performed under budget and time constraints. Nevertheless, the fact remains that neither ASCOT nor the British traffic-responsive control strategies have materially improved traffic flow in the cities where they were tested. Preliminary experience with the UTCS second generation control strategy, which has undergone both simulated and real-life testing, has produced similar results. Thus, this paper was written as an attempt to present an objective explanation for these results.

Ross stated that the random delay formula used in the paper exaggerates the effects of random delay. The basis on which this statement is made is a sentence by Robertson (11) taken out of context: "This curve... is seen to exaggerate the mean random delay at the higher saturation levels." As can be seen from the curve presented in this reference and reproduced here (Fig. 6), the higher saturation levels referred to are above 80 percent. Yet the example that Ross claims exaggerates the result uses a saturation level of 75 percent, a value that was selected specifically to avoid the possibility of exaggerated results.

Ross indicated that other delays are more significant than random delay and as evidence again referenced Robertson (11, Fig. 9). The accuracy of this depends on signal offset, degree of intersection saturation, and the ratio of primary to secondary vehicle flows. The use of an arbitrary example to support such a statement is hardly conclusive. Furthermore, for the Ross statement to be correct, the effects of the split errors discussed in the paper would have to be more pronounced than those that

Figure 6. Variation of random delay with saturation.



were presented in the example because the effects of nonrandom delay are additive with the random delay. Thus, the results presented here can be considered understated rather than exaggerated results as Ross implies.

Ross referred to research performed in Toronto as evidence that the UTCS prediction volume errors are quite large. The UTCS predictor was tested by using data from Toronto. The results of these tests were errors in the same 4 to 5 percent range experienced by Lam of Corporation of Metropolitan Toronto. However, volume errors with a historically based predictor depend on the daily variation in vehicle volume. The Toronto system has the relatively smooth repeatable traffic-flow characteristic of suburban arterials. In the UTCS network the opposite is true, and a degradation in predictor performance results. Therefore, the errors presented in the paper must be considered typical of those that would be experienced in the central business district of a major U.S. city.

Although Ross is correct in stating that the cycle length chosen for the 75 percent-saturated case is long compared to cycle lengths generally used in coordinated signal systems, his assumption that the cycle length at every intersection must satisfy Webster's equation for "optimum" cycle length is not correct in all cases. In a coordinated signal system, cycle length is selected to satisfy the intersection requirements of longest cycle length and minimum green times as dictated by pedestrian crossing times. Furthermore, Webster's equation produces optimum results only at an isolated intersection with random arrivals and is not applicable for a network with platooned arrivals.

Perhaps the most important point is the fact that the effect of the 2-sec timing error was the purpose of the example and was not the particular set of circumstances that produced it. For example, a similar 2-sec error could have been produced by a cycle length of 80 sec with A-phase time of 50 sec, A- and B-phase volumes of 500 and 300 vehicles per hour respectively, and an A-phase volume error of 53 vehicles per hour (less than 1.2 vehicles per cycle).

In conclusion, it must be stated that none of the specific points raised by Ross in any way detracts from the content or conclusions of the paper. Although there is still a place for traffic-responsive strategies in cities that plan to install additional hardware as a substitute for manual updating of traffic signal timing, the potential of these strategies for improvements of traffic flow is far from assured. On the basis of available information, it is the responsibility of every research organization to avoid raising the false hopes that traffic-responsive strategies in their current form can provide a major improvement in urban traffic flow.

AN EVALUATION OF SIGNING TO WARN OF POTENTIALLY ICY BRIDGES

Fred R. Hanscom, BioTechnology, Inc., Falls Church, Virginia

This field study examined driver responses to a potentially slippery bridge during periods of possible preferential icing. Study objectives were to examine motorists' general awareness of the hazard and to assess the relative effectiveness of various warning sign treatments. Measures of signing effectiveness were motorists' speeds at critical bridge approach locations and questionnaire responses regarding motorists' observations and interpretations of the signs. Two bridge approaches were signed with combinations of activated and nonactivated signs at the bridge and 1,000 ft (305 m) before the bridge during periods of possible preferential icing. Significant speed reductions on the bridge and at the bridge entry were elicited by activated signing. The most effective signs were (a) activated, before the bridge and (b) activated, at the bridge during hours of darkness. Activated signing used at the bridge was observed to have a greater impact than activated signing used before the bridge. Drivers were more responsive to the signs during hazardous periods. Bridge-approach roadway geometry was seen to affect motorists' observation of and response to the signing. Improved results were obtained on a short sight-distance approach where the bridge did not visually compete for driver attention.

•LOCALIZED bridge icing poses a severe threat to motorists' safety throughout much of the year in many regions. Traffic engineering problems associated with remedying this threat are compounded by many diverse issues: unreliability of ice detection devices, complexities of legal liability, and credibility of motorists' warning devices. The first of these issues is the subject of much past and current research; the second is the subject of many past and current tort liability suits; and the third represents a critical research need. The purpose of the work described in this paper is the development of effective signing to warn motorists of an icy bridge hazard. The subject experiment is a field evaluation of 8 signing schemes that were derived from a review of the literature, a survey of current operational practice, and a preference test.

PREFERENTIAL ICING HAZARD

Preferential icing of bridge decks, a well-known but elusive safety hazard, is the formation of ice on bridge decks when approach roadways may not be icy. Ice will form on any surface when the temperature of that surface is 32 F (0 C) or lower and moisture is applied. There are 2 basic ways in which atmospheric moisture can be applied to a surface: condensation and precipitation. Condensation will occur on a surface if that surface is at the saturation temperature of the surrounding air; the rate at which condensation occurs when the temperature is 32 F (0 C) or lower results in formation of frost on the surface (13). Icing of the surface is caused by precipitation when the surface temperature is 32 F (0 C) or lower. Preferential icing of bridge decks occurs when bridge surface temperature is at or below freezing and the approach

Publication of this paper sponsored by Committee on Traffic Control Devices.

roadway is warmer (because of the earth's heat). The most conducive environmental conditions are moderate daytime temperatures, high relative humidity, and subfreezing night temperatures.

Research has been undertaken to correlate the variables of weather, geographic location, and bridge-deck thermal properties, which lead to preferential icing (2). A study of ice and snow detection and warning system feasibility provides much detail on the physical and meteorological aspects of the problem, highway department maintenance and warning policies, and legal aspects (8).

PRIOR STUDIES

Considerable detailed literature describes ice-detection and warning systems that generally include a warning sign as 1 component of the system (1, 2, 3, 4, 5, 7, 8, 10, 12, 14). However, relatively little effort has been devoted to evaluating motorists' responses to the sign (1, 6, 8, 14). Ice-detector and sign use remains undocumented in a number of states, and research is in progress in others.

A summary of documented ice warning sign evaluations is shown in Figure 1. A Colorado study (1) examined the operation of 2 ice warning systems, but it did not study motorists' responses to the signs. However, a noteworthy observation in the report was that static signing "inconsistent with prevailing conditions" generally was disregarded by motorists. This observation is compatible with one made in the California study that asserted that static ice or frost warning signs are ineffective because they are continuously visible to the motorist (14). These 2 observations, though subjective, substantiate the well-documented fact that motorists are more likely to respond to a warning sign in the presence of a perceived hazard (9, 11, 15). The California study also examined motorists' responses to an activated, flashing, ICY BRIDGE warning sign. Measures of sign effectiveness were the activation of brake lights and vehicle decelerations as recorded by manual observers. The authors

Figure 1. Documented motorist responses to icy bridge warning signs.

Researcher	Measure	Finding
Ballinger, 1966	Subjective Observation	Static signing ineffective — Motorists disregard ice warning signs which are continuously displayed.
Stewart and Sequeira, 1971	Subjective Observation	Static signing ineffective — Motorists do not respond to static signing in place year round.
	Brakelight Application Vehicle Deceleration	Disappointing response — Less than 50% of motorists responded to a flashing "ICY BRIDGE" sign.
Culp and Dilhoff, 1970	Accident Rates	Significant accident reduction — Before and after study of static "WATCH FOR ICE ON BRIDGE" sign at 24 pairs of test and control locations.
Glauz and Blackburn, 1971	Vehicle Speeds Traffic Lane Volume Weaving Brakelight Application Driver Interviews	Varied response — "ICY BRIDGE" with flasher — Sign accounted for average speed reduction of 7 mph. — 65% interviewed motorists saw sign. — Better overall response in presence of hazard.
Kentucky Dept. of Highways, not dated (unpublished)	Vehicle Speeds	Flashing sign effective — 85th percentile speeds were substantially reduced by flashing "REDUCE SPEED-ICE ON BRIDGE" sign.
Arizona Highway Dept. 1971 (unpublished)	Vehicle Speeds	Illuminated sign ineffective — Static "BRIDGE AHEAD" sign combined with illuminated "ICE" panel had little, if any, effect on 85th percentile speeds.

Note: 1 mile = 1.6 km.

stated that the results were disappointing: Motorists' responses ranged from 23 percent to 66 percent. It is noteworthy that significantly higher responses were obtained during conditions of fog and its accompanying reduced visibility.

An accident study of static WATCH FOR ICE ON BRIDGE signing was conducted in Ohio (5). Before-and-after accident reduction rates were analyzed for 24 site pairs, each of which comprised a test and control bridge location. Signing was placed at test sites during winter months for 3 consecutive years. Reductions in accident rates were realized at 41 of the 48 study sites; however, significantly greater reductions at test locations were evidence that "driver awareness, attributable to the signing," reduced accidents (5). Significant reductions were noted for wet and dry as well as icy conditions. The signing used was identical to the WATCH FOR ICE ON BRIDGE used in this study and similarly was located before bridge locations.

A field evaluation of motorists' responses to an ICY BRIDGE AHEAD sign was conducted as part of the Glauz et al. study (8). The fixed message sign, an advisory speed limit panel, and an amber flashing light were mounted on a rotating frame so that they could be displayed to motorists when conditions warranted. Data collected during the experiment included: (a) vehicular speeds, (b) traffic volume by lane, (c) lane change frequency, (d) brake-light occurrences, and (e) motorist interviews. The principal measure, speed reductions between the bridge approach and the upstream location, showed a statistically significant increase during 3 of 4 periods when the sign was displayed. Average speed reductions of 7 mph (11.3 km/h) were attributed to the signing; larger reductions occurred during periods of localized icing. The data showed no significant effect of the sign on lane change distribution at the bridge, although there was a suggestion that the sign caused some weaving from the right lane to the center of 3 lanes. The data also suggested that during the localized icing the warning sign did not increase braking activity on the bridge approach. In fact, drivers were observed to wait and brake after they were on the bridge. However, when ice or packed snow was on the approach, the warning sign appeared to increase the amount of braking on the bridge approach. The study included interviewing 43 motorists downstream from the bridge. Sixty-five percent said that they had seen an ice warning sign.

Unpublished studies by 2 state highway departments have demonstrated seemingly conflicting results using 85th percentile vehicular speeds. The Kentucky Department of Highways conducted an in-house evaluation of an alternating message sign—REDUCE SPEED, ICE ON BRIDGE. Activation was provided by an ice detection system, and each of the messages was displayed alternately for 2 seconds at a time. Speed-check studies at the sign location, about 1 mile (1.6 km) from the bridge, showed 85th percentile speeds to be reduced from 65 to 35 mph (104.6 to 56.3 km/h) when the sign was activated. However, no information is available on either the novelty effect of the sign or its effect on speeds at the bridge. The study concluded that the sign was effective in warning motorists. The Arizona Highway Department evaluated an illuminated ICE panel mounted on a standard BRIDGE AHEAD sign. Simultaneous sets of speed data were taken on the bridge approach (before the point where the sign was readable) and on the bridge to assess motorists' reactions. The observed speed reductions, noted when the panel was illuminated, were attributed to normative speed variations, and the study concluded that the sign had little, if any, effect on motorists' driving speeds.

To assess the documented effectiveness of ice warning signs based on the reviewed studies, one should examine the common measures used and conclusions drawn. Subjective observations of sign effectiveness by Ballinger (1) and Stewart and Segueira (14) jointly establish that icy bridge signing should be responsive to the immediate hazard. A common inference from the 2 studies is that activated signing is necessary for desirable motorists' responses. Driver brake-light indications were used as a measure of sign effectiveness by Stewart and Segueira (14) and Glauz et al. (8). Both studies indicated that many motorists wait until they reach the bridge before they apply their brakes. As a tool to determine response to the sign the measure appears marginal, as evidenced by the 2 following points. Stewart and Segueira (14) show a significantly higher percentage of brake-light activation for poorer weather conditions.

Glauz et al. (8) point out that considerable speed reduction takes place without brake-light indications and that higher braking frequencies prevailed on certain days both with and without sign use. It appears from these 2 studies that brake-light applications are a response to environmental conditions rather than to signing.

Vehicular speed data obtained by Glauz et al. (8), Kentucky, and Arizona exhibit both conflicts and similarities. The most marked speed reductions were noted in the Kentucky study; however, because no data were collected at the bridge itself, the results are not compatible with the other 2 speed studies. The upstream and bridge observations of the Glauz et al. (average speeds) and Arizona (85th percentile speeds) studies are compared in Table 1. The data are similar in appearance, but the study conclusions conflict. Glauz et al. (8) found speed reductions at the bridge due to signing to be significant at the 0.01 confidence level. Although no formal statistical test was applied in the Arizona study, observed speed differentials were interpreted to have no meaning because of variations observed in the upstream data. However, it should be noted that the reduction of 6.7 mph (10.78 km/h) observed at the bridge between signing conditions is similar to those recorded by Glauz et al. (8). The reviewed studies comprise virtually all available documentation examining motorists' responses to icy bridge warning signs. Because the efforts were aimed at remedying a severe hazard and provided conflicting results, it is evident that more research is needed.

DEVELOPMENT OF EXPERIMENTAL SIGNING

A review of the literature revealed a rather limited use of sign wording and formats to advise motorists of an icy bridge hazard. It was therefore apparent that further surveys should be conducted before designation of the specific signing to be used. Letters of inquiry were sent to numerous highway departments to seek out representative sign characteristics. Responses and information gathered during the literature review provided 24 different sign messages and a diversity of formats. They are shown in Figure 2. It was felt, based on the literature review, that activated signing would be more effective than nonactivated. However, after one considers the financial constraints of highway agencies, the most promising signs of both types remained as candidates for evaluation. Selected signing concepts from those listed in the table were pretested on the basis of the preference rating of 20 subjects, some of whom were knowledgeable in highway sign design.

The signs shown in Figure 3 were selected for field evaluation. Primary sign characteristics studied were activation type and location. Eight combinations of the 4 signs were used at 2 bridge approaches to permit comparisons of activated and non-activated signs, at-the-bridge and before-the-bridge locations, and short and long sight-distance approaches. All signs were displayed both singularly and in combination on both approaches.

The standard diamond 36-in. (91.4-cm) sign with 6-in. (15.2-cm) black lettering on yellow reflective backing was used. Activated advance signing had ICE steadily displayed in brightly illuminated, red, 6-in. (15.2-cm) letters. The activated sign at the bridge location used two 8-in. (20.3-cm) beacons flashing alternately at a rate of 50 times per minute. At-the-bridge and before-the-bridge signs were located 100 and 1,000 ft (30.5 and 305 m) respectively before the bridge.

SELECTION OF TEST SITES

Site selection involved seeking candidate sites that met certain criteria related to the bridge environment and traffic characteristics. The bridge had to represent a potential ice hazard. That is, it had to be in a region where the temperature frequently fell below freezing in winter. Certain other bridge characteristics that would enhance its ice-proneness were sought. The bridge had to be high enough to allow rapid cooling beneath the deck, and it had to be over water that flowed throughout the year. Also certain traffic characteristics were necessary for a meaningful evaluation of signing.

Table 1. Upstream and at-the-bridge speed observations.

	Glauz Study (mph)			Arizona Study (mph)
Location	Lane 1	Lane 2	Lane 3	All Lanes
Upstream				
With sign	70.6	75.1	79.1	70.5
Without sign	71.7	75.1	78.1	73.9
At bridge				
With sign	56.1	60.2	62.7	59.7
Without sign	63.3	68.3	69.4	66.4
Difference				
With sign	14.5	14.9	15.4	10.8
Without sign	8.4	6.8	9.3	7.5

Note: 1 mile = 1.6 km.

Figure 2. Characteristics of signs that warn of icy bridges.

Message	Format	Activation	Documented Usage	Non-Documented Usage
BRIDGE ICY AHEAD	Diamond, black BRIDGE AHEAD on yellow	"ICY" activated by ice detector	Illinois, Michigan, Virginia	Arkansas
BRIDGE ICY WHEN FLASHING	Diamond, black on yellow	Amber flashers ice detector	Virginia	
BRIDGE FREEZES BEFORE ROADWAY	Diamond, black on yellow	Static		Tennessee
BRIDGE FREEZES BEFORE ROAD SURFACE	Rectangular, black on yellow	Static		Pennsylvania
BRIDGES FREEZE BEFORE PAVEMENT	Rectangular, black on yellow	Static		Kentucky
BRIDGES FREEZE BEFORE ROAD	Diamond, black on yellow	Static		Vermont
BRIDGES ICE BEFORE HIGHWAYS	Rectangular, black on silver	Static		Delaware
BRIDGE MAY BE SLIPPERY	Diamond, black on yellow	Static		New Jersey
BRIDGES MAY BE ICY	Diamond, black on yellow	Static		Idaho, Wyoming, Colorado, Nebraska
	With amber flasher	Manual or Ice Detector		North Dakota
CAUTION-BRIDGE FREEZES BEFORE PAVEMENT	Diamond, black on yellow	Static		Connecticut
ICE	Diamond, black on yellow	Static		Arizona
	Rectangular, red neon letters on black	Activated		Oregon
ICY BRIDGE	Rectangular, 8" florescent flashing letters	Ice detector	California	
	Diamond, black on yellow	Manual, folding		South Carolina
ICY-BRIDGE AHEAD	Diamond, black BRIDGE AHEAD on yellow	ICY panel activated by ice detector	Arizona	
ICY BRIDGE AHEAD — 65 mph	Diamond, black on yellow	Amber flashers manually	Missouri	
ICE ON BRIDGE	Diamond, black on yellow	Manual, folding		North Carolina, Missouri, Georgia, Texas, Louisiana
ICY ROAD	Diamond shape neon letters amber flasher	Manually or ice detector	Colorado	
REDUCE SPEED ON ICE ON BRIDGE	Rectangular, 12" letters alternating messages; two seconds each	Ice detector	Kentucky	
SAFE SPEED 25 ICE AHEAD	Overhead illuminated	Manually or ice detector	District of Columbia	
SLIPPERY WHEN FROSTY	Diamond, black on yellow	Static		Minnesota
SLIPPERY WHEN WET OR FROSTY	Diamond, black on yellow	Flare pot or nonactivated	California	
WARNING-ICY SPOTS NEXT MILES	Rectangular, black on orange/yellow	Static		Arizona
WATCH FOR ICE	Diamond, black on yellow	Static		Washington
	Diamond, black on yellow	Manual, folding		Arkansas
WATCH FOR ICE ON BRIDGE	Diamond, black on yellow	Static	Ohio	Mississippi, West Virginia, Indiana, Kansas Montana
WATCH FOR ICE ON BRIDGE	Diamond, black on yellow	Static		South Dakota
	Rectangular, black on yellow	Static		Virginia

Note: 1 in. = 2.54 cm. 1 mile = 1.6 km.

The bridge had to be on a well-traveled interregional route to obtain a sizable population of unfamiliar motorists in the early morning hours (maximum likelihood of preferential icing). However, the vehicle detection sensors that were used function best under low to moderate traffic volumes. Therefore, the desirable type of road was deemed to be a primary 2-lane route that had no parallel Interstate route.

Twenty candidate bridge sites in Virginia, West Virginia, and western Maryland were considered for inclusion into the study. The selected bridge was the US-340 bridge over the Potomac River, which is 2 miles (3.2 km) east of Harper's Ferry, West Virginia. This location is noted for frequent freezing temperatures because of its elevation. The bridge is 2 lanes, approximately 0.4 miles (0.64 km) in length, and about 40 ft (12.2 m) above the river. US-340 at that point has sufficient average daily traffic for data collection beginning at 6 a.m. and a suitable number of nonfamiliar motorists. Fortunately, the location was not affected by the reduction of speed limits imposed by the early 1974 energy shortage.

Two data collection sites were designated as the long sight-distance (westbound) and short sight-distance (eastbound) approaches. Another bridge on US-340 that crosses the Shenandoah River 2 miles (3.2 km) farther west was used as a control site for data collection on the eastbound approach. Identical approach geometry on that bridge made it a well-suited control site. Because the eastbound approach on the control bridge was instrumented, the same motorist sample was used for testing experimental signing effects at the eastbound study site.

DATA COLLECTION PROCEDURES

Two primary data collection techniques were employed. Vehicle performance data were gathered by using the traffic evaluator system, and driver characteristic data were obtained by questionnaire. Figure 4 shows each technique.

The traffic evaluator system consisted of road sensors, manual code switches, and a digital tape recorder. The equipment was small, portable, and easily concealed. The system was a powerful data collection technique that allowed precise measurements of driver-vehicle behavior over large areas of highway. The traffic sensors (tape switches) were extruded plastic devices about ⅛ in. (3.2 mm) high by ½ in. (12.7 mm) wide. The tape switch was an unobtrusive sensor that caused little vibration and noise to a vehicle when it crossed it. The system allowed monitoring of all vehicles in lanes instrumented with the sensors. Associated with the traffic evaluator system was provision for manual code inputs. Thus, randomly selected vehicles were coded to be interviewed, and their speed data were matched with appropriate questionnaire responses.

Interviewing of motorists was conducted during the testing of all experimental signing conditions. Speed data for each vehicle were matched to questionnaire responses for analysis. Interview locations were beyond driver sight-distances from the speed sensors. In this way, unbiased speed data were obtained. Vehicles selected for motorists' interviews were those with sufficient headways that their speeds were not influenced by others in the traffic stream. An interviewing strategy was adopted that permitted certain driver characteristics data to be obtained before the drivers knew that the study related to potential skid hazard. After a brief introduction that advised the motorists that a safety study was being conducted, general questions were asked to derive their familiarity with the site and the level of their driving practice. More specific questions were then asked regarding their assessment of safe speed during possible icing conditions and whether the bridge was always sanded when icy. By this time, the motorist knew the study pertained to potential skidding. The driver was then asked whether the bridge was a potential hazard and, if so, what their cue of the hazard was. In cases in which the experimental sign was not cited as the cue, the drivers were asked if they had seen a warning sign. If they had, they were asked to identify the sign by describing its appearance and message and to rate the sign as being helpful or not helpful.

Figure 3. Icy bridge warning signs that were evaluated.

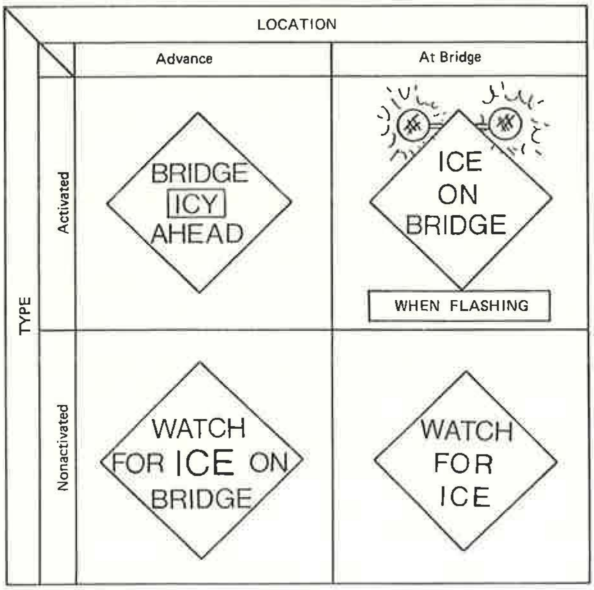
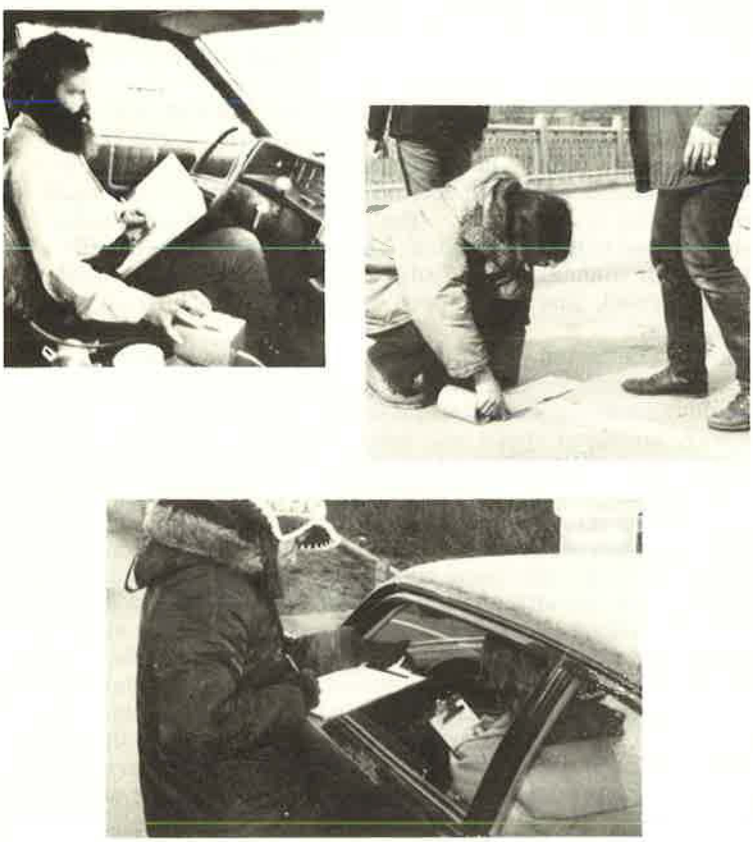


Figure 4. Traffic evaluator system and interviewing.



ANALYSIS OF SPEED DATA

Two approaches to the bridge were used to gather data revealing motorists' responses to the series of warning signs. One approach was characterized by long sight-distance and the other was characterized by short sight-distance. Data were collected on each approach at distances of 1,200 and 600 ft (365.8 and 182.9 m) from the bridge, at the bridge entrance point, and on the bridge at a distance of 150 ft (45.7 m) beyond its entry point. Two sign locations were studied: at-the-bridge [100 ft (30.5 m) before the bridge] and before-the-bridge [1,000 ft (305 m) before the bridge]. Both activated and nonactivated signing were tested during daylight hours and periods of predawn darkness. Ambient conditions were conducive to preferential bridge icing, and frost occurred during some periods of data collection. Data collection could not be accomplished during periods of extreme icing because of the hazard associated with stopping vehicles to conduct interviews.

Experimental sign conditions consisted of the signs shown in Figure 3 (singularly and in combination). Eight schemes were used to determine the effects of activation type and sign location. One day's baseline data were gathered on each bridge approach to permit a sign versus no sign comparison for all experimental sign conditions. Times of data collection were 2 hours before sunrise and 2 hours after sunrise.

Westbound Approach

Motorists' responses to signing in the long sight-distance approach were generally not as favorable as those later observed on the short sight-distance approach. Two reasons related to approach roadway geometry contribute to this effect. First, the relative positioning of the signing with respect to motorists' field of view was less conducive to their observing the signing on this long tangent approach where wide roadway shoulders necessitated a substantial lateral displacement of the signs. Second, the bridge itself was a major competitor for motorists' attention as it came into view before their reaching the advance sign.

An attempt to compare the effects of all signs is shown in Figure 5. An hour-for-hour comparison of each experimental signing condition and its corresponding time period in the baseline data reveals the relative effects of each signing condition. This figure shows mean speed differences ranked so that the most effective signing condition is at the top; statistically significant reductions are indicated. The result is somewhat suspect in that reductions in mean speed were noted for most signing conditions, which is unlikely and contradicts effects that have been shown in the literature. It is likely that normative speeds were higher during the baseline data collection. However, the relative implied effects are noteworthy. A clear differential reduction in mean speeds is seen for the case of an activated, at-the-bridge sign used in combination with a nonactivated, before-the-bridge sign. Promising effects are also evident from other activated signs used singularly and in combination with nonactivated signs. The combination of activated signs at both locations did not perform well during daylight hours. Questionnaire results confirm that fewer motorists saw signing during daylight hours.

To verify or refute the cited differential effects, a more detailed examination was made of the driving samples. Figures 6 and 7 show plots for darkness and daylight observations of mean speeds for both the total and highest quartile samples. Generally, high speeds at the 600-ft (182.9-m), before-the-bridge location are seen to result from the approach grade. No consistent effect on speeds at that location was exerted by the presence of either activated or nonactivated before-the-bridge signing.

Interesting contrasts can be noted in the behaviors of the 2 samples, especially during hours of darkness. Although total-sample mean speeds were generally lowest at the bridge approach, the highest quartile group was still decelerating as it reached the bridge. The faster motorists exhibited greater variability in speeds as they reached the bridge, the greater were their overall approach decelerations, and they generally exhibited greater differential decelerations in response to various signing

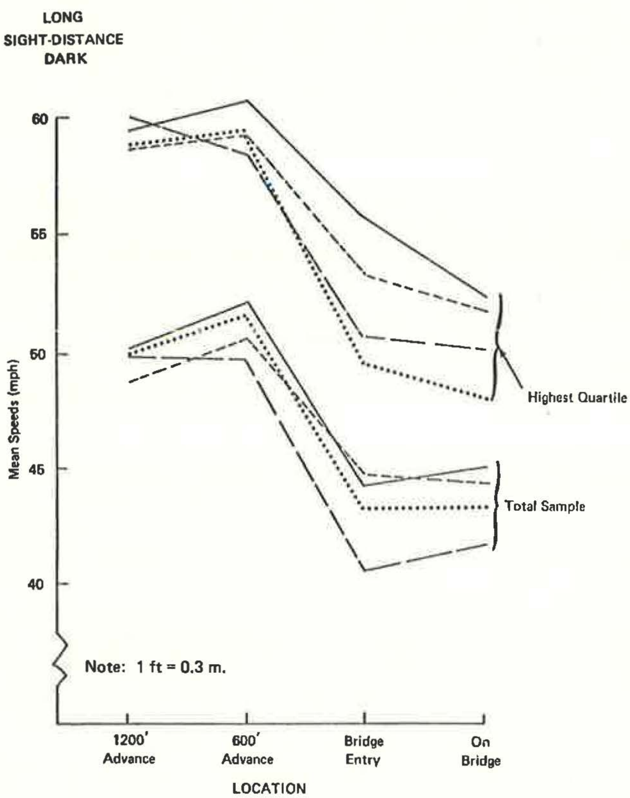
Figure 5. Effects of all signs.

Advance Location	At Bridge	Ambient Condition	Speed Reduction: Bridge Entry	Speed Reduction: On Bridge
WATCH FOR ICE ON BRIDGE	ICE ON BRIDGE WHEN FLASHING	Dark	6.0*	5.9*
BRIDGE ICY AHEAD	WATCH FOR ICE	Dark	4.6*	4.7*
No Sign	ICE ON BRIDGE WHEN FLASHING	Dark	4.4*	4.5*
BRIDGE ICY AHEAD	No Sign	Daylight	4.2*	3.6*
WATCH FOR ICE ON BRIDGE	No Sign	Daylight	3.0*	2.6*
BRIDGE ICY AHEAD	ICE ON BRIDGE WHEN FLASHING	Daylight	2.4*	2.4*
WATCH FOR ICE ON BRIDGE	WATCH FOR ICE	Daylight	.9	1.5
No Sign	WATCH FOR ICE	Dark	9	1.3

Note: 1 mile = 1.6 km.

*Significant reduction from normal condition: a < 0.05.

Figure 6. Mean speeds during predawn hours on the long sight-distance approach.



	Advance	At Bridge
——	No Sign	WATCH FOR ICE
----	BRIDGE ICY AHEAD	WATCH FOR ICE
— —	No Sign	ICE ON BRIDGE WHEN FLASHING
•••••	WATCH FOR ICE ON BRIDGE	ICE ON BRIDGE WHEN FLASHING

Figure 7. Mean speeds during daylight hours on the long sight-distance approach.

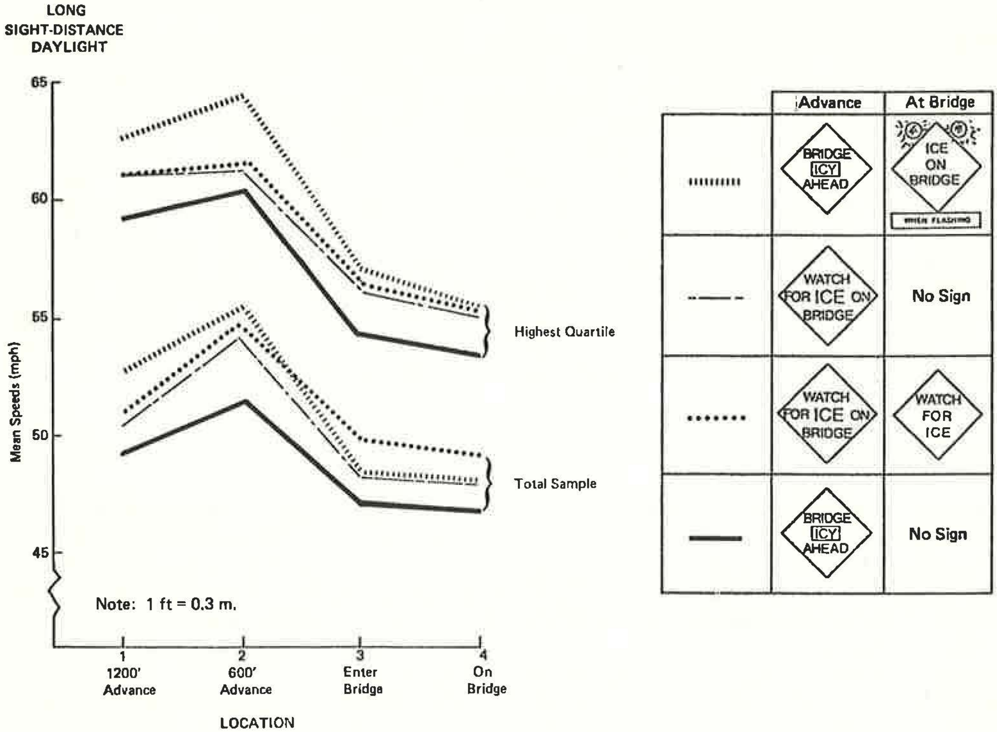
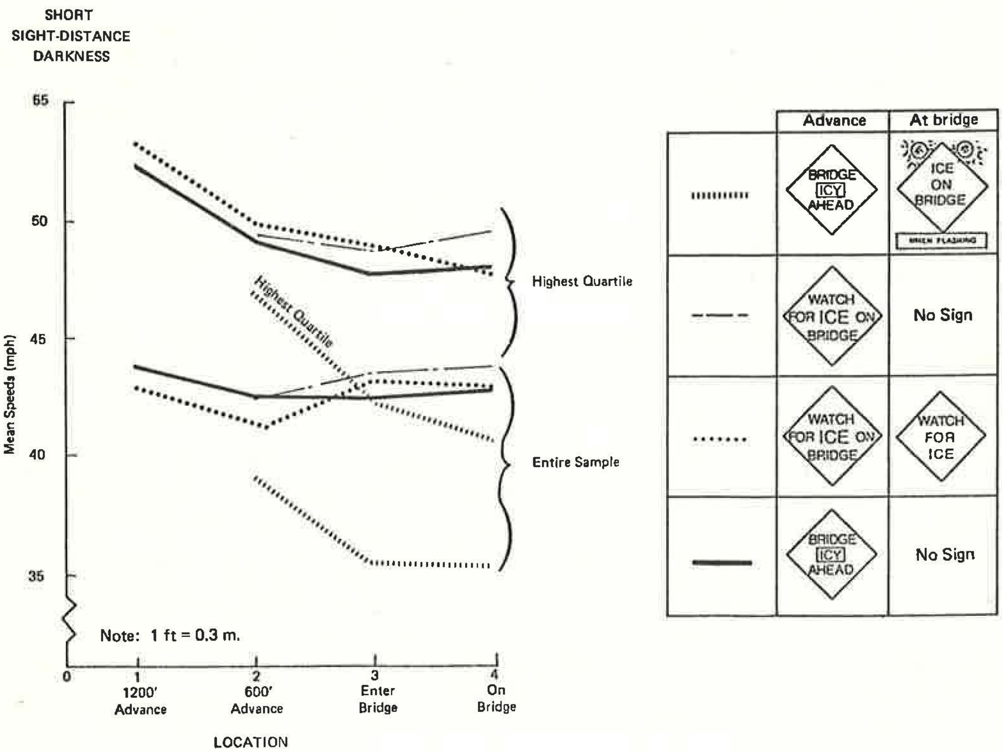


Figure 8. Mean speeds during hours of darkness on the short sight-distance approach.



conditions. The sharpest approach decelerations were observed in response to the activated sign located at the bridge during hours of darkness.

Certain inferences relating to signing effects can be gained from the data shown in Figures 6 and 7. Lowest speeds were obtained in response to activated signing located at the bridge and displayed during predawn darkness. Although the mean speeds were lower for the bridge sign used alone, the highest quartile group slowed more when the accompanying nonactivated advance sign was displayed. The nonactivated, before-the-bridge sign performed well when it was displayed by itself. But direct speed comparison is not the best effectiveness measure in this case because of possible day-to-day variations that could not be accounted for because no control site was available. To eliminate spurious effects, a final judgment of results is based on overall speed reductions obtained for each sign between the 1,200-ft (365.8-m), before-the-bridge location and the bridge during each condition of darkness and daylight. The 2 signing schemes that gave the best performance were the at-the-bridge sign activated by itself and the combination of activated signs at both locations.

Eastbound Approach

An improved experimental method for examining icy bridge warning sign effects was applied at the short sight-distance approach; this was possible because a suitable control site was available. A bridge similar to that of the test site was located 2 miles (3.2 km) upstream on US-340. The identical approach geometry of the 2 bridges created a well-suited experiment site pair. Because there were no major intervening access or egress routes, virtually the same sample of motorists who passed the control bridge was used as the test sample at the experimental site.

Vehicle performance data were gathered in a way similar to that used for the long sight-distance approach to permit determination of the effects resulting from the sight-distance change. Speed data collection points and sign locations were at identical distances from the bridge. The only procedural variation was to reverse the hourly data collection schedule used at the westbound approach so that the effects of darkness versus daylight could be examined for each signing condition.

Control site data were limited to the bridge entry location because it was the most critical point at which to examine motorists' sign responses. Direct speed comparisons were made between bridge entry points of the 2 sites for sign evaluation purposes because no-sign speed data at both sites indicated compatibility between the locations. It follows that the most illustrative indication of relative sign impact is the bridge entry speed difference between the sites. Observed values show that the use of 2 activated signs results in maximum speed differential. This signing scheme performed better than that observed for the long sight-distance approach because it was used during hours of darkness. Signing offering the next best effect was the at-the-bridge, activated sign. This confirms its observed result on the long sight-distance approach.

After comparing bridge approach speed data on an hour-for-hour basis, I noted 3 expected findings: (a) activated, before-the-bridge signing produced significantly greater speed reduction than did nonactivated, before-the-bridge signing; (b) 2 activated signs produced better results than did nonactivated signs; and (c) for activated and nonactivated signs used together, activation of the at-the-bridge sign produced better results than did activation of the before-the-bridge sign. Findings a and b were observed during conditions of darkness, and finding c was observed during daylight. These observations were based on mean speeds for the total vehicle sample. Keep in mind that the fastest motorists are most suitably designated as the target sample. Figures 8 and 9 provide mean speed plots for both the total and highest quartile samples. Data are somewhat incomplete because the 1,200-ft (365.8-m) before-the-bridge tape switch failed to adhere to the pavement on 1 morning. However, the lost data did not prove to be critical. Most notable in Figure 8 is the extreme speed reduction resulting from use of activated signs at both before-the-bridge and at-the-bridge positions. Ambient conditions were highly conducive to such a response to the

Figure 9. Mean speeds during daylight hours on the short sight-distance approach.

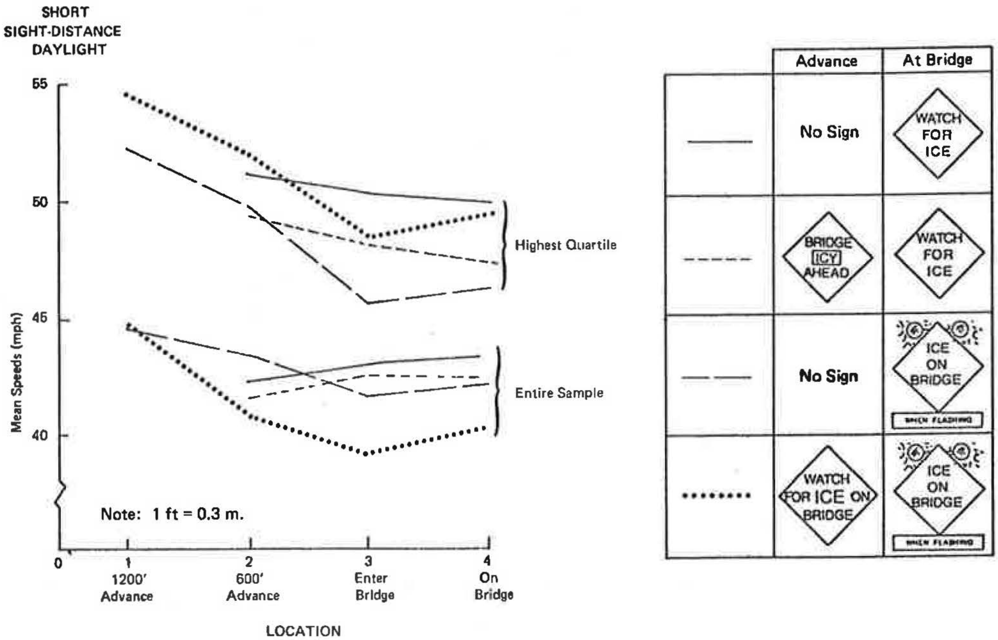
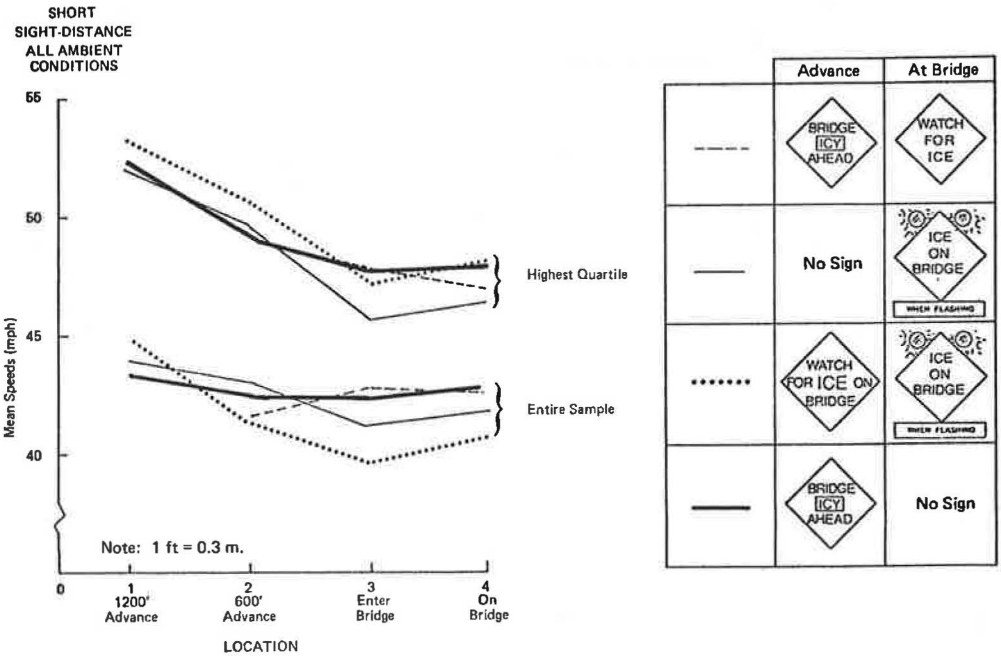


Figure 10. Adjusted mean speeds for all signing conditions with 1 activated sign.



signing. Their conspicousness was increased during darkness, and their perceived credibility was enhanced by the occurrence of moderate frost. Further note must be made of concurrent speeds at the control site, which had no signing. Observed higher-than-average speeds, representing a slight increase from the preceding hour, indicated that motorists were not concerned about the potential hazard in the absence of signing.

Relatively closer grouping of mean speeds was noted for the remainder of the sign conditions tested during hours of darkness; the nonactivated before-the-bridge sign afforded the lowest speed reduction. In fact, average speeds for the total sample increased at the bridge approach during use of both nonactivated signing schemes. The highest quartile motorists slowed slightly for the nonactivated advance sign, but then they increased speed. The combination of 2 nonactivated signs caused continual slowing on the part of the highest quartile motorists, yet speeds remained high. The activated before-the-bridge sign did cause motorists to slow down but to an insignificant degree compared to both nonactivated signs.

Response to signing under daylight conditions (Fig. 9) shows the poor results obtained with the nonactivated sign. Better results generally were obtained with the at-the-bridge, activated sign rather than with the before-the-bridge, activated sign. The single exception is the highest quartile response to the at-the-bridge, activated sign used together with the nonactivated warning.

In view of the criticality of distinguishing between the relative merits of activated signs located at the bridge and before the bridge (because of the cost of providing both), we took a further analytic step. Figure 10 shows plots of adjusted mean speeds for both the total and the highest quartile samples for all signing conditions containing a single activated sign. Adjustments were based on speed differences at the control site in an attempt to correct for any spurious speed effects. As seen from the figure, closer groupings were obtained for both average and highest quartile speeds. Lower bridge entry speeds were observed for both the average and highest quartile speed samples with the use of at-the-bridge, activated signing rather than with before-the-bridge, activated signing.

Speed Data Results

In all compatible instances, activated signing elicited greater speed reductions than did nonactivated signing. Of the nonactivated signing observed, the WATCH FOR ICE ON BRIDGE sign before the bridge provided better results than did the WATCH FOR ICE sign at the bridge. Undoubtedly, the bridge competed for driver attention and negated any effect of the latter sign. Improved responses to signing were obtained at the short sight-distance bridge approach. Better overall responses were obtained during the hours of darkness.

The sign condition eliciting the maximum speed-reducing effect consisted of activated signing at both before-the-bridge and at-the-bridge locations during hours of darkness. At-the-bridge, activated signs elicited larger speed reductions than did before-the-bridge, activated signs during both daylight and darkness.

QUESTIONNAIRE RESPONSES

Motorist reaction to both the hazard and experimental signing was examined through a regression analysis of data obtained from 168 questionnaires.

Signing Type

Activated and nonactivated signs were tested at each site for 2 approach locations: at the bridge and 1,000 ft (305 m) before the bridge. The effects of each type, location, and combination were studied. Because activated signing was found to have a greater effect on motorists, specific attention was given to the effect of its location. The advantages of activated over nonactivated signing were seen through correlations obtained among numerous variable-pair comparisons. At both sites, the use of activated signing increased the tendency for motorists to (a) see the signing, (b) notice both signs when 2 were displayed, and (c) properly identify both the sign's appearance and wording. Motorists were more prone to acknowledge the possibility of bridge icing when at least 1 activated sign was displayed. Because the data were collected during periods of virtually dry pavements, an inference from this last finding is that motorists would be more aware of icing possibilities when activated signing serves as a reminder. A comparison of those signing schemes incorporating before-the-bridge signs and those incorporating at-the-bridge signs showed no significant differences with respect to any questionnaire variables.

Ambient Condition

Two ambient condition comparisons revealed an effect on driver responses of icy bridge warning signs. Effects of daylight versus darkness and dry pavements versus light frost were notably different.

Darkness

During hours of darkness, a significantly higher proportion of interviewed motorists reported seeing the signs and properly identified their appearance and wording. The increased conspicuousness of activated signing because of darkness was undoubtedly responsible for the difference because no significant change in the nonactivated sign observation rate was noted.

Frost

During periods of light frost more motorists acknowledged the possibility of ice formation on the bridge. The motorists' cue of frost was predominantly its accumulation on their windshields.

Sign Observation

Significant increases in the proportions of motorists observing signs were noted with the use of activated signs. Improved responses were obtained when the activated sign was located at the bridge rather than before the bridge. A higher proportion of motorists noticed the signing during hours of darkness. Motorists who were more familiar with the sites were more prone to notice signs. Those motorists who noticed the signs were more likely to acknowledge the possibility that ice might be on the bridge. Greater speed reductions and lower overall speeds were observed for drivers at both sites who had observed the signs. Highest observation rates were obtained for the ICE ON BRIDGE—WHEN FLASHING sign at the bridge.

Observation of Both Signs

When there were 2 signs, data were maintained on which of the signs was observed by motorists. Motorists were more likely to see both signs for conditions when at least 1 activated sign was in use. Both signs were more often seen during hours of darkness and periods of frost. Drivers who thought that the bridge was not regularly sanded were more prone to see both signs. Motorists seeing both signs were more likely to exhibit greater speed reductions than those seeing 1 sign. Speed reductions throughout the entire approach for those motorists seeing both signs were greater on the long sight-distance approach. Both signs were observed more frequently on the short sight-distance approach.

Proper Identification of Sign Appearance

A significantly higher proportion of motorists properly identified activated over non-activated signs, and their performance improved during hours of darkness. Drivers who properly identified a sign's appearance were more likely to identify its wording and to rate the sign as being helpful. Lower approach and bridge speeds were observed for those motorists.

Proper Identification of Sign Wording

Motorists were more likely to properly identify the wording of activated signs, and a higher proportion of correct responses was obtained during hours of darkness. It stands to reason that drivers who properly identified sign appearance were more prone to correctly identify the wording and to rate the sign as being helpful. At 1 site, wording was more often correctly identified by older drivers. The sign correctly identified most often was the WATCH FOR ICE—WHEN FLASHING activated sign located at the bridge.

Driver Characteristics

Relationships among selected driver characteristics and signing responses were examined to provide a better understanding of icy bridge warning sign requirements.

Familiarity With Site

A greater proportion of familiar motorists was observed at both sites during hours of darkness because of commuter traffic. Familiar motorists were more likely to observe the experimental signing; however, their recognition of specific sign characteristics did not differ from those of unfamiliar drivers. As expected, familiar motorists were more prone to report prior skidding experience on the bridge. Familiar motorists drove more slowly as they reached the bridge than did unfamiliar motorists, and they maintained lower speeds as they continued on the bridge.

Prior Skidding Experience on Bridge

Motorists who reported prior skidding experience on the bridge exhibited greater speed reductions as they approached the bridge. A speed reduction is defined here as the difference between the greater speed recorded on either of the advance traps and the lesser of the speeds recorded at the bridge entry or on the bridge.

Knowledge of Bridge Maintenance

Motorists were asked if they knew whether the bridge was salted or sanded when it was icy. The intent was to ascertain the effect on the speeds of those drivers who were confident of maintenance activity, but no speed differences were observed. Those motorists with more driving practice felt that the maintenance was not regularly performed. Drivers who felt that the bridge would probably not be sanded were more prone to observe both signs when 2 were displayed.

Driving Practice

The most significant finding based on driving practice, measured by miles (kilometers) per year currently driven, was that higher speeds were observed for those with more driving practice. As mentioned, motorists with more practice were less likely to feel that the bridge was regularly salted or sanded when it was icy. Interviewed motorists who drove more miles (kilometers) per year were the younger and male drivers.

Assessment of Possible Icing

Because interviewing was conducted during marginal occurrences of bridge icing, motorists were asked if they thought the bridge might be icy. Responses correlated with a number of variables. More motorists acknowledged the possibility of icing during periods when activated signing was being used. Increased responses were noted during periods when 2 activated signs were displayed. The inference from this finding is that motorists were made more aware of the icing probability as a result of the cue afforded by the signing. However, it should be noted that motorists also responded to actual ambient conditions because more acknowledgments were noted during predawn hours and during the presence of frost.

Those drivers who acknowledged the possibility of bridge icing exhibited lower speeds throughout the array of speed data collection points. The most notable speed reductions at both sites occurred at the bridge entry location—the critical slowing point for motorists concerned about bridge icing. The second highest speed reduction occurred on the bridge, which confirmed the motorists' concern about bridge icing. That signing was largely responsible for the speed reductions of those motorists who suspected bridge icing is evident from the locations of the speed decreases as well as from the sign observation responses. Speed reductions were not significant at the most advanced tape switch pair on the short sight-distance approach where the at-the-bridge warning sign was not visible.

Age

The mean age for motorists at both sites was 41. Younger drivers at both sites were observed to drive faster and to have less driving practice. The only location at which no age-related speed difference was noted was the 1,200-ft (365.8-m) before-the-bridge location on the short sight-distance approach. Another finding that confirms that younger drivers have less regard for the icy bridge hazard is that they were significantly less likely to recognize sign wording at 1 site.

Sex

Two observations were made regarding differences according to sex: (a) at both sites, interviewed females drove significantly fewer miles (kilometers) per year; and (b) at 1 site, females were more likely to acknowledge the possibility that the bridge was icy.

SUMMARY AND FINDINGS

An examination of 8 experimental signing combinations made up of activated, nonactivated, before-the-bridge, and at-the-bridge signs was conducted at 2 bridge approaches. Activated signing elicited greater speed reductions than did nonactivated signing. The sign condition eliciting the maximum speed-reducing effect consisted of activated signing at both before-the-bridge and at-the-bridge locations during hours of darkness. At-the-bridge, activated signs elicited larger speed reductions than did before-the-bridge, activated signs during both daylight and darkness.

Motorist interviews were used to expand and clarify reactions to icy bridge warning signs. It was found that activated signing elicited significantly higher responses than did nonactivated signing in terms of drivers' observing, recognizing, and reading test signs. Interviewed motorists who had observed the signing exhibited lower speeds on the bridge and its approach. Better overall responses were elicited by activated signing located at the bridge rather than 1,000 ft (305 m) before the bridge. Activated warning signs were effective as a hazard cue because more drivers acknowledged the possibility of bridge icing when activated signs were displayed whether frost was present or not.

Two sign conditions employing activated signing produced promising results and are recommended for further study based on field observation at other locations to establish the general nature of these results. The sign scheme eliciting the best response was an activated, BRIDGE ICY AHEAD sign 1,000 ft (305 m) before the bridge together with an ICE ON BRIDGE WHEN FLASHING sign incorporating activated hazard identification beacons. An effective, less costly alternative was observed when the nonactivated WATCH FOR ICE ON BRIDGE was substituted at the 1,000-ft (305-m) before-the-bridge location. The effectiveness of the signing would be dependent on a reliable ice detection system for its activation.

ACKNOWLEDGMENTS

The research presented in this paper was performed under a Federal Highway Administration contract. Contract managers were Alfred Klipple and Michael Freitas. Their permission to publish this work is most appreciated.

REFERENCES

1. C. A. Ballinger. Automatic Icy Road Sign Study. Planning and Research Div., Colorado Dept. of Highways, Final Rept, Aug. 1966.
2. C. Birnie, Jr., and W. E. Meyer. Prediction of Preferential Icing Conditions on Highway Bridges. HRB Spec. Rept., 115, 1970, pp. 27-35.
3. L. Bourget. Pavement Icing Warning Systems, Laboratory Evaluation of Econolite and Nelson Prototype Devices. California Dept. of Public Works, Sacramento, Jan. 1967.
4. M. F. Ciemochoswki. Highway-Bridge and Airport Runway Detection Systems—Predicts Frost and Ice Before They Occur. Society of Automotive Engineers, Paper 680117, 1968.
5. M. F. Ciemochoswki. A Detection System for Frost, Snow, and Ice on Bridges and Highways. Highway Research Record 298, 1969, pp. 17-24.
6. T. B. Culp, and R. L. Dillhoff. Watching for Ice on Bridge—Does It Reduce Accidents? Ohio Dept. of Highways, Oct. 1970.
7. C. E. Forbes, C. F. Stewart, and D. L. Spellman. Snow and Ice Control in California. HRB Spec. Rept. 115, 1970, pp. 177-195.
8. W. D. Glauz et al. An Ice and Snow Detection and Warning System Feasibility Study. Midwest Research Institute, Final Rept., MRI Project 3394-E.
9. A. Howard. Traffic Sign Recognition. Proc., Canadian Good Roads Assn., Oct. 1964.
10. M. Inoue, K. Baba, and Y. Takada. Ice Detection, Prediction, and Warning System on Highways. HRB Spec. Rept. 115, 1970, pp. 17-26.
11. W. T. Jackman. Driver Obedience to Stop and Slow Signs. HRB Bull. 161, 1957, pp. 9-17.
12. Reduce Speed—Ice on Bridge Sign. Kentucky Dept. of Highways.
13. M. E. Roussel. Ice Detection and Warning Systems for Roadways. Oklahoma Div., FHWA, Working Paper, May 1973.

14. C. F. Stewart and A. Segueira. Bridge Deck Frosting. California Div. of Highways, Nov. 1971.
15. D. I. Williams and M. D. Van Der Nest. The Human Factor in Road Traffic Signs: The View of the Road User. Council for Scientific and Industrial Research, Johannesburg, South Africa, Rept. CSIR PER 113, May 1969.

FREEWAY-STYLE DIAGRAMMATIC SIGNS IN NEW JERSEY

A. W. Roberts, E. F. Reilly, and M. V. Jagannath,
New Jersey Department of Transportation

Ninety-four freeway-style diagrammatic signs were installed on I-287 in New Jersey to replace 120 conventional signs at 22 interchanges. Traffic volumes and specific unusual maneuvers were counted and categorized for a 7-day period at each of 10 exit sites before and after guide-sign changes. Evaluation of 5 of the sites included statistical comparisons of correlated, hourly, unusual-maneuver rates. The results of comparisons between conventional and diagrammatic signs showed that (a) conventional signs produced a lower rate of critical maneuvers than did diagrammatic signs at a split for parallel roadways; (b) diagrammatic signs at cloverleaf interchanges effected a lower rate of stopping and backing maneuvers but a higher rate of exit gore weaving maneuvers; and (c) diagrammatic signs at right-hand, T-ramp exits approached by either an auxiliary lane or a regular deceleration lane resulted in a lower rate of stopping and backing maneuvers as well as fewer or the same number of exit gore weaving maneuvers. Seasonal and yearly variations in unusual maneuvers were derived from the results of seven 7-day studies performed at 1 site from 1969 to 1974. The adequacy of the diagram standards was reviewed, and recommendations for further research were made.

•SINCE the trial of diagrammatic guide signs at freeway interchanges has been advocated (1), there have been several major developments in the evaluation of these signs. The Federal Highway Administration sponsored a national program to design and evaluate U.S.-style diagrammatic signs to reduce unusual maneuvers at high-incidence exits. Several research studies worthy of mention came out of this effort. The effects of diagrammatic signs were shown in a field study at a left-hand exit gore by means of a narrowly defined unusual maneuver measure and a matched pairs statistical analysis of unusual maneuver rates (2). Diagrammatic sign concepts were evaluated in a laboratory setting, and recommendations were made for specific applications and graphic design (3). Inclusion of graphic lane lines in signed diagrams was evaluated in another field study (4). Relevance of erratic maneuvers at several zones upstream and downstream of the exit gore and results of driver interviews were shown in a diagrammatic sign evaluation (5). The utility of unusual maneuvers, lane changes, speed changes, and headways in evaluating diagrammatic signs was demonstrated by information gathered by automated tape switches in several zones approaching exits as well as from driver interviews and learning-curve detection (6).

In pilot studies conducted at I-287 and US-22 in New Jersey (4), signs with both diagrams and lane lines were considered to be more effective than conventional signs. However, some doubt remained about whether the improved performance found in unusual maneuvers was due to the effects of novelty or sudden importance given to the signs. Thus a study of diagrammatic signs over a continuous, 22-mile (35.4-km) section of I-287 was conducted that included a variety of geometric, exit-ramp configurations.

Publication of this paper sponsored by Committee on Motorist Information Systems.

AREA CHARACTERISTICS

I-287 in Somerset and Middlesex Counties was chosen for study because of its varied geometry and its importance as part of a road system that encircles New York City.

The route in the study area is not yet connected to the rest of the encircling system although extending sections of the route in the study area were constructed between the 1971 and 1973 studies. At the southern end, entrances from the Garden State Parkway and US-9 were opened during the 2 main study years, which extended the roadway approaching US-1 from 2 to $3\frac{1}{2}$ miles (3.2 to 5.6 km). Other existing entrances included those from US-1, the New Jersey Turnpike, and county Route 514. At the northern end, 6 miles (9.7 km) of roadway were added, which gave access to NJ-24 and county Route 510. Seven miles (11.3 km) of a connecting Interstate route, I-78, were constructed to the east of I-287 during the 1971 studies. A total of 12 miles (19.3 km) were constructed before the 1973 studies. The study site locations are shown in Figure 1.

A noticeable amount of growth in shopping and industrial areas took place during the study years in the study area, and the section between US-22 and county Route 527 was resurfaced and concrete center barriers were reconstructed immediately before and during the study in 1973. The repaving contract was coordinated with the diagrammatic sign study in such a manner as to keep studies and construction as far apart as possible. Although cooperation was good, studies and repaving work were not a great distance away from each other at a few sites and were bound to affect motorist behavior.

STUDY AND SIGN DESIGN PROCEDURES

Scope

This study includes an evaluation of 30 signs at 10 exit sites within a 22-mile (35.4-km) section of I-287 and I-95 in Somerset and Middlesex Counties. The exit sites had both simple and complex geometric situations among which are left- and right-hand exits; semidirect, indirect, and direct ramps; left- and right-turn connections; and interchanges that have both 1 and 2 ramps exiting from a single direction. Exit sites were selected originally for their potential for unusual maneuvers. All the exit sites in the study route were studied to determine this potential.

A limitation that no new structures for mounting signs could be built was placed on the study. Within that limitation, several signs were modified to upgrade existing sign messages before formal data collection began. Unusual-maneuver and volume data were collected at each site during July and August 1971 and at site 5 during August 1972 after an additional sign had been erected at that location. Figure 2 shows a project activity bar chart.

All structures were analyzed for design wind load to determine maximum panel size for the increased area that is often required for diagrams on ground-mounted signs. Less than standard layouts had to be resorted to on some ground-mounted signs. Less overall area usually was needed when 1 diagrammatic sign replaced 2 or more conventional signs.

Diagrammatic signs with breakaway posts on ground mounts replaced conventional signs in the spring of 1973 and studies at all 10 sites were again made in July and August 1973. One hundred twenty conventional sign panels were replaced on the entire study section of I-287 with 94 diagrammatic sign panels. Twenty-two interchanges were involved in the change.

Letters, numbers, diagrams, and lane lines were reflectorized by using cube-corner reflex buttons; high-intensity, beaded reflex sheeting was used for shield backgrounds. Both cube-corner buttons and beaded sheeting had been tested for diagram visibility and recognizability at night from up to 1,000 ft (305 m). Panel backgrounds were made from extruded aluminum coated with polyvinylidene fluoride paint.

Figure 1. Study site locations.

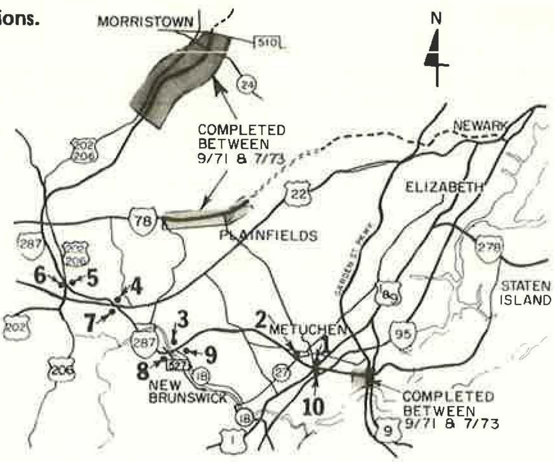


Figure 2. Project activity bar chart.

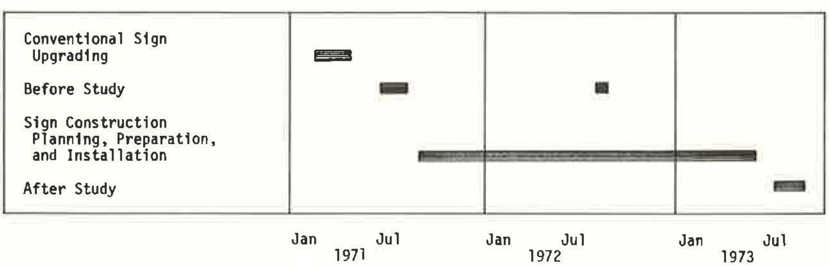
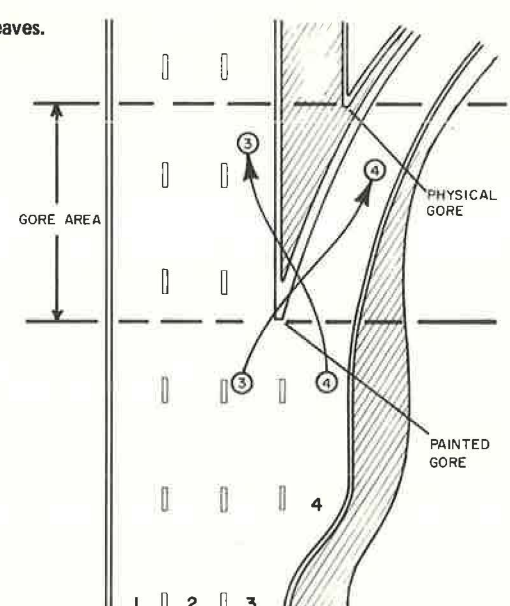


Figure 3. Typical gore weaves.



Sign Design Procedure

The diagram design methodology was aimed at satisfying the following requirements:

1. Greater visibility to compensate for increased information content,
2. Simplicity for ease of interpretation,
3. Road-diagram and diagram-message relatability,
4. Diagram continuity within each interchange,
5. Message redundancy for continued confirmation within exit approaches, and
6. Uniform application among all interchanges for reliable driver recognition.

Diagram designs and sign layouts were developed with the use of a photographic inventory film. Frames of signs on location were projected against a large screen. Alternate diagram and message layouts were superimposed on the scenes to judge their adequacy from a driver's visual point of view. Opinions from engineering and nonengineering persons were solicited on an informal basis.

Both plan view (3) and a more symbolic style of diagram were simulated and reviewed. A practical set of standards that offered a consistent and reliable basis for designing a diagram for any geometric condition was arrived at by using the more symbolic style of plan view. Overhead signs were often easier to design than were ground-mounted signs because less space was needed after messages from several signs were organized onto that with a diagram.

Data Collection

During each site study unusual maneuvers and volumes were counted by hidden observers. Three observers counted all unusual maneuvers while 2 observers counted all volumes. For this study, an unusual maneuver was any stopping or backing up in the exit gore section or any driving on the gore line between the physical gore and a predetermined point upstream of it (Fig. 3). At each site, unusual maneuvers and volumes were collected for both through and exit movements as well as for 2-axle and 3- or more axle vehicles in each hour from 2:00 to 5:00 p.m. on 7 consecutive days. On days when data were collected, before and after conditions were matched to the closest date from year to year.

Analysis

The rates of unusual maneuvers (in maneuvers per thousand vehicles) for conventional and diagrammatic sign conditions were compared at each site. Only the rates for 2-axle vehicles were compared because vehicles with extra axles were considered to be in a different group. Differences between studies were tested statistically by using the conservative, nonparametric, Wilcoxon, matched pairs, signed ranks test (8). The following formula was used:

$$Z = \{T - [N(N + 1)]/4\}/\sqrt{[N(N + 1)(2N + 1)]/24}$$

where

T = sum of positive ranks,
N = number of qualified ranks, and
Z = normal standard deviation.

Rates were paired by hour, day, and movement. Rates paired by movement are referred to in the following abbreviated form:

1. EUR—exiting unusual maneuver rate: a proportion of exiting volume such as a 3 to 4 maneuver (Fig. 3),
2. TUR—through unusual maneuver rates: a proportion of through volume such as a 4 to 3 or a 3 to 3 maneuver,
3. SBR—stopping and backing unusual maneuver rates for both exit and through movements as a proportion of combined volumes, and
4. UR—exiting and through maneuvers for a combined volume.

A change was considered significant when it could be accepted at greater than the 95 percent level of confidence when a 1-tailed test was used. The analysis of data for each site included:

1. Statistical comparisons pairing unusual maneuver rates for each sign condition,
2. Sufficiency or validity of the data, and
3. Possible or probable effects produced by extraneous factors such as construction activity, new access, changed volume, accidents, and rain.

A stopping or backing maneuver was considered to be more critical than a gore weaving maneuver, but SBRs were included with EUR and UR totals as well as separated for before and after comparisons. EUR or UR may be regarded as gore weaving rates because SBRs were quite infrequent.

Nonexperimental Factors Control

Repaving and restriping, independent signing programs, new access, changing traffic volumes, changing travel patterns, and new industrial development can bias comparisons of before and after data.

The method of analysis used minimized effects from factors that were dependent on hour, day, and season. Loss of visibility during heavy rain is rare, but 1 hour of affected data was eliminated by the analysis.

Although control groups could not be used effectively in this kind of study, a great effort was made to reduce bias from outside influences. For a repaving project that was conducted simultaneously with the after study, plans were prepared for repaving and restriping work so that the result on exit striping would be the same in both studies. This was accomplished by documenting the dimensions of original gore stripes and entering them in restriping plans. The contractor also was scheduled to conduct work away from planned study sites. Requests were made to other agencies to refrain from making their own sign changes at locations relating to planned studies. Needed changes in destination names and exit numbering generally were postponed. The only exception was at site 6 where gore-mounted supplemental signs were removed and a standard exit gore sign with a number was installed.

Despite all efforts to minimize bias, the accuracy of the data at sites 2, 4, 6, 7, and 8 was in doubt, and we were left without a reasonable method of adjustment.

Sign changes made in relation to exit geometry at sites 1, 3, 5, 9, and 10 are shown in Figures 4, 5, 6, 7, and 8.

RESULTS

Table 1 summarizes the data collected and the results of statistical analysis. Vehicles with 3 or more axles are not included. The significance of the rate change is shown in a Wilcoxon, nonparametric, matched pairs test. The change in rate from before to after conditions is shown to be significant or not significant at the 95 percent level of confidence in the direction of the change.

Study sites 3 and 5 were the only sites at which both the direction of SBR and UR change agreed and there were no observable sources of bias. Conclusions without qualification are based on the results from these sites.

Figure 4. Study site 1.

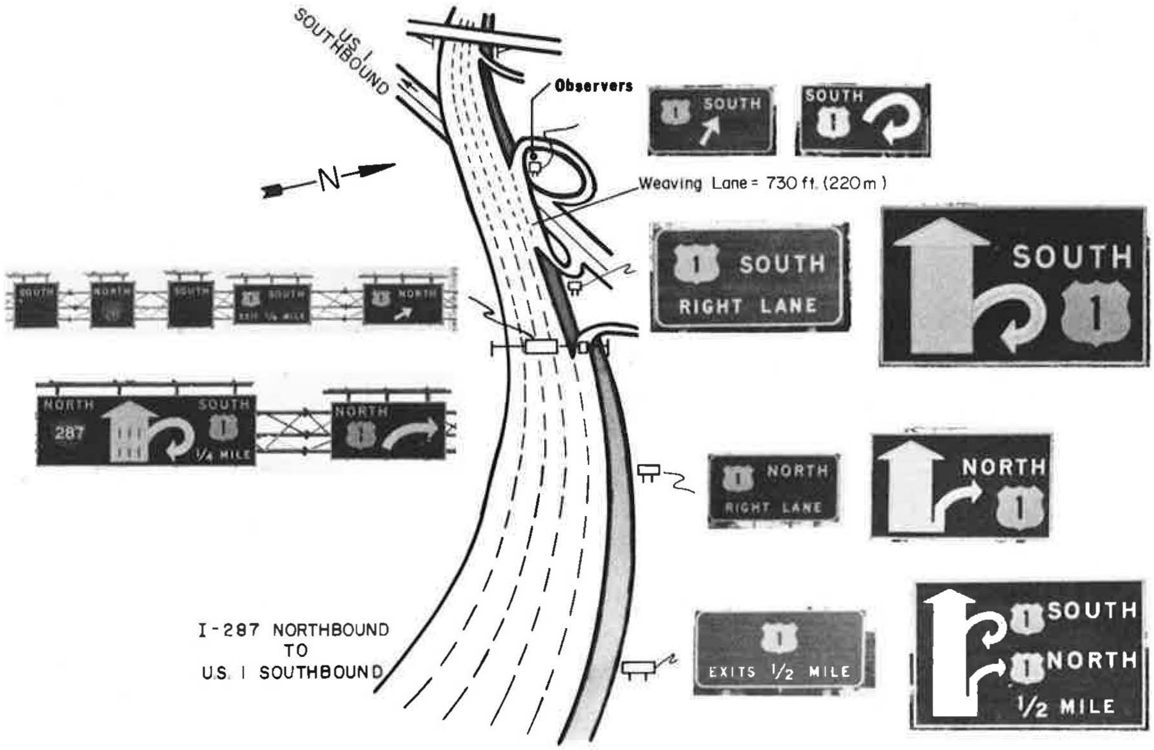


Figure 5. Study site 3.

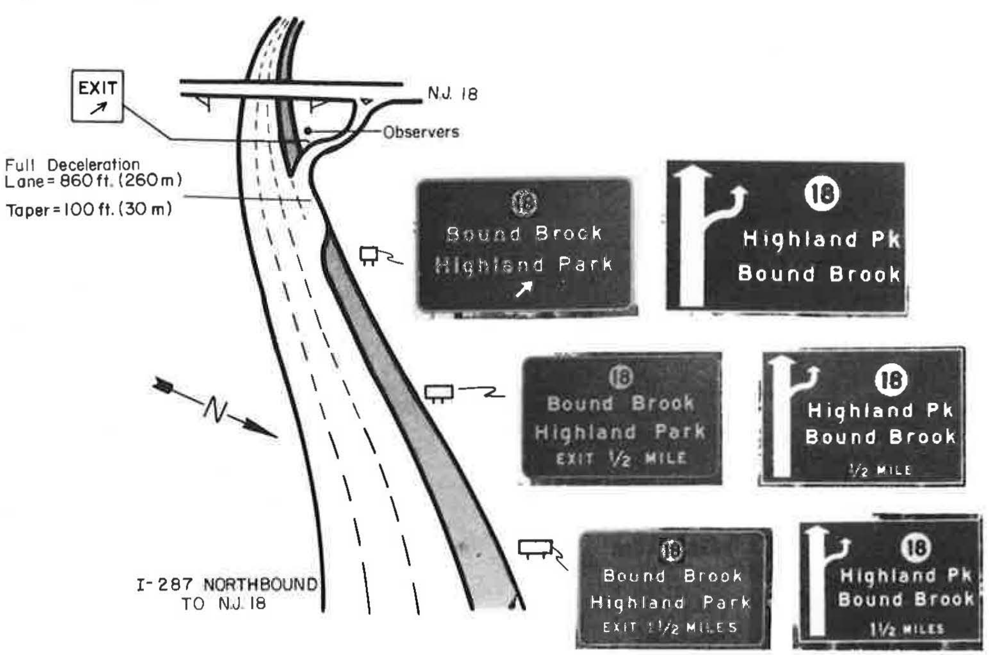


Figure 6. Study site 5.

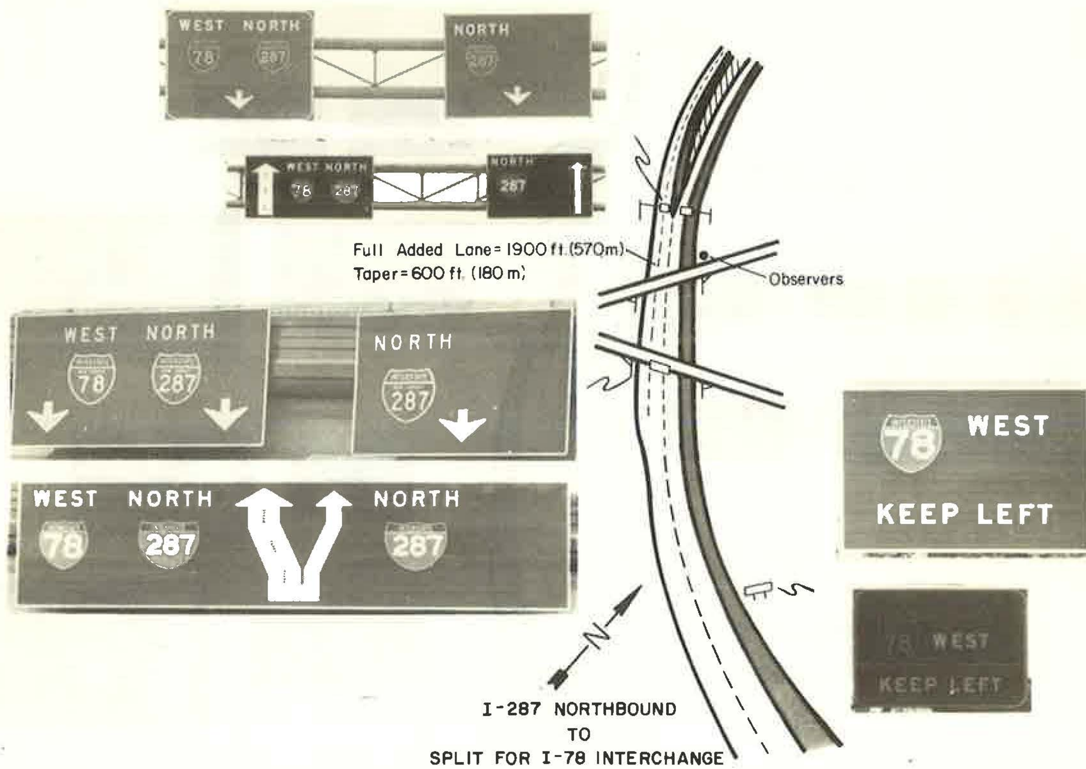


Figure 7. Study site 9.

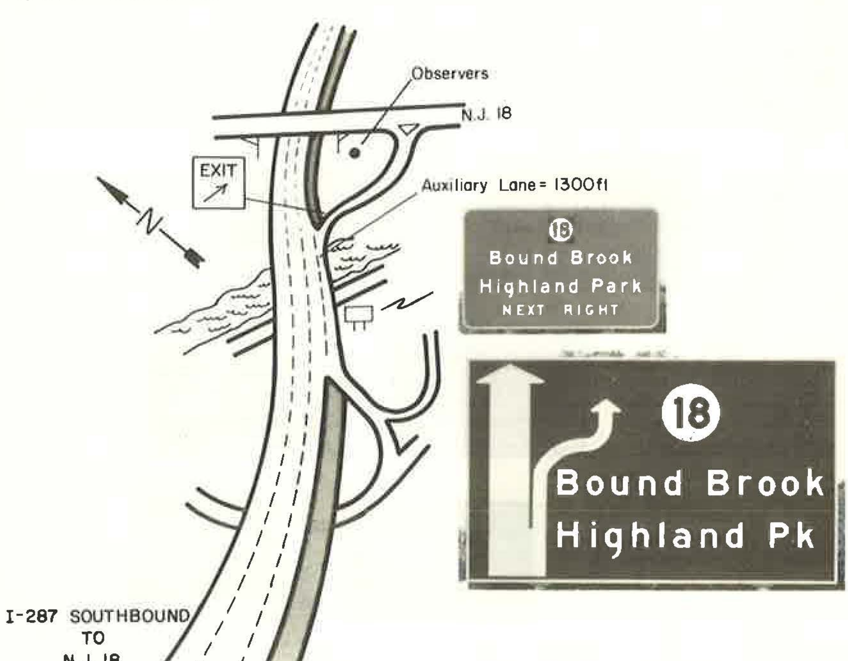


Figure 8. Study site 10.

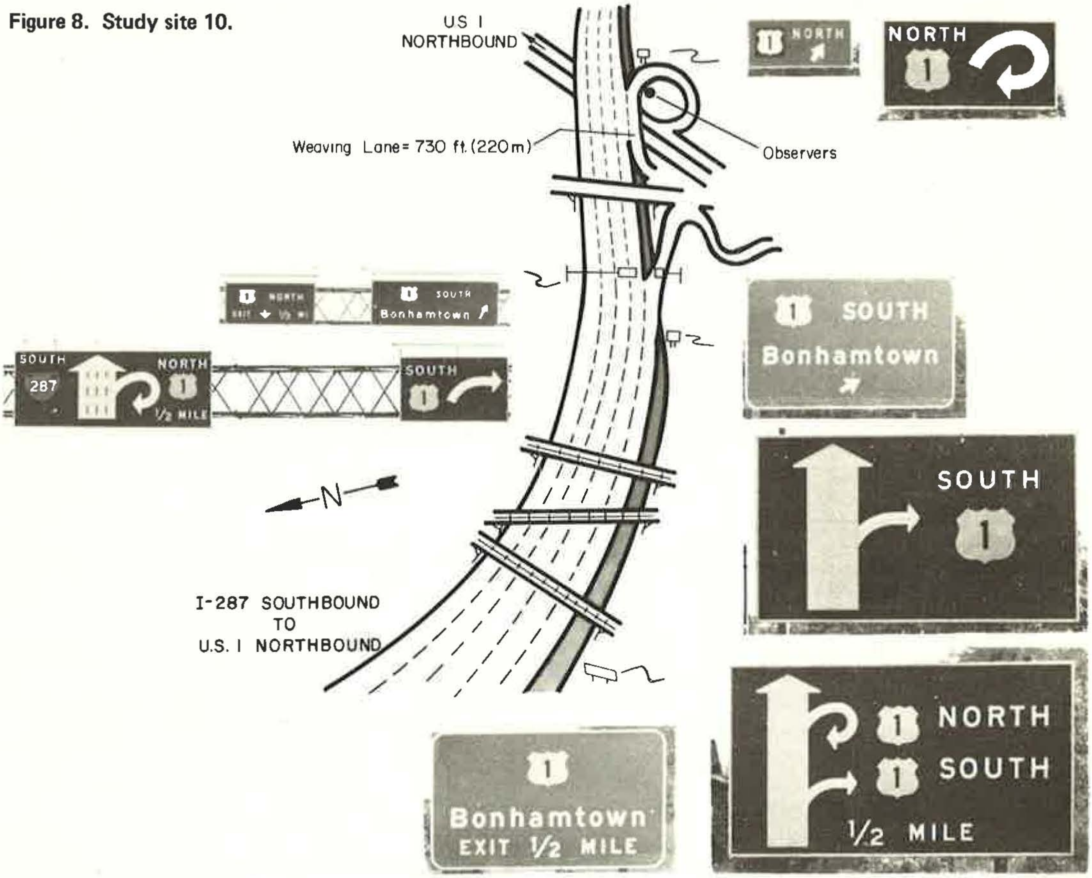
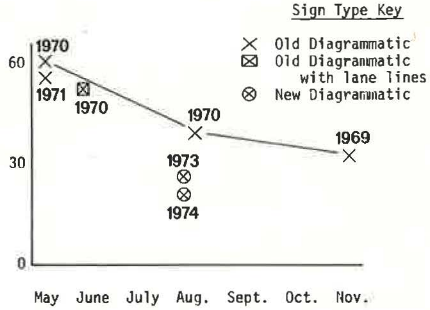


Table 1. Unusual maneuvers and statistical summary by study site.

	Maneuvers									
	SBR (per 100,000 vehicles)		UR (per 1,000 vehicles)		Avg Hourly Volume		Exit Volume (percent)		Z-Value[a] (hourly rates)	
Site	Before	After	Before	After	Before	After	Before	After	SBR	UR
1	19.0	6.4	4.3[b]	30.8[b]	1,002	1,483	7	10	1.21[c]	3.75[d]
3	16.2	3.3	1.0	0.8	1,232	1,528	10	8	1.75[d]	1.04[c]
5	23.0	60.5	38.7	46.1	826	865	50	51	1.78[d]	2.85[d]
9	23.6	3.1	1.1	0.6	1,209	1,550	19	17	2.02[d]	2.50[d]
10	134.6	39.0	23.2[b]	22.7[b]	1,309	1,833	38	24	3.33[d]	0.14[c]

[a]Wilcoxon test. [b]Includes unusual exit gore maneuver rates for exiting vehicles only. [c]Not significant. [d]Significant.

Figure 9. Unusual-maneuver rates for site 4 studies by month of year.



Agreement between the SBR and UR direction of change in rates is shown for study sites 9 and 10, but a bias factor was known to have been operating during the after studies at each of these sites. At site 10, the effect of bias from a lower percentage of approach volume that exited would make the reduction in rates appear to be greater than it would have been without bias because bias would make the after study rates superficially low. The opposite was true at site 1. At site 9, bias from the effects of upstream construction activity would tend to have made the reduction in rates given in Table 1 less than it would have been without bias because the after study rates tend to be superficially high. For this reason the reduction of rates should be regarded as conservative at site 9.

Estimated results of studies at sites 1 and 10 are made on the basis that the data may be combined for analysis because the sites have geometrically similar loop exits and the same type of interchange. The bias effects from a change in percentage exiting are opposite and should cancel when combined. Applying the Wilcoxon test for 1 direction at the 95 percent level of confidence, we found a significant decrease in SBR and a significant increase in EUR after diagrammatic signs were erected at sites 1 and 10.

CONCLUSIONS

Some conclusions can be drawn from the results of comparing the performance between conventional and diagrammatic signs.

1. At a split for parallel roadways, conventional signs were shown to be more effective in reducing critical maneuvers than were diagrammatic signs.
2. At 2 loop exits within a cloverleaf interchange, a conversion to diagrammatic signs effectively reduced stopping and backing maneuver rates but resulted in either no change or an increase in unusual exit gore maneuver rates. Grouping these data for a combined before-and-after analysis resulted in a decrease of stopping and backing maneuver rates and an increase in unusual exit gore maneuver rates.
3. At a right-hand T-ramp, a conversion to diagrammatic signs reduced stopping and backing maneuver rates and resulted in no change in unusual exit gore maneuver rates.
4. At a right-hand T-ramp terminating an auxiliary lane, a diagrammatic sign was more effective than a conventional sign in reducing critical unusual maneuvers.

FINDINGS FROM STUDIES AT SITE 4

In Figure 9, 7 studies performed for 6 years at site 4 are plotted by month according to average unusual maneuvers per 1,000 vehicles for each study. A base curve is drawn among 3 studies performed with the same signs in different seasons within a 12-month period to illustrate seasonal variation. Variation between years is also shown among studies made with the same signs on the closest dates of the same month. Several conclusions can be made from these comparisons.

1. Seasonal variation can be greater than annual variation.
2. Declining gore weaving rates may be found for at least a year after installation of a diagrammatic sign.

Seasonal variation of unusual-maneuver rates may have a marked effect on before and after comparisons with a month between studies. The actual rate reduction found when lane lines were added on the old diagrams in June 1970 may be seen in Figure 9 by directly comparing the point for old diagrammatic signs with lane lines with the point directly above it for old diagrammatic signs without lane lines. A lower rate than the previous year when there was no change in signing was found on 2 occasions as can be seen in Figure 9 by comparing 1970 with 1971 in May and 1973 with 1974 in August. The reduced rates have not been found to be related to volume but may be related to driver familiarization. Comparisons of before and after conditions in future studies should take these possible variations into account.

DISCUSSION AND RECOMMENDATIONS FOR FUTURE RESEARCH

The method developed to design diagrams for I-287 could be improved to benefit the motorist's perception. We recommend the following modifications to the diagrams:

1. Complete separation of graphic movements for exit-lane drops and splits (Fig. 10),
2. Extension of leading ends on gore-located, independent arrows that curve (Fig. 11),
3. Establishment of standard stem lengths for ground-mounted advance-guide exit-direction sign diagrams (Fig. 12),
4. Widening of lane lines to 3 in. (7.6 cm) instead of 2 in. (5 cm) (Fig. 13), and
5. Heavier overhead arrowheads (Fig. 14).

In general, more panel space should be considered for advance-guide, diagrammatic, ground-mounted signs than that that was used on I-287 because the addition of a diagram requires more space for the given letter standards. In addition to needing more space, diagrammatic signs require a greater degree of overall graphic organization than do conventional signs. In future work, design standards should be developed to minimize incorrect message-to-diagram associations, maximize correct associations, and organize diagram-to-message and message-to-message interfaces in predictable locations. With the exception of sign panel space, these goals were accomplished on I-287 by standardization.

There is some uncertainty about the specific design of diagram symbols beyond the long, narrow plan view for I-495 (6) and the short, wide symbolic type for I-287 (4) as shown in Figure 15. The plan view type is a truer reproduction of the gore area approach because main roadway curvature is shown. Although the plan view type shows exiting sides, it does not show exit-ramp turn directions. The symbolic type reproduces the gore area approach, but all approach roadways are shown vertically regardless of approach curvature. In the more symbolic type, exiting sides as well as exit-ramp turn direction are displayed.

Although all drivers should profit from knowing the exiting sides in advance, and, although the information should not be too hard to understand, there is some doubt about the ability of all drivers to perceive the symbolic value of an exit-ramp turn direction when it is semidirect or indirect. Further research is suggested to document the values of these differences in types of signs in terms of exit performance because this area has not yet been adequately documented.

ACKNOWLEDGMENTS

The authors wish to acknowledge D. W. Gwynn for initial project impetus; J. Powers for codesigning the signs, supervising in the field; R. L. Hollinger who along with J. Powers coprogrammed the statistical analysis and cosupervised the writing of sign specifications for the sign replacement contract; and R. Strizki who determined the wind-load capability of the sign supports. Support from the Federal Highway Administration and the Bureaus of Structural Design, Contract Administration, Traffic Engineering, and Maintenance of the New Jersey Department of Transportation, and the New Jersey State Police Department of Law and Public Safety also is gratefully acknowledged.

The cooperation given to us for the timely replacement of signs by the Pennwalt Corporation, Porce-Len Company, PPG Industries, Amerace Corporation, 3M Company, Interstate Highway Sign Company, and Whitmyer Brothers Corporation also is appreciated.

Figure 10. Exit-lane drops and splits.

Figure 11. Extension of gore-located arrows.

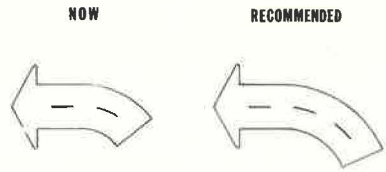


Figure 12. Stem lengths.

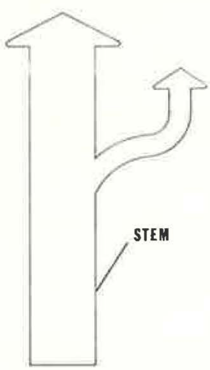


Figure 13. Lane lines.

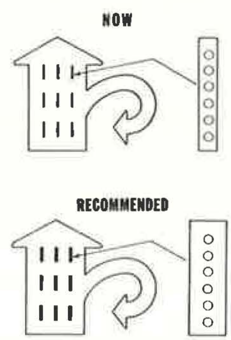


Figure 14. Overhead arrows.

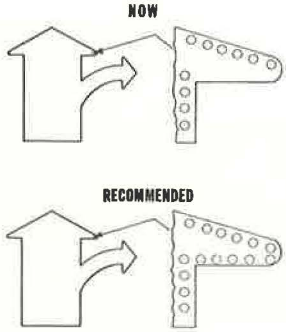
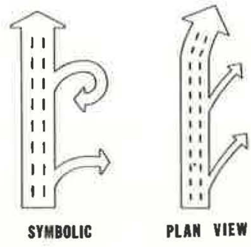


Figure 15. Symbolic and plan view arrows.



REFERENCES

1. Highway Safety, Design and Operations: Freeway Signing and Related Geometrics. Hearings before 90th Congress, 2nd Session, U.S. Govt. Printing Office, 90-39, 1968.
2. A. W. Roberts. Diagrammatic Sign Study: Phase I Report. New Jersey Department of Transportation, May 29, 1970.
3. J. W. Eberhard and W. G. Berger. Criteria for Design and Development of Advanced Graphic Guide Signs. Highway Research Record 414, 1972, pp. 24-29.
4. A. W. Roberts. Diagrammatic Sign Study. Highway Research Record 414, 1972, pp. 42-49.
5. F. R. Hanscom. Evaluation of Diagrammatic Signing at Capital Beltway Exit 1. Highway Research Record 414, 1972, pp. 50-58.
6. G. S. Kolsrud. The I-495 (Capital Beltway)/I-70S Field Study. In Diagrammatic Guide Signs for Use on Controlled Access Highways, Federal Highway Administration, Vol. 3, Part 2, Dec. 1972.
7. G. J. Alexander, G. F. King, and M. S. Warskow. Development of Information Requirements and Transmission Techniques for Highway Users. Texas Transportation Institute, College Station, Research Rept. 606-1, 1969.
8. G. A. Ferguson. Statistical Analysis in Psychology and Education. McGraw-Hill Book Co., New York, 1966.
9. Manual on Uniform Traffic Control Devices for Streets and Highways. Federal Highway Administration, 1971.
10. A. W. Roberts and E. F. Reilly. Study of the Left-Hand Exit From Southbound Express I-287 to U.S. 202-206. New Jersey Department of Transportation, 1969.

MOTORISTS' PREFERENCES IN ROUTE DIVERSION SIGNING

J. W. Hall and L. V. Dickinson, Jr., University of Maryland

A questionnaire survey was conducted in the Baltimore-Washington corridor to determine motorists' preferences in route diversion signing. The study site was located near the southern end of the corridor at the interchange of I-95 and I-495. The objectives of the questionnaire study were to determine 3 types of information concerning the motorist and the diversion-signing system. In addition to determining the characteristics of drivers that use a major intercity corridor, the study sought to identify the type of real-time information desired by the road user and determine the format for conveying this information. Two different questionnaires were developed and tested, and over 6 thousand were distributed to motorists during 3 interview periods. On the basis of the motorists' responses, the study found 3 sign messages worthy of further consideration for real-time route diversion. These signs contained the following information: (a) length of congestion, (b) cause of congestion and exit instructions, and (c) alternate route information. The first type was preferred because of its conciseness. The other 2 were preferred because they conveyed a sense of authority and presented the motorist with an alternative to travel delay.

•THE MAJORITY of highway facilities operate satisfactorily for most of the day, but during the morning and evening peak periods some facilities become extremely overloaded, and heavy congestion, bottleneck conditions, delays in travel, and increased travel times result. These conditions are further aggravated during both peak and off-peak periods when other incidents are introduced on the roadway, and even lower levels of service are then given by the highway system. If specific problem locations can be identified, means can be used to divert some of the traffic to other facilities and thus increase the quality of flow for all motorists on the surrounding highway system.

Some information exists in the literature concerning the development and implementation of procedures for the real-time diversion of traffic (2). Several studies have investigated both the hardware and software problems in determining when traffic should be rerouted and have recommended policies to be followed for such programs. Numerous parameters, including traffic speed, traffic density, occupancy, traffic volume, travel time and delay, rate of motion, acceleration noise, and conflict analysis, have been suggested and used to quantify the level of service for highway facilities. In a typical program of real-time route diversion, 1 or more of these parameters would be monitored on major routes and on suitable alternate routes to evaluate level of service. When the level of service on a route falls below or is in the range of a specified threshold value, traffic would be rerouted to other facilities to alleviate the problems caused by a particular incident or heavy traffic.

Local motorists traveling on a facility that is operating under a lower than normal level of service because of some incident may elect to divert to some other facility and avoid the congested area. In other words, drivers, as decision-makers, evaluate the situation and make individual determinations on the best alternative to select. Motorists generally will not have knowledge of or interest in the system parameters used by the engineer to evaluate an operation. From the point of view of real-time diversion, mo-

Publication of this paper sponsored by Committee on Motorist Information Systems.

torists on intercity journeys or not familiar with the area create a more complex problem. Some means must be developed to accurately convey the highway situation and explain in an understandable manner what should be done. This was one of the objectives of this study—to determine what information concerning degraded system operation and the availability of alternate routes should be conveyed to motorists on intercity travel.

Of the 4 component areas found in the literature concerning real-time route diversion (quality of operation, incident detection, diversion policy, and communication), communication—the manner in which information is conveyed to motorists—has received the least attention. It may involve messages describing the nature of the situation or advice on the need to change to another facility because of an incident. If the facility is operating at a level of service low enough that it is advisable to divert traffic to another route, it is essential to know how to convey this real-time information to the motorists. In general, a real-time route diversion program advises motorists of the existing roadway conditions so that they can make an informed choice concerning their route of travel. It is therefore necessary to provide for conveying proper information to the drivers. Several forms of communication are applicable for diversion of intercity traffic. These can be either audio or visual and can use either stationary or in-vehicle equipment.

According to the literature, the most promising method of communicating with motorists is variable message signs. The 5 prominent types of variable message signs are as follows:

1. Lamp matrix,
2. Drum sign,
3. Blank-out sign,
4. Roller screen, and
5. Flipping disk.

Each of these types of signs has been employed in system surveillance and control programs. Although the physical type of the sign is important, the sign message and arrangement are of greater importance. The real-time information from these formats is divided into 2 general categories.

1. Descriptive information, which indicates the general roadway condition, includes CONGESTION, ACCIDENT AHEAD, KEEP LEFT, and ICY CONDITION signs.
2. Quantitative information, which specifies numerical values for some traffic parameter that the driver is able to comprehend, includes variable speed limit signs and estimates of delay time such as 5 MIN TO GORMAN ROAD.

Studies have found that motorists prefer messages that convey real-time information. Heathington, Worrall, and Hoff (4) found through a questionnaire study that freeway motorists in the Chicago metropolitan area preferred signing that presented speed information or descriptive terms over signing that presented quantitative travel-time or delay information when there was heavy, moderate, and no congestion. This seems to indicate that motorists consider and choose alternate routes not only on the basis of travel time and delay during all types of driving conditions but also on the basis of information to which they can clearly relate. Other studies have found that motorists prefer a message design that distinguishes real-time visual displays from other types of freeway signs and employs unique features such as color to distinguish unusual or abnormal traffic conditions (3). It has been suggested that a uniform and recognizable color-coded message would stand out from the static signs and would help the driver to quickly perceive freeway conditions. For example, green could mean that the traffic conditions ahead are favorable; flashing amber might suggest that the driver should use extra caution (6).

In designing a signing format to be used to encourage diversion from one facility to another, the attitudes and behavior of motorists on route diversion must be considered. A recent report outlines the results of a study in Chicago with regard to voluntary diversion from a normal route to work to some other route (5). Generally, the respondents were more receptive to diversion to avoid delay or to save travel time on the trip from work to home. Furthermore, the respondents indicated that they would be more likely to divert to avoid a delay than they would be to save travel time. Even though these studies were concerned with trips from home to work and from work to home, we felt that some of the findings might be useful in this study concerning motorists' preferences for signing for intercity travel.

Although there have been notable exceptions (1, 8, 9), work reported previously on real-time route diversion has concentrated on urban corridors. And although comparatively little has been done with respect to intercity diversion, the elements of traffic-flow evaluation and incident detection are basically analogous for urban and intercity diversion. The policies associated with intercity diversion would be established initially through simulation techniques on the proposed corridor and then would be tested experimentally. However, special attention is warranted for the signing aspects of intercity diversion because of the differing needs of local and through motorists.

The problems related to the design of a message format are specific. It is generally agreed that motorists choose routes on the basis of travel time and cost, safety, comfort, and convenience. The signing for route diversion must indicate when one or more of these factors exist at a degraded level. However, it was unclear as to exactly what format and what traffic parameters should be used to convey this information.

STUDY DESIGN AND FIELD INVESTIGATION

The type of information to be determined was motorists' preferences with respect to the wording and presentation of a sign message. To accomplish this, it was decided that a questionnaire survey of the road users offered the best opportunity to collect the necessary information. The questionnaire had to satisfy 3 objectives. It had to determine

1. The characteristics of the driver, including information about the number of years driven, distance driven annually, and frequency of freeway use;
2. What type of real-time information the road user desires; and
3. In what manner the road user wanted the information presented.

Because of the amount of information required and the number of questions asked, it was necessary to use 2 questionnaires. Questionnaire A allowed for the selection of a best and a worst sign and asked the respondents to indicate their reasons for making their choices from among 3 available signs. Questionnaire B was designed so that motorists could choose pairs from a total of 5 different signs, which would allow making 10 choices. During the interview, motorists were given either questionnaire A or B, together with an explanatory letter and a postage-paid return envelope. Both questionnaires contained identical questions to determine driver characteristics and information about daily commuting to work. These questionnaires are described in detail in a separate report (7).

Because the focus of this study on route diversion was the Baltimore-Washington corridor, it was necessary to choose a field location where traffic is oriented toward corridor rather than local usage. It was also desirable to have a representative level of through traffic in the sample of drivers interviewed. A third important criterion in selecting the site was to have a location where traffic could be stopped safely to distribute the questionnaires.

In conjunction with the Maryland Highway Administration, the decision was made that the most logical location for the field survey was the southern terminus of I-95 in Maryland at its interchange with the Capital Beltway, I-495. This 8-lane facility was opened to traffic in the summer of 1971. It runs approximately parallel to the 3 other major, north-south facilities in the corridor:

1. US-29, a 4-lane highway located approximately 3 miles (4.8 km) west of I-95 that

connects I-495 and I-70N (west of Baltimore);

2. US-1, a 4-lane highway with numerous access points located approximately 2 miles (3.2 km) east of I-95 that connects I-495 to I-695 (the Baltimore Beltway); and

3. Baltimore-Washington Parkway, a 4-lane facility with access control located approximately 5 miles (8.0 km) east of I-95 that provides the most direct connection between the centers of Washington and Baltimore.

The locations of these routes within the corridor are shown in Figure 1. There are 7 completed interchanges on the 22-mile (35.4-km) section of I-95 between the 2 beltways. The major portion of traffic on I-95 is not local, and a sizable percentage of the truck traffic, which is excluded from the Baltimore-Washington Parkway, uses this facility.

Figure 2 shows the location of the study site and the 3 stations where the questionnaires were handed out to the motorists. The researchers cooperated with officials of the Maryland Highway Administration and the Maryland State Police in the design and implementation of the field study. Personnel from these organizations together with researchers from the University of Maryland directed the traffic into the appropriate interview lanes by means of traffic cones and flares.

The advance warning signs used to advise the motorists of the traffic survey are shown in Figure 3. The signs were mounted on portable barricades along the roadway shoulder and were turned to face approaching traffic just before the beginning of the interviews. Table 1 gives some information on the interviews that were conducted on 3 different days in 1973. The interview process was quite smooth at stations 1 and 3, and the queue length never exceeded 10 vehicles per lane. The failure of the motorists to exit the ramp and approach interview station 2 at a suitable rate coupled with the poor operating characteristics of trucks on the upgrade approaching this station produced an unstable operation. The queue became excessive almost immediately after the distribution of questionnaires was begun. Because of the potentially hazardous situation created by this backup, this survey was halted after 12 min. More detailed information concerning the procedures used for this study and other pertinent facts suitable for future studies of this type are documented in a separate report (7).

A total of 6,593 questionnaires was distributed at the 3 stations with a breakdown per station as follows:

Station	Questionnaires
1	3,136
2	212
3	3,245

Peak questionnaire distribution rates of 8 per minute per lane were observed for lanes having 1 interviewer, and 13 per minute per lane for lanes having 2 interviewers.

ANALYSIS OF RESPONSES

An equal number of A and B questionnaires was distributed, and, of these, 2,896, or nearly 44 percent, were returned. In response to the first question, approximately 72 percent of the respondents indicated that, in general, current signing on the highway system was adequate. However, there may be an inherent bias in this question because those who responded negatively were asked to elaborate on their answer. Of those who indicated that the signing was inadequate, the most frequently cited complaint involved the lack of sufficient advance warning to permit proper route choice at interchanges and intersections. Ambiguity of sign messages, especially for motorists unfamiliar with the area, was the second most frequently cited complaint. Numerous respondents indicated a desire for more real-time information, especially with respect to traffic

Figure 1. Washington-Baltimore corridor.

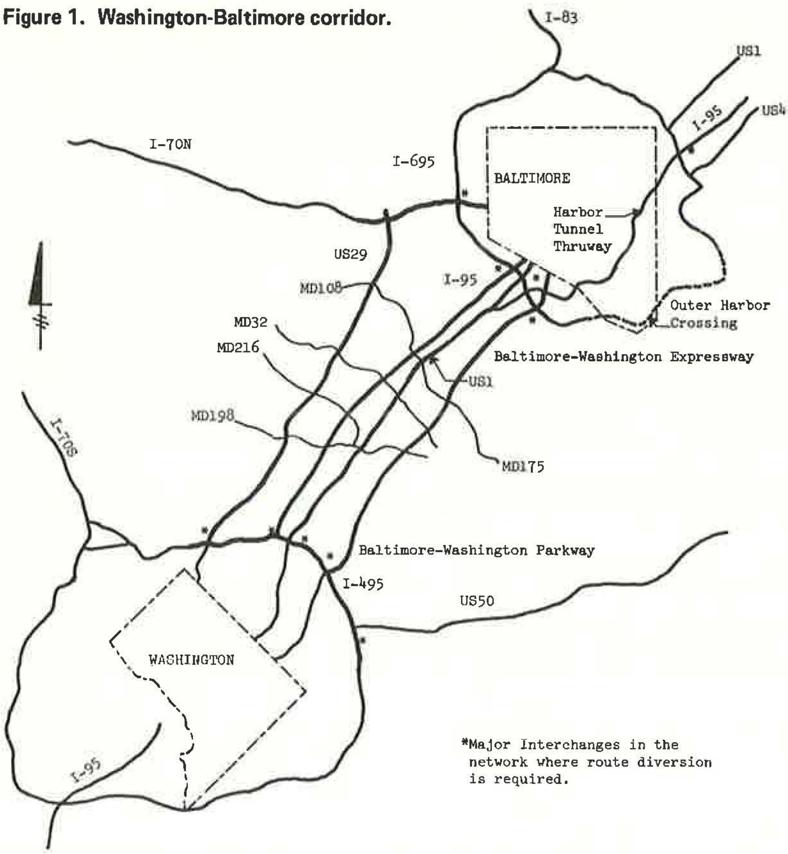


Figure 2. Operation of the I-95/I-495 interchange.

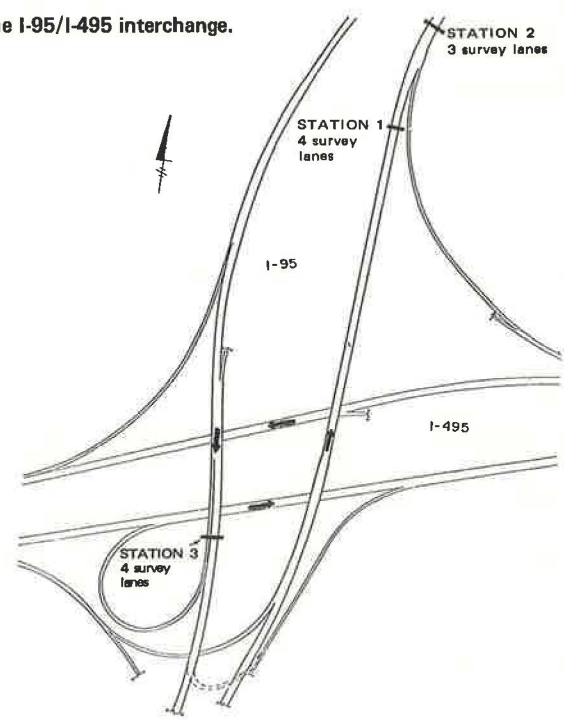


Figure 3. Signing for survey approach.

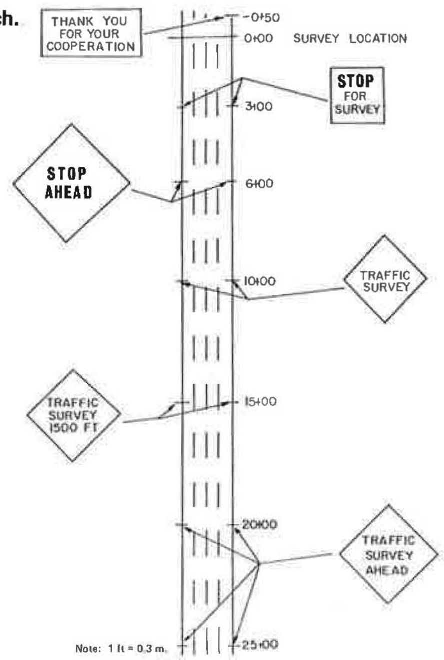
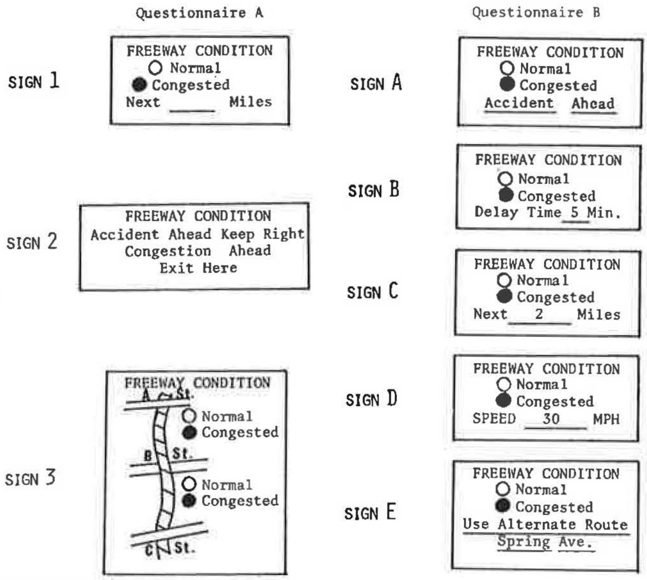


Table 1. Interview dates and locations.

Station	Movement	Date	Time
1	I-495 eastbound to I-95 northbound	Oct. 23	4:00 to 6:00 p.m.
2	I-495 westbound to I-95 northbound	Oct. 24	4:00 to 4:12 p.m.
3	I-95 southbound to I-495 eastbound	Oct. 31	7:00 to 9:00 a.m.

Figure 4. Signs in questionnaires.



Note: 1 mile = 1.6 km.

operations and weather conditions.

With few exceptions, complaints regarding current signing centered on guide signs rather than on regulatory or warning signs. This was expected because guide signs predominate on most highway systems. Also the degree of standardization that is inherent in the design, use, and placement of regulatory and warning signs tends to make them more commonly understood. On the other hand, the uniqueness of guide signs, coupled with only generalized criteria on their design and use, is conducive to motorists' misunderstanding and confusion.

Proper signing techniques are essential for a real-time program of intercity route diversion. The message must be presented at the proper location, be clearly and quickly understood by motorists, and be credible. That a small but significant percentage of drivers feel that current signing, which is primarily static, is inadequate suggests that special attention would be warranted in the development of variable message signs.

Information on the general characteristics of the motorists using I-95 was obtained through a set of 10 questions; this information was used primarily to examine possible relationships between these characteristics and sign preferences. The characteristics served to create a profile of the typical peak-period motorist on this facility. The typical driver on this facility had been driving for 19 years and had covered an annual driving distance of 21,000 miles (33 800 km). The mean number of vehicles in the family was 2, and the freeway was used approximately 10 times per week. The only average characteristic that seemed unusual was the reported 21,000 miles (33 800 km) of travel per year. This may be explained by the fact that drivers tend to overestimate their annual travel and because there was a rather high percentage of commercial drivers in the sample [17 percent of the drivers reported more than 30,000 miles (48 300 km)].

As mentioned previously the 2 questionnaires that were distributed provided different methods for the respondents to indicate their preference for various types of diversion signing. Questionnaire A presented 3 signs, and asked the drivers to choose the best, the worst, and to select from a list the reason for their choices. Questionnaire B presented 5 signs and asked the respondents to indicate their preference when the signs were compared in pairs. The signs are shown in Figure 4. A summary of the responses from the 2 questionnaires is presented in an appendix.[1]

Questionnaire A

Sign 1 on questionnaire A had color indications for normal or congested flow together with information concerning the length of the congested section. Sign 2 employed a word message advising of travel conditions. This message would be changeable in response to conditions but would not use color or picture indicators. Sign 3 combined a pictorial representation of the area with color indications of normal and congested flow.

The sign most frequently chosen as best was sign 2. Approximately 46 percent of the time it was chosen as the best sign; 41 percent of the time sign 1 was selected. The distribution is comparatively small and does not suggest a real difference between a simple, colored-coded sign containing minimal information and a word-message sign giving more detailed messages. The sign most frequently identified as worst was the pictorial, color-coded sign. This sign was chosen as worst 74 percent of the time. About 15 percent of the time sign 2 was chosen as the worst sign. The conciseness of sign 1 was most frequently cited as its major advantage. The authoritative nature of sign 2 (EXIT HERE) was appreciated by many motorists who were concerned about what action to take. The major problem noted for sign 3 was that it took too long to find the desired information on the sign.

---

[1] The original manuscript of this paper included an appendix, Traffic Survey Summary. The appendix is available in Xerox form at cost of reproduction and handling from the Transportation Research Board. When ordering, refer to XS-54, Transportation Research Record 531.

Analysis of the responses to questions that solicited drivers' opinions showed that drivers desired that sign messages, in order of importance, (a) be brief and concise, (b) indicate the nature of the situation, (c) suggest appropriate driver response, and (c) provide supplementary information. When the respondents were asked to supplement their reasons for selecting a sign, few did so. The only substantial comments were related to the use of the words normal and congested. Several motorists noted that the terms were ambiguous. It was not clear to some drivers whether normal meant free flow, average flow, or typical conditions for the specific time of day when the message was displayed. Although the color code itself was comparatively straightforward, the interpretation of the message must be clarified if this method of presenting information is to be useful.

Further analysis of the signs on questionnaire A indicated that sign 2, the most popular sign, was increasingly more popular among those motorists who drove great annual distances. On the other hand, those that indicated lower annual driving distances tended to prefer the graphic representation given on sign 3. Using an analysis of variance test to compare the sign selected as best versus the annual distance driven, we found (at the 5 percent level) that these 2 factors were interrelated.

The hypothesis that the number of years of driving experience was unrelated to the choice of the best sign was tested in a similar manner. Preliminary analysis suggested that preference for sign 2 increased with an increase in driving experience. Sign 1 was clearly most popular with those who had less than 8 years of driving experience. Enthusiasm for sign 3 tended to decrease among drivers who had more experience. These apparent trends were confirmed by statistical testing, which concluded that sign preference was related to years of driving experience.

Questionnaire B

Five different signs were presented on questionnaire B, and the respondents were asked to indicate their preference among signs that were compared 2 at a time. Each sign included a color-coded indication for normal or congested conditions along with 1 of the following types of variable message:

Sign	Message
A	Cause of congestion
B	Expected delay time
C	Length of congestion
D	Variable speed limit
E	Alternate route information

Sign C, which indicated length of congestion, was the most popular. It compared favorably with each of the other 4 signs. The runner-up was sign E, which suggested an alternate route of travel to avoid the congestion. In response to the question comparing signs C and E, sign C was selected 54 percent of the time. Sign B ranked third, and was followed by sign A. In comparison with each of the other signs, D was always judged to be the worst.

Rather consistent patterns were found when sign choice was related to annual distance driven. Sign C was identified as best in all of the distance-driven groupings. Those traveling less than 12,000 miles (19 200 km) per year selected sign B as their second choice; those reporting higher levels of annual travel selected sign E as their second choice.

Analysis of sign preference as a function of years of driving experience produced somewhat mixed results. Those citing experience in the 3 middle ranges (4-45 years) selected sign C as the best sign although the comparatively small sample of respondents (4 percent of the total) with less than 4 years or more than 45 years of experience preferred sign E. The variable speed limit message on sign D was judged worst by those with less than 20 years of driving experience although the remaining drivers assigned the cause-of-congestion message on sign A to this category.

SUMMARY

On the basis of this analysis, there are 3 types of signs that are worthy of consideration for real-time route diversion. These are

1. Congestion length (questionnaire A, sign 1; questionnaire B, sign C),
2. Congestion cause and exit instruction (questionnaire A, sign 2), and
3. Alternate route information (questionnaire B, sign E).

The congestion-length message is preferred because of its conciseness. It provides motorists with information they should be able to evaluate and translate into effective action. The latter 2 signs by nature are not concise, but they do convey a sense of authority. In addition, they give motorists an alternative to being unnecessarily delayed on the planned travel route.

The comparatively close ranking of these 3 signs precludes a judgment at this time on which sign is truly the best. Based on this sample of intercity freeway drivers, it is not possible to recommend signs employing schematic representations or those indicating speed or length of delay. The former seem to require too much time to locate the intended message, and the latter apparently do not satisfy drivers' needs for meaningful information.

Although the signs tested in this study are representative of those that others have suggested (and have used in urban corridors), there is no assurance that the optimal sign message and design is included among the signs presented on the questionnaires. This, in fact, is an inherent problem in trying to select an optimal alternative. However, judicious field testing and evaluation of the recommended signs may suggest ways in which they can be modified to achieve the most suitable design.

The generally recognized need for uniformity and consistency tends to support the concept that 1 type of route diversion sign should be used. It is easy to appreciate, however, that the sign should be suited to the circumstance, and that this may require 2 or more types of signing along a route. The placement of signs advising the motorist to exit or to use a specified alternate route requires that the diversion network include a proper route for diverting at least moderate volumes of traffic at that location. When the only available alternate at a particular interchange is a local road, the more subtle advice given by the congestion-length message might be more appropriate. It would be more likely to prompt the early exit from the freeway of a few local motorists who had originally planned to exit within a short distance. On the other hand, the diversion of through motorists at major diversion points is probably enhanced by signs advising of alternate routes.

REFERENCES

1. J. C. Brisson and J. A. Ville. Traffic Control for the 10th Winter Olympic Games of Grenoble. Traffic Engineering, Vol. 39, No. 1, Oct. 1968.
2. P. F. Everall. Urban Freeway Surveillance and Control. U.S. Department of Transportation, June 1973.
3. F. R. Hanscom. Changeable Sign Messages. Rural and Urban Roads, Vol. 9, No. 1, Jan. 1970.
4. K. W. Heathington, R. D. Worrall, and G. C. Hoff. An Analysis of Driver Preferences for Alternative Visual Information Displays. Highway Research Record 303, 1970, pp. 1-16.
5. K. W. Heathington, R. D. Worrall, and G. C. Hoff. Attitudes and Behavior of Drivers Regarding Route Diversion. Highway Research Record 363, 1971, pp. 18-26.

6. The Changeable-Message Concept of Traffic Control. HRB Spec. Rept. 129, 1972, 48 pp.
7. Research on Route Diversion Parameters. Maryland Highway Administration, Final Rept. AW 74-131-46, June 1974.
8. Sperry Systems Management Division. Traffic Surveillance Study, Eight Northeast New Jersey Counties, Topics Program T4001 (29)—Phase I Final Report. New Jersey Department of Transportation, Dec. 1972.
9. Sperry Systems Management Division. Diversion of Intercity Traffic at a Single Point. Federal Highway Administration, Final Design Rept., Draft, Sept. 1973.

GENERALIZED COMBINATION METHOD FOR AREA TRAFFIC CONTROL

Nathan Gartner and John D. C. Little,
Operations Research Center,
Massachusetts Institute of Technology

A simple generalization of the British combination method is given for optimizing offsets in synchronized, traffic-signal networks of a general structure. The method then is used in a recursive procedure to determine values for the offsets along each street, the splits of green time at each intersection of the network, and the common cycle time of the controlled area. The signals' cost to travelers is evaluated as the sum of 2 components: one associated with a deterministic traffic-flow model and the other associated with randomness in traffic behavior. The deterministic component is a function of the coordination among the signals in the network and generally increases with cycle length. The stochastic component depends on the expected overflow queue at each traffic light and decreases with cycle length. It is shown that optimal settings are determined at the equilibrium point of minimum total cost resulting from the combined effect of the 2 components.

•THE PRIMARY objectives of an areawide traffic-control system are to provide smooth flow conditions for all traffic streams through the area and to reduce the delay, or travel time, incurred by users of the system. The variables of each signal program that affect the traffic flow are cycle time, splits of green time, and offsets. A coordinated traffic-signal network requires a common cycle time for all signals in the network or a cycle that is a submultiple of a master cycle. In some cases it is advantageous to partition the network into subnetworks that operate with nonsynchronous cycle times.

The conventional procedure for determining control variables is a sequential decision process. First, a common cycle time is selected for the network. Second, the splits at each intersection are determined according to the proportions of demand-capacity ratios on conflicting approaches. Third, linking of the signals is achieved by an appropriate method for selecting a fundamental set of offsets throughout the network.

Experience of researchers and practitioners in the urban traffic-control field has shown that cycle time may well be the most important control variable in a synchronized traffic-signal network (1). The approaches for selecting a cycle time can be divided into 2 classes. The first class is the node approach. Because through capacity increases with cycle length, this approach is based on analyzing the capacity requirements of each intersection in the network. Common cycle time is determined according to the requirements of the most heavily loaded intersection—the intersection with the highest sum of demand-capacity ratios on conflicting signal phases. A procedure that is used for a single intersection, such as Webster's method (2), is then used to calculate cycle length. This approach has been primarily used in conjunction with offset optimization methods such as COMBINATION and TRANSYT (3). The main deficiency in this approach is that the interaction of flows in the spatial road network structure of the area is disregarded. A formula devised for an isolated intersection, assuming that arrival

Publication of this paper sponsored by Committee on Traffic Control Devices.

times of cars are randomly distributed, is not necessarily valid in a network situation in which flows are fed from adjacent intersections. The result is generally a cycle time that is too long, which causes excessive delays (4, discussion). The second class is the network approach. In this case an attempt is made to select a cycle time that satisfies the capacity requirements at each intersection and is congruent with the particular network structure at hand. Simple examples in this category are the arterial progression schemes in which a cycle that produces maximal bandwidths is selected according to distance and speed data (5, 6, 7). The underlying principle is that optimal progression (offsets between signals) for a given block-length pattern is strongly dependent on cycle time. In a general network this approach is used principally by SIGOP (8). A predetermined number of cycle times are scanned in this method. For each cycle, offsets are optimized by the OPTIMIZ subroutine and performance is evaluated by a coarse simulation of traffic flow through the network (VALUAT subroutine). The optimal set of cycle and offsets is selected according to the results obtained by VALUAT. TRANSYT also indicates the possibility to iterate on cycle time in conjunction with the hill-climbing procedure for offset selection (9). However, the extensive computational requirements of this method seem to rule this out in practice. Two deficiencies of the network approach in SIGOP are apparent. First, the offset optimization procedure determines a local optimum rather than a global optimum. Second, stochastic effects on link performance are ignored. These effects do not affect the selection of offsets at a fixed cycle time, but they are of prime importance in evaluating a range of cycle times. They become pronounced as a signalized intersection approaches its capacity and, in an optimal procedure, would deter the cycle time from assuming values close to the minimum. One typical study has shown that the lower bound on cycle time was consistently selected as the optimal value. Stochastic effects conceivably would have shifted the result upward (10).

In this paper, network settings, including cycle, splits, and offsets, are determined in conjunction with a rigorous synchronization procedure (that is, one capable of determining the global optimum) that is an extension of the British combination method (CM). The combination method is an offset optimization procedure applicable to series-parallel networks; it was first introduced by Hillier (11). It was then applied by Allsop (12) to networks of a more general structure. The method was later formulated in terms of dynamic programming optimization and applied in conjunction with a computationally efficient network partitioning algorithm (13). The dynamic programming procedure for the general network is presented in this paper as a set of 2 network operation rules that are a straightforward generalization of the combination method rules for series-parallel networks. The procedure is further used as a tool in determining optimal network settings that take into account costs attributable to both the deterministic traffic-flow model and the stochastic fluctuations inherent in the traffic process.

TRAFFIC-FLOW MODEL

To illustrate the key features of the traffic-flow process, we should consider an idealized model. The discrete nature of vehicular movement would be disregarded and traffic would be thought of as continuously fluid. The following assumptions would be made:

1. All cars travel with uniform speed between adjacent intersections; and
2. Traffic flow is saturated; that is, traffic volume at each intersection equals serving capability.

Let i and j denote 2 adjacent signalized intersections in the network; cars can travel from i to j along the link connecting them. The following are definitions of the parameters shown in Figure 1:

g_j = effective green time of signal j,
r_j = effective red time of signal j,
$C = g_j + r_j$, network common cycle time,

ϕ_{ij} = offset time between signals i and j,
t_{ij} = travel time from i to j,
$F_{ij}(t)$ = instantaneous traffic flow in vehicles per unit of time, and

$$F_{ij} = \frac{1}{C}\int_0^C f_{ij}(t)dt, \text{ average traffic flow.}$$

When traffic is assumed to have a periodic arrival pattern of rectangular shape as shown in Figure 2a, it can be easily verified that the rate of delay or delay per unit of time, $d_{ij}(\phi_{ij})$, on the link i, j, is

$$d_{ij}(\phi_{ij}) = \begin{cases} F_{ij}\dfrac{r_j}{g_j}(t_{ij} - \phi_{ij}) & \text{if } t_{ij} - g_j \le \phi_{ij} \le t_{ij} \\ F_{ij}(\phi_{ij} - t_{ij}) & \text{if } t_{ij} \le \phi_{ij} \le t_{ij} + r_j \end{cases} \tag{1}$$

and is similarly periodic with respect to ϕ_{ij} (Fig. 2b). Examination of Figure 1 indicates that the offset ϕ_{ij} can be expressed as follows:

$$\phi_{ij} = mC + \theta_{ij} \tag{2}$$

Figure 1. Link and signal parameters.

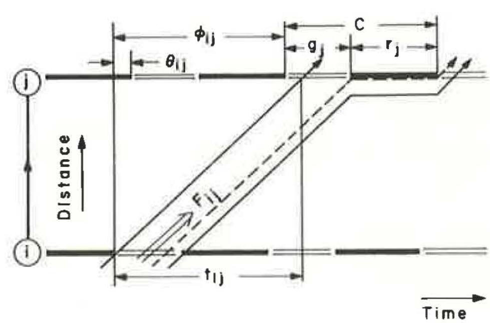
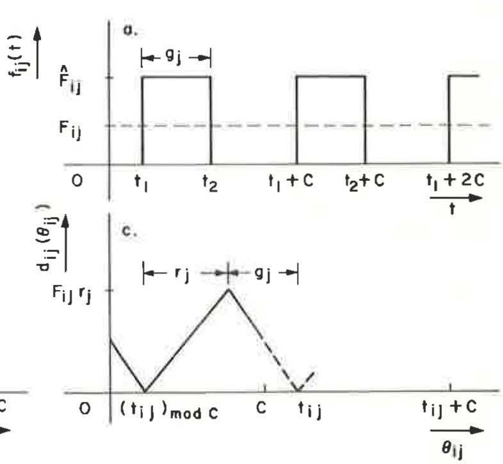
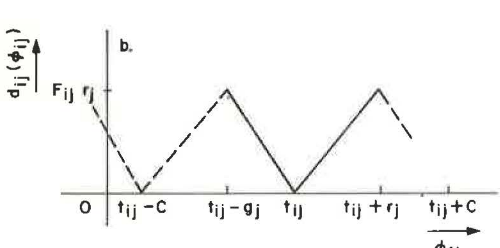


Figure 2. Idealized traffic-flow pattern and link delay functions.





where m is an integer number (in Fig. 1, n = 1). Thus we can confine offset variations to a single cycle time—$0 \leq \theta_{ij} \leq C$—by introducing the transformation

$$\theta_{ij} = \phi_{ij} - mC = (\phi_{ij})_{\bmod C} \tag{3}$$

The resulting delay function $d_{ij}(\theta_{ij})$ is shown in Figure 2c. More elaborate models can be used to more closely approximate traffic conditions by taking into account secondary flows, platoon dispersion, and the like (14, 15, 16). An example of an actual traffic-flow pattern that has been measured directly by detectors on the street in the Toronto traffic-control system is shown in Figure 3a. The link delay function associated with this pattern is obtained by applying elementary queuing relationships (17) and is shown in Figure 3b. This paper primarily considers delays, but the same optimization methods can be used with a more general link performance function combining costs of delays, stops, acceleration noise, or other measures of effectiveness by using appropriate weighting factors. Huddart (18) and Chung and Gartner (19) discuss additional measures of effectiveness.

CRITERION OF OPTIMIZATION

The objective of the network optimization procedure adopted here is to determine signal settings (cycle time, splits, and offsets) that minimize total delay. In a recent report (20) it was shown that total delay in the network, D, can be regarded as a sum of 2 components as follows:

$$D = D_d + D_s \tag{4}$$

The first component, D_d, is the delay time resulting from the deterministic traffic-flow model previously described. In a network context it is obtained by summing all the individual link delay functions such as those represented by Eq. 1 or Figure 3b.

$$D_d = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}(\theta_{ij}) \tag{5}$$

where n = the number of intersections (nodes) in the network. $d_{ij} \equiv 0$ if the link i, j does not exist. For given cycle and splits this delay is a function of offsets only. The second component of delay, D_s, is due to the stochastic nature of traffic flow. It is taken to be independent of the choice of offsets in the network but is of primary importance for evaluating the best choice for cycle time because a change in cycle time involves a change in the degree of saturation at the intersection.

The procedure for optimization consists of scanning a number of cycle times that are usually in 10-sec intervals in the range of 40 to 120 sec. For each cycle time, splits at each node are calculated according to proportions of conflicting traffic streams (2), and offsets throughout the network are optimized by the generalized combination method (GCM). Another approach would be to formulate the problem in terms of an existing optimization code such as mixed-integer linear programming and to simultaneously have all the signal timings as decision variables (20).

A physical requirement of the system is that the sum of offsets around any closed loop must be equal to an integral number of cycle times. The maximum number of offsets, θ_{ij}, that can be assigned independent values in a network of n nodes is n - 1,

Figure 3. Actual platoon profile and link delay function.

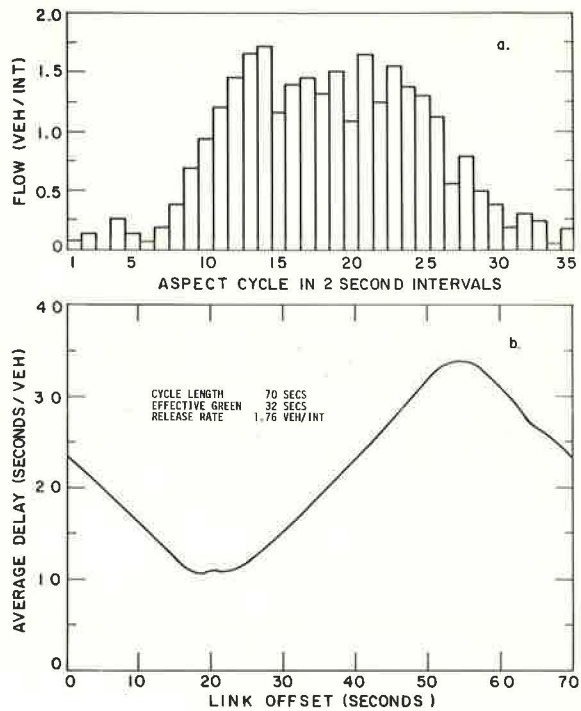


Figure 4. Series-parallel network reduction rules.

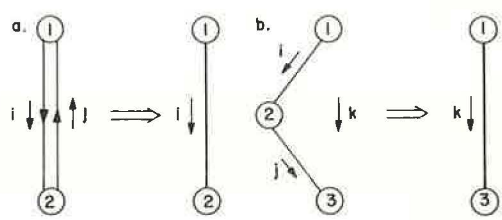
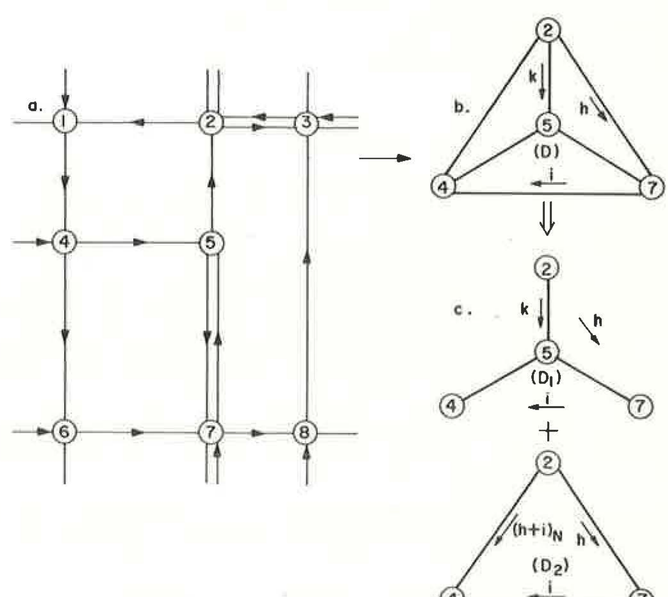


Figure 5. Signal network for example 1.



and the links across which they are defined must be in a tree pattern; in other words, they must have no loops (21). It has been shown that offset variations can be confined to the range of a single cycle's time. The computational procedure for minimizing D with respect to offsets involves the division of this range into equal N intervals. It is convenient to consider link delays to be a function of an integer number k, where k = 0, 1, ..., N - 1, which represents the offsets at which delay is to be evaluated. To simplify notation we also should adopt the convention $(x)_{\bmod N} \equiv (x)_N$.

COMBINATION METHOD

The combination method determines offsets that minimize delays in series-parallel networks (11). The method applies a network reduction sequence to yield a total delay function for the complete network that can be represented by a single equivalent link. The optimizing offsets of the network are determined by minimizing this function. An efficient procedure for determining this sequence was developed by Robertson (22). The reduction sequence is based on 2 rules.

1. The first rule, CM 1, is the reduction of parallel links, which states that, when 2 or more links occur in parallel and join a pair of nodes, the delay functions of the individual links are added with respect to the same offset to yield a combined delay function represented by a single link between the 2 nodes. Application of this rule is shown in Figure 4a. Given $d_{12}(i)$ and $d_{21}(j)$, the combined delay, $D_{12}(i)$, for the equivalent link is

$$D_{12}(i) = d_{12}(i) + d_{21}(N - i) \tag{6}$$

for each offset i = 0, 1, ..., N - 1.

2. The second rule, CM 2, is the reduction of series links, which states that, whenever a node is connected by 2 links to 2 other nodes, it is deleted and the 2 links are replaced by a single link. The equivalent delay function for this link is computed by minimizing the total delay for each offset between the extremities of the 2 links. At each step the procedure involves a search of all the possible offsets between one of the extremal nodes and the common node and a selection of the minimum.

Given $d_{12}(i)$ and $d_{23}(j)$, the delay function for the equivalent link 1, 3 in Figure 4b is obtained by eliminating from further consideration the offset of node 2 with respect to node 1 (offset i) through the following minimization:

$$D_{13}(k) = \min_{i} \{d_{12}(i) + d_{23}(j)\} \tag{7}$$

for all offsets k, i = 0, 1, ..., N - 1. Because the 3 offsets i, j, and k form a closed loop, they must add up algebraically to an integral number of cycle times:

$$i + j - k = mN \tag{8}$$

or equivalently

$$j = (k - i)_{\bmod N} \tag{9}$$

Therefore, Eq. 7 can be rewritten as follows:

$$D_{13}(k) = \min_{i} \{d_{12}(i) + d_{23}[(k - i)_N]\} \tag{10}$$

for $k, i = 0, 1, \ldots, N - 1$.

GENERALIZED COMBINATION METHOD

This method relieves the series-parallel restriction imposed on the structure of networks by the ordinary combination method. By generalizing the rules stated in the preceding section it is possible to optimize networks of arbitrary layout (subject to computational considerations only).

1. The first generalized rule, GCM 1, is the combination of partial networks. Delay functions that pertain to separate parts of a network and depend on offsets between the same set of nodes are added to produce an equivalent delay function for the combined parts of the network.
2. The second generalized rule, GCM 2, is the elimination of interior nodes. An equivalent delay function for a partial network is calculated for all offsets between the boundary nodes (the nodes that disconnect a part of the network from the remainder of the network) by eliminating from the optimization process the offsets related to the interior nodes. The values of the function are determined by minimizing the total delay of the partial network for all offsets between the boundary nodes. At each step the calculation is effected by searching over all possible offsets associated with the interior nodes and selecting the minimum.

Rules CM 1 and CM 2 are special cases of rules GCM 1 and GCM 2. Recursive application of these rules defines a total delay function for the complete network for offsets between a certain final set of nodes. Optimizing offsets are determined by minimizing this function. Application of the generalized combination method is illustrated in the following 2 examples.

Example 1

The network to be optimized is illustrated in Figure 5a. Series-parallel combination produces the ∇ - Y configuration shown in Figure 5b that cannot be further reduced by these simple operations. At this stage the network is disconnected into 2 parts and a delay function is calculated for each separately (Fig. 5c). Following rule GCM 2 we obtain

$$D_1(h, i) = \min_{k} \{d_{25}(k) + d_{75}[(k - h)_N] + d_{54}[(h + i - k)_N]\}$$

This partial minimization also yields the relation $k^*(h, i)$ where k^* is the optimizing value of offset k for each combination of h and i. This relation is stored for subsequent use. Now applying rule GCM 1 we obtain

$$D_2(h, i) = d_{27}(h) + d_{24}[(h + i)_N] + d_{74}(i)$$

and D is

$$D(h, i) = D_1(h, i) + D_2(h, i)$$

Minimization of D(h, i) with respect to offsets h and i determines optimizing values h* and i*. Backtrack computation via the stored relation k*(h, i) and loop constraints yields the optimizing offsets for all links of the original network.

Example 2

The original signal network is shown in Figure 6a. After series-parallel reductions the compressed network of Figure 6b is obtained. Optimizing offsets are calculated by staged partitioning of this network and recursive application of the GCM rules at each stage. A partitioning plan that minimizes the number of operations and storage requirements for this network is given in the following table:

Stage Number	Disconnecting Nodes	Eliminated Interior Nodes
1	2, 3, 4	1
2	5, 3, 4	2
3	5, 6, 4	3
4	5, 6, 7	4, 8

The detailed minimization process is given as follows and shown in Figure 6c:

$$D_1(h, i) = \min_{j} \{d_{12}[(j - h)_N] + d_{13}(j) + d_{14}[(i + j)_N]\} + [d_{23}(h) + d_{34}(i)]$$

$$D_2(\ell, i) = \min_{j} \{d_{25}(k) + D_1[(k + \ell)_N,\ i]\}$$

$$D_3(n, p) = \min_{m} \{d_{36}(m) + D_2[(n - m)_N,\ (m + p)_N]\}$$

$$D_4(n, r) = \min_{q} \{d_{47}(q) + D_3[n, (r - q)_N]\} + [d_{56}(n) + d_{67}(r)] + \min_{s} \{d_{58}[(n + s)_N] + d_{68}(s) + d_{78}[(s - r)_N]\}$$

The delay function obtained at stage 4 represents total delay in the network for each possible combination of offsets n and r. The terminal optimization stage consists of minimizing this function with respect to n and r and calculation, by backtracking, of an independent set of optimal offsets (in this case, offsets j*, k*, m*, q*, n*, r*, s*).

NETWORK CYCLE TIME

The traffic-flow pattern on a signalized link can be regarded as the combination of a periodic component imposed by the preceding signal and a random component arising from variations in driving speeds, marginal friction, and turns. The latter component causes additional delay because of the occurrence of an overflow queue at the signal's stop line. The overflow queue represents the number of vehicles that were not cleared during the preceding green phase. Although this effect is negligible at low degrees of saturation, its predominance at high values has been proved in several studies (17, 23, 24).

Using Webster's notation for traffic-signal settings (2), we have at each node in the network the following relation

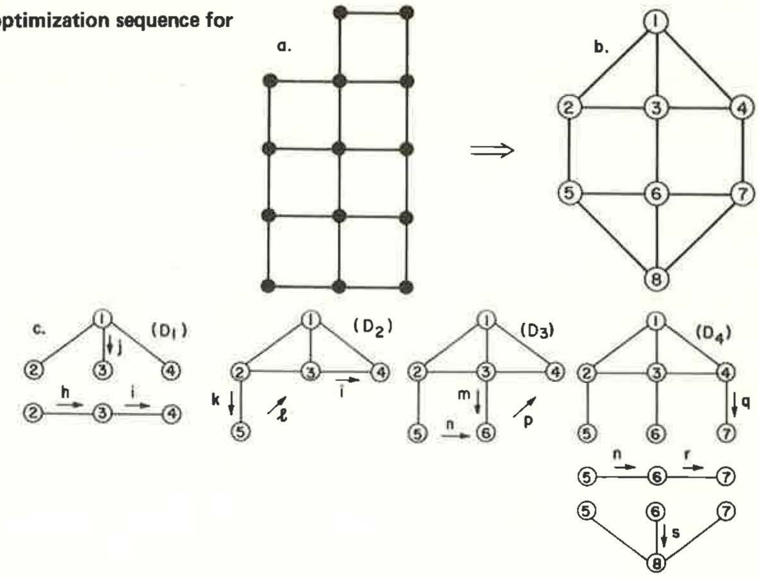


Figure 6. Signal network and optimization sequence for example 2.

Table 1. Expected overflow queue.

Signal Capacity (vehicles per cycle)	Degree of Saturation						
	0.20	0.40	0.60	0.80	0.90	0.95	0.975
5	0.00	0.02	0.20	1.15	3.50	8.41	18.36
15	0.00	0.00	0.04	0.70	2.81	7.61	17.50
25	0.00	0.00	0.01	0.47	2.41	7.08	16.91
35		0.00	0.00	0.34	2.11	6.68	16.45
45			0.00	0.23	1.88	6.34	16.05
55					1.68	6.02	15.67

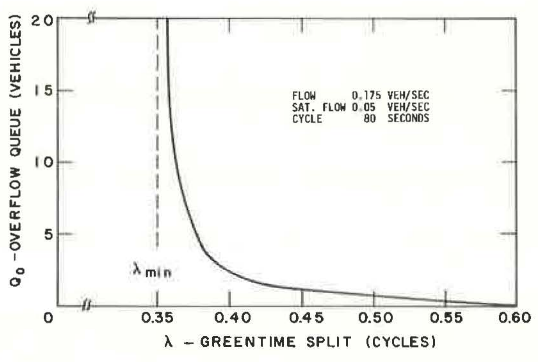


Figure 7. Overflow queue versus green-time split.

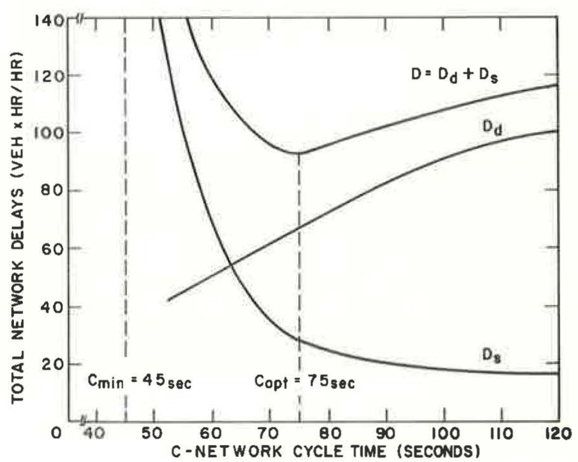


Figure 8. Variation of total network delays with cycle time.

$$\sum_{j} g_j = C - L \tag{11}$$

That is, the sum of effective green times on all phases equals the net green time available for movement through the intersection (cycle time less lost time). Rearranging this we obtain

$$\sum_{j} \lambda_j = 1 - \frac{L}{C} \tag{12}$$

where $\lambda_j = \frac{g_j}{C}$ denotes the green split j (fraction of cycle time allotted to phase j). The split, in turn, is determined as follows:

$$\lambda_j = \frac{y_j}{Y}\left(1 - \frac{L}{C}\right) \tag{13}$$

$y_j = \frac{F_{ij}}{s_j}$ is the representative ratio of flow (F_{ij}) to saturation flow (s_j) of a particular phase and $Y = \sum_{j} y_j$ is the sum of y-values over all phases of the intersection. The y-values depend only on flow and saturation flow, but not on the signal settings themselves. The total lost time, L, is usually a fixed quantity at a particular intersection (3 to 5 sec for each phase). Therefore, a change in C alters the total net green time available for passage through the intersection and, consequently, its allotment to the phases—the green splits. This eventually brings about a change in the degree of saturation and, with it, the size of the overflow queue.

An estimate of the expected overflow queue, based on the capacity of the signal's approach and the degree of saturation, was calculated by Wormleighton (25) and is given in Table 1. Following field studies in Toronto, he developed a model describing the traffic behavior along a signalized link as a nonhomogeneous Poisson process with a periodic intensity function. A typical relationship between expected overflow queue and split time in this model is shown in Figure 7. Similar characteristics are used by Webster (2) in the case of the single intersection and by Robertson (9) in the TRANSYT network model.

Let us denote the expected overflow queue on link i, j with a downstream green split λ_{ij} by $Q_o(\lambda_{ij})$. The delay incurred by these queuing vehicles is simply $Q_o(\lambda_{ij})$ for any time unit that is used, such as [vehicles × hour/hour] or [vehicles × sec/sec]. The networkwide expected delay associated with the overflow queue thus will be

$$D_s = \sum_{i=1}^{n} \sum_{j=1}^{n} Q_o(\lambda_{ij}) \tag{14}$$

This provides the second component of the network objective function given in Eq. 4.

Recursive application of the GCM for different cycle times, taking into account both deterministic and stochastic effects, produces typical results as shown in Figure 8. These curves were calculated for the network shown in Figure 6a. Input links also must be included in the calculation. Although they do not affect signal coordination (calculation of offsets), they play an important role in evaluating total delay for selecting the proper cycle time.

It is evident that optimal cycle time for the network constitutes an equilibrium point between delays caused by deterministic effects and delays caused by stochastic effects. Although the former usually increase with cycle length, the latter decrease with it because of the decrease in the degree of saturation (load factor). They are asymptotic to the minimal cycle time for the network, which is the theoretical minimal cycle time for the most heavily loaded intersection that would still provide capacity if all flows were deterministic. These characteristics are completely analogous with the behavior of delay with respect to cycle time at a single intersection as studied by Webster (2). However, the results are significantly different and an analysis of a single intersection would virtually never give the optimum cycle time for the network. In the example shown in Figure 8, the optimum cycle time for the critical intersection in the network is approximately 90 sec. If this cycle time were adopted for the whole network, delay would be about 10 percent higher than optimum.

SUMMARY AND CONCLUSIONS

A systematic procedure was developed to determine signal settings (including offsets, green splits, and cycle time in a network). The basic building block of the procedure was the generalized combination method, which extended the applicability of the original combination method to networks of a general structure.

The traffic-flow model consisted of deterministic and stochastic components. The deterministic component represented periodic platoons of similar shape and size generated in a synchronized signal network. The stochastic component accounted for the variability in the characteristics of these platoons as observed in practice. A travel-cost function was associated with each component. The deterministic component cost function tended to drive cycle time down and minimized its value. On the other hand, the stochastic component cost function deterred the signal timings from approaching saturation levels at any intersection of the network and thus drove the cycle time upward. This interplay between the 2 functions was of fundamental importance in analyzing the performance of area traffic-control systems. Optimal settings in a network were determined by the least-cost equilibrium point reached as a result of this interplay.

Preliminary results obtained by applying this method to test networks indicated a potential for significant improvements in the performance of traffic-signal systems compared with other techniques in current use. As with any new model or methodology, further testing and evaluation are necessary and implementation studies are planned.

REFERENCES

1. F. A. Wagner, F. C. Barnes, and D. L. Gerlough. Improved Criteria for Traffic Signal Systems in Urban Networks. NCHRP Rept. 124, 1971.
2. F. V. Webster. Traffic Signal Settings. Road Research Technical Paper No. 30, Her Majesty's Stationery Office, London, 1958.
3. J. Holroyd. The Practical Implementation of Combination Method and TRANSYT Programs. Colloquium on Area Traffic Control, Institution of Electrical Engineers, England, 1972.
4. P. K. Munjal and Y. S. Hsu. Comparative Study of Traffic Control Concepts and Algorithms. Highway Research Record 409, 1972, pp. 64-80.
5. T. M. Matson, W. S. Smith, and F. W. Hurd. Traffic Engineering. McGraw-Hill Book Co., New York, 1955.
6. J. E. Baerwald, ed. Traffic Engineering Handbook, 3rd Edition. Institute of Traffic Engineers, Washington, D.C., 1965.
7. J. D. C. Little. The Synchronization of Traffic Signals by Mixed-Integer Linear Programming. Operations Research, Vol. 14, 1966, pp. 568-594.
8. SIGOP: Traffic Signal Optimization Program, A Computer Program to Calculate Optimum Coordination in a Grid Network of Synchronized Traffic Signals. Traffic Research Corp., New York, PB 173 738, 1966.

9. D. I. Robertson. TRANSYT: A Traffic Network Study Tool. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, RRL Rept. 253, 1969.
10. F. A. Wagner. SIGOP—TRANSYT Evaluation: San Jose, California. Dept. of Transportation, Research Rept. DOT-FH-11-7822, July 1972.
11. J. A. Hillier. Appendix to Glasgow's Experiment in Area Traffic Control. Traffic Engineering and Control, Vol. 7, No. 9, 1966, pp. 569-571.
12. R. E. Allsop. Selection of Offsets to Minimize Delay to Traffic in a Network Controlled by Fixed-Time Signals. Transportation Science, Vol. 2, 1968, pp. 1-13.
13. N. Gartner. Optimal Synchronization of Traffic Signal Networks by Dynamic Programming. Proc. 5th International Symposium on the Theory of Traffic Flow and Transportation, American Elsevier Publishing Co., New York, 1972, pp. 281-295.
14. F. A. Wagner, D. L. Gerlough, and F. C. Barnes. Improved Criteria for Traffic Signal Systems on Urban Arterials. NCHRP Rept. 73, 1969.
15. K. W. Huddart and E. D. Turner. Traffic Signal Progressions—GLC Combination Method. Traffic Engineering and Control, Vol. 11, 1969, pp. 320-327.
16. P. A. Seddon. The Prediction of Platoon Disperson in the Combination Methods of Linking Traffic Signals. Transportation Research, Vol. 6, 1972, pp. 125-130.
17. N. Gartner. Microscopic Analysis of Traffic Flow Patterns for Minimizing Delay on Signal-Controlled Links. Highway Research Record 445, 1973, pp. 12-23.
18. K. W. Huddart. The Importance of Stops in Traffic Signal Progressions. Transportation Research, Vol. 3, 1969, pp. 143-150.
19. C. C. Chung and N. Gartner. Acceleration Noise as a Measure of Effectiveness in the Operation of Traffic Control Systems. Operations Research Center, M.I.T., Cambridge, Working Paper OR 015-73, 1973.
20. J. D. C. Little, N. Gartner, and H. Gabbay. Optimization of Traffic Signal Settings in Networks by Mixed-Integer Linear Programming. Operations Research Center, M.I.T., Cambridge, Technical Rept. g1, March 1974.
21. N. Gartner. Constraining Relations Among Offsets in Synchronized Signal Networks. Transportation Science, Vol. 6, 1972, pp. 88-93.
22. D. I. Robertson. An Improvement to the Combination Method of Reducing Delays in Traffic Networks. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, RRL Rept. 80, 1967.
23. J. A. Hillier and R. Rothery. The Synchronization of Traffic Signals for Minimum Delay. Transportation Science, Vol. 1, 1967, pp. 81-94.
24. W. Jacobs. A Study of the Averaged Platoon Method for Calculating the Delay Function at Signalized Intersections. Dept. of Civil Engineering, Univ. of Toronto, B.A.Sc. thesis, 1972.
25. R. Wormleighton. Queues at a Fixed Time Traffic Signal With Periodic Random Input. Journal Canadian Operations Research Society, Vol. 3, 1965, pp. 129-141.

NO-STOP-1 VERSUS SIGPROG

Helmut R. Leuthardt*, Stuttgart, West Germany

This paper discusses the theory and fundamentals of progressively timed signal systems, whose objective is the achievement of maximum possible bandwidth. The relationship of bandwidth to speed, cycle length, splits, and distances is defined. Two conditions concerning the centers of green and their respective eccentricities are discussed, and their relationship to the bandwidth is demonstrated. Based on these relationships, a model is developed that satisfies the conditions of the Tschebysheff theorem. This theorem serves as the optimization model for maximizing bandwidth; in it speed and cycle length are varied within reasonable and defined limits. NO-STOP-1 is an outgrowth of this model. It includes a table printout that identifies all data necessary to describe a time-space diagram as well as a computer-plotted, time-space diagram that takes into consideration various options. NO-STOP-1 then is compared to SIGPROG. NO-STOP-1 yields larger bandwidths than those produced by SIGPROG. NO-STOP-1 also has a greater variety of options. NO-STOP-1 is also briefly compared to other programs.

•MORE sophisticated traffic engineering tools are needed to lessen the burden of the increasing numbers of automobiles on already overcrowded city streets. Coordination of signals along arteries is an important means of reducing delays and unwanted stops. SIGPROG is a computer program capable of providing the traffic engineer with the data required to achieve progressive movement. This program, however, has some shortcomings. NO-STOP-1 was developed as a more sophisticated computer program to solve complex problems related to progressively timed signal systems.

FUNDAMENTALS

Relationship at an Arbitrary Signal

Figure 1 shows the constant through band for 1 direction at arbitrary signal i. It is assumed that signal i is part of a larger system that has maximized bandwidths and that this signal is critical, which means that either the upper or lower limit of the through band touches red. From Figure 1, it can be shown that

$$b_i = t_{gi} - 2E_i \qquad (1)$$

where

b = bandwidth,
t_g = green time, and
E = eccentricity.

Publication of this paper sponsored by Committee on Traffic Control Devices.

*Mr. Leuthardt was with Tippetts-Abbett-McCarthy-Stratton, New York, when this research was performed.

Normally, the split at an intersection is known and relates to the traffic volumes at that intersection. When the split is known and the cycle length is given, t_g is also known, which leaves the bandwidth a function of E.

$$b = f(E) \tag{2}$$

which is to say that the location of the band axis determines the bandwidth. E may be interpreted as the offset difference of the band axis from the center of green (CG). In ideal cases, E is zero, which means that the band axis passes exactly through CG.

Relationship in a System

The primary objective in designing a progressively timed signal system is to determine the maximum constant bandwidth for both directions. Within a system, cycle length, speed, distances, and splits are the independent variables that influence the location of the band axis, and from that, the bandwidth. Distances and splits are set values for a given system because of the volumes and the physical layout of the system. Bandwidth is a function of cycle length and speed.

$$b = f(c) * f(v) \tag{3}$$

where

c = cycle length and
v = speed.

Figure 2 shows the time trajectories of 2 vehicles departing and arriving at the first intersection at the same time. Speed and cycle length are arbitrary and may be within reasonable limits. The trajectories represent the band axes of their respective through bands. Two conditions have to be met for a symmetric system to achieve the maximum possible constant bandwidth for both directions.

1. Two successive CGs must either occur simultaneously or be offset a half cycle length from each other. The band axis in 1 direction is the mirror image of the band axis in the opposite direction (Fig. 2). Therefore it is necessary to calculate the band axis for 1 direction only.
2. Under ideal conditions, the band axis passes at each intersection exactly through the CG, thereby providing the maximum possible bandwidth. Because conditions are seldom ideal, one must try to place the CGs as close as possible to the band axis so that the maximum E becomes as small as possible (Fig. 2). This condition may be expressed as

$$\max |E| = \min \tag{4}$$

Cycle length and speed as variables that influence the bandwidth are, of course, subject to reasonable limits. The range within which they may vary has to be defined. For cycle length, a range of 40 to 120 sec is commonly used. For desired progression speed, the range or progression of speed tolerance acceptable to the driver is approximately 15 percent from the desired progression speed, as studies by Leutzbach (1) and Desrosiers and Leighty (2) have shown.

By varying the cycle length and the desired progression speed, Es and locations of CGs are affected as shown in Figure 3. Figure 3 shows a band axis with a desired

Figure 1. Relationship at arbitrary signal i.

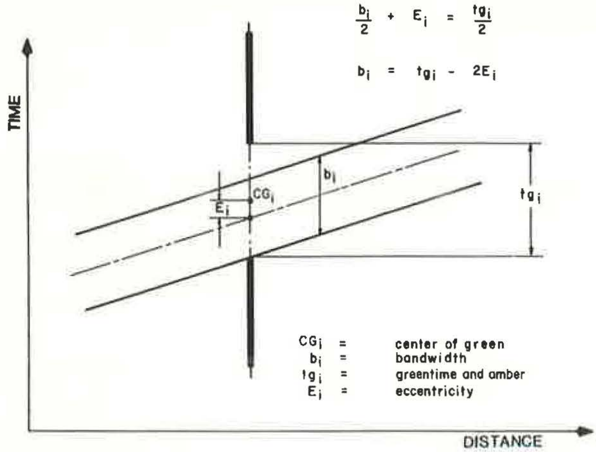
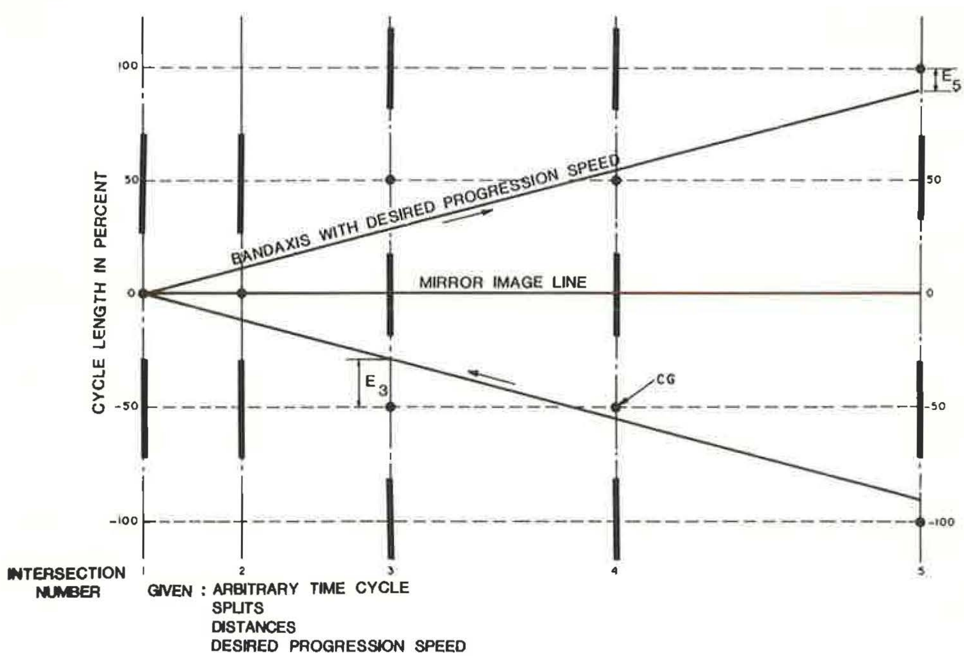


Figure 2. Relationship in a system.



progression speed and allowed tolerances (defined here as plus or minus) for an arbitrary cycle length. At intersection 5, for example, the CG for the progression speed plus the allowed tolerance would be offset 50 percent to the zero base line; the CG for the progression speed minus the allowed tolerance would be offset 100 percent (or zero) to the zero base line. To affect CGs at both ends of the signal system to the same degree when varying the speed, the progression speed in the reverse direction must also be considered (Fig. 4). Figures 3 and 4 are combined in Figure 5 to show the limits within which the band axis can be shifted and rotated to produce maximum bandwidths. For each change in cycle length, this shifting and rotating of the band axis must be repeated.

Optimization

For the largest possible bandwidth to be found, the axis of the band has to fulfill the 2 previously stated conditions. The first is rather easy to accomplish—after choosing cycle length and speed, simply place the CGs according to condition 1, as close as possible to the axis whose slope represents the speed. This in turn yields a set of eccentricities, a set of offsets, and a bandwidth. Whether this bandwidth is the largest possible remains to be seen. Of critical importance, therefore, is the second condition, which states that the largest of all eccentricities has to become as small as possible. The locations of the band axis and CGs have to be optimized with respect to each other. To find the best location of the band axis, we used the Tschebysheff theorem as the optimization model. A detailed outline of the Tschebysheff theorem is given elsewhere (5). This theorem states that the approximation of a cluster of points (in this case, the CGs) by a straight line (the band axis) yields 3 points (α, β, γ) out of the cluster of points located alternately above and below the line whose deviations, Es, from the straight line are equal. The largest of all deviations are as follows:

$$E_\alpha + E_\beta = 0$$

$$E_\beta + E_\gamma = 0 \tag{5}$$

$$|E_\alpha| = |E_\beta| = |E_\gamma| = \max |E| \tag{6}$$

Any other straight line will result in a larger maximum deviation and a failure to meet conditions of Eqs. 5 and 6. Thus the Tschebysheff theorem meets the requirement noted earlier that $\max |E| = \min$. Figure 6 shows an example of a case in which the cluster of points represents the set of CGs and the straight line represents the band axis. E_2, E_4, and E_7 meet conditions of Eqs. 5 and 6.

In order to be able to compare bandwidths of different cycle lengths, Bleyl (3) defined efficiency as follows:

$$\text{Efficiency} = \frac{\text{bandwidth}}{\text{cycle length}} * 100$$

This efficiency will vary as cycle length is varied, and the objective is to find the cycle length that yields the highest efficiency together with the largest bandwidth.

The significance of the offset should be understood. The Traffic Engineering Handbook (4) defines offset as "the number of seconds or percent of the cycle length that the green indication appears at a given traffic control signal after a certain instant used as a time reference base." Offset, like bandwidth, centers of green, and eccentricities,

Figure 3. Band axis with desired progression speed and allowed tolerances from left to right.

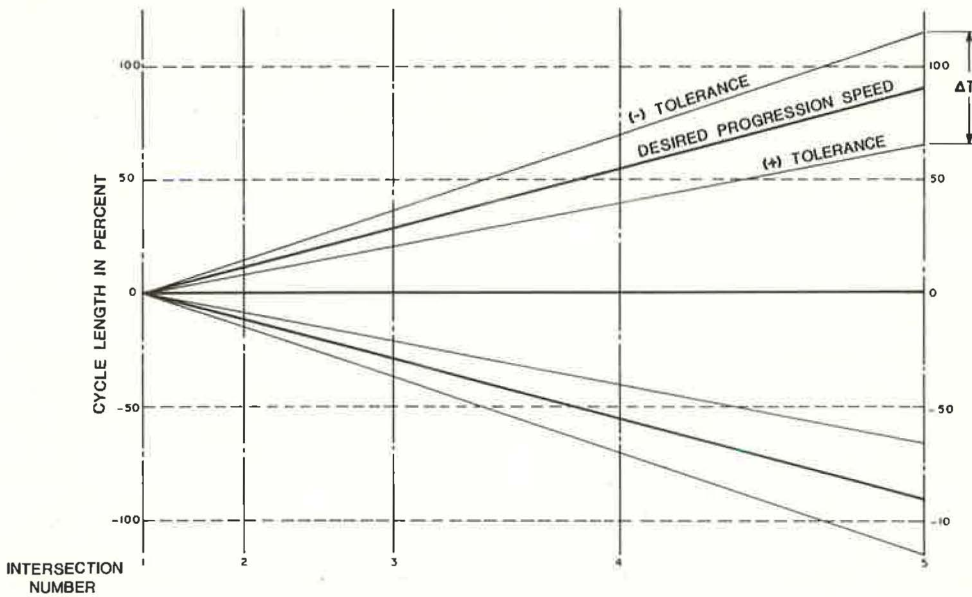


Figure 4. Band axis with desired progression speed and allowed tolerances from right to left.

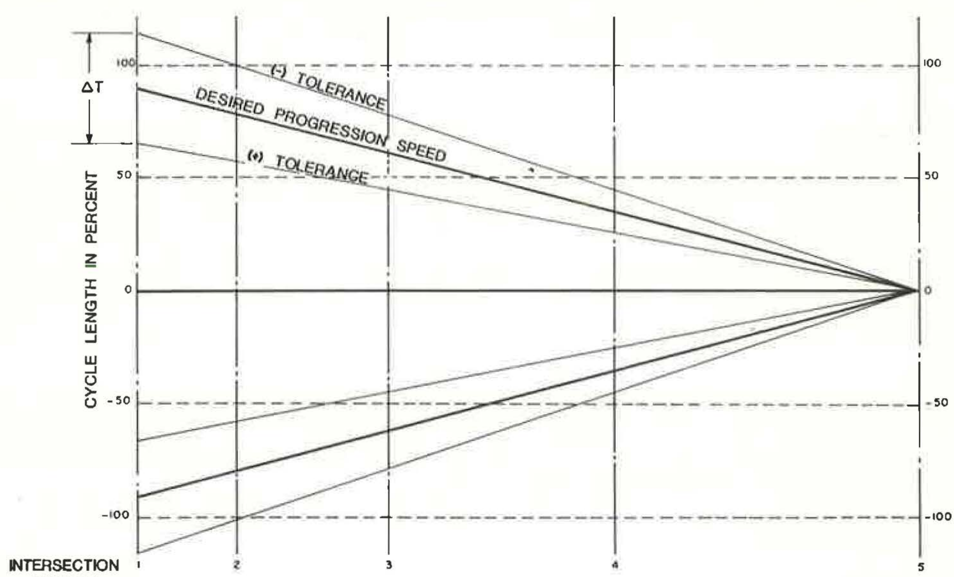


Figure 5. Array in which the band axis can be shifted and rotated within the allowed tolerances.

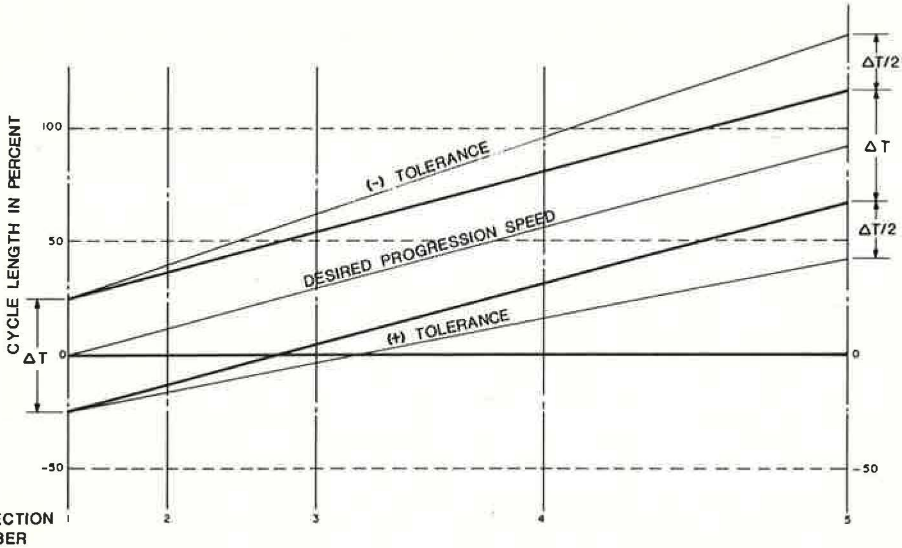


Figure 6. Approximation of a cluster of points by the Tschebysheff approximation.

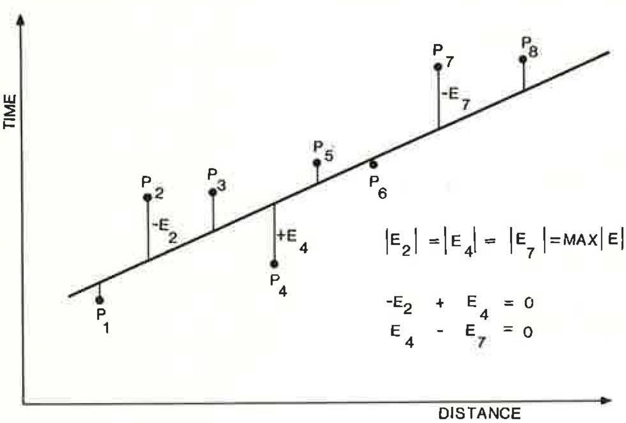


Figure 7. NO-STOP-1 computer printout of time-space diagram data.

```
COMMONWEALTH  AVE  BOSTON  MASS.

AM PEAK

*******************************
*CYCLE = 78.0 SEC             *
*BAND LEFT-RIGHT = 26.4 SEC   *
*BAND RIGHT-LEFT = 31.8 SEC   *
*EFF  LEFT-RIGHT = 33.9 PERCENT*
*EFF  RIGHT-LEFT = 40.8 PERCENT*
*******************************
```

INTERSECTION NAME	NO	DISTANCE IN FEET	SPLIT PERC.	ALL-RED SEC	RED-LR SEC	RED-RL SEC	SPEED-LR MPH	SPEED-RL MPH	OFFSET PERC.	BEG-AMBER PERCENT	BEG-RED PERC.	BEG-AMB PERC.
BABCOCK ST	1	0.0	73.0	1.0	34.0	22.0	33.0	33.0	41	69	73	96
PLEASANT ST	2	780.0	73.0	1.0	22.0	42.2	33.0	33.0	92	69	73	96
ST.PAUL ST	3	1510.0	73.0	1.0	39.0	22.0	33.0	33.0	92	69	73	96
AMORY ST.	4	2060.0	99.9	0.0	16.0	15.0	33.0	33.0	78	95	99	96
B.U.BRIDGE	5	2835.0	53.0	0.0	36.6	36.6	33.0	33.0	53	49	53	96
CARLTON ST	6	3235.0	43.0	0.0	44.4	44.4	27.5	27.5	56	39	43	96
MIDBLOCK	7	3515.0	99.9	0.0	17.0	0.0	27.5	27.5	29	95	99	96
ST.MARY'S ST	8	4045.0	99.9	0.0	17.0	17.0	27.5	27.5	78	95	99	96
CUMMINGTON	9	4485.0	70.0	1.0	24.4	24.4	27.5	27.5	94	66	70	96
GRANBY ST	10	5865.0	68.0	1.0	25.9	25.9	27.5	27.5	46	64	68	96
BLANFORD ST	11	6715.0	73.0	1.0	39.0	22.0	27.5	27.5	43	69	73	96

is a dependent variable subject to change of cycle length and progression speed. Splits and distances are assumed here to be known values. As speed or cycle length or both are varied, a set of centers of green subject to condition 1 is selected that yields a set of eccentricities, a bandwidth, and a set of offsets. If the eccentricities fulfill condition 2, the largest possible bandwidth and the best set of offsets have been found.

DESCRIPTION OF NO-STOP-1

NO-STOP-1 was developed on the fundamentals I have just discussed. A system such as that shown in Figure 5 was developed, and the band axis was shifted and rotated within the allowed tolerances. The best solution for the specific cycle length was found when the band axis (eccentricities) fulfilled condition 2 after condition 1 was observed. The cycle length was varied within specified limits, and the cycle length that yielded the highest efficiency is printed out as shown in Figure 7. Input to the program consists of 15 data cards, which are described in detail elsewhere (6). The data consist of titles, street names, cycle range with increments, speed, distances, splits, all-red clearances, and other data related to speed tolerances, amber time, metric or customary units, multiphase operations, 1-way or 2-way street systems, and bandwidth proportionment. Output consists of a table (Fig. 7) that lists all data related to a time-space diagram. Of special importance for the worker in the field are the last 4 columns, which list offsets, dial settings for begin-amber on the main street, begin-red on the main street, and begin-amber on the side street. (The begin-green setting for the main street always starts at zero percent.) If a plotter is available, the time-space diagram can be plotted directly from the computer.

The following options are available in NO-STOP-1:

1. Balanced system,
2. Unbalanced system,
3. Different speeds from segment to segment,
4. Different directional speeds per segment,
5. Multiphase operations,
6. T-intersection,
7. Midblock operation for pedestrian signals,
8. One-way street system, and
9. Completely nonconcurrent mainline green.

DESCRIPTION OF SIGPROG

The SIGPROG program has been well-defined by Bleyl (3), whose description I will use.

> The approach used by SIGPROG in determining traffic signal system timing plans converts all speed and distance units to travel time units. The diagram is then constructed in terms of time along both axes; the distance axis being replaced by an average-travel-time axis. From a base signal which is the signal having the shortest green interval, the two progressive bands for this interval width are created. The interferences to these bands resulting from both the 0 percent and 50 percent offset of centers of green conditions are determined for each signal. The total interferences to the bands is then selected in such a way that it is a minimum; hence, the bandwidth is a maximum.
>
> The input cards (to the computer program) consist of 12 general control cards and a series of sets of 2, 3 or 4 signal cards. These various cards contain the basic information needed to define the system and its variability.
>
> The printed output consists of three tables. The first table is a listing of the parameters and controls transmitted to the program from the input card deck. The second table contains the results of an incremental cycle scan between the minimum and maximum cycle lengths to find the maximum efficiency obtainable at each increment. The third tabulation indicates the timing elements that yield the greatest efficiency under the specified conditions. If desired, the program

will punch a deck of data processing cards containing all the parameters necessary for a supplemental computer program to plot, draw or print a timespace diagram.

COMPARISON OF NO-STOP-1 AND SIGPROG

Differences

The main difference between SIGPROG and NO-STOP-1 is the fact that NO-STOP-1 varies speed and SIGPROG does not. SIGPROG uses the speed tolerance given as input only as a parameter for further interaction of cycle length. From this follows another major difference in the approach of selecting the centers of green. SIGPROG determines the best combination of centers of green out of a definite number of combinations derived from 1 desired progression speed. NO-STOP-1 determines from each variation in speed (change in location of band axis) a set of centers of green according to condition 1 and a set of eccentricities. If the eccentricities of the band axis fulfill condition 2, the best combination of centers of green yielding the largest possible bandwidth has been found. If they do not, the speed is varied (the band axis is shifted or rotated or both), and new sets of centers of green and eccentricities are determined until condition 2 is fulfilled.

Examples

Two examples are given to illustrate the differences between NO-STOP-1 and SIGPROG. SIGPROG time-space diagrams for Mohawk Street and Genesee Street were prepared during the TOPICS study for Utica, New York. Subsequently, the NO-STOP-1 program was run for each street.

Figures 8 and 9 show the time-space diagrams produced by SIGPROG and NO-STOP-1 respectively for Mohawk Street. Similarly, Figures 10 and 11 show the time-space diagrams for Genesee Street. Dashed vertical lines indicate intersecting streets. Percentage of cycle is on the vertical axis, and distance is on the horizontal axis. Table 1 gives the results of the comparison. NO-STOP-1 showed an increase in efficiency over SIGPROG of 5.3 percent for Mohawk Street and 14.3 percent for Genesee Street. Each method produced a different best cycle length for the same range of cycle lengths. Even when SIGPROG and NO-STOP-1 had the same cycle lengths the efficiencies of NO-STOP-1 were 5.2 and 6.7 percent higher for Mohawk and Genesee streets respectively.

For Mohawk Street, the speed yielded by the NO-STOP-1 program was 1.4 mph (2.3 km/h) or 5.6 percent lower than the originally desired progression speed. For Genesee Street, NO-STOP-1 produced the originally desired speed as the optimum speed. Both methods produced the same speed for Genesee Street; this illustrates the importance of proper selection of the centers of green. The centers of green differed at 6 locations, which explains the large difference in efficiencies in both methods. The centers of green for Mohawk Street were the same for both methods, but the speeds were different. This is why NO-STOP-1 has higher efficiency than SIGPROG has.

Other Differences

There are other differences between the SIGPROG and NO-STOP-1 programs. NO-STOP-1 can handle multiphase operations, completely nonconcurrent mainline green, T-intersections at divided and undivided highways, and midblock pedestrian crossings. SIGPROG cannot handle these things. In the time-space diagram, NO-STOP-1 can plot directional speeds for each section, and SIGPROG can plot only 1 speed.

Figure 8. Time-space diagram for Mohawk Street produced by SIGPROG.

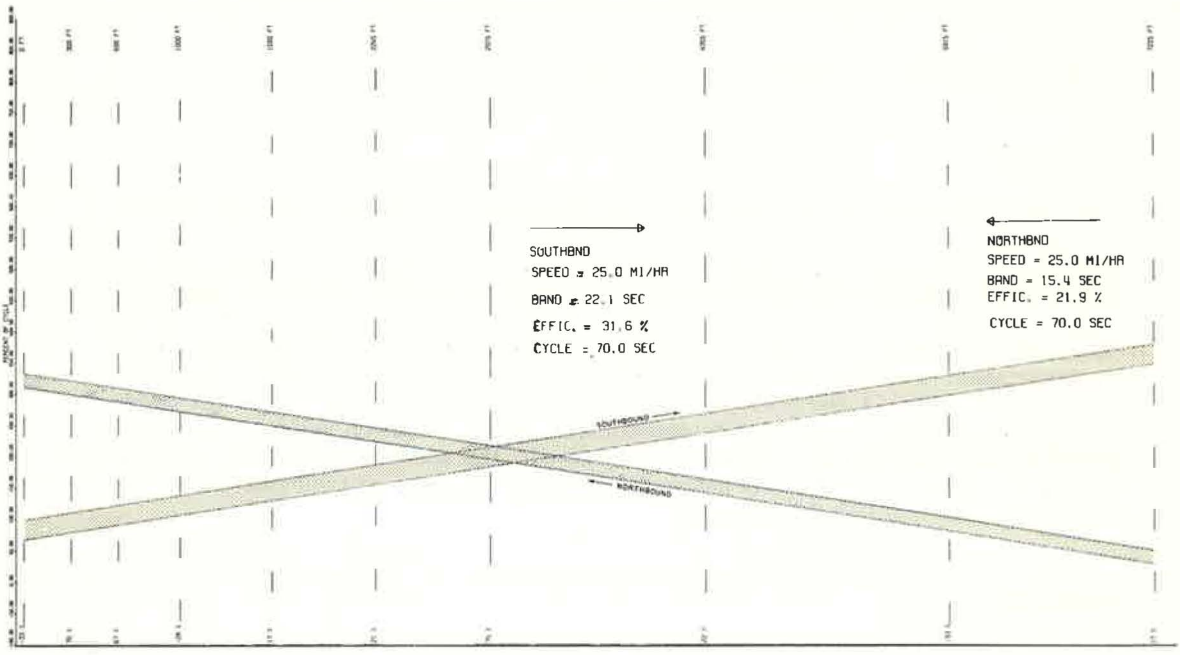
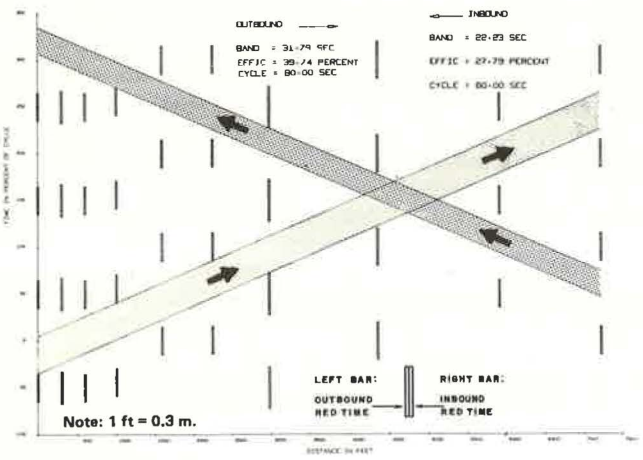


Figure 9. Time-space diagram for Mohawk Street produced by NO-STOP-1.



Note: 1 ft = 0.3 m.

Figure 10. Time-space diagram for Genesee Street produced by SIGPROG.

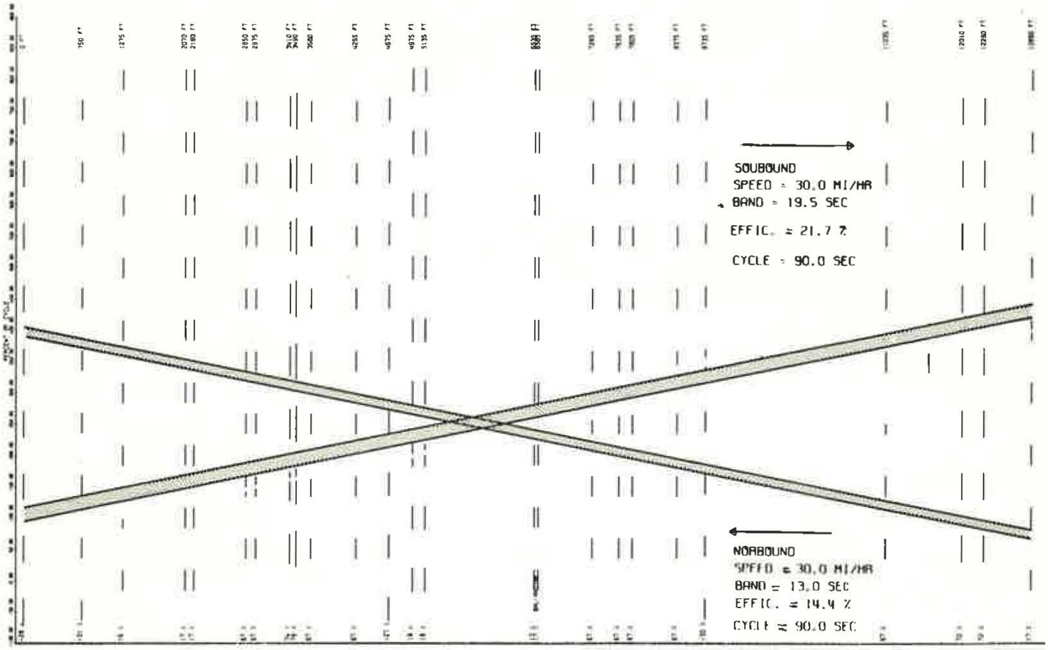
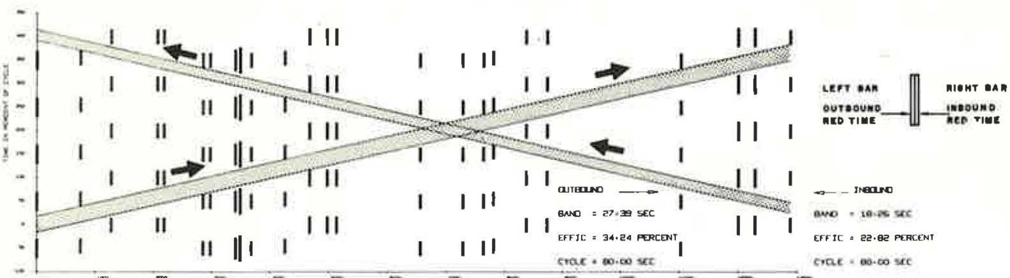


Figure 11. Time-space diagram for Genesee Street produced by NO-STOP-1.



Note: 1 ft = 0.3 m.

Table 1. Comparison of 2 SIGPROG and NO-STOP-1 examples.

Category	SIGPROG	NO STOP-1: Cycle Length Different From SIGPROG	NO STOP-1: Cycle Length Same As SIGPROG
Mohawk Street			
Cycle range, sec	50 to 80	50 to 80	
Chosen cycle length, sec	70	80	70
Speed, mph	25	23.6	26.9
Bandwidths[a], sec	43.5	53.9	47.1
Efficiency (total)[a], percent	62.1	67.4	67.3
Genesee Street			
Cycle range, sec	60 to 100	60 to 100	
Chosen cycle length, sec	90	80	90
Speed, mph	30	30	28.0
Bandwidths[a], sec	38.5	45.7	44.5
Efficiency (total)[a], percent	42.8	57.1	49.5

Note: 1 mile = 1.6 km.

[a]Sum for both directions.

COMPARISON OF NO-STOP-1 TO OTHER METHODS

No practical comparison has been made between NO-STOP-1 and SIGOP (7), the maximal bandwidth program developed by Little, Martin, and Morgan (8), and TRANSYT (9). SIGOP uses the least squares method as the optimization model to determine the best set of offsets. Because the least squares method minimizes the sum of the squares of the differences of the set of offsets to their respective ideal offsets, it cannot at the same time minimize the maximal amount of those differences or meet the requirement that max $|E|$ = min, which can only be accomplished by the Tschebysheff approximation. The program by Little, Martin, and Morgan cannot vary the speed within allowed tolerances whereas SIGPROG can. Furthermore it cannot handle multiphase operations, T-intersections, and completely nonconcurrent mainline green. TRANSYT uses a "hill-climbing" process to optimize the offsets. However, a characteristic of hill-climbing methods is that the optimum they find is not necessarily the best because the offset, which is a dependent variable and a function of the cycle length and progression speed, is used in the optimization process as an independent variable.

CONCLUSIONS

The objectives of this paper were to develop the fundamentals of progressively timed street signal systems, to demonstrate the capabilities of 2 programs—NO-STOP-1 and SIGPROG—to find the best timing plan for progressively timed signal systems, and to compare the results of both programs.

It was shown that the cycle length and the desired progression speed are the parameters that have to be varied within defined limits to achieve the maximum constant bandwidth, and that the Tschebysheff theorem is an appropriate optimization model to achieve that objective. Based on these fundamentals, the NO-STOP-1 program was developed, which included a wide variety of options from multiphase operations to midblock pedestrian crossings. By contrast, the SIGPROG program does not vary the desired progression speed, and has a rather limited variety of options.

The results of both programs were compared, and the NO-STOP-1 program yielded results that were better by up to 15 percent. NO-STOP-1, when compared to SIGOP, the program by Little, Martin, and Morgan, and TRANSYT, also proved superior. Based on these findings, the NO-STOP-1 program is an improved tool for the traffic engineer. Its general versatility recommends itself for widespread use.

ACKNOWLEDGMENT

The author wishes to extend his gratitude to Thomas F. Sergi of Tippetts-Abbett-McCarthy-Stratton whose help in developing the NO-STOP-1 program was essential.

REFERENCES

1. W. Leutzbach. Probleme der Kolonnenfahrt. Strassenbau und Strassenverkehrstechnik, Bonn, West Germany, Heft 44, 1966.
2. R. D. Desrosiers and C. H. Leighty. Traffic Flow Responses to Unannounced Increases in Progression Speed of Signal Systems. Public Roads, Vol. 34, No. 1, April 1966.
3. R. L. Bleyl. A Practical Computer Program for Designing Traffic-Signal-System Timing Plans. Highway Research Record 211, 1967, pp. 19-33.
4. J. E. Baerwald, ed. Traffic Engineering Handbook. Institute of Traffic Engineers, Washington, D.C., 1965.
5. G. Meinardus. Approximation of Functions: Theory and Numerical Methods. Springer-Verlag Inc., New York, 1967.
6. H. R. Leuthardt and T. F. Sergi. Manual for the Application of the NO-STOP-1

Program. Tippetts-Abbett-McCarthy-Stratton, New York, May 1, 1972.
7. SIGOP—Traffic Signal Optimization Program. Peat, Marwick, Livingston and Co., Washington, D.C., 1968.
8. J. D. C. Little, B. V. Martin, and J. T. Morgan. Synchronizing Traffic Signals for Maximal Bandwidth. Highway Research Record 118, 1967, pp. 21-47.
9. D. I. Robertson. TRANSYT: A Traffic Network Study Tool. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, RRL Rept. 253, 1969.

SPONSORSHIP OF THIS RECORD

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

Committee on Traffic Control Devices
J. R. Doughty, Pennsylvania Department of Transportation, chairman
Zoltan A. Nemeth, Ohio State University, secretary
John L. Barker, Robert L. Bleyl, Frank Carroll, Robert E. Conner, Charles E. Dare, John O. Elstad, Roy D. Fonda, Arthur Freed, Albert L. Godfrey, Sr., Alan T. Gonseth, Robert L. Gordon, John Gray, Rudolph J. Israel, Ken F. Kobetsky, Hugh W. McGee, Robert D. McMillen, Joseph A. Mickes, Juri Raus, James I. Taylor, Francis E. Twiss, Earl C. Williams, Jr., Jason C. Yu

Committee on Motorist Information Systems
Gerson J. Alexander, Federal Highway Administration, chairman
Terrence M. Allen, Herbert J. Bauer, Wallace G. Berger, Robert Dewar, A. C. Estep, Eugene Farber, Fred R. Hanscom, John C. Hayward, Robert S. Hostetter, Larry L. Jenney, Gretchen Schabtach Kolsrud, F. G. Lehman, Victor L. Lindberg, Harold Lunenfeld, J. Larry Madsen, Truman Mast, James J. McGrath, Peter B. Moreland, Rudolf G. Mortimer, Robert M. Nicholson, Richard A. Richter, Arthur W. Roberts, III, Bob L. Smith, Myron Michael Zajkowski

James K. Williams, Transportation Research Board staff

Sponsorship is indicated by a footnote on the first page of each report. The organizational units and the chairmen and members are as of December 31, 1974.