# DELAY-VOLUME RELATIONS FOR TRAVEL FORECASTING: BASED ON THE 1985 HIGHWAY CAPACITY MANUAL

## ALAN J. HOROWITZ

# DEPARTMENT OF CIVIL ENGINEERING AND MECHANICS UNIVERSITY OF WISCONSIN - MILWAUKEE

# MARCH 1, 1991

Prepared for the Federal Highway Administration U.S. Department of Transportation

#### ABSTRACT

This report discusses the 1985 Highway Capacity Manual in relation to travel forecasting models. It was found that important incompatibilities exist between the HCM and most travel forecasting models; ways of reconciling these incompatibilities are suggested.

This report suggests parameters for speed/volume functions for uncontrolled road segments. For controlled facilities, the reports suggests values for link speed and link capacity to be used prior to network calibration. These speeds and capacities depend upon the type and manner of traffic control.

The report also provides sample specifications for delay relationships that can make a travel forecasting model consistent with the HCM. Separate specifications are provided for signalized intersections, all-way stop controlled intersections, some-way stop controlled intersections, and two-lane roads.

#### TABLE OF CONTENTS

Introduction

Deficiencies in and Problems with the HCM from the Standpoint of Travel Forecasting

Typical Limitations of Travel Forecasting Models

Data Limitations How HCM Violates Model Limitations Minimum Requirements of Forecasting Models to Reasonably Approximate HCM Delay

Procedures

Sample Specifications for Models of Intersection Delay

Traffic Assignment

Available Techniques

A Test of Equilibrium/Incremental Assignment

Advantages and Possible Problems
Delay Functions for Uncontrolled Road Segments
Functions and Standards
Definition of Capacity
Parameter Estimation
Application to Delay-Volume Relations at Signalized Intersections
Calculating Intersection Delay According to HCM Procedures
Results of Signalized Intersection Simulations
Methods of Approximating Capacity
Estimating Delay from Volume and Capacity
Generalized Adaptive Intersections
Nature of a Generalized Intersection
Levels of Adaptation
Two-Lane Roads
Initial Settings for Capacities and Free Speeds
Initial Capacities
Assumptions and Extensions for Initial Capacity
Adjusting Initial Capacity for Old BPR Parameters
Initial Free Speeds
Discussion of Initial Free Speeds
Conclusions
Recommendations
References
Appendix A: Sample Specifications for Intersection Delay
Signalized Intersections
Some-Way Stop Intersections
All-Way Stop Intersections
Appendix B: Best Fit Speed/Volume Functions
Appendix C: Delay/Volume Relationships for Signalized Intersections
Appendix D: Generalized Intersection Data for Two-Way and Four-Way Stops

#### **INTRODUCTION**

The 1985 Highway Capacity Manual provides delay relations for a wide variety of highway facilities. Travel forecasting models also must calculate estimates of delay. Delay is required for determining the shortest paths through networks, the spatial distribution of trips throughout the region, and the relative advantages of one travel mode over another. It has often been suggested that travel forecasting models should incorporate delay relations found in the **HCM**. Potentially, travel forecasts would be more accurate and forecasted volumes would be more consistent with operations-level traffic models and with accepted principles of highway design.

Unfortunately, incorporating **HCM** delay relations into travel forecasting models is not easy. Not only are the **HCM** delay relations too complex for existing software packages, but they also are inconsistent with available theory and algorithms. To properly accommodate the delay relations, both software and theory would require substantial revision.

The purpose of this report is to find ways to make travel forecasts more consistent with the **HCM**. Both preferred and alternative approaches are recommended.

This report identifies properties and requirements of existing travel forecasting models; it then lists deficiencies and problems with the HCM procedures. Full specifications are developed for incorporating HCM-type delay relations into travel forecasting models. These specifications are illustrated by a complete test forecast. Simple delay/volume functions are recommended where possible. Finally, advice is given to planners who must cope with existing software, particularly during the network calibration process.

# DEFICIENCIES IN AND PROBLEMS WITH THE HCM FROM THE STANDPOINT OF TRAVEL FORECASTING

The 1985 Highway Capacity Manual is seriously incompatible with traditional travel forecasting models. The principal reason for this incompatibility is the complexity of many of the delay relations, particularly those relations which compute delay as a function of more than a single link volume or more than a single turning movement.

## **Typical Limitations of Travel Forecasting Models**

There are many travel forecasting packages; their capabilities vary greatly. The most popular packages have the following characteristics, which greatly limit users' ability to determine realistic estimates of delay.

- 1. Delay on a link may be a function of volume only on that link. Models that can calculate delay for a turn do so by looking only at the volume for that single turn.
- 2. The most preferred method of equilibrium traffic assignment, Frank-Wolfe decomposition, cannot handle delay as a function of many link volumes. Furthermore, the delay function must not contain discontinuities, must be monotonically increasing (i.e., strictly increasing with volume), and must be able to be analytically integrated.

- 3. Many models permit only one functional form for delay and only one set of parameters for that function. This one functional form (typically the BPR function) is built into the model and cannot be easily user-modified; however most models permit all the parameters to be varied.
- 4. Many models do not provide the ability to calculate turn penalties as a function of turning volumes.
- 5. Traffic assignment algorithms tentatively estimate volumes greatly exceeding ultimate capacity (LOS E), particularly in early iterations of the calculation. Consequently, delay formulas must be capable of estimating delay for volume-to capacity ratios far beyond 1.0.
- 6. It is very difficult to introduce user judgment during the assignment process. Delay formulas must be entirely self-contained.
- 7. Some models recommend setting "capacity" on a link to the service flow at LOS C, sometimes referred to as the design capacity.
- 8. Depending upon the nature of the path building algorithm, the existence of turn penalties or turning delay functions within a network can greatly increase computation times.

Relative to other parts of travel forecasting models, the calculation of delay is not particularly time consuming. If turn penalties can be avoided, additional complexity in delay relationships should not cause unreasonable increases in computation time.

#### **Data Limitations**

Networks can have thousands of links and intersections, so there are severe limits to the amounts of data that can be economically provided for each. A typical model now requires only two pieces of information about each link for the purposes of delay calculations: capacity and free travel time. It is important not to burden the user with additional data requirements, unless the need has been firmly established through appropriate sensitivity tests of realistic delay relationships.

By their nature forecasts are done for future years; planners do not have very precise information about many of the important traffic characteristics affecting delay. For example, a planner doing a long-range forecast would have little knowledge about .the type of traffic control at any given intersection. The signal timing for signalized intersections would be essentially unknown, and there would be only vague information about the presence of pedestrians, bus operations, and parking maneuvers. Clearly, it would be inappropriate to construct delay relationships requiring data that cannot be obtained.

### How the HCM Violates Model Limitations

The following list of violations does not include assessments of the accuracy of the estimates of delay. It is likely that more realistic and more transferable models of delay can be devised, given sufficient time and resources.

Basic Freeway Sections and Multilane Highways

1. The shapes of the speed/volume functions for basic freeway sections and multilane highways differ by facility type.

#### Two-Lane Roads

- 1. Complete delay relations are not available for two-lane roads. Only a sketchy speed/volume function is presented. This speed/volume function differs significantly from those of other road types or from those of traffic flow theory. Approximate speeds are given for each level of service (HCM Table 8-1). These approximate speeds indicate that a different speed/volume function would be required for each category of percent-no-passing and for each category of terrain.
- 2. The capacity of a two-lane road is a function of the directional split, which complicates the comparison of volume and capacity. A volume-to-capacity ratio could be calculated, but it requires knowledge of traffic volumes in both the subject and opposing directions.
- 3. No mention is made about the applicability of the two-lane road procedures to lower-speed urban facilities, including road segments between traffic controlled intersections. The HCM does not discuss the effects of low-speed passing, turning at driveways, on-street parking, loading, etc. Better estimates of two-lane road capacity may be necessary on suburban arterials, especially where signal spacing is greater than 1 mile.

#### Weaving Sections

1. Delay in a single weaving section is a function of up to four types of movements within the section.

#### All-Way Stop Controlled Intersections

1. The 1985 HCM provides, at most, rough guidelines for the capacity of all-way stop controlled intersections. Delay relations are not presented. More complete all-way stop models have been developed (Richardson, 1987; Kyte, 1989) but have not yet been adopted.

#### Some-Way Stop Controlled Intersections

- 1. The HCM provides procedures for calculating one-way and two-way stop capacity, but does not include delay relationships. Delay relations have been proposed (see Appendix A for an example).
- 2. The relationship between potential capacity and conflicting traffic (Figure 10-3 in the HCM) does not span a sufficiently wide range of traffic conditions. No mathematical form or derivation is provided for this relationship.
- 3. Capacity of any one approach is a function of turning and through volumes on all other approaches.
- 4. No provision is made for traffic distribution across multilane approaches.
- 5. The subprocedure for determining gaps in platooned traffic streams is not well integrated with other parts of the procedure.

#### Signalized Intersections

- 1. The HCM provides conventional guidelines for setting cycle lengths and determining the lengths of green phases, but does not incorporate these principles into its delay procedures.
- 2. The HCM provides only a sketchy discussion about the appropriateness of protected left turns; it does not indicate when a left turn should be protected, nor does it indicate how the protection should be accomplished.
- 3. The HCM does not give a clear indication of how left-turning traffic should be split between protected and permitted phases for all possible cases. The Highway Capacity Software, for example, sometimes asks the user to determine this split.
- 4. No guidance is given on how to allocate right turns to red phases.
- 5. There are discontinuities in the estimates of delay; i.e., small increases in volume can cause abrupt increases or decreases in delay. A major discontinuity is introduced by the subprocedure for determining whether a shared left lane is operating as an exclusive left lane.
- 6. Delay at an approach is affected by the amount of turning at this approach. Furthermore, delay at an approach is affected by the amount of left turns at the opposing approach.
- 7. The delay function can become undefined for volume-to-capacity ratios only slightly greater than 1.0. This is due to the denominator of the d1 term (uniform delay), which can become negative for large values of g/C (ratio of green time to cycle length). This property of the HCM delay function is unlikely to cause problems for practicing traffic engineers, but it can cause computational difficulties in travel forecasting models.
- 8. The time period for oversaturated flow has been set at 15 minutes (Akcelik, 1988); travel forecasting is typically done for a minimum time period of one hour. The HCM does not indicate how the time period may be changed for the purposes of travel forecasting.
- 1. No explicit provision is made for acceleration and deceleration delays. These are included in the 1.3 factor between total and stopped delay. Consequently, acceleration delay is insensitive to the speed of traffic.
- 10. Under some circumstances, the procedure gives separate delays for the left, through, and right moments. Under other circumstances, it does not.
- 11. No mention is made of delay at freeway ramp meters.

#### General Issues

A more general problem concerns the definition of LOS C, often taken as the definition of "design capacity" in forecasting models. LOS C is largely subjective and is determined by different methods, depending upon the type of facility or type of traffic control. Thus, there no longer exists a simple method of relating LOS C to LOS E (ultimate capacity) that works across the full range of facilities or traffic controls.

For example, LOS C on freeways is determined by traffic density, while LOS on two lane roads is determined by percent time delay. The volume-to-capacity ratio for LOS C varies between 0.77

(freeway basic segment, 70 mph design speed) to 0.16 (two-lane road, mountainous terrain, 1 00% no passing).

# Minimum Requirements of Forecasting Models to Reasonably Approximate HCM Delay Procedures

As indicated in the preceding paragraphs serious incompatibilities exist between the HCM and existing travel forecasting models. The incompatibilities can be fully resolved only by extensive revisions to the forecasting models. The amount of effort necessary to make these revisions depends upon the structure of the existing computer code.

- 1. The model must be capable of calculating intersection delay for each approach separately from delay on the link that includes the approach. For some models, this delay could easily be expressed as a turn penalty, but there would probably be a significant increase in computation time. A better but more complicated solution is to add the intersection delay, once calculated, to the delay for the approach link.
- 2. At traffic-controlled intersections and at weaving sections, delay must be calculated considering all the movements. For example, delay for an approach at a four-way signalized intersection is related to all 12 possible movements at the intersection.
- 3. Delay on two-lane roads must be calculated from both subject and opposing volumes.
- 4. Different delay functions must be available for freeways at various design speeds, multilane highways at various design speeds, two-lane roads, and urban streets. If a sufficiently general functional form is available (for example, see Spiess, 1990), the differences between facility types could be accommodated with alternate sets of parameters.
- 5. A method other than Frank-Wolfe decomposition must be available for calculating equilibrium traffic assignment.

### Sample Specifications for Models of Intersection Delay

In order to better understand the implications of the HCM delay procedures for travel forecasting, a set of sample specifications was developed. Separate specifications were written and programmed for delay at signalized intersections, all-way stop intersections, and some-way stop intersections. These specifications were directly incorporated into a travel forecasting model. An attempt was made to stay as close as possible to HCM procedures while providing routines that could successfully be interfaced with the travel forecasting model. Parts of HCM procedures that appeared to have little effect on delay were abridged. Otherwise, the specifications follow the HCM quite closely.

The specifications are used later in this report (1) to develop delay/volume relationships for forecasting models that cannot be modified, (2) to demonstrate the feasibility of directly incorporating HCM procedures into a travel forecasting model, and (3) to suggest values for link capacity and free speed to be used prior to network calibration.

The sample specifications are fully described in Appendix A.

#### TRAFFIC ASSIGNMENT

#### **Available Techniques**

The HCM delay relationships are discontinuous, nonmonotonic, and nonintegratable. The only method of equilibrium traffic assignment known to be able to handle similarly difficult delay relationships is most often referred to as "one-over-kay" assignment or "equilibrium/incremental" assignment or "method of successive averages". The method finds an unweighted average of many all-or-nothing assignments, where the delay found prior to any iteration (k+l) is calculated from the average of volumes from the preceding (k) assignments. Equilibrium/incremental assignment produces identical results to Frank-Wolfe decomposition (LeBlanc, et. al, 1975) on networks with simple (such as the BPR) delay relationships (Powell and Sheffi, 1982; Horowitz, 1990); however, convergence is slightly slower.

This algorithm has not yet been extensively tested on networks where delay can be a function of several volumes.

#### A Test of Equilibrium/Incremental Assignment

The UTOWN network, originally created for testing UTPS, was modified by incorporating signalized intersection and two-way stops, primarily at freeway off-ramps. The modified UTOWN network is shown in Figure 1.



Figure 1: UTOWN Network with Traffic Control

Convergence to an equilibrium solution needs to be checked, but the standard methods derived from Frank-Wolfe decomposition will not work in this case. We are looking for a *user-optimal* assignment. In such an assignment each trip is assigned to a shortest path between its origin and destination. Therefore, it is possible to determine when equilibrium has been achieved by checking whether the used paths are indeed the shortest paths. A simple test can be devised that compares total travel time between two assignments.

**Step 1.** Run the assignment algorithm through the desired number of iterations. Obtain estimates of volumes. Recalculate the link travel times. Compute total travel time with the estimates of link volumes and the new travel times.

**Step 2.** Using the new travel times and averaged trip table from Step 1, perform an all-or-nothing assignment. Do not recalculate link travel times. Compute total travel time.

**Step 3.** Compare the total travel times from Steps 1 and 2. The total travel time from Step 2 will always be the smallest. If they are nearly the same, convergence to an equilibrium solution has been achieved.

If they differ significantly, there could be two causes: (1) more iterations are required; or (2) the algorithm failed.

This test is similar to one ("S1 - S2") found in UTPS.

The test was performed on the UTOWN network (containing HCM delay relationships) for varying numbers of iterations of equilibrium/incremental assignment. As seen in Table 1, the equilibrium/incremental assignment algorithm will produce an equilibrium solution on a network with traffic controls. After 200 iterations the difference between Steps 1 and 2 was inconsequential. Equilibrium was effectively achieved after about 20 iterations. This rate of convergence is similar to Frank-Wolfe decomposition.

A significant body of research is being assembled on "asymmetric" traffic assignment problems, which include assignments where delay is a function of several link volumes. It is likely that even faster (and perhaps surer) algorithms will be developed within the next few years.

#### Table 1. Convergence of Equilibrium/incremental Assignment on the UTOWN Test Network



#### **Total Travel Time**

An inspection of the assigned volumes revealed that similar results would have been difficult to obtain with conventional delay/volume relationships. The assigned volumes on approximately half of the links in the original UTOWN network (without traffic controls) were considerably different from those of the modified UTOWN network (Figure 1). For example, the volumes for one particular freeway link differed by a factor of more than two. The other half of the links had surprisingly similar volumes across the two networks. One striking difference between the two assignments was the *higher* arterial volumes on congested links in the modified network. The algorithm gave these links more green time, thus more capacity. The original network, of course, had to provide equal signalization priority to each approach, regardless of need.

The UTOWN network is artificial and exaggerates problems with assignment algorithms. Still, it adequately demonstrates the importance of having precise estimates of intersection capacity.

### **Advantages and Possible Problems**

A traffic assignment involving complex intersection delay relationships, such as those in the HCM, is *adaptive* in the same sense as an actuated signal, which can adjust itself to the existing traffic volumes. The algorithm allocates capacity to an approach according to its volume and competing volumes. Approaches with relatively large volumes receive more green time, and thus capacity, than approaches with small volumes. Theoretically, the maximum capacity of an approach is its saturation flow rate, less any possible flow lost during phase changes. In practice, however, a small amount of green time must be given to conflicting approaches, even when there is very little traffic.

Such an assignment is quite realistic, but there is one unfortunate side effect - the solution may not be unique. It is entirely possible for an adaptive traffic assignment to have two or more equally valid

equilibrium solutions. Under such circumstances, one cannot judge which solution is the correct one. Indeed, all solutions may be correct. Differences would be due to small variations in signalization - something that is impossible to predict.

#### DELAY FUNCTIONS FOR UNCONTROLLED ROAD SEGMENTS

#### **Functions and Standards**

The most widely used delay function for both controlled and uncontrolled road segments is the BPR

function:

where X is the volume-to-capacity ratio, to is the free travel time, and a and P are empirical coefficients. Many practitioners recommend that capacity be taken as the design volume for the link, normally LOS C. Other practitioners recommend computing X with ultimate capacity. When X is calculated with ultimate capacity, it is possible to approximate a from the free speed, so, and the speed at capacity, Sc.

#### <u>That is,</u>

thereby effectively reducing this function to one with a single parameter, P.

Spiess (1990) has identified seven standards for speed volume functions:

- 1. The function should be strictly increasing with volume; i.e., it is monotone.
- 2. The function should yield the free travel time for zero volumes and twice the free travel time for volumes at capacity.
- 3. The derivative of the function should exist and be strictly increasing; i.e., the original function is convex.
- 4. The function should have only a few and well defined parameters.
- 5. The function should be finite for all volumes.
- 6. The function should have a positive derivative at zero volume.
- 7. The evaluation of the function should require less computation time than the BPR function.

If these standards are met, then it is assured that an equilibrium can be found with Frank-Wolfe decomposition, that the model is easily calibrated, and that the computational effort will be modest. The BPR function meets the first six standards.

Standard 2 assumes that speed at capacity is always one-half of free speed. Unfortunately, Spiess ignored the rest of the speed/volume function, so standard 2 should be revised to read:

2. The function should provide realistic values of delay across the range of volumes from zero to capacity, especially at zero volume and at capacity.

The revised second standard is required to retain realistic assignments and to provide good path travel times for the trip distribution and mode split steps. Spiess' third and seventh standards are unnecessary and would be inhibiting, if accuracy is of paramount importance.

 $t = t_0 \{2 + [\alpha^2 (1 - X)^2 + \beta^2]^{1/2} - \alpha (1 - X) - \beta\}$ where  $\beta = (2\alpha - 1)/(2\alpha - 2)$ 

Spiess proposed an alternative to the BPR function,

which may fit the various HCM delay/volume relationships more closely:

and X is the volume-to-capacity ratio. This function always yields a travel time at capacity of twice the free travel time - something which may not always be desirable. This function has the general shape of a hyperbola, and is referred to by Spiess as a conical delay function. It is very similar to a delay function developed by the Traffic Research Corporation in 1966 (Branston, 1976).

Still another alternative function with a single parameter has the form:

Like the BPR function, Equation 5 is assured to exactly fit the delay/volume curve at zero volume and capacity. This equation was proposed by Overgaard (1967). It meets Spiess' first six standards.

## **Definition of Capacity**

Networks originally prepared for Planpac and UTPS largely relied on the default coefficients of the BPR function ( $\alpha$ =0.15 and  $\beta$ =4.0). With these coefficients, link capacity was set to design capacity, normally taken to be LOS C in earlier editions of the Highway Capacity Manual. More recent travel forecasting packages have generally retained these traditional coefficients and definition of link capacity. Technically, design capacity should be interpreted as the volume that causes free speed to drop by 15 percent. There are valid reasons for trying to retain this definition of capacity in previously calibrated networks.

Unfortunately, the 1985 Highway Capacity Manual does not provide a similarly simplistic relationship between service flow at LOS C and speed. In order to continue using the "design capacity" definition of link capacity, it would be necessary to establish a set of procedures to (1) find it and (2) assure that it yielded reasonable estimates of speed (or delay) at all feasible volumes.

It is possible to develop new parameters for the BPR curve (or another speed/volume function) using any reasonably consistent definition of capacity. There would be little difference in the quality of fits to speed and volume data. Consequently, the choice of a definition for capacity must be made on the grounds of convenience. There are four important arguments for defining link capacity to be ultimate capacity (LOS E for most facilities).

- 1. Ultimate capacity has a consistent meaning across all facility types, while design capacity does not. For example, it is a relatively simple matter to relate the capacity of an intersection to the capacity of the street approaching that intersection.
- 2. Ultimate capacity is always easier to compute than design capacity. Finding the design capacity of a signalized intersection is especially difficult.
- 3. Ultimate capacity can be more easily related to traffic counts than design capacity, which would also require estimates of density, percent time delay, reserve capacity or stopped delay.
- 4. Ultimate capacity is the maximum volume that should be assigned to a link by the forecasting model. Design capacity does not give such firm guidance during calibration and forecasting.

#### **Parameter Estimation**

All three delay functions (Spiess', BPR, Overgaard's) were fit to the speed/volume relationships contained in the Highway Capacity Software, Version 1.5, which closely approximate those in the HCM. The coefficient,  $\alpha$ , in the BPR function was determined by forcing the curve to fit the speed/volume data at zero volumes (free speed) and at capacity (LOS E). The second coefficient,  $\beta$ , was found by nonlinear regression. The single coefficients of Spiess' function and of Overgaard's function were also found by nonlinear regression. Table 2 summarizes the best coefficients.

It is seen that all three functions performed well, as judged by the standard deviation of the residuals,  $\sigma v$ , and the percent of variance explained, R<sup>2</sup>. The quality of the fit varied with the facility type and design speed. In general, it was easier to fit speed/volume functions when the design speed was 50 miles per hour. Spiess' function produced the most consistent results, explaining about 97% of the variance for all six facilities. It is likely that Spiess' function would yield even better results if the assumption about speed at capacity (Spiess' original standard 2) could be improved. Appendix B shows the HCM speed/volume functions for each facility and the best

		Sun Speed/Vo	nmary o lume F 4-Lane	of Leas Junctio Rural	able 2 it-Squa ns for 6 Divideo	Fits Lane F Highw	to HCM Freeways ays	and		
			BPI	REUN	CTION					
FAC	ILITY		a		ß	0	~	<b>B</b> <sup>2</sup>		-
	70 n	nph	0.88		9.8	1.5	90	91.8%		31
Freeways	60 n	nph	0.83		5.5	1.5	93	91.2%		31
	50 n	nph	0.56		3.6	0.7	70	98.4%		29
	70 n	nph	1.00		5.4	2.1	78	87.3%		21
Multilane	60 n	nph	0.83		2.7	1.5	50	95.8%		21
	50 n	nph	0.71		2.1	0.7	77	98.3%		19
				DIESS	ELINIC	TION				
	FACI		0	a	. 0.40	<i>a</i> .	82			
	1 40.	70 mph		9.8		90	97.9%	22	31	
Free	wavs	60 mph		8.5	1	21	96 5%	22	31	
		50 mph		7.5	0	.71	97.5%	10	29	
		70 mph		7.1	1	.34	97.0%	10	21	
Mult	ilane	60 mph		4.0	1	.17	97.5%		21	
		50 mph		4.0	0	.98	97.3%		19	
			OVER	GAAR	D'S FU	NCTION	J			
	FACI	LITY		a		0.	R2		n	
		70 mph		9.0	1	.99	92.9%		31	
Freeways		60 mph		4.5	1	.68	93.3%		31	
		50 mph		3.3	0	64	98.7%		29	
		70 mph		4.3	2	.42	90.4%		21	
Mult	lane	60 mph		2.3	1	.20	97.4%		21	

fitting functions.

The HCM provides three slightly different speed/volume curves for freeways with 70 mph design speeds - one each for 4-lane, 6-lane, and 8-lane segments. The curves for 4-lane and 8-lane segments differ from the one for 6-lane segments (used here) by at most 1 mile per hour.

Consequently, there is little advantage to having three separate speed/volume functions for 70 mph segments.

## **Application to Delay/Volume Relations at Signalized Intersections**

It is possible to estimate delay at traffic controlled intersections with any of the three curves discussed in the previous section. Instead of fitting a speed/volume relationship, it is necessary to fit a travel-time/volume relationship, where travel-time is taken from the HCM signalized intersection delay formula. Examples of some nonlinear least-squares fits to HCM's delay formula are seen in Figure 2. The HCM delays are for an intersection with a 90 second cycle length, a 60 second green time, and a saturation flow rate of 5400 vph. It is seen that the BPR and Overgaard's functions can reasonably approximate the HCM formula, but Spiess' formula performs badly. (The BPR function parameters were a = 5.0 and P = 3.5.)

Although it is possible to fit a BPR curve to the HCM delay function, doing so would be undesirable for the following reasons:

- 1. A different set of parameters would be required for every combination of cycle length, green time, saturation flow rate, and arrival type.
- 2. The BPR curve differs substantially from the HCM delay function for oversaturated conditions; i.e., when the volume-to-capacity ratio exceeds 1.0.
- 3. Network coding would be more difficult, because an additional link would be required for each approach.
- 4. Acceleration delays are ignored.

A better approach, but one that requires considerable rewriting of software, is to calculate intersection delay directly from the HCM procedures, as described in previous sections and in Appendix A.



Figure 2: Least Square Fits to the HCM Delay/Volume Function

## CALCULATING INTERSECTION DELAY ACCORDING TO HCM PROCEDURES

## **Results of Signalized Intersection Simulations**

The signalized intersection delay specification, described in Appendix A, was implemented in a travel forecasting model (a specially modified version of QRS II) and tested. An attempt was made to extract the implied delay/volume relationship while letting the model determine the phasing and green times. Since green times are no longer exogenous variables, the possibility exists for a simpler means of calculating delay.

Figure 3 shows three delay/volume curves for the same intersection. The curves show the delay on all approaches (subject, opposing, and conflicting) when the volume on just one subject approach is varied. This intersection has a high percentage of turns (25% lefts and 25% rights at all approaches). It is readily seen that the delay on any approach depends on the volumes for the other approaches. For instance, the delay for both the subject and conflicting approaches are nearly the same, even though the conflicting volume was held fixed at 800 vph. The delay on the opposing approach is more complex - first rising gradually, peaking at 2400 vph on the subject approach, and then declining. The reason for the declining delay is the increasingly ample green time available to handle the 800 vph on the opposing approach.

# Figure 3: Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach

(25% Right Turns, 25% Left Turns, 800 VPH at Opposing and Conflicting Approaches, No Exclusive Lanes, 3600 VPH Ideal Saturation Flow Rate, 20 mph



Figure 4 is similar to Figure 3, except that there are no turning vehicles. The subject and conflicting delay curves have similar shapes, but do not coincide. It is again seen that the delay on the opposing approach declines, in this case after 800 vph on the subject approach. Figure 4 also shows that the delay on the subject approach is not necessarily monotonic (i.e., steadily increasing with volume). The delay rises to a local maximum at 800 vph (the fixed volume on the conflicting again.

The delay curves of Figures 3 and 4 are very consistent with the theory and procedures of Chapter 9 of the Highway Capacity Manual. Consequently, it can be concluded that the results are realistic. However, these results could cause difficulties for traditional travel forecasting models. Delay cannot be a declining function of volume without introducing the possibility of multiple, equally valid, equilibrium solutions. Whether multiple equilibria could occur in real, full-scale networks has not yet been established.

# Figure 4: Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach





The signalized intersection delay specification was extensively exercised, varying the percentage of turns, the cycle length, the approach type, the presence or absence of exclusive lanes, and the levels of opposing and conflicting volumes. A selection of these delay/volume curves are shown in Appendix C. A review of these curves indicate that no simple relationship, such as the BPR formula, can accurately estimate intersection delay.

## **Methods of Approximating Capacity**

*Flow Ratio Method.* The best that can be offered for models dependent on the BPR formula is a weak approximation to these simulation results. Assumptions must be made about the amount of traffic at all approaches, the cycle length, the number of phases, and the saturation flow rate of all



approaches, including the effects of turns. A capacity, c, for the approach is approximately,

A practical use of Equations 6 and 7 would require capacities to be computed after volumes have been assigned to the network, rather than given as data.

*Equal Greens Method.* In the absence of information about opposing and conflicting volumes, it would be necessary to assume that the flow ratios are identical at all approaches. Under such a situation the green times would be approximately equal on all approaches. Equations 6 and 7 reduce to a single equation,

$$c = Ss(1/2)(C - L)/C$$

Equation 8 is similar to methods currently used by planners prior to network calibration. Because Equation 8 ignores signal timing, it is not a desirable method for estimating capacity.

*Graphical Method.* A related method of calculating the capacity of an approach is to use the information such as that contained in Appendix C and in Figures 3 and 4. The first parameter of the BPR formula would be set so that delay at capacity is exactly twice delay at zero volume ( $\alpha = 1.0$ ). As seen previously, this setting for a is approximately correct for most uncontrolled road segments. The capacity would then be defined at the volume on the subject approach that exactly doubles delay. This capacity can be directly read from one of the graphs, or interpolated from two or more graphs.

For example, in Figure 3 the delay for the subject approach at zero volume is 18 seconds. "Capacity" would therefore be slightly less than 1200 vph (Figure 3 shows the delay at 1200 vph to be about 38 seconds). In Figure 4, "capacity" is seen to be slightly more than 2400 vph. This result can be compared with Equation 21, assuming Vs 2400 and L = 6,

c = 3600 [0.667/(O.667 + 0.222)] (90 - 6)/90 = 2524

The results of these methods appear to be reasonably consistent. The graphical method could best be viewed as an aid to hand calibration of networks.

*Drawbacks.* All three methods are clumsy. They require prior assumptions about volumes and require a considerable amount of user intervention, especially for the calculation of saturation flow rates. Furthermore, the three methods deviate to varying extents from the HCM.

## **Estimating Delay from Volume and Capacity**

Once capacity has been calculated, it is possible to estimate delay from the BPR or a related function. Figure 5 shows the best fits of the BPR, Spiess' and Overgaard's functions to the subject approach delay from Figure 4 (Ss = 3600, 0% turns). As described in the last section, capacity was taken to be the volume that doubles delay. Therefore, the value of a was set to 1.0 in the BPR function; no changes were required of Spiess' function. It is seen that the BPR and Spiess' functions fit well; the Overgaard function misses badly at volumes exceeding capacity. The best fit of the BPR curve was obtained with = 5.3; the best fit of Spiess' curve was obtained with a = 7.4.



Figure 5: Least-Squares Fit to Signalized Intersection Specification

## **GENERALIZED ADAPTIVE INTERSECTIONS**

## Nature of a Generalized Intersection

An adaptive intersection is one in which the capacity of all approaches can be adjusted to provide better or fairer traffic flow. In reality, all signalized intersections are somewhat adaptive, because signal timing can at least be manually adjusted to better serve existing volumes. At very low volumes, a signalized intersection would impose greater delays than a stop-controlled intersection or an uncontrolled intersection. Therefore, if the assignment is completely adaptive, it also should be able to change the nature of the traffic control (such as adding or removing signals and signs, changing to four-way flash, etc.) Such a highly adaptive assignment algorithm would design the traffic controls as it loads traffic to the network. Although it would be significantly slower, this type of algorithm would not be particularly difficult to accomplish. The computer code written for the tests in the above paragraphs could be easily so modified. The question of whether a highly adaptive assignment is desirable cannot yet be completely answered.

*Estimating the Effects of Adaptation.* Planners, however, may choose to modify the nature of the traffic control after they see the assigned volumes - in essence adapting their networks. To do this properly, they would need information about delays at stop-controlled intersections. Figure 6 shows the relationship between volume and delay at a two-way stop-controlled intersection, a fourway stop-controlled intersection, and a signalized intersection. The lane geometry and volumes were the same in all three cases. In this figure, the subject and opposing volumes were varied together, while the conflicting volumes were held constant. The delays at each approach are shown in Appendix D.

Figure 6 shows that the three types of traffic control perform almost equally well at a volume of 400 vph on the subject and opposing approaches. Below 400 vph the two-way stop is superior; above 400 vph the signal is superior. Other tests show that the point at which all controls are equally effective varies with the amount of conflicting volume. This point is at about 100 vph when the conflicting volume is a 600 vph; it is at about 200 vph when the conflicting volume is 400 vph. In no circumstances did the four-way stop outperform the combination of the signal and the two-way stop, suggesting that the four-way stop need not be considered any further. Rules, similar to the signal warrants in the Manual on Uniform Traffic Control Devices, could be used to select the type of traffic control.

In a highly adaptive network, low volumes on one or more approaches might indicate a need for a two-way stop. The effect on the delay/volume curve depends upon whether the subject approach is signed or unsigned. At very low volumes, a vehicle at a signed approach experiences a delay consisting of about 2 to 4 seconds plus any time lost to acceleration (typically 4 to 7 seconds; see Equation A.1 in Appendix A). Vehicles at unsigned approaches experience almost no delay.

The concept of a generalized intersection implies that the delay values in Appendix C for signalized intersections are excessively large for very low volumes on the subject approach. Planners need to be aware of this possibility while calibrating their networks and performing forecasts.

#### Figure 6: Total Delay on All Approaches for a Four-Way Stop, a Two-Way Stop and a Signal (Opposing Volume Same as Subject Volume, Conflicting Volumes at 200 vph, 25% Right Turns, 25% Left Turns, One Lane at All Approaches, 20 MPH



## **Levels of Adaptation**

Planners need to seriously consider the appropriate amount of adaptation for their networks. Even if their assignment algorithm is not formally adaptive, planners indirectly introduce adaptation as they calibrate their networks or choose their assignment algorithms. Although the Highway Capacity Manual does not discuss adaptive assignment, it does indicate how adaptation can occur. The following levels of adaptation could be invoked, to various degrees, for any given network.

Level 0. No adaptation. Capacity is rigidly fixed on all streets and intersection approaches.

Level 1. Low cost traffic engineering improvements for isolated intersections without changing the *type of traffic control.* Capacity varies with the amount and nature of conflicting and opposing traffic. (Examples: signal timing; conversion of a through lane to an exclusive lane.)

*Level 2. Major traffic engineering improvements for isolated intersections.* Capacity varies with the amount of and nature of conflicting, opposing, and subject approach traffic. (Examples: installation of signals, rearrangement of signs, relocation of bus stops.)

*Level 3. Traffic engineering improvements involving a system of intersections.* Capacity and delay vary with the nature of traffic at surrounding intersections.

(Example: signal coordination.)

*Level 4: Geometric changes at isolated intersections.* Capacity varies principally with volume on the subject approach. (Examples: adding exclusive lanes, removal of on-street parking, increasing curb radii.)

Only Level 1 has been tested here (see the previous discussion of the UTOWN network). Any combination of the levels of adaptation could be mixed in a single assignment.

Levels 1, 2, and 3 are now included in forecasts through the process of network calibration. Because these levels reallocate resources between facilities, inclusion of one or more of them can result in multiple equilibrium solutions.

Level 4 is now included in forecasts by proposing alternative projects. If all levels of adaptation are included in the forecast, the assignment would be constrained only by cost or operational limitations.

All long term forecasting should be adaptive to the extent that obvious design flaws in the highway system are eliminated. A good working assumption is that continuing efforts will be made to eliminate bottlenecks due to poor geometry or operations, especially those with low-cost solutions. An important implication of adaptation is that planners may be able to ignore many small and isolated reductions in capacity when building and calibrating their future year networks.

## **TWO-LANE ROADS**

Most two-lane streets in urban areas operate well below their uncontrolled capacity, so delay relationships for this type of facility are not critical to a forecast. Nonetheless, it is possible to make a simple change to the BPR formula (or a similar relationship) to obtain better estimates of delay.

With no opposing volume, the HCM states the capacity to be 2000 pcph. However, the capacity of a subject direction on a two-lane road depends upon its opposing volume. With a 50/50 directional

split, the capacity drops to 1400 pcph. The HCM does not indicate whether this dependence on

directional split holds for urban streets.

where  $V_s$  is the volume in the subject direction,  $V_o$  is the volume in the opposing direction, and  $\tau$  is an empirical constant. The adjusted volume,  $V_a$ , would then be used in the BPR formula when finding the volume-to-capacity ratio. The capacity would be taken to be slightly less than 2000 pcph (appropriately adjusted for heavy vehicles, terrain, narrow lanes, restricted-width shoulders, and other local circumstances).

Based on Table 8-4 in the HCM, a value of  $\tau = 0.4$  is approximately correct for rural roads. Further research is required to properly determine this constant for urban streets.

## **INITIAL SETTINGS FOR CAPACITIES AND FREE SPEEDS**

### **Initial Capacities**

Ideally capacities should be set according to those obtained from the Highway Capacity Manual or from the Highway Capacity Software or similar programs. However, separately setting capacities on every link or on every intersection approach can be quite tedious, especially considering that many of the values may change during network calibration. Many planners prefer to start with rough estimates of capacities and then to refine these estimates during calibration.

Depending upon the forecasting software, the capacities can be entered in a variety of ways. For example, UTPS and similar packages require that capacities be computed as a function of area type, facility class and number of lanes. A look-up table must be prepared giving the maximum lane volume as a function area type and facility class. The software determines the capacity of the link by multiplying the looked-up maximum lane volume by the number of lanes. Other software packages allow capacities to be set for individual links, thereby providing the user with more flexibility during calibration.

The following capacities are recommended for starting values. Where they are given as total directional capacities, they can be divided by the number of through lanes to obtain maximum lane volumes. These values should not be varied by more than  $\pm 20\%$  unless justified by abnormal deviation from ideal conditions.

			60, 70 MPH	50 MPH
	Divided	Level	1800	1700
Rurat		Rolling	1350	1250
	Undivided	Level	1700	1600
		Rolling	1250	1200
	Divided	Level	1600	1500
Suburban		Rolling	1150	1100
	Undivided	Level	1450	1350
		Bolling	1050	1000

Table 3. Initial Capacities for Multilane Highways, Each Lane - Ultimate Capacity

		Table 4		
Initial Capa	cities for Multila	ne Highways, Ea	ach Lane — Desig	gn Capacity
			60, 70 MPH	50 MPH
	Divided	Level	1250	1200
Rural		Rolling	950	875
	Undivided	Level	1200	1125
		Rolling	875	850
	Divided	Level	1125	1050
Suburban		Rolling	800	750
	Undivided	Level	1025	950
		Rolling	750	700
Initial C	apacities for Fr	Table 5 eeways, Each Li	ane — Ultimate C	apacity
		60, 70 MPH	5	OMPH
Level Terr	ain	1800		1700
Rolling Ter	rain	1350		1250
Initial C	apacities for Fr	Table 6 eeways, Each L	ane — Ultimate C	apacity
		60, 70 MPH	5	OMPH
Level Terr	ain	1250		1200
Bollina Ter	rain	950		875

		Table 7	
10	itial Capacities for Two-	Lane Roads — Ultimate	Capacity
		Level	Rolling
Peak	Little No Passing	1500	1050
	Extensive No Passing	1500	950
Off Peak	Little No Passing	1200	800
	Extensive No Passing	1200	750
		Table 8	
le le	nitial Capacities for Two-	Lane Roads — Design	Capacity
		Level	Rolling
Peak	Little No Passing	1050	750
	Extensive No Passing	1050	675
Off Peak	Little No Passing	850	550
	Extensive No Passing	850	525
		Table 9	
Initial C	apacities for Single-Lan Ultima	e, Signalized Intersection ate Capacity	on Approaches
		Low Turns	High Turns
	Low Priority	550	350
No Exclusive	Left Medium Priority	825	550
	High Priority	1100	900
	Low Priority	550	550
Exclusive L	eft Medium Priority	825	825
	High Priority	1100	1100

	Table	e 10	
Initial Capaciti	ies for Single-Lane, S Design C	ignalized Intersectio	on Approaches
		Low Turns	High Turns
	Low Priority	350	250
No Exclusive Left	Medium Priority	550	350
	High Priority	725	600
	Low Priority	350	350
Exclusive Left	Medium Priority	550	550
	High Priority	725	725
Initial Capaci	Table	e 11	Approaches
inda Capaci	Ultimate 0	Capacity	Approacties
		Low Turns	High Turns
	Low Priority	1100	650
No Exclusive Left	Medium Priority	1650	900
	High Priority	2200	1400
	Low Priority	1100	850
Exclusive Left	Medium Priority	1650	1300
	High Priority	2200	2000



Table 13. Initial Capacities for Each Lane Beyond Two, Signalized IntersectionApproaches Ultimate Capacity (not available at this time)



## Assumptions and Extensions for Initial Capacity

The initial capacities for uncontrolled road segments assume 14% trucks, 4% RV's and 0% buses, as suggested for default by the HCM for two-lane roads. The forecast period is one hour. Otherwise, ideal conditions were assumed.

Priority of signal controlled intersections relates to percent of available green time for the approach as follows: low=33%; medium=50%; high=67%. Turns relate to the percentage of traffic: low turns = 0%; high turns = 25%. The lane count does not include exclusive lanes, if applicable.

Consistency of priority should be maintained for all approaches at any given intersection. For example, it would be inappropriate to have more than two high priority approaches at an intersection.

Initial capacities for a medium amount of turns may be interpolated from the values for low and high turns.

Additional ultimate capacity for a exclusive right lane should be provided as follows for each through lane: 0 vph for low turns; 75 for medium turns; and 150 for high turns. Additional design capacity for a exclusive right lane should be provided as follows for each through lane: 0 vph for low turns; 50 for medium turns; and 100 for high turns. For example, the initial ultimate capacity for an approach with two through lanes, both exclusive left and right lanes, high priority and high turns should be 2300 (i.e.; 2000 + 2xl5O).

For signalized approaches with three or more lanes, it is necessary to extrapolate from the data for one and two lanes. For example, the initial capacity for a three lane approach with high turns, medium priority, and an exclusive left lane may be computed as follows:

Two lanes, exclusive left, med. priority, high turns	1300
One lane, exclusive left, med. priority, high turns	825
Additional capacity for each lane beyond the first	475
Total capacity of three lane approach	1775

Some-way stops are seldom included in region-wide networks. For signed approaches at a someway stops capacity varies greatly with the amount of conflicting traffic. Ultimate capacity for each lane should not exceed 1000 vph. See Chapter 1 0 of the HCM for more information about some-way stops.

For travel forecasting packages which explicitly allow signs and signals in the network, consult the software reference manual. For example, QRS 11 requires that the capacity be set to the total saturation flow rate of the through lanes at the approach, without adjusting for signalization priority (amount of green) or amount of turning.

For links containing multiple intersections, choose the smallest capacity.

## Adjusting Initial Capacity for Old BPR Parameters

in order to obtain design capacities. The exponential term takes the fourth root of the expression in brackets; this is easily accomplished on a hand calculator by taking two successive square roots. In this equation a is between 0.56 and 1.0, depending upon the facility type (see previous discussions, Table 2 and Equation 2). This translates into values Of fold of between 0.72 and 0.62. A value of a of 0.63 (yielding a value Of fold Of 0.70) was used to construct the initial design capacities contained in the preceding sections.

## **Initial Free Speeds**

The other important link attribute is the free speed. The following free speeds would be approximately correct for uncontrolled highway segments.

Two-lane roads	
level terrain	58
rolling terrain	57
Freeways and rural multilane highways	
50 mph	48
60 mph	55
70 mph	60

Free speeds should not be set higher than observed speeds under uncongested conditions (LOS A).

It has frequently been observed that drivers in smaller communities choose routes as if freeways were slower than their actual speeds. Consequently, it may be necessary to reduce free speeds for freeways by a significant amount to obtain good agreement with ground counts.

The initial free speed for a long segments of uncontrolled urban streets should be set to no higher than the speed limit, unless evidence to the contrary has be obtained through spot speed studies. The initial free speeds for links containing traffic controlled intersections must be calculated from

the time necessary to travel across the link and the amount of intersection delay. Perform the following steps.

**Step 1.** Determine the length of the link in miles, the average speed of free flowing traffic (speed limit or speed of progression, whichever is applicable), the cycle lengths of signals, and the quality of signal coordination. Express signal coordination as an "arrival type" between 1 to 5, with 5 corresponding to perfectly good progression and 1 corresponding perfectly bad progression (refer to the HCM's definitions for "arrival types"). Assume values for signalization priority according to the expected share of available green time (low=33%; medium=50%; high=67%).

Step 2. Calculate the free flow travel time in seconds. That is,

tf = (3600) (link length)/(free flow speed)

**Step 3.** Choose a value for intersection delay in seconds, tg, from Table 17 for each signalized intersection. Use between 10 and 14 seconds for all-way stops, depending upon the amount of conflicting traffic.



**Step 4.** Find the total intersection delay for signalized intersections only, ts, by totaling the values of tg and multiplying by the progression factor, as indicated below.

Arrival type 1 (poor coordination)	1.85
Arrival type 2	1.35
Arrival type 3 (no coordination)	1.00
Arrival type 4	0.72
Arrival type 5 (excellent coordination)	0.53

Choose a value for the progression factor of 1.00, if the arrival type is unknown or if the forecast is long-term. Be sure that the signalization priority and arrival type are consistent with one another. For example, it would be unusual to have low priority for green time while also having good coordination.

Step 5.	Find total intersection delay, $t_i$ , by adding unsignalized otal signalized delay from Step 4.	I delay from Step 3 to the
Step 6.	Find total link free travel time by summing the results o	f Steps 2 and 5.
	$t_t = t_i + t_f$	(12)
Convert the trave	t total link free travel time to the appropriate units (e.g., el forecasting software.	, minutes) as required by
Step 7.	Compute link free speed in mph:	
	Free speed = (3600)(link length)/t	(13)
where t	has units of seconds and link length is in units of miles	9
has mo second	for example, consider a link that is 1.5 miles long, has a derately good progression, has 3 signals, each of which s and high priority for green time.	a 30 mph free flow speed, h has a cycle length of 90
	Step 2: $t_{p} = (3600)(1.5)/(30) = 180$ seconds. Step 3: $t_{0} = 17$ seconds. Step 4: $t_{0} = (0.72)(17 + 17 + 17) = 37$ seconds. Step 5: $t_{1} = 37$ seconds. Step 6: $t_{1} = 37 + 180 = 217$ seconds or 3.62 min Step 6: $t_{1} = 37 + 180 = 217$ seconds or 3.62 min	autes.

## **Discussion of Initial Free Speeds**

*Signal Timing.* If signal timing is essentially unknown, then assume each signal adds 20 seconds of delay to free travel time. For different values of green time, g, and cycle length, C, the following equation from the HCM can be used to estimate delay when traffic volumes are low: **(not available at this time)** 

*Some-Way Stops.* Consistency should be maintained between the capacity of a single lane at some-way stops and the delay under low volume conditions. Intersection delay is approximately,

tg = 3600/(lane capacity) + acceleration delay,

when there is little traffic approaching the sign.

## **CONCLUSIONS**

Current travel forecasting models are quite limited in their ability to estimate delay on links or at intersections. It is unlikely that good delay estimates can be calculated without substantial rewriting of software.

The 1985 Highway Capacity Manual was not developed for the purpose of travel forecasting, so many important relationships were omitted. Furthermore, HCM's delay relationships violate strict mathematical requirements that are necessary for the most widely adopted equilibrium traffic assignment algorithm, Frank-Wolfe decomposition.

For uncontrolled, multilane road segments, link delay can be adequately calculated with the BPR speed/volume function or with alternative functions proposed by Spiess and Overgaard.

Some models, including UTPS, calculate link capacity from a preset capacity for each lane, which can vary only by location in the region and by facility type. The complexity of the HCM procedures suggest that it is not possible to accurately calculate capacity within this type of modeling framework.

Complicated delay relationships are required for signalized intersections, unsignalized intersections, weaving sections, and two-lane roads. For these situations, delay on a single link is a function of volumes on two or more links.

It is possible to build a travel forecasting model that contains intersection delay relationships very similar to those in the HCM. One algorithm, sometimes referred to as equilibrium/incremental assignment, is available for finding an equilibrium solution. Strict application of the HCM procedures would result in networks with multiple equilibrium solutions. It is likely that the burdens of network calibration will be considerably reduced with such a model.

Levels of adaptation are important to the results of travel forecasts. Adaptation is a principal justification calibrating a network. The HCM provides sufficient information about the relationships between volume, capacity and delay to build assignment algorithms that are highly adaptive.

### **R**ECOMMENDATIONS

The BPR function fits the various delay/volume relations in the HCM with good consistency. If only one curve can be chosen, the BPR function is preferred to Spiess' and Overgaard's.

Capacity is the most important variable when estimating volumes on congested highways. Since the definitions of levels of service vary greatly by facility type, "capacity" in delay/volume functions should be set at LOS E, ultimate capacity. Design capacity should be phased-out as a variable in delay/volume functions.

Because of the large number of factors affecting capacity of uncontrolled road segments, capacity should be separately determined for each link. The Highway Capacity Manual provides procedures for most types of facilities, and these procedures should be followed.

If only one set of parameters can be chosen for the BPR function, then the volume-to-capacity multiplier,  $\alpha$ , should be approximately 0.83 and the volume-to-capacity exponent,  $\beta$ , should be approximately 5.5.

Additional research is needed on capacity of two-lane streets in urban areas.

Travel forecasting software should contain procedures, similar to those in the HCM, in order to achieve more precise estimates of capacity and delay at intersections.

In the absence of such software, planners can still improve their forecasts while calibrating their networks. Planners should adopt one of the methods presented in this report to better specify capacity at intersection approaches.

During calibration, planners need to achieve consistency between their assigned volumes and the nature of traffic control at intersections. This can be done by referencing signal warrants from the Manual on Uniform Traffic Control Devices or by comparing total delay from alternative traffic control strategies. Planners need not consider the possibility of all-way stop controlled intersections, unless this form of traffic control is required for purposes other than minimizing delay.

Network calibration, as now practiced by planners, appears to be a means of overcoming deficiencies in existing delay/volume relationships. It is important that the same calibration process, which is applied to the base network, also be applied to future-year networks. Specifically, planners need make sure that their values of capacity are consistent with the distribution of traffic at intersections, at weaving sections, and at two-lane roads. It is not possible to assume that values of capacity set for the base-year network also hold for future-year networks.

#### REFERENCES

- Akcelik, Rahmi, "The Highway Capacity Manual Delay Formula for Signalized Intersections", *ITE Journal*, Vol. 58, No. 3, March 1988, pp. 23-28.
- Branston, David, "Link Capacity Functions: A Review", *Transportation Research, Vol.* 1 0. pp. 223-236, 1976.
- Baass, Karsten G., "The Potential Capacity of Unsignalized Intersections", *ITE Journal*, pp. 43-46, October 1987.
- "Highway Capacity Manual", Transportation Research Board, Special Report 209, Washington, DC, 1985.
- Horowitz, Alan J., "Convergence Properties of Some Iterative Traffic Assignment Algorithms", *Transportation Research Record*, No. 1220, 1990, pp. 21-27.

- Kyte, Michael, "All-Way Stop Controlled Intersections: Draft Procedures for Capacity and Level of Service Analysis", University of Idaho, December 17, 1989.
- LeBlanc, L., E. Morlok, and W. Pierskella, "An Efficient Approach to Solving the Road Network Equilibrium Traffic Assignment Problem", *Transportation Research*, Vol. 9, 1975, pp. 309-318.
- Powell, W. B., and Y. Sheffi, "The Convergence of Equilibrium Algorithms and Predetermined Step Size", *Transportation Science*, Vol. 16, 1982, pp. 45-55.
- Richardson, Anthony J., "A Delay Model for Multiway Stop-Sign Intersections", *Transportation Research Record*, No. 1 1 12, pp. 107-112, 1987.
- Spiess, Hans, "Conical Volume-Delay Functions", *Transportation Science*, Volume 24, Number 2, May 1990, pp. 153-158.)
- Overgaard, K. R. (1967) "Urban Transportation Planning: Traffic Estimation", *Traffic Quarterly*, pp. 197-218.)
- Hansson, Arne, "Swedish Highway Capacity Manual: Part 2. Capacity of Unsignalized Intersections", *Transportation Research Record*, No. 667, pp. 4-11, 1978.

### APPENDIX A: SAMPLE SPECIFICATIONS FOR INTERSECTION DELAY

The following specification of intersection delay models assumes prior knowledge of the HCM. References are made to equations, tables, and figures from Chapters 9 and 1 0 of the HCM.

### Signalized Intersections

When a signalized intersection is included in a network, the model should only require information about:

- a. the cycle length;
- b. the saturation flow rate for the through lanes of each approach;
- c. the existence of exclusive lanes at each approach;
- d. the link's arrival type; and
- e. the link's speed.

The model should be able to calculate all other intersection information that normally would be part of a capacity/delay analysis.

The signalized intersection specification follows the HCM, except as noted here.

*Adjustment Factors.* The model not does not necessarily have to make adjustments for lane width, grade, parking, buses, heavy vehicles, and/or area type. For example, deviations from ideal conditions can be incorporated by the user into the saturation flow rate for the through lanes at the approach.

*Green Times.* The model should determine whether protected left phases are required and should determine the amount of green time to be allocated to each phase. When a protected phase is warranted the model should always adopt the phase sequence [(L + L), (LTR + LTR)], sometimes referred to as dual leading lefts with overlap. The model should not determine optimal green times. Rather, the model adheres to standard traffic engineering practice by

allocating time to a phase in proportion to the maximum flow ratio (ratio of volume to saturation flow rate) during that phase.

*Protected Lefts.* The model should introduce a protected left phase, if there is insufficient capacity to process all left-turning vehicles without one. In ascertaining this capacity, the model should consider the number of gaps available during the unblocked green time and the number of sneakers. The protected left phase is given only sufficient time to process vehicles that cannot be handled during the LTR phase of the worst approach. The model then divides left turning traffic between the L and LTR phases for all approaches, nearly filling the protected left phase with traffic. The saturation flow rate for the LTR lane group includes the left lane capacity, if the left lane can be shared.

*Left Lane Saturation Flow Rate.* The left turn factor for exclusive lanes should be calculated according to Cases 1 or 2 from Table 9-12. The model should be able to modify the saturation flow rate for left turn lanes by using the implied reduction from the ideal saturation flow rate for the through lanes (e.g., for heavy vehicles and grades).

*Shared Left Lanes Acting as Exclusive Lanes.* To avoid discontinuities in delay, the model should create an exclusive left lane from a shared LT lane, only if a protected phase is warranted. The HCM's procedure for determining defacto left lanes should not be used.

*Exclusive Right Lanes.* The model need not create a separate lane group for an exclusive right turn lane. Rather, the saturation flow rate for the LTR or TR lane group can be adjusted upward to reflect the additional lane. The model should add sufficient capacity to just accommodate the right turning vehicles, with a maximum adjustment equal to a single lane's saturation flow rate.

*Right Turns from Shared Lanes.* The model need not provide for pedestrians. Consequently, the right turn adjustment factor would be calculated according to Case 4 on Table 9-1 1.

*Period of Analysis.* Because the model forecasts travel during whole hours, the peak-hour-factor is unnecessary. For multihour assignments, the model should take a volume-weighted average of the delay in each hour.

*Delay Function.* The model should calculate stopped delay from the HCM delay function (i.e., total delay divided by 1.3). The HCM delay function can become undefined for volume-to-capacity ratios only slightly greater than 1.0. Consequently, the model can use the HCM delay function only up to a volume-to-capacity of 1.0. Beyond 1.0, delay should be calculated as a linear extrapolation of the delay at a volume-to-capacity ratio of 1.0.

*Acceleration Delay.* The model should estimate the fraction of stopping vehicles and add acceleration delays for those vehicles. The fraction of stopping vehicles depends upon the arrival type and the volume-to-capacity ratio. The acceleration delay depends upon the link speed. For stopping vehicles,

Acceleration Delay

(Speed/2)(l/Acceleration Rate + 1/Deceleration Rate)

As a convenience, the speed can be taken from the link constituting the approach. For the simulations of this report, acceleration rate was set at 3.5 mph/second and deceleration rate was set at 5.0 mph/second.

*Fraction of Stopped Vehicles.* The model can determine the number of stopped vehicles by interpolating between 1.0 (at the value of the volume-to-capacity ratio, X, where all vehicles are assumed to have stopped, e.g., 1.2) and the fraction assumed to stop when the volume-to-capacity

ratio is zero. This latter value will be referred to as the lowerbounds, L. There are separate lowerbounds for each possible arrival type. For an arrival type of 1 (least favorable progression),

L <sub>1</sub> = 1		(A.2)
For an arrival type of 3 (random arrivals) the lowerbound	d is	
$L_3 = (C - g)/C$ .		(A.3)
where C is the cycle length and g is the green time. favorable progression), no vehicles stop. Therefore,	For an ar	rival type of 5 (most
		(A.4)

all vehicles must stop. So,

The lowerbound for arrival type 2 is found from averaging the lowerbound for arrival types 1 and 3. Similarly, the lowerbound for arrival type 4 is found from averaging the lowerbound for arrival types 3 and 5.

Regardless of the arrival type, all vehicles are assumed to stop when the volume-to-capacity ratio exceeds the user-specified value of the volume-to-capacity ratio, X.

It should be noted that the fraction of vehicles stopping at a signalized intersection under arrival type 3 can be easily derived from elementary traffic flow theory. The resulting nonlinear relationship is closely approximated by application of Equation A.3, above. A linear relation was chosen for consistency with the other arrival types.

*Lane Utilization.* Because the model calculates average delay across all lanes, a lane utilization factor is not needed.

*Progression Adjustment.* Like the HCM, the model should adjust delay as a function of the arrival type and the volume-to-capacity ratio. To avoid discontinuities, the model should use a set of linear equations to estimate the adjustment factor - one equation for each arrival type. The linear equations range from a volume-to-capacity ratio of 0.0 to a volume-to-capacity ratio of 1.2 (or another user-supplied parameter value), where the progression adjustment factor always becomes 1.0 (equivalent to no adjustment). Beyond a volume-to-capacity ratio of 1.2, no adjustment to delay is made. No adjustment is made to delay for vehicles in exclusive left-turn lanes.

Define F as the lowerbound value of the progression factor, i.e., when X is zero. For an arrival



type of 1 (least favorable progression) the value of delay must be increased. Consequently,

For values of the volume-to-capacity ratio less than the user-specified maximum, the model interpolates between the lowerbound, F, and 1.0. The progression factor when the arrival type is 2 is found by averaging those for 1 and 3. The progression factor for a arrival type of 3 is found by averaging those for 3 and 5.

*Overflow Time Period.* Unlike the HCM, the model must allow the user to vary the overflow delay time period, T, fixed at 0.25 hours in the HCM. In addition, it should be possible to vary the ratio of total to stopped delay,  $\eta$ , fixed at 1.3 in the HCM. These changes affect the three

constants in Equation 9-18. (See Akcelik, 1988, for a technical analysis of the HCM delay function.) The constant leading the first term (seen as 0.38) is found from:

First Constant = 0.5/r7

The constant leading the second term (seen as 173) is found from:

Second Constant =  $90OT/\eta$ 

The last constant appears within the radical (seen as 16), and is calculated from:

Third Constant = 4/T

## **Some-Way Stop Intersections**

In order to calculate delay at some-way stop intersections, the specification requires information about the locations of stop signs and the lane geometry at approaches with signs. Three types of lane configurations can be readily handled: one LTR lane; one LT and one R lane; and one LT and one TR lane. The model also needs the speeds of traffic on all links at the intersection.

The some-way stop model is consistent with the unsignalized model in the HCM, except as follows.

*Potential Capacity Curves.* The curves for potential capacity as a function of conflicting volume, Figure 10-3 in HCM, must be extended to handle any amount of conflicting volume (Baass, 1987). Figure 10-3 suggests that there should be a minimum capacity of 33 vehicles per hour, regardless of the amount of conflicting volume. The user should be able to change this minimum for all intersections or for any given intersection.

*Treatment of Left Turns.* The model need not make a distinction between left and through vehicles at signed approaches. Consequently, a left-turning vehicle would not impact the capacity of its opposing approach. However, the model should be consistent with the HCM in its treatment of left turns from unsigned approaches.

Acceleration Delay. The specification provides for acceleration delay for all vehicles at signed approaches and for left-turning vehicle at unsigned approaches. The acceleration delay depends upon the link speed.

*Right-turn Lane Geometry.* The model can consider right-turn lane geometry. For example, the user should be able to make adjustments to the acceptable right-turn gap at signed approaches.

*Number of Lanes for the Major Street.* The number of lanes for the major street can be determined by observing the capacity (or saturation flow rate) of the unsigned approaches. The number of lanes may be found by dividing the capacity by the ideal saturation flow rate and rounding to a whole number. The number of lanes is taken to be the maximum over all unsigned approaches.

*Capacity.* Capacity of a movement is computed by the German method as summarized by Baass (1987). This method produces almost exactly the same results as the HCM, but permits any value for the critical gap and any value for conflicting traffic.

*Stopped Delay.* The HCM provides relationships for estimating the capacity of some-way stops, but does not provide relationships for estimating delay. The specification includes queuing delay for all vehicles at signed approaches and for left-turning vehicles at unsigned approaches.

Following the Swedish Highway Capacity Manual (Hansson, 1978), the model estimates delay, D, for any lane assuming Poisson arrivals and exponential service times:

D = 1 / (VI - c)

where D is measured in seconds, VI is the lane volume (in vehicles per second), and c is the lane capacity (in vehicle per second). Equation A.11 is used for volume-to-capacity ratios less than or equal to 0.9. For greater volume-to-capacity ratios the model should compute delay from the tangent to Equation A.11 at a volume-to-capacity ratio of 0.9. Thus, delay can still be calculated even when volume exceeds capacity.

*Distribution of Through Vehicles Across Lanes.* At signed approaches with two shared lanes, the model must divide the through traffic between the LT and TR lanes. An attempt should be made to equalize the volume-to-capacity ratios of the two lanes. To do this, the model calculates the



proportion of through to be allocated to the right lane, PR

If  $P_{R}$  is greater than 1 or less than 0, all through vehicles are allocated to either the right or left lanes, respectively.

# **All-Way Stop Intersections**

The HCM does not contain methods for estimating capacity or delay at all-way stop intersections. Consequently, the model must adopt other procedures for delay at all-way stop intersections. An enhanced version of Richardson's M/G/1 queuing model is chosen. Unlike Richardson's original formulation, the specification considers delays due to turning and delays caused by the need for coordination between drivers on the same and opposing approaches.

Definition of Processing Time and Service Time. The M/G/1 model estimates delay at an approach from the rate of arriving vehicles and from the mean and variance of the amount of time it takes for vehicles to pass through the intersection, referred to as the service time. The service time for an approach is equal to the sum of the time necessary to process a vehicle through the subject approach and the time necessary to process a vehicle through a conflicting approach, provided there is a vehicle at the conflicting approach. Both of these processing times (subject and conflicting) are computed by the same method, although they will have different values because of differing traffic characteristics. A typical processing time is about 4 seconds, so a service time is either about 4 seconds or about 8 seconds, depending upon the absence or presence of a conflicting vehicle.

*Capacity in Relation to Service Time.* The capacity of an intersection is inversely related to service time. For example, a single-lane approach at an intersection with heavy traffic in all directions would have a uniform service time of about 8 seconds, because there will always be conflicting vehicles. The capacity of such an approach would be 1/8 vehicle per second or 450 vehicles per hour.

*Factors in Processing Time.* For single lane approaches, the processing time depends upon (1) the presence or absence of right and left turning vehicles on the subject or opposing approaches and (2) the presence or absence of any vehicle on the opposing approach. This is handled by adding and subtracting constants for each effect. In general, left turns increase processing time, while right turns decrease processing time. For two lane approaches, the processing time also depends upon the presence or absence of a second vehicle on either the subject or opposing approaches. These additional vehicles introduce a need for coordination among drivers and, therefore, tend to increase processing time.

*Lane Distribution.* Each vehicle arriving at an approach has a different service time, but the average service time is assumed to be the same for all vehicles, regardless of their turning behavior. Consequently, traffic is distributed across lanes, at multilane approaches, as evenly as possible (taking into consideration the required lane assignments for left and right turning vehicles).

*Lane Configurations.* Possible lane configurations for approaches at all-way stops are the same as for some-way stops.

Acceleration Delay. Since all the vehicles stop, the model must add an acceleration delay to the queuing delay found from the M/G/1 model.

Stopping Delay. One of two delay relations could be used, depending upon user preference.



First, delay can calculated from the following relation for each lane (Kyte, 1989),

Equation A.15 differs from Richardson's (1987) by including terms for coordination of vehicles on the subject and opposing approaches. This expression for variance is an approximation because it only includes variation due to the presence or absence of conflicting traffic, ignoring variation due to turning and due to the presence or absence of other vehicles on the subject approach or opposing approach.

Delay for a lane is computed by the following equation for values of less than or equal to 0.9:

For values of X greater than 0.9, the model should take the delay from the tangent to Equation A.16 at a value of X of 0.9. This second method was used for the simulations in this report.

*Parameters.* The parameters of the all-way stop model consist of "waits" in units of seconds. The following "waits" affects processing time.

a. Subject Unit Wait = 3.6

(Processing with no other vehicle present.)

b. One Left Wait = 1.

(Additional processing time if there is exactly one left turning vehicle on the subject or opposing approaches.)

c. Two Lefts Wait = 1.

(Additional processing time if there is exactly two left turning vehicles on the subject and opposing approaches.)

d. One Right Wait = -0.5

(Additional processing time if there is exactly one right turning vehicle on the subject or opposing approaches.)

e. Two Rights Wait = -1.

(Additional processing time if there is exactly two right turning vehicles on the subject and opposing approaches.)

f. Another Lane Wait = 1.

(Additional processing time if there is a second vehicle at the subject approach.)

g. One Opposing Lane Wait = 0.25

(Additional processing time if there is exactly one vehicle on the opposing approach.)

h. Two Opposing Lanes Wait = 1.

(Additional processing time if there is exactly two vehicles on the opposing approach.)

The remain "waits" affect service time, only if there is a vehicle at an conflicting approach.

i. One-Lane Added Wait = -0.5

(Additional service time when the subject approach has one lane.)

j. One+ Right Added Wait = 0.

(Additional service time when the subject approach has one left/through lane and one right lane.)

1. Two-Lane Added Wait = 0.5

(Additional service time when the subject approach has two lanes.)

These parameters were selected to match data collected by Kyte (1 989).

**APPENDIX B: BEST FIT SPEED/VOLUME FUNCTIONS** 



Figure B.1: Best Fit Speed/Volume Curves for Freeways, 70 MPH Design Speed



Figure B.2: Best Fit Speed/Volume Curves for Freeways, 60 MPH Design Speed



Figure B.3: Best Fit Speed/Volume Curves for Freeways, 50 MPH Design Speed

Figure B.4: Best Fit Speed/Volume Curves for Rural Divided Multilane, 70 MPH



Figure B.5: Best Fit Speed/Volume Curves for Rural Divided Multilane, 60 MPH



Figure B.6: Best Fit Speed/Volume Curves for Rural Divided Multilane, 50 MPH



Speed

# APPENDIX C: SELECTED DELAY/VOLUME RELATIONSHIPS FOR SIGNALIZED INTERSECTIONS

**Figure C.1:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 1 000 VPH at Opposing and Conflicting Approaches, No Exclusive Lanes, 3600 VPH Ideal Saturation Flow Rate, 20 MPH Speed,



Arrival Type = 3, 90 Second Cycle)

**Figure C.2:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 600 VPH at Opposing and Conflicting Approaches, No Exclusive Lanes, 3600 VPH Ideal Saturation Flow Rate, 20 MPH Speed,





**Figure C.3:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 200 VPH at Opposing and Conflicting Approaches, No Exclusive Lanes, 3600 VPH Ideal Saturation Flow Rate, 20 MPH Speed, Arrival Type = 3, 90 Second Cycle)



**Figure C.4**: Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 1 000 VPH at Opposing and Conflicting Approaches, Exclusive Left, 3600 VPH Ideal Saturation Flow Rate, 20 MPH Speed, Arrival



Type = 3, 90 Second Cycle)

**Figure C.5:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (O% Right Turns, 0% Left Turns, 1 000 VPH at Opposing and Conflicting Approaches, No Exclusive Lanes, 3600 VPH Ideal Saturation Flow Rate, 20 MPH Speed,



Arrival Type = 3, 90 Second Cycle)

**Figure C.6:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 600 VPH at Opposing and Conflicting Approaches, No Exclusive Lanes, 1800 VPH Ideal Saturation Flow Rate, 20 MPH Speed, Arrival Type = 3, 90 Second Cycle)



**Figure C.7:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 600 VPH at Opposing and Conflicting Approaches, Exclusive Left, 1800 VPH Ideal Saturation Flow Rate, 20 MPH Speed, Arrival



Type = 3, 90 Second Cycle)

**Figure C.8:** Delay on All Approaches of a Signalized Intersection as a Function of Volume on a Single Approach (25% Right Turns, 25% Left Turns, 600 VPH at Opposing and Conflicting Approaches, Exclusive Right, 1800 VPH Ideal Saturation Flow Rate, 20 MPH Speed, Arrival



Type = 3, 90 Second Cycle)

# APPENDIX D: GENERALIZED INTERSECTION DATA FOR TWO-WAY AND FOUR-WAY STOPS

**Figure D.1:** Delay on Subject and Conflicting Approaches for a Four-Way Stop (Opposing Volume Same as Subject Volume, Conflicting Volumes at 400 vph, 25% Right Turns, 25% Left Turns, One Lane at All Approaches, 20 MPH Speed)



**Figure D.2:** Delay on Subject and Conflicting Approaches for a Four-Way Stop (Opposing Volume Same as Subject Volume, Conflicting Volumes at 600 vph, 25% Right Turns, 25% Left



Turns, One Lane at All Approaches, 20 MPH Speed)



Figures D.3 and D.4

**Figure D.5:** Delay on Subject and Conflicting Approaches for a Two-Way Stop (Opposing Volume Same as Subject Volume, Conflicting Volumes at 400 vph, 25% Right Turns, 25% Left Turns, One Lane at All Approaches, 20 MPH Speed)



**Figure D.6:** Delay on Subject and Conflicting Approaches for a Two-Way Stop (Opposing Volume Same as Subject Volume, Conflicting Volumes at 600 vph, 25% Right Turns, 25% Left



Turns, One Lane at All Approaches, 20 MPH Speed)