Late-Night Traffic Signal Control Strategies for Arterial Systems

John Z. Luh and Kenneth G. Courage

The late-night, low-volume arterial roadway presents a specific signal control problem. The control decision involves a trade-off between the motorists on the artery and those on the cross street. The motorists on the artery are irritated by frequent stops if signals are not coordinated, whereas those on the cross street are annoyed by long waiting time if signals are coordinated. The choice between coordination and free operation is often subjective, especially when semiactuated signals are involved. A method is described for facilitating the choice between coordination and free operation on arterial roadways controlled by semiactuated signals when traffic is light during off-peak hours. The decision is made on the basis of a disutility function that is a combination of the number of stops on the artery and the average cross-street waiting time. A case study was performed to demonstrate the application of this methodology under the closed-loop signal system in the city of Gainesville, Florida. The results indicated that this method provides a promising tool for arterial control with semiactuated signals during late-night hours.

Coordinating the timing of adjacent signals to promote progressive traffic movement has been recognized as one of the most effective means for reducing vehicular stops, delay, fuel consumption, and exhaust emissions. Early efforts on the subject of signal control almost always indicated the need to interconnect signals into a single system and to work toward maximizing progressive movement during peak periods.

However, free (i.e., uncoordinated) operation may be preferable when traffic is light, such as during late night hours (*1*). The reason is evident. Coordination to reduce arterial stops is often accomplished at the expense of longer waiting time of the cross-street traffic. When the arterial traffic decreases during off-peak periods, the benefits of fewer stops gained on the artery may not offset the losses of longer delays on cross-street traffic. This phenomenon has been observed by Riddle and Hazzard (*2*). From a series of simulation and field tests, they concluded that coordination is superior to free operation under all conditions where volumes exceeded 350 vehicles per hour (vph).

This finding suggests that it is preferable for signals in proximity to be coordinated during peak periods. During off-peak periods when traffic is light, however, free operation may be more efficient from the standpoint of system performance. The choice between coordination and free operation is often subjective, because the literature offers little quantitative support for the decision. Therefore, a method was developed to promote a more effective choice on low-volume arterial signal operations.

J. Z. Luh, Langan Engineering Associates, Inc., River Drive Center 2, Elmwood Park, N.J. 07407. K. G. Courage, Transportation Research Center, University of Florida, 512 Weil Hall, Gainesville, Fla. 32611.

Ideally, the decision should be made on the basis of which operation offers better performance. The decision is easier to make when only pretimed signals are involved, because an ample choice of models is available to evaluate their performance. When semiactuated signals are involved, however, the decision is difficult, because the literature offers few models dealing exclusively with them. Semiactuated signals are emphasized because they are the dominant choice in modern arterial signal systems.

A common approach for semiactuated signal evaluation is to use coarsely estimated averages for cycle lengths and green times in a model for pretimed signals. The use of pretimed models to evaluate semiactuated signals should produce reasonable approximations where traffic volumes are moderate to high. In this case, each actuated phase appears in almost every cycle, and the semiactuated signal functionally acts like a pretimed signal except that the phase lengths may vary from cycle to cycle. Heavier traffic demand produces longer times and lighter demand produces shorter times. The use of average phase times to reflect average demands is therefore valid under these conditions.

When traffic is light, however, this method is no longer applicable. In low-volume situations, both phase skipping caused by lack of traffic demand and dwelling on the nonactuated phase waiting for cross-street demand occur frequently. Phase skipping and dwelling make it difficult to determine the average cycle length and green times. Also, the models for pretimed phases are not valid for actuated phases because phase skipping and dwelling are not considered.

Semiactuated signal evaluation may be performed in the field, or by either analytical models or simulation. A number of methods are available for making the field evaluation (*3,4*). But field evaluation, which is costly and time consuming, cannot be applied to nonexistent situations, such as a projected signal control plan.

If available, analytical models can easily determine the actual characteristics of the evaluation. However, the dynamic characteristics of semiactuated signal operation do not lend themselves well to analytical treatment. A limited number of models have been developed for semiactuated signals (*5–9*). However, these models either have limited applications or rely on information that is difficult to obtain, or are restricted to moderate to high volumes. Luh (*10*) provided a new entry to estimate stop probabilities and average delays for semiactuated signals under low-volume conditions. But his method is still complicated to apply when signals have more than three phases or have special control capabilities, such as permissive periods.

In contrast, simulation provides a convenient tool for semiactuated signal evaluation. The dynamic characteristics of semiactuated signal operation can be easier simulated than analyzed. Simulation also makes it possible to study alternative control strategies of a system in compressed time. Moreover, simulation is generally better at comparing alternative control strategies than at optimization (*11*). For these reasons, a simulation technique is used for identifying which operation (coordinated or free) offers better performance.

ASSUMPTIONS

Arrival and Departure Times

As in many delay models, vehicles are regarded as identical in size and performance. They are assumed to arrive and depart the intersection at a constant speed, and delay is regarded as time spent at the stop line. In other words, stopped delays are of concern, and deceleration and acceleration are not considered. Yellow and lost times are not considered.

Vehicle Arrivals

Moreover, vehicles of the actuated phases are assumed to have Poisson arrivals (*12,13*), which means that the probability of n arrivals in a time interval has a Poisson distribution (*13*). (It is unnecessary to assume Poisson arrivals for the nonactuated phase because the nonactuated phase is not controlled by vehicle detections.)

Steady State

It is assumed that the system has operated for a sufficiently long time with the same average traffic volume and vehicle departure rate to have settled into a steady state (*14*).

Signal Operations

Consider a typical semiactuated signal with the following characteristics (*1*):

1. Detectors are located only on actuated-phase approaches.
2. Detectors are placed near the stop line (presence detectors).
3. The nonactuated phase receives at least the minimum green interval each cycle.
4. The actuated phases receive green on actuation, provided that the nonactuated phase has completed its minimum green interval.
5. A signal under coordinated control has a background cycle length, but a signal under free operation does not.
6. The following are preset: the background cycle length for coordinated control, the minimum green time for the nonactuated phase, and initial green intervals for actuated phases.

Actuated phases are assumed to terminate their green times after the initial green intervals elapse, because the probability of requiring green extensions after the completion of the initial green interval is usually small under low volumes. The probability of an actuated phase's having arrivals more than the departure capacity of the initial green interval such that green extensions are required can be estimated as follows:

$$\text{Probability} = 1 - \sum_{n=0}^{\text{capacity}} \frac{u^n e^{-u}}{n!} \qquad (1)$$

where

u = (red time + green time) * volume/3,600,
capacity = (green time)/(saturation headway), and
saturation headway = 2 sec.

It can be easily verified that in most cases the probabilities are small when volumes are less than 300 vph per lane (vphpl).

Traffic Progression

Signal performance at individual intersections is influenced by the quality of traffic progression from their neighbors. When the signal progression is favorable to the subject traffic movement, performance will be considerably better than those for random arrivals. Conversely, when signal progression is unfavorable, performance can be considerably worse than those for random arrivals (*3*). When signals are under free operation, no progression is considered. When signals are under coordinated control, the nonactuated phase (which controls the arterial traffic) is considered under the influence of progression, but actuated phases (which control the cross-street traffic) are not.

METHODOLOGY

The methodology for choosing between coordination and free operation on arterial roadways controlled by semiactuated signals when traffic is light during off-peak hours is based on system performance in terms of disutility. To decide which operation is more efficient for given traffic conditions, the system performance measures both for coordinated and free operation are computed and compared.

Disutility

The common measures of effectiveness for evaluating traffic control system performance are stops, delay, fuel consumption, and other road user cost estimators that may be used in a cost-benefit analysis. However, the problem facing the motorist at night is not excessive accumulated delay. The total number of vehicle-hours of delay that accumulate under late-night, low-volume conditions is hardly worth computing when compared to the peak periods. The decision between the coordinated and isolated modes might have to be made on the basis of differences of a few vehicle-minutes of delay.

On the other hand, there is a substantial perceived disutility associated with late-night, low-volume operations. It is important to minimize this disutility to retain the confidence of the motoring public and to reduce signal violations by irate motorists. So, the first step is to define late-night motorist disutility.

The arterial traffic experiences little disutility in a coordinated system under low volumes provided that the signals are spaced to provide reasonable progression. As the volumes decrease, more green time is assigned to the artery because less is required by the cross streets. In this process, individual waiting times on the cross streets may approach a full cycle because of the need to begin the cross-street green at a specific time to promote arterial progression. The combination of longer cross-street waiting times, combined with the perceived absence of arterial traffic, is the main source of disutility on the cross streets.

Isolated semiactuated operation eliminates this problem, because the right-of-way is assigned to the cross street immediately on the arrival of the cross-street vehicle provided that the minimum arterial green interval has elapsed. Under these conditions, the cross-street disutility is low. On the other hand, the lack of coordination under these conditions may require the arterial traffic to stop at several intersections on the route. It is not difficult to argue that the arterial motorist has little tolerance for frequent stops that are apparently not necessary.

Therefore, there exists a numerical trade-off between the interests of the motorists on the artery and those on the cross street. The arterial disutility is measured by the number of intersections at which a vehicle may be stopped because of lack of progression. The cross-street disutility is expressed in terms of the length of time for which a vehicle must wait for the right-of-way.

It is reasonable to assume that the degree of disutility is not linear with respect to each of these measures. A combination of two successive stops is more than twice as annoying as a single stop. Similarly, doubling the cross-street waiting time more than doubles the frustration felt by the cross-street motorist.

The assumption proposed, more or less arbitrarily, is that the disutility increases parabolically with each of the measures. For example, two stops on the same route produce four times the irritation of one stop; three stops produces nine times as much, etc. The same holds true for cross-street waiting times. The parabolic relationship is incorporated into the study technique through the use of root-mean-square (RMS) values for the disutility measures.

System Performance Function

On the basis of the previously stated reasons, the performance function is defined as a combination of arterial stops and cross-street waiting times, as follows:

$$PI = P_a * K \sum_{j=0}^{N} P_j * j^2 + (1 - P_a)W \qquad (2)$$

where

- PI = system performance index,
- P_a = arterial proportion of total volume,
- P_j = probability of j stops on the artery,
- W = average cross-street waiting time, and
- K = stop penalty that equates one stop to K sec.

The performance function has two weighting factors. The stop penalty (K) represents the waiting time equivalent to one stop. A default value of 20 sec per stop is used. The arterial and cross-street proportions of total volume (P_a and $1 - P_a$) account for the weights of traffic flow rates. A lower performance index indicates reduced motorist operation costs, and thus is more desirable.

Simulation

As stated previously, simulation is used for the purpose of semiactuated signal evaluation. It would naturally be preferable to use an existing simulation program. NETSIM (*15*) could be a reasonable choice for this purpose, because it is the most widely used simulation program in the traffic engineering community. Although NETSIM has established some credibility in simulating traffic control system operations, it does not provide the measures of effectiveness that are appropriate for late-night, low-volume operation. In order to provide these measures, a simple simulation model had to be developed specifically for this study.

The simulation is conducted intersection by intersection. In the simulation process, an arrival table that consists of a sequence of vehicle arrival times for each actuated phase with exponentially distributed interarrival times is first generated. These arrival times represent the times of vehicle arrivals at the stop line.

The mechanics of the semiactuated signal operation are incorporated as a deterministic submodel into the program to determine the start times of each actuated phase. Coordination and free operation are simulated separately using the same arrival table.

For example, for coordinated operation the program first determines whether the first actuated phase has any arrivals up to the yield point. If it does, the actuated phase is given green time right after the yield point, and the program checks if the second actuated phase registers any arrivals up to the end of the green time of the first actuated phase.

On the other hand, if the first actuated phase does not have any arrivals up to the yield point, it is skipped and the program determines whether or not the second actuated phase has any arrivals up to the yield point. If the second actuated phase also has no arrivals, the next actuated phase is checked. After all the actuated phases have been checked, one cycle is completed, and the program starts the next cycle.

System Performance Computations

The system performance index considers both the cross-street waiting times and the arterial multiple-stop probabilities. Their computations are as follows:

1. Cross-Street Waiting Times. The cross-street waiting time for each actuated phase is the time between the arrival of the first vehicle on the red phase and the beginning of the green phase. The waiting time experienced by the following vehicles (if any) of the actuated phase is neglected because it contributes little to the perceived disutility for cross-street traffic.

The average RMS waiting time is calculated as follows:

$$W = \left[\frac{\sum_{j=1}^{O} \sum_{k=1}^{M} d_{ij}^2}{O * M} \right]^{1/2} \quad (3)$$

where

W = RMS value of the cross-street waiting times,
d_{ij} = cross-street waiting time for Phase j in Cycle k,
M = total number of cycles simulated, and
O = total number of actuated phases.

W is then entered into Equation 2 to compute the performance index. The cross-street delay computations are the same both for isolated and coordinated operations.

2. Arterial Multiple-Stop Probabilities for Isolated Operation. For isolated operation, the arterial multiple-stop probabilities are computed as follows:

Step 1: Determine the stop probability for the arterial traffic at each intersection as

$$\mathrm{Ps}_i = \frac{R_i + \mathrm{Gs}_i}{R_i + G_i} \qquad i = 1, \ldots, N \quad (4)$$

where

Ps_i = stop probability for the arterial traffic at Intersection i,
R_i = average red time (sec) at Intersection i,
G_i = average green time (sec) at Intersection i,
Gs_i = $V_iR_i/(s - v_i)$ = average saturation green time (sec) at Intersection i, (5)
s = saturation flow rate (veh/sec),
v_i = flow rate (veh/sec) at Intersection i, and
N = number of intersections.

The average red and green times required in Equations 4 and 5 are from the simulation. Because the arterial traffic is controlled by a nonactuated phase that does not depend on traffic demand, the average red and green times are determined by simply taking the averages of all the red and green times of individual cycles.

Step 2: Calculate the average stop probability (Ps) for the arterial traffic along the route.

$$\mathrm{Ps} = \frac{\sum_{i=1}^{N} \mathrm{Ps}_i}{N} \quad (6)$$

Step 3: Use the binomial distribution to compute the probability of j stops on the artery.

$$P_j = c(N,j)\mathrm{Ps}^j(1 - \mathrm{Ps})^{(N-j)} \qquad j = 0, \ldots, N \quad (7)$$

where P_j = the probability of j stops on the artery and

$$c(N,j) = N!/(N - j)!j! \quad (8)$$

3. Arterial Multiple-Stop Probabilities for Coordinated Operation. For coordinated operation, the arterial stop probabilities depend on the time-space relationship. Because an average time-space diagram can be constructed using given offsets and average red and green times from simulation, the probability of j stops on the artery (P_j) is then determined by vehicular trajectories through the time-space diagram at a time interval of 1 percent of the background cycle length.

DELVACS Program

As shown in Figure 1, this methodology has been implemented as a computer program called "Delay Estimation for Low-Volume Arterial Control Systems" (DELVACS) to facilitate the choice between coordination and free operation. The program is intended to provide a tool for arterial control with semiactuated signals during late-night hours. The program has been designed to be compatible with the input coding scheme of the Arterial Analysis Package (AAP) (*16*). This compatibility serves the dual purpose of expanding the scope of the AAP and simplifying the data entry process.

CASE STUDY

A case study was performed to demonstrate the application of this methodology. Because the decision between coordination and free operation is a traffic-responsive problem, a

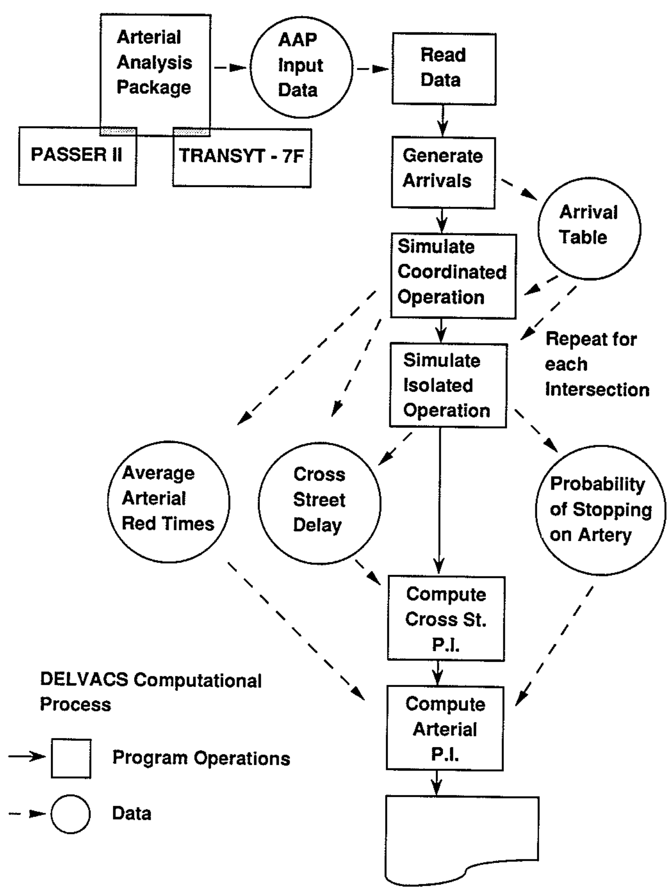

FIGURE 1 Block diagram of the DELVACS program.

traffic-responsive signal control system was chosen for this purpose.

The study was performed in Gainesville, Florida, using a newly installed closed-loop system. A PC-compatible microcomputer monitors a total of 17 on-street master controllers by telephone dial-up communication. Each master controller supervises its own system of local controllers. Most of the systems control a single artery. The system changes the timing plans in the time-of-day and day-of-week (TOD/DOW) modes according to a preset schedule. The timing plans were developed off-line, using TRANSYT–7F when the system was installed. Signals in most subsystems are coordinated during late-night hours.

The arterial system on West University Avenue from 15th Street to 22nd Street was selected for this demonstration. The system contains five signalized intersections, all controlled by semiactuated signals. The artery layout and signal phasing sequences are shown in Figure 2. For convenience, the signals are numbered 1 to 5 from east to west. Among the five intersections, Intersections 1, 4, and 5 are T-type. The signals at these three intersections are single ring with three phases, and others are dual-ring operation. The arterial through movements are controlled by a nonactuated phase, whereas others, including arterial left turns and cross-street traffic, are controlled by actuated phases. For convenience, the term "cross-street traffic" will here include all of the actuated phase movements, including the arterial left turns.

To perform the before-and-after study, a separate program called "Signalized Intersection Monitor" (SIMON) was developed. SIMON was designed to collect local intersection data from the central site personal computer through the on-street master. The status of each detector (active or inactive) and the status of the phase (green, yellow, or red) are recorded second by second. The flow rates and cross-street traffic waiting times can then be derived from this information. Because on-street masters and the central personal computer are connected by telephone lines, only one local controller can be dialed from the central computer at one time. In other words, the five intersections have to be monitored individually.

SIMON was run for three nights, started from staggered times at 10, 10:30, and 11 p.m., both under coordinated and free operations. The five intersections were monitored for 30 min each, sequentially, for a total of 8 hr per night. Permissive left turns and right turns on red were removed from the recorded arrivals, because DELVACS does not deal with them.

The moving-vehicle method was also used to determine the arterial flow rates because there were no system detectors in this subsystem. In each direction, 20 test runs were made over the route under study. The eastbound arterial flow rate was computed from the following equation (*4*):

$$V_e = \frac{3{,}600\ (M_w + O_e - P_e)}{T_e + T_w} \tag{9}$$

where

V_e = eastbound volume per hour,
M_w = opposing traffic count of vehicles met when the test car was traveling east,

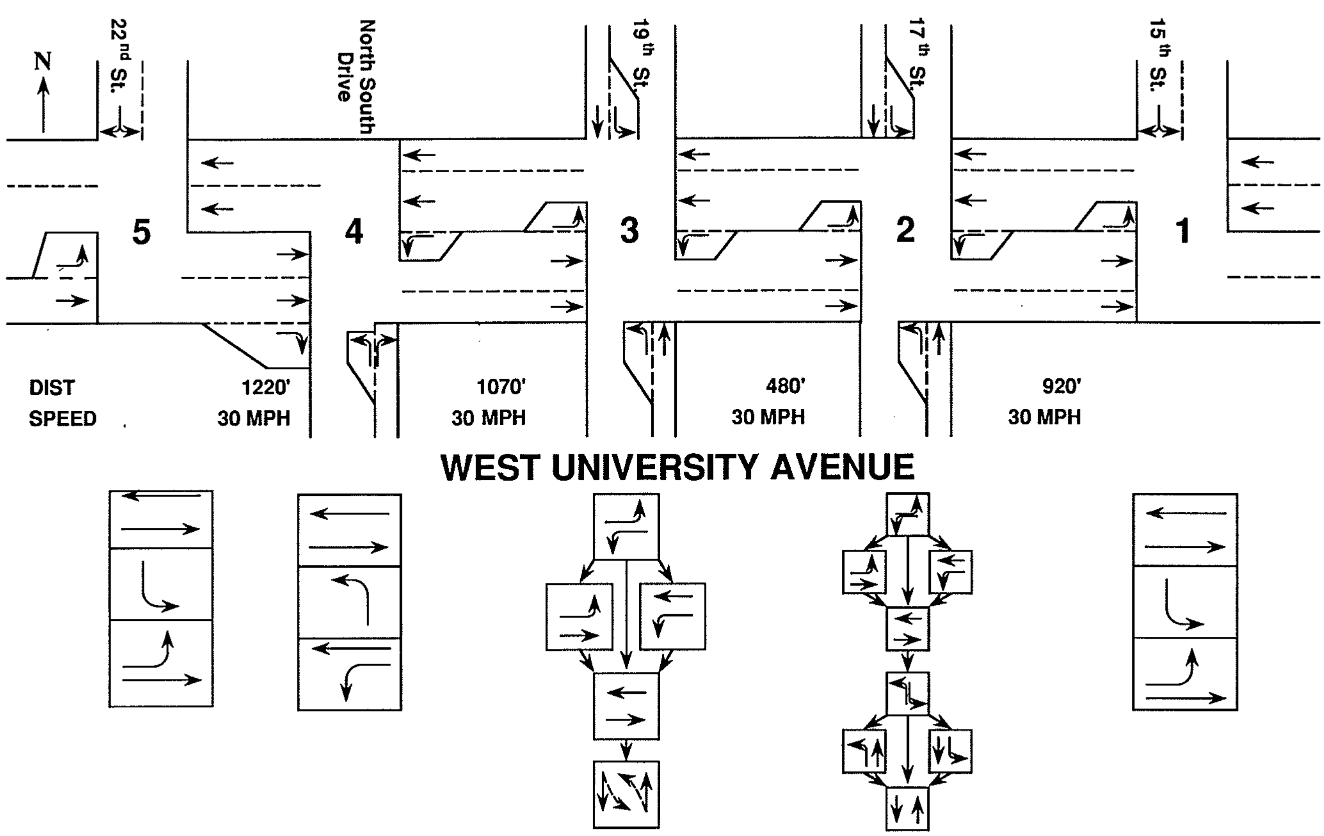

FIGURE 2 Artery layout and phase sequences of the case study.

O_e = number of vehicles overtaking the test car while traveling east,
P_e = number of vehicles passed by the test car while traveling east,
T_e = travel time when traveling east (sec), and
T_w = travel time when traveling west (sec).

For westbound arterial flow rate, the same model applies with all subscripts reversed.

The arterial flow rates from Equation 9 and the average flow rate for each cross-street movement from SIMON are then entered into DELVACS to determine which operation is more efficient. DELVACS uses the same input data structure as the AAP.

The output from DELVACS is shown in Figure 3. In this figure, the multiple-stop probability includes the probabilities of zero stops, one stop, two stops, etc., of the arterial traffic both under free and coordinated controls. Under free operation, 85 percent of the arterial traffic experiences no stop, and 14 percent experiences one stop. In contrast, more than 90 percent of arterial traffic experiences no stop under coordinated control. Following the stop probabilities is the RMS value of the number of stops, which indicates the central tendency of the stop probability distribution. The RMS value of stops under free operation is slightly higher than that for stops under coordinated control.

LOW VOLUME PERFORMANCE COMPARISON

ARTERIAL: WEST UNIVERSITY BY: DELVACS

NO. OF INTERSECTIONS: 5 STOP PENALTY: 20

AVERAGE VOLUMES (VPH): ARTERY: 258 CROSS STREET: 5

NO. OF STOPS	MULTIPLE STOP PROBABILITY ISOLATED	COORDINATED WESTBOUND	COORDINATED EASTBOUND
0	0.850	0.910	0.920
1	0.141	0.090	0.070
2	0.009	0.000	0.010
MORE THAN 2	0.000	0.000	0.000
RMS OF STOPS	0.43	0.30	0.33

MAXIMUM CROSS STREET: WAITING TIME (SEC)	WAITING TIME DISTRIBUTION (%) ISOLATED	COORDINATED
0 - 20	100	32
20 - 40	0	25
40 - 60	0	18
60 - 80	0	25
AVERAGE CROSS STREET: WAITING TIME (RMS IN SECONDS)	11.7	50.5
PERFORMANCE INDEX	378	300

FIGURE 3 Disutilities from DELVACS in the case study.

The waiting time distribution shows the frequency of the maximum waiting times for cross-street traffic in each range. The length of the range is taken as one stop penalty. Under free operation, all of the cross-street traffic experiences a delay less than 20 sec. In contrast, the cross-street waiting times scatter widely from 0 to 80 sec under coordinated control. The RMS value of the waiting times under free operation, as expected, is much less than that under coordinated control.

The performance indices under the two control alternatives indicate that coordination has a lower disutility, and thus is more desirable. This result is attributable to the significant difference between the arterial and cross-street flow rates. As shown in Figure 3, the average arterial volume is 258 vph, whereas that for cross-street traffic is only 5 vph. In other words, although the arterial traffic is low compared to that in daytime periods, it is still much heavier than the cross-street traffic. Therefore, under coordinated control, the benefits of fewer stops gained on the artery can offset the losses of longer waiting times on cross-street traffic. However, it would be reasonable to expect that free operation could be desirable if the arterial traffic is lighter and the cross-street traffic is heavier.

In order to validate that coordination is actually more efficient in the field, multiple stops of the arterial traffic and cross-street waiting times were also collected in the field. Multiple stops were sampled by driving a car on the artery back and forth 60 times, both under coordinated and free operations. Cross-street waiting times were obtained from the data collected from SIMON. The waiting time was calculated as the beginning of the green time minus the arrival time of the first vehicle. The observed arterial multiple stops and cross-street waiting times are shown in Figure 4.

In Figure 4, less than 43 percent of arterial traffic experiences no stop under free operation, in contrast to more than 70 percent under coordinated control. The RMS values of arterial stops in the two directions are more than 1.00 under free operation, whereas they are less than 0.54 under coordinated control. The RMS value of the cross-street waiting times is 16 sec under free operation, but 43 sec under coordinated control. On the basis of this information, coordinated operation shows a smaller performance index.

An inspection both of Figures 3 and 4 suggests that the results in the two figures follow the same pattern. The arterial traffic experiences more stops under free operation and the cross-street traffic suffers longer waiting times under coordinated control. But the significant difference in the two figures regarding the multiple stops deserves a further investigation. Under free operation, for example, DELVACS predicts that 85 percent of the arterial traffic experiences no stops, but the survey shows that only 43 percent in westbound and 37 percent in eastbound experience no stops. The difference could be attributable to pedestrian movements. DELVACS does not consider cross-street pedestrians, but there were some pedestrians at the intersection of 17th Street because of the nearby restaurants and night clubs. Pedestrian activity causes an increase in the effective cross-street volume as well as the cross-street minimum green time.