TRANSPORTATION RESEARCH RECORD 667

Highway Capacity, Measures of Effectiveness, and Flow Theory

TRANSPORTATION RESEARCH BOARD

COMMISSION ON SOCIOTECHNICAL SYSTEMS
NATIONAL RESEARCH COUNCIL

NATIONAL ACADEMY OF SCIENCES
WASHINGTON, D.C. 1978

Transportation Research Record 667
Price $5.00

subject areas
51 highway safety
53 traffic control and operations
54 traffic flow
55 traffic measurements

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

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

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

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

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

Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Highway capacity, measures of effectiveness, and flow theory.

(Transportation research record; 667)
Reports given at the 57th annual meeting of the Transportation Research Board.
Includes bibliographical references.
1. Traffic flow–Congresses. 2. Traffic safety–Congresses. 3. Traffic engineering–Congresses. I. Title. II. Series.
TE7.H5 no. 667 [HE336.T7] 380.5'08s [387.3'1]
ISBN 0-309-02807-8 79-309

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

TRANSPORTATION RESEARCH BOARD
National Research Council
ERRATA 1979

Special Report 175

page 67, column 2, line 36
Change "because it has a larger size aggregate" to "if it had had a larger-size aggregate"

Special Report 176

page 3, caption for Figure 1.3
Change "Segeto" to "Shigeto"
page 13, column 2, line 4
Change "Figure 2.6" to "Figure 2.5"
page 15, Figure 2.10, Zone A
Change heading "E" to "C"
page 17, column 2, line 37
Change "fluids, as shown in Figure 2.1p5." to "fluids."
page 18, column 1, line 11 from bottom
Change "site" to "side"
page 18, caption for Figure 2.17
Change "(7 to 6 ft)" to "(7 to 16 ft)"
pages 92 and 93, Figures 4.8 and 4.9
Reverse captions for figures
page 231, column 3, line 18
Change "13, 17" to "17"

Transportation Research Record 667

page 41, column 1, line 9 from bottom
Change "versus" to "per unit of"
page 41, column 2, line 3 from bottom
Change "versus" to "and"
page 42, column 2, line 17 from bottom
After "All scenarios" insert "chosen"
page 43, column 1, lines 24-26
Change no. 4 to read: "4. A table of fuel consumption versus emission rates is accessed each second for each vehicle by using the vehicle's speed-acceleration couplet (8)."
page 43, column 1, line 31
Change "fuel consumption" to "fuel efficiency"
page 43, column 2, last line
Change Equation 6 to read: "FE = 0.695 + 0.182*(average speed) - 0.0023*(average speed)^2"
page 44, Figure 3, caption and vertical axis
Change "fuel consumption" to "fuel efficiency"
page 45, Figure 8, caption and vertical axis
Change "fuel consumption" to "fuel efficiency"
page 47, column 1, line 1
Change "fuel consumption" to "fuel efficiency"
page 47, column 1, line 3
Change Equation 7 to read: "HC = 8.342 - 0.240*(average speed) + 0.007 25*(average speed)^2"
page 47, column 1, line 6
Change Equation 8 to read: "HC = 1.36 + 34.522 ÷ (average speed)"
page 47, column 1, line 8
Change Equation 9 to read: "CO = 171.71 - 9.222*(average speed) + 0.164*(average speed)^2"
page 47, column 1, line 11
Change Equation 10 to read: "CO = 16.03 + 766.0 ÷ (average speed)"
page 47, column 1, lines 16-17
Change "the cycle length, which minimizes delay at an isolated intersection, also" to "a cycle length that minimizes delay at an isolated intersection also"
page 47, column 1, line 4 from bottom
After "is not considered" insert "in these other papers."

Transportation Research Record 673

page 52, column 2, line 12 from bottom
Change "90 and 100" to "90 and 10"

Transportation Research Record 675

page 19, column 1
Insert the following before the last paragraph:

Hydrated Portland Cement Pastes

Hydrated portland cement pastes were used to test the hypothesis discussed above. The second-intrusion method was applied to hydrated portland cement pastes, which were mixed in vacuum with a water-cement ratio of 0.4 and cured for 3 and 60 days in saturated calcium hydroxide solution at 20°C. The pastes were oven dried at the end of curing periods, and duplicate specimens (1-3 g each) were tested. The specimens were taken out of the sample cell at the end of the first pressurizing-depressurizing cycle and placed and tested in the porosimeter again as a new specimen. The surface tension of mercury was taken as 484 dyne/cm^2 and the contact angle as 117° for all cycles. Thus, the Washburn equation becomes

$$P = 127.500/D(\mu m) \quad (3)$$

where P is in pounds force per square inch.

Winslow and Diamond (3) determined the contact angle of mercury on oven-dried hydrated portland cement pastes by observing the penetration pressure of mercury into small boreholes of known diameters. The assumption of equal contact angle in first and second intrusions is supported by Winslow and Diamond's procedure, since they intruded and extruded mercury into the small holes in cycles and observed no change in penetration pressure.

Conventional and uniform pore-size distributions of hydrated portland cement pastes are shown in Figures 6 and 7. The first-intrusion curves are similar to those reported in early investigations (3). There is practically no intrusion up to the threshold diameter D_t, and a major fraction of the total pore volume is within a small range of sizes smaller than D_t; finally, the slope tends to become smaller and smaller as smaller diameters are intruded. Thirty-six percent of the total mercury intruded is retained in the paste hydrated for 3 days and 40 percent in the paste hydrated for 60 days at end of depressurizing.

Transportation Research Record 679

page 1, column 2, line 13
Change "4895 N/s (1100 lb/s)" to "4895 N·s (1100 lbf·s)"
page 16, column 2, line 20 from bottom
Change "10.05 kN" to "1025 kg" and "62.27-kN" to "6350-kg"

page 16, column 2, line 15 from bottom and lines 10-11 from bottom
Change "19.93 kN" to "2032 kg"
page 16, column 2, lines 5-6 from bottom
Change "19.88 kN" to "2028 kg"
page 18, column 1, lines 6 and 12
Change "10.05-kN" to "1025-kg" and "1.913-kN" to "195-kg"
page 18, column 1, line 20, and column 2, lines 20 and 23
Change "62.27-kN" to "6350-kg"
page 26, column 1, lines 13-15
Change each "nt·s" to "N·s" and each "lb-s" to "lbf·s"
page 26, column 2, lines 4-6
Change each "nt" to "N" and each "lb" to "lbf"
page 27, column 1, lines 1-2 and following table
Change to "production lot (1 kN = 225 lbf):

Test Piece	Axial Load (kN)
CT 7-11-1	136.3
CT 7-11-2	115.6
CT 7-11-3	116.1
CT 7-12-3	119.7
CT 7-12-4	122.1
Average	121.9"

page 27, column 1, lines 9-11 from bottom and following table
Change to "table (1 kN = 225 lbf):

Test Piece	Shear Load (kN/coupling)
CT 7-11-4	28.0
CT 7-11-5	17.3
CT 7-11-6	23.6
CT 7-12-1	17.3
CT 7-12-2	19.6
Average	21.2"

page 27, column 2, line 13
Change "227 kg (500 lb) and 4536 kg (10 000 lb)" to "2.2 kN (500 lbf) and 44.5 kN (10 000 lbf)"
page 28, column 1, lines 3, 14-15, 18, and 22
Change each "nt·s" to N·s" and each "lb·s" to "lbf·s"
page 29, Abstract, line 15
Change "362 kg·s" to "3.6 kN·s"
page 30, column 2, line 21
Change "1145, 1105, and 1060 kg·s" to "11.2, 10.8, and 10.4 kN·s"
page 30, column 2, line 17 from bottom
Change "492, 487, 500, and 464 kg·s" to "4.93, 4.89, 5.02, and 4.65 kN·s"
page 31, column 1, line 11
Change "350, 360, 350, and 357 kg·s" to "3.43, 3.53, 3.43, and 3.50 kN·s"
page 31, column 1, line 39
Change "338, 349, and 388 kg·s" to "3.31, 3.43, and 3.80 kN·s"
page 31, column 2, line 35
Change "91 kg·s" to "0.89 kN·s"
page 31, column 2, line 44
Change "Tunnel momentum change, kg·s 504 351 338 452" to "Tunnel momentum change, kN·s 4.93 3.43 3.31 4.43"
page 31, Table 3
Change the momentum change values from kg·s to kN·s for each category: "Speed Trap Measurement: NM, 4.75, 3.51,–,4.96"; "Integration of Tunnel Acceleration: 11.17, 4.93, 3.42, 3.31, 4.50"; "Integration of Rear-Deck Acceleration: 10.8, 4.89, 3.53, 3.43, 4.10"; "High-Speed Film Analysis: 10.4, 4.89, 3.43, 3.80, 4.48"
page 32, column 1, lines 7 and 10
Change "91 kg·s" to "0.89 kN·s" and "457, 418, 457, and 506 kg·s" to "4.43, 4.11, 4.25, and 4.97 kN·s"
page 33, column 1, lines 7-8
Change "500 kg·s" to "4.89 kN·s" and "350 kg·s" to "3.34 kN·s"
page 33, column 1, text table
Change the momentum change values from kg·s to kN·s for each test: "Test 1:–, 11.22, 10.83, 10.39"; "Test 2: 4.75, 4.94, 4.89, 5.02"; "Test 3: 3.51, 3.44, 3.53, 3.43"; Test 4:–, 3.31, 3.42, 3.80"; "Test 5: 4.96, 4.48, 4.10, 4.48"
page 33, column 1, line 31
Change "350 kg·s" to "3.34 kN·s"
page 33, column 2, lines 5 and 7
Change "500 kg·s" to "4.89 kN·s" and "91 kg·s" to "0.89 kN·s"

Transportation Research Record 681
page 19, column 2, line 22
Change "frequently" to "infrequently"

Transportation Research Record 720
page ii, Library of Congress data
Change "[666'.89]" to "[66';893]"

Transportation Research Record 721
page ii, Library of Congress data
Add "National Research Council. Transportation Research Board.
Rail and motor carrier reports.
(Transportation research record; 721)
1. Railroads–Addresses, essays, lectures. 2. Railroads–United States–Addresses, essays, lectures. 3. Transportation, Automotive–Freight–Addresses, essays, lectures. 4. Transportation–Law and legislation–United States–Addresses, essays, lectures. I. Title. II. Series.
TE7.H5 no. 721 [HE1031] 380.5s [385]
ISBN 0-309-02971-6 79-607922

Transportation Research Record 734
page ii, column 1
Change publication data to
Transportation Research Record 734
Price $3.40
Edited for TRB by Mary McLaughlin

mode
1 highway transportation

subject areas
24 pavement design and performance
31 bituminous materials and mixes
34 general materials
35 mineral aggregates
62 soil foundations

Change Library of Congress data to
Library of Congress Cataloging in Publication Data
National Research Council. Transportation Research Board.
Copper mill tailings, incinerator residue, low-quality aggregate characteristics and energy savings in construction.

(Transportation research record; 734)

1. Waste products as road materials—Addresses, essays, lectures. I. Title. II. Series.
TE7.H5 no. 734 [TE203] 380.5s [625.7'35] 80-12762
ISBN 0-309-02989-9 ISSN 0361-1981

page ii, column 2

Change sponsorship data to

Sponsorship of the Papers in This Transportation Research Record

GROUP 2–DESIGN AND CONSTRUCTION OF TRANSPORTATION FACILITIES
Eldon J. Yoder, Purdue University, chairman

Bituminous Section
Moreland Herrin, University of Illinois at Urbana-Champaign, chairman

Committee on Characteristics of Bituminous Paving Mixtures to Meet Structural Requirements
Bernard F. Kallas, Asphalt Institute, chairman
Grant J. Allen, Charles W. Beagle, Oliver E. Briscoe, H. W. Busching, J. T. Corkill, Jon A. Epps, Charles R. Foster, William O. Hadley, Erling Henrikson, R. G. Hicks, Rudolf A. Jimenez, Ignat V. Kalcheff, Thomas W. Kennedy, Charles V. Owen, Charles F. Parker, David W. Rand, R. M. Ripple, Donald R. Schwartz, Jack E. Stephens, Ronald L. Terrel, David G. Tunnicliff, James D. Zubiena

General Materials Section
Roger V. LeClerc, Washington State Department of Transportation, chairman

Committee on Mineral Aggregates
T. C. Paul Teng, Mississippi State Highway Department, chairman
Grant J. Allen, Gordon W. Beecroft, John C. Cook, J. T. Corkill, Ludmila Dolar-Mantuani, Karl H. Dunn, Stephen W. Forster, Richard D. Gaynor, Donn E. Hancher, Robert F. Hinshaw, William W. Hotaling, Jr., Richard H. Howe, Eugene Y. Huang, William B. Ledbetter, Donald W. Lewis, Charles R. Marek, John E. McChord, Jr., Andy Munoz, Jr., Jerry G. Rose, N. Thomas Stephens, Hollis N. Walker, Richard D. Walker

Compaction and Stabilization Section
Eugene B. McDonald, South Dakota Department of Transportation, chairman

Committee on Soil-Portland Cement Stabilization
J. M. Hoover, Iowa State University, chairman
Ara Arman, H. H. Duval, Jr., Donald G. Fohs, K. P. George, Dennis A. Grylicki, John B. Heagler, Jr., Lynne H. Irwin, Thomas W. Kennedy, Manuel Mateos-de-Vicente, James L. Melancon, Peter Nussbaum, J. Frank Redus, E. Guy Robbins, Sam I. Thornton, Anwar E. Z. Wissa, Toyotoshi Yamanouchi

Committee on Soil-Bituminous Stabilization
Jon A. Epps, Texas A&M University, chairman
David D. Currin, Jack N. Dybalski, K. P. George, J. M. Hoover, Edwin H. Jones, Larry L. Kole, Sidney Mintzer, Raymond K. Moore, James A. Scherocman, Charles G. Schmitz, Leonard J. Stern, B. A. Vallerga, Thomas D. White, Anwar E. Z. Wissa, Walter H. Zimpfer

Task Force on Optimizing the Use of Materials and Energy in Construction
William B. Ledbetter, Texas A&M University, chairman
Grant J. Allen, James Douglas, Woodrow J. Halstead, A. A. Lahna, Oscar E. Manz, Charles R. Marek, Vaughn Marker, William F. McFarland, K. H. McGhee, W. S. Mendenhall, Jr., Lyndon H. Moore, E. Guy Robbins, Robert D. Schmidt, Richard W. Smith, R. R. Stander, T. C. Paul Teng, Marshall R. Thompson, Larry G. Walker, William A. Yrjanson

Bob H. Welch and John W. Guinnee, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each report. The organizational units and officers and members are as of December 31, 1978.

Transportation Research Circular 180

page 14, column 2, line 3 from bottom

Change "must be applied" to "must not be applied"

page 15, column 1, line 3 from bottom

Change "and 2 percent lime gives" to "but at least two consecutive percentages of lime give"

page 15, column 2, line 1

Change "If the highest pH is 12.30 and only 1 percent lime gives" to "However, if only the highest percentage checked gives"

NCHRP Synthesis of Highway Practice 62

page 34, Table A-7

Add as footnote "Note: In Alaska, Delaware, New Jersey, New York, and Rhode Island, highway-user revenues are placed in the state's general fund, from which the state's highway appropriations are taken."

Contents

Swedish Capacity Manual

Part 1. Objectives, Scope, and Arrangement of the Manual

Bo E. Peterson, National Swedish Road Administration

Different methods for calculating the capacity of a traffic facility have given different solutions to the same problem. At the same time, it has been clearly shown that the actual capacity is higher than that indicated by calculation methods. The Swedish Road Administration has therefore completed a research and development project that produced a new capacity manual, which should improve old methods, produce new ones for types of facilities and efficiency factors not previously treated, and achieve uniform application of the methods in Sweden. The new manual treats three categories of traffic: motor vehicle traffic, bicycle traffic, and pedestrians. The main efficiency factors are capacity, queue length, delay, and proportion of stopped vehicles. Explanatory models based on the queue theory of motorist behavior have been chosen to limit the empirical evidence to parameters such as critical time headway. The main new types of facilities are unsignalized intersections and bicycle and pedestrian facilities. For signal-controlled intersections, new developments have been made for left-turning traffic with opposing conflict, right-turning traffic with pedestrian conflict, various lane divisions, and calculation of cycle length. Another objective was to systematize the calculation of different measures of efficiency. The methods are reported as a series of steps in a computation.

The purpose of this and the following two papers is to present the new Swedish capacity manual, a comprehensive study undertaken by the National Swedish Road Administration and published early in 1977. This first part of the paper gives an overview of the objectives, scope, and arrangement of the manual. The second part by Arne Hansson presents the theoretical developments, field studies, and recommended methods for unsignalized intersections, and the third part by Karl-L. Bång deals with the same aspects for signalized intersections.

BACKGROUND

There was no Swedish traffic manual until the Swedish Transport Research Commission published the Academy of Engineering Sciences Bulletin 39 on the capacity of streets and roads in 1958. This report contained an analysis by Stig Nordqvist of the capacity problem and was based on available foreign literature—especially the 1950 Highway Capacity Manual (HCM)—and a number of Swedish studies. The manual adapted calculation methods and recommended capacity values for sections and different types of intersection. However, it was intersections with signal controls that were dealt with in most detail.

The publication of the 1965 HCM has meant, at least in Sweden, that a number of methods for the calculation of the capacity of a traffic facility have been in existence for the last few years. In certain cases, however, different methods have given different solutions to the same problem. At the same time measurements in several countries have clearly shown that actual capacity is higher than that indicated by the methods.

During the last 10 years there has been considerable research in this field in the United States and Europe as well as in Australia. Recently gained knowledge, however, has only been applied to formulating new calculation methods to a limited extent. Some proposals for improving the old methods have been made, but there has been no coherent overview or critical analysis.

In light of the above, the National Swedish Road Administration at the end of 1971 initiated a study aimed at developing methods to calculate the discharge capacity of road facilities. The purpose was to produce methods for types of facilities and efficiency factors not previously treated and also to achieve a uniform application of the methods throughout Sweden.

The first stage of the investigation was a study of the literature. The report itself (1) includes a review and analysis of about 900 references from numerous countries and is to my knowledge the most thorough overview now in existence.

The second stage involved working out the actual calculation methods on the basis of requirement specifications. At this stage Bång and Hansson developed a number of methods for calculating capacity, queue length, delay, and proportion of stopped vehicles for different road facilities. This work required that the geometry and traffic dependence of certain variables under Swedish conditions be explained, which demanded rather extensive measurement (2, 3).

Some of the results from these field surveys on critical time headway, for instance, have received international attention. Great interest in the survey's actual technique and equipment has also been shown.

As a final result of the second stage, proposals for calculation methods for each type of facility were received in the form of internal memoranda, which were then distributed to various selected Swedish traffic engineers in order to canvas opinion on the choice of method. The methods were tested in a number of case studies, and several different examples were checked.

The experience gained from seminars and the tests mentioned above were taken into consideration before the final manual was edited early in 1977 (4). Some computerization of the methods was carried out by Arne Hansson, who helped work out a dialogue computer program—CAPCAL—that calculates capacity and delay in an intersection for which the design and traffic load are known. Four different types of control—yield sign, stop sign, traffic signal, and traffic circle—can be studied.

OBJECTIVES

A main objective was to create a handbook, rather than a textbook of the discursive and explanatory type, that dealt with descriptions of consequences, known design, degree of accuracy, computational steps, and samples. One aim was to produce methods for describing consequences rather than for dimensioning alone. This also means that no recommendations for design standards, such as the 1965 HCM service levels, are given.

Highway and traffic engineering measures aim, among other things, at improving road facilities' discharge capacity, which is measured in such efficiency factors as capacity, queue length, delays, and proportion of stopped vehicles. Decisions on measures to be taken on a road network or parts of it should be based on a judgment of the consequences of these measures for society. This judgment should be the result of wide-ranging inquiry into not only discharge capacity but also factors such as road safety, cost, and environment. Existing calculation methods only provide knowledge of the effects of different measures on the discharge capacity and thus do not provide the complete basis required for decisions on dimensioning road facilities or on dealing with a road network.

Because methods for dimensioning were not developed, the design of a facility must be known or assumed. Traffic flow must also be known or assumed. If a facility is required to meet some specific demands expressed in one of the efficiency factors, an iterative calculation has to be made. When the design of the facility has been assumed, the efficiency factor can then be calculated for this design. Thereafter the design has to be altered and new efficiency factors recalculated until the design meets the requirements.

Resources in the road and street sector are today very limited. Each decision made must therefore be preceded by careful calculation to arrive at the amount of detail to be included in the calculation method. In the process, methods of a more schematic nature give results that are too approximate. Thus, considering the economic consequences of incorrectly dimensioned roads and streets, the amount of work needed for calculations according to the new methods is fully justified. This can also be expressed as choosing the degree of accuracy that is right both socially and economically.

Another objective was systematizing the calculation of different measures of efficiency—capacity, queue length, delay, and proportion of stopped vehicles. This is in line with the aim of producing methods for analyzing social and other consequences rather than for dimensioning alone. This systematization also makes it possible to add different delays together for the total delay of an intersection or street section. The methods are thereby reported as a series of computational steps for each efficiency factor.

Calculations for vehicle traffic facilities are made on special forms. For each type of facility, a complete sample calculation for a standard case is given, by which it is possible to ensure that one has completely understood the methods.

SCOPE

Three categories of traffic lie within the scope of the manual: motor vehicle, bicycle, and pedestrian. The types of facilities included and their efficiency factors that can be calculated are listed below. All common traffic facilities, such as streets, ramps, weaving sections, and different types of intersection, are covered, as are pedestrian sidewalks and bicycle paths. The primary efficiency factors treated are capacity, queue length, proportion of stopped vehicles, and delay. In some cases speed, travel time, and number of stops can also be calculated.

Type of Traffic	Type of Facility	Efficiency Factor
Motor vehicle	Road sections	Capacity
	Street sections	Travel time
	Ramps	Capacity
	Long weaving sections	Capacity
	Traffic circles and short weaving sections	Capacity, queue length, proportion of stopped vehicles, delay
	Unsignalized intersections	Capacity, queue length, proportion of stopped vehicles, delay
	Signalized intersections	Timing, capacity, queue length, proportion of stopped vehicles, delay
	Street network without signal synchronization	Capacity, number of stops, travel time
	Street network with signal synchronization	Effect on capacity, queue length, number of stops, travel time
Pedestrian	Sidewalks	Capacity, speed
	Steps	Capacity, speed
	Unsignalized crossings	Queue space, delay
	Signalized crossings	Timing capacity, queue space, delay
Bicycle	Streets with mixed traffic	Capacity
	Bicycle paths	Capacity, speed
	Unsignalized crossings with bicycle paths	Capacity, queue length, proportion of stopped bicycles, delay
	Signalized crossings with bicycle paths	Timing, capacity, queue length, proportion of stopped bicycles, delay

INNOVATIONS

One of the most important innovations in the calculation methods is that these are based not only on the results of regression analyses from a number of field surveys but also on the queue theory of motorist behavior.

The advantages of a theoretical model are better consistency, fewer required observations, and facility of updating.

The empirical evidence was thus limited to certain parameter values, for example, critical time headway. It should be pointed out in this connection that it is precisely the measurement of critical time headways that has, in the opinion of many, become more accurate than was previously the case.

The main new types of facilities are unsignalized intersections and facilities for bicycle traffic and pedestrians. The studies of the literature and the development of the methods for unsignalized intersections are presented in part 2 of this article. A new method has been produced for weaving sections of up to 40-60 meters. The model is similar to that developed for unsignalized intersections. It has been found that traffic circles operate in much the same way as series of T-intersections. Longer sections are calculated according to the 1965 HCM.

The methods for pedestrian and bicycle facilities are mainly applicable to those facilities that are separated from motor vehicle facilities.

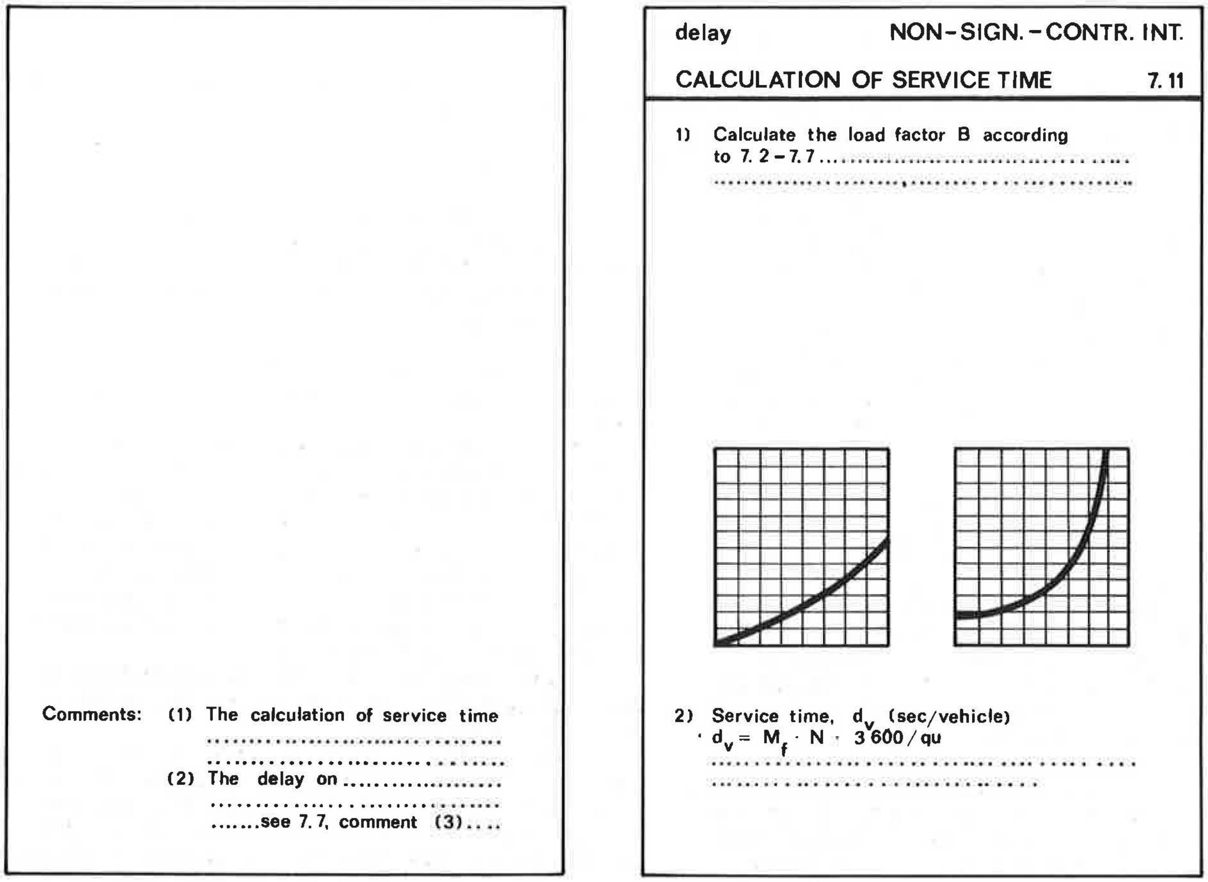
For pedestrian crossings, the methods are primarily developed for crossings between intersections. In some cases calculations can also be made for crossings at intersections, but no reliable method has been found for calculating the capacity of a pedestrian crossing when the pedestrian flow is in conflict with the vehicle flow.

The recommended methods for measuring the results of bicycle traffic in intersections are primarily intended to be used for intersections with special cycleways. All calculations for signalized intersections are based on the assumption that there are no conflicts with traffic (or pedestrians) during the green phase. For unsignalized intersections, however, instructions on how to calculate the effects of bicycle traffic mixing with vehicle traffic are given.

It may generally be stated that the empirical basis for the calculation method for pedestrian and bicycle facilities is less reliable than that for motor vehicle traffic. For instance, it has not been possible to provide any reliable calculation method for capacity at unsignalized pedestrian crossings. It should be stated that the results for bicycle traffic should be used with great caution where mixed traffic with many bicycles is concerned.

For signalized intersections, which are treated in part 3 of this article, four aspects of considerable importance have been distinguished: left-turning traffic with opposing conflict, right-turning traffic with pedestrian conflict, various lane divisions, and calculation of cycle length. Using a development of previous

Figure 1. Arrangement of the manual (right- and left-hand pages).



methods with these elements has meant that all the various designs of signalized intersections, from the very simple to the highly complex, can be dealt with.

Street networks are dealt with in two parts, those with and those without synchronized traffic signals. The efficiency factors in networks without such traffic signals are calculated by using the methods developed for other facilities. For networks with synchronized signals, a rough description of how coordination affects efficiency factors is given.

ARRANGEMENT

The methods are reported as series of computational steps, on the one hand for each efficiency factor (capacity, queue length, delay, proportion of stopped vehicles), and on the other within each efficiency factor (division into subapproaches for whole intersections; determination of critical time headway, service time, and part-capacity ratios for each vehicle stream in each subapproach; and determination of capacity ratio and capacity for each subapproach).

These steps are reported on the right-hand pages of the report, while notes, where necessary, are given on the facing left-hand pages (Figure 1). Thus a complete computational stage can be seen at a glance upon opening the report.

The purpose of this arrangement is for the right-hand pages to give the necessary instructions for carrying out the calculations, omitting deviations and explanatory arguments. The idea behind this is that the calculation process not be interrupted or the train of thought broken, but that easily accessible notes be provided. Insofar as possible, diagrams and tables have been provided as support material to the actual calculations.

Great importance has been attached to design, and great care has therefore been devoted to editing the methods. A secondary advantage may be that certain diagrams are of a size smaller than that demanded for ease of reading. This possible disadvantage is justified on the grounds that ease of understanding, on the other hand, has been achieved.

REFERENCES

1. Kapacitetsutredning: Literaturstudier och analys. Statens Vägverk, TV 118, 1973.
2. Trafikavlastning i usignalreglerade korsningar. Institute for Traffic Engineering, Univ. of Lund, 1975.
3. A. Hansson. Fältmätningar avseende fordons- och gångtrafikens framkomlighet i signalreglerade korsningar. Statens Vägverk, TV 127, Feb. 1977.
4. Beräkning av kapacitet, kölängd, fördröjning i vägtrafikanläggningar. Statens Vägverk, TV 131, 1977.

Part 2. Capacity of Unsignalized Intersections

Arne Hansson, Vattenbyggnadsbyrån (VBB), Malmö, Sweden

This paper presents a method for calculating capacity, queue length, and delay at unsignalized intersections that was developed for and recommended by the new Swedish capacity manual. The method is based on a queuing model that considers each lane in the approaches controlled by yield or stop signs as a service position. Service time is calculated, for each stream of vehicles in the lane, as a function of primary road flow rates and gap acceptance parameters. The lane capacity is assumed to be the reciprocal of the mean of these service times. Starting from the actual flow-to-capacity ratio, the queuing model estimates queue length distribution and mean delay. The method is believed to apply to most intersections controlled by yield or stop signs. The most important parameter in the model is critical headway, which was measured at 18 Swedish intersections. The results range between 3.3 and 7.5 s, depending primarily on direction, speed limit, and type of control; they generally agree well with those of previous studies. A model similar to the one for unsignalized intersections was developed for traffic circles with short weaving sections.

Intersections without traffic signals were rather summarily treated in the previous Swedish capacity manual (5), and in the 1965 HCM (6). Nevertheless, a considerable number of total traffic delays in urban areas occur in such intersections; thus their performance and design have attracted increasing interest from planners and engineers. Simultaneously, a number of field studies and theoretical works have facilitated the development of useful calculation procedures (1); some countries have already developed such procedures (7, 8).

In the work on the new Swedish capacity manual (4), whose background and objectives are described in part 1 of this article, the subject of unsignalized intersections was considered a priority area, and investments were made in theoretical developments as well as in field studies. This paper presents the results of some of those efforts.

As explained in part 1, the calculation method is descriptive in nature. The model chosen is theoretical and based on road-user behavior. It is believed that, for the problem being considered, such a model ought to achieve greater consistency, with regard to the numerous influencing factors, than a purely statistical empirical approach.

Intersections here mean at-grade intersections, excluding weaving sections and those ramp connections common to grade-separated interchanges. Traffic circles, which are attracting attention in Sweden now, are treated in a separate chapter in the manual. The method recommended by the manual for signalized intersections is presented in part 3.

OVERVIEW

The method recommended for calculating capacity, queue length, and delay is based on a simple queuing theory model. Each lane in the minor road is represented by an M/G/1 queuing system with Markov-type arrivals and a general distribution of service times.

Models of this type easily lend themselves to the generalizations required to analyze any real intersection with interactions among a number of traffic flows. Similar approaches, using either the M/M/1 or the M/G/1 model familiar to queuing theorists, have been used by Surti (9) and others. More sophisticated models, such as Tanner's (10), that do not require assuming independent service times cannot be generalized to the same extent. Although some of the simpler assumptions inherent in the suggested model can be corrected for by choosing suitable parameter values, some accuracy has thus been sacrificed in order to gain comprehensiveness and consistency with regard to the many influencing factors.

PROCEDURE

The basic method procedure is described in Figure 2, where only the basic interaction between one minor and one major road flow is considered; the latter proceeds uninterrupted. The scheme can be explained in the following terms.

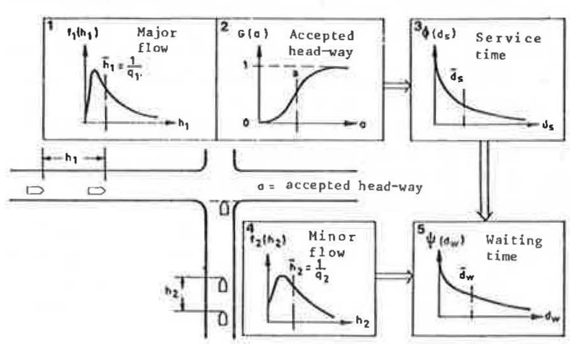
1. Frequency function $f_1(h_1)$ describes the frequencies of headways of different size (h_1) in the major road flow.
2. Distribution function $G(\alpha)$ expresses the probability that a headway (gap or lag) of size $h_1 = \alpha$ will be accepted by a waiting vehicle on a minor road.
3. Statistical distribution of the service time, d_s, depends on frequency of headways and probability of acceptance of occurring headways. The capacity, C, in vehicles per second, of the minor road approach must approximately equal the reciprocal of the mean service time $\bar{d}_s$, in seconds.
4. Frequency function $f_2(h_2)$ describes the frequencies of headways of different size (h_2) in the arriving minor road flow.
5. Statistical distribution of waiting time, d_w, i.e., delay minus a retardation and acceleration component, depends on service times for vehicles first in queue and on frequency of intervals between arrivals. The same applies to queue lengths: the mean queue length, $\bar{N}$, in vehicles is related to mean waiting time, $\bar{d}_w$, in seconds and arrival rate, q_2, in vehicles per second through

$$\bar{N} = \bar{d}_w \times q_2 \tag{1}$$

Input parameters for the model in Figure 2 are headway distributions in the two flows, with means $1/q_1$ and $1/q_2$, and distribution of accepted headways. In the calculation procedure, the function $f_2(h_2)$ is assumed to be exponential, while $f_1(h_1)$ depends on number of lanes and mean flow rate.

The gap-acceptance distribution, $G(\alpha)$, is simplified to a step function and defines a critical gap as $\bar{\alpha}$ seconds for one vehicle and assumes a constant increment for each additional vehicle. The values of the critical gap and the increment depend on variables describing intersection design and control and do, in

Figure 2. Analysis procedure for interaction between a major and a minor road flow.



fact, to a large extent, summarize the influence of such variables.

In an actual intersection, a number of such interactions are superimposed on each other; each interaction describes the crossing or converging movement between one major (uninterrupted) and one minor (interrupted) flow from a pair of lanes. The recommended procedure for calculating capacity, queue length, and delay for minor road approaches to an intersection considers these interactions in the specific order (Figure 3) that follows.

A. Minor road approaches are divided into lanes, combining two or more lanes from the same approach that have one or more directions in common. Left-turning vehicles from the major road have to be considered as well.

B. For each traffic stream from the lanes considered, a corresponding total primary flow (q_1) is calculated. This flow depends on the number of interactions in which vehicles from the secondary stream must yield. The influence of primary flows in all these interactions is reduced to a single value.

C. For each secondary stream, as above, a critical gap is calculated and will depend on direction of movement for the secondary stream.

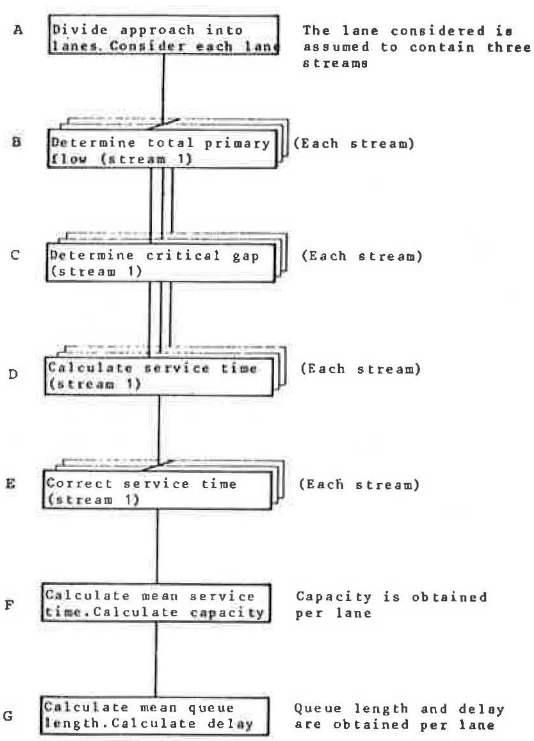
D. For each secondary stream, theoretical mean service time is calculated by using the results of B and C and an assumption regarding the headway distribution of the total primary flow.

E. Some service times thus calculated will have to be corrected, because some streams, such as a left-turning stream from the major road, are primary in some interactions but secondary in others. This corresponds to preemptive priority in queuing theory.

F. Mean service time and hence capacity can now be obtained for each lane. In case two or more streams share one lane, their mean service times are weighed according to their respective arrival rates. The capacities of adjacent unreserved lanes are simply added.

G. Mean queue length and mean waiting time are obtained for each lane by using the flow-to-capacity ratio and the queuing theory model. This model also gives the probability of stopping (at a yield-signed intersection) and then the total delay, as mean waiting time plus acceleration losses for stopped vehicles.

Figure 3. Steps in calculating capacity and delay for a minor road approach.



Such a calculation procedure of necessity contains a number of assumptions, some verified through extensive measurements or calculations, many not. A few of the more important of these assumptions will be discussed.

Conflict Patterns

One basic aspect of the manual is the unambiguous hierarchy of different vehicle streams in an intersection. It is assumed that signed or other legally stated conventions are observed, so that (for example) a vehicle traveling straight ahead on a major road approach is not affected by minor road vehicles.

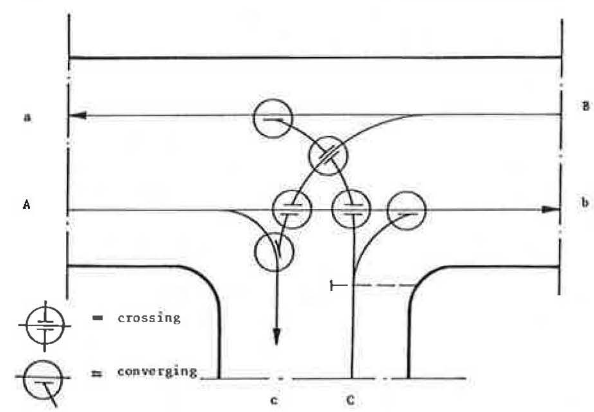
Measurements taken in Sweden indicate that this is true when the load factors (flow to capacity ratio) of the minor road approaches do not exeed 0.8-0.9. As load factors increase, some minor road vehicles tend to accept very short gaps, which forces the major road flow to slow down or even halt.

The manual does not apply to such conditions, which would normally, if they were frequent, warrant the installation of traffic signals.

For the T-intersection in Figure 4, it is thus assumed that stream Cb yields to Ab, stream Ca to Ab, Ba, and Bc, and stream Bc to Ab and Ac. The manual requires calculating an equivalent total primary flow for each secondary stream. In the case of one secondary stream's crossing different primary streams from one or more lanes, the flows of these streams are simply added together for total flow. Theoretically, this result would give a correct total headway distribution only when all the individual distributions are exponential. The mean will always be correct, though, and, as shown by Tanner (11) for the M/D/1 headway-generating model, the effect of "reasonable aberrations" is probably negligible.

In the case of converging conflicts, such as the one between Cb and Ab in Figure 4, it is generally unclear to what extent the stream with legal priority (in this case Ab) should be included as primary flow with re-

Figure 4. Interactions at a yield- or stop-signed T-intersection.



spect to the secondary stream (Cb). Where the exit has only one lane and no acceleration lane or broad shoulders used as such, all vehicles in the priority stream must be included, just as for a crossing conflict. Note, however, that the critical gap might be different.

When there is more than one lane in the exit where the convergence takes place, the priority flow effective in the conflict will be determined by the lane distribution of the priority flow and by a number of design features such as number of lanes in the approach, lane markings, and islands.

Rather than trying to consider all these factors, which, at least in the case of lane distributions, is virtually impossible in urban areas, the manual recommends that priority flow in a converging conflict be divided by the number of exit lanes before adding it to the total primary flow. This general rule may be modified or disregarded for special designs or known lane distributions.

Other assumptions have to be made regarding the effect of streams turning off the primary road from some secondary streams where no direct conflict occurs but where some influence is possible. In Figure 4, for example, it is likely that some right-turning vehicles from approach A prevent some gaps from being used by vehicles from approach C (by not signaling their intention to turn).

On the basis of the field studies described below, different hypotheses on the effects of such streams were tried. Although their presence did have some effect at smaller intersections, it was decided to exclude streams not in conflict with the secondary flow. Accordingly, critical gaps and lags were defined and measured at the point of conflict. It might be noted here that the German manual (8) differs in this aspect, including such streams as Ac in Figure 4 with a factor of 0.5 (also independent of geometry).

As an application of the principles outlined above, the total priority flows presented below are calculated for the example in Figure 4.

Secondary Flow	Corresponding Primary Flow
$q_2 = q_{Bc}$	$q_1 = q_{Ab} + q_{Ac}/n_c$
$q_2 = q_{Cb}$	$q_1 = q_{Ab}/n_b$
$q_2 = q_{Ca}$	$q_1 = q_{Ab} + q_{Bc} + q_{Ba}/n_a$

where

q_2 = secondary flow in vehicles per second;
q_1 = corresponding primary flow in vehicles per second;
q_{xy} = vehicle stream between approach x and exit y; and
n_y = number of lanes in exit y (or multiples of 3 m if lanes are not marked).

Critical Gaps

By the critical gap for vehicles in a certain secondary stream we mean a time headway in the total primary flow that is just acceptable to the average driver for passing through the intersection. This parameter affects capacity in a nearly exponential manner and further summarizes the influence of several interesting design variables (from a capacity and delay point of view). As will be described below, extensive measurements were taken to determine this parameter and its dependence on different variables.

The values recommended in the manual are based on these measurements for Swedish conditions, as well as on previous studies elsewhere (1, the bibliography). The most important determinants are stream or direction of travel, speed limit, and type of priority (stop or yield sign), the effects of which are expressed by the base values in Table 1.

Corrections for the following factors are added to or subtracted from the base values (Table 1): width of main road, existence of a median on main road, radius for right-turning vehicles, angle between major and minor roads, major road one-way or not, percentage of heavy vehicles, and size of urban area. Some of these factors affect the critical gap for some streams only.

The German capacity manual (8) considers the same factors as Table 1 plus main-road width for estimating critical gap. Although the influence of the different factors is similar, the recommended values are generally 0.5-1.5 s above those given in Table 1. This difference has not been verified as real or originating from differences in definitions and measurement techniques.

The use of a single parameter to describe gap-acceptance behavior merits some comment.

First, no distinction is made between lag (the first headway when arriving in an empty approach lane) and gap. The term "critical gap" as used in the manual implies some weighted average between critical gap and critical lag, as measured without distinguishing between the two.

Second, the concept of move-up time (the increment in headway a second queuing vehicle requires to follow the first vehicle) has not been mentioned. This parameter is as important as critical gap in estimating capacity or mean service time. From the measurements, we may conclude that move-up time comprises about 60 percent of the critical gap and depends on different factors in much the same way. This relation is implied in the manual without explicitly calculating move-up time.

Service Time and Capacity

Assuming that gap-acceptance behavior and any given total primary flow are as outlined, the capacity of a secondary stream can be calculated analytically thus:

$$C = q_1 \sum_{i=1}^{\infty} \left\{ \int_{a+(i-1)a_m}^{(a+i)a_m} f_1(h_1)\,dh_1 \right\} \tag{2}$$

where

C = capacity in vehicles per second;
q_1 = total primary flow in vehicles per second;
a = critical gap in seconds;
a_m = 0.6 a or move-up time in seconds; and
$f_1(h_1)$ = headway distribution of the primary flow, with mean $1/q_1$s.

Inverting this capacity value yields what may be called the mean service time during queuing conditions.

Table 1. Recommended base values for critical gaps.

Speed (km/h)	Sign	Primary Approach Stream: Left Turn	Secondary Approach Stream: Right Turn	Straight	Left Turn
50	Yield	5.0	4.8	5.2	5.3
	Stop	5.0	5.5	5.8	6.0
70	Yield	5.8	6.0	6.0	6.2
	Stop	5.8	6.5	6.5	6.8
90	Stop	6.5	7.2	7.0	7.5

If several streams share one lane, their mean service times have to be weighted according to arrival rates.

Several expressions for the headway distribution $f_1(h_1)$ in Equation 2 have previously been suggested and tried (1, 12, 13). Comparisons between these results and some from the measurements described below indicate that the shape of the distribution depends on several factors, some of which cannot be quantified in an urban area (platoon dispersions from nearby traffic signals or circles, curb frictions, and so forth).

Of the quantifiable factors, the number of streams or lanes was found to be the most important. Fewer lanes for the primary flow, as well as the presence of nearby traffic signals, increase platooning and thereby the coefficient of variation of the headway distribution.

The mean service time flowcharts in the manual are based on three different headway distributions with different coefficients of variation representing three classes of such conditions. The choice of these particular functions can be disputed, and further studies on this point are being undertaken. Regarding the effect of platooning, it should be noted that traffic signal densities are generally lower in Sweden than in the United States, for example.

The use of inverse capacity as mean service time will give an overestimate for nonqueuing conditions, because the constant move-up time is always included. The service time during nonqueuing conditions, which then equals total waiting time, can similarly be calculated for various headway distributions. The difference obtained equals move-up time, at zero primary flow, and then decreases as the primary flow (and hence the mean service time) increases.

In order to represent capacity correctly, this difference is disregarded in the calculation of mean service times but is compensated for on an approximate, average basis in the ensuing calculation of delay. Figure 5 shows an example of the resulting service time flowcharts for the case of primary flow in two lanes and no nearby traffic signals.

Where the secondary stream must cross or turn left onto the main road and a median allows this in two steps, the mean service time may be considerably reduced (and capacity increased). This effect has been calculated theoretically by Tanner (10), whose results are used in the manual.

Now we shall consider the case of one lane used by two or more secondary streams. If several streams jointly use several unreserved lanes, the calculations are performed with averages per lane. Instead of computing a weighted average of the mean service times, the manual directly computes a load factor defined as the ratio of actual flow to theoretical capacity. Remembering that mean service time, as defined here, equals the reciprocal of capacity, we obtain load factor as

$$B = \sum_k q_k (\bar{d}_s)_k \tag{3}$$

where

- B = load factor for a lane, or capacity ratio;
- k = index for streams utilizing the lane;
- q_k = flow rate in vehicles per second and per lane of stream k; and
- $(\bar{d}_s)_k$ = mean service time in seconds for stream k.

The products within the sum of Equation 3 represent the degrees of capacity use of individual streams in the lane and are called "part-capacity ratios" in the manual. Before adding these ratios in Equation 3, however, two types of corrections generally have to be made.

First, corrections for queuing in primary flows must be made. Some streams that are primary to the secondary stream are in turn secondary in some other conflicts; this prevents some otherwise acceptable gaps from being used by the secondary stream. The effect for the latter is the same as if the vehicles from this primary approach were, on arrival, immediately placed first in the queue of secondary stream vehicles. The effect can be estimated with the help of the theory for queues with preemptive priority. The size of the correction depends on the part-capacity ratio of those primary streams where queuing may occur. A similar correction is used in the German capacity manual, although with some differences in the approximations (8, 13).

Second, corrections for short lanes are made. Lanes of a limited length are initially ignored or, rather, considered part of the adjacent lanes. If the adjacent lane is thus used by two or more streams, the capacity ratio is reduced by an amount that depends on the degree of possible use of the short lane. This degree of use is estimated by elementary probability theory.

The capacity of the lane can now be obtained from actual lane flow and the load factor of Equation 3. It should be noted that this capacity is a theoretical concept and not necessarily the maximum lane flow. Increasing lane flow to such a level would entail assuming that other (primary) flows remain constant and may exceed the validity of some assumptions such as constant critical gap. Nevertheless, the load factor of

Figure 5. Calculation of mean service time from primary flow and critical gap in two lanes.

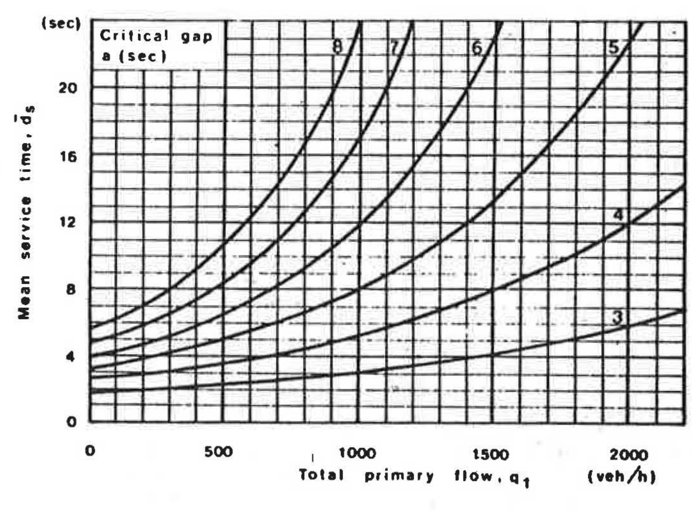


Figure 6. Relation between number of queuing vehicles and load factor for one lane.

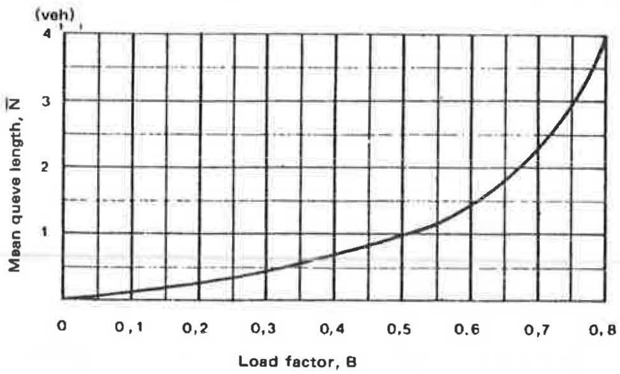
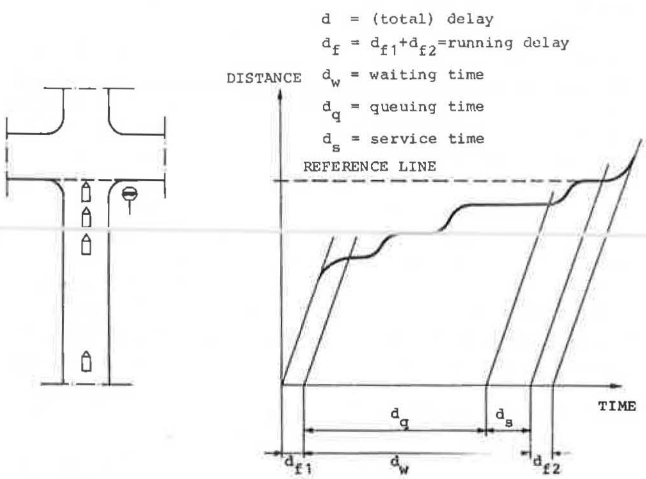


Figure 7. Delay illustrated.



Equation 3 is a valuable concept for analyzing intersection performance and can be further used for calculating queue lengths and delay.

Queue Lengths and Delay

Assuming independent service times from a distribution with a certain mean and variance and a Markov arrival process, the mean queue length in a lane can be obtained by Pollaczek's formula,

$$\bar{N} = B + [B^2/(1 - B) \times (1 + C^2)/2] \tag{4}$$

where

- $\bar{N}$ = mean queue length in vehicles, including the vehicle being served;
- B = load factor (see Equation 3); and
- C^2 = Var $(d_s)/(\bar{d}_s)^2$, or the coefficient of variation for the distribution of service times.

For C^2 = 1, the M/M/1 queuing model is obtained.

In a preliminary version of the manual, the variance of the service time distribution was estimated theoretically as a function of critical gaps and total primary flows in much the same way as the mean service time.

In the present version, however, the manual has been simplified in this respect, and the coefficient of variation, C^2 in Equation 4, is assumed to be a function only of the load factor. This gives a unique relation between mean queue length and load factor. This relation was calibrated partly by a simulation model, partly with the help of data from the measurements described later on. The final function is shown in Figure 6.

Compared to the M/M/1 model, Figure 6 sometimes gives slightly lower queue lengths, although the theoretical coefficient of variation is always larger than one. This is because service times are not quite independent; a shorter lag is required by a queuing vehicle (the move-up time is shorter than the critical gap). The differences are quite small, however.

A larger error results from assuming the coefficient of variation within a lane to be dependent only on average service time rather than on the mean service times of all the individual streams in the lane. This will cause some underestimates for lanes used by equal proportions of streams with widely different mean service times and some overestimates for reserved lanes.

The effect of this approximation has been accepted in view of the rather extensive simplification in procedure it makes possible. A slight improvement, substituting a family of curves representing different types of lanes for Figure 6, is being considered for future versions of the manual.

When designing an intersection, the higher fractiles of the queue length distribution, exceeded only by a few percent of arriving vehicles, are more interesting than the mean. These fractiles have been calculated as for the M/M/1 model, and mean queue length is obtained from Figure 6. Finally, queue lengths are converted from number of vehicles to meters.

The delay incurred by a minor road vehicle when passing through the intersection can be divided into different components, as shown by the time-space diagram of Figure 7.

Waiting time includes service time as well as time spent standing in queue and can be thought of as the flow-dependent part of delay. Its mean for a certain lane with one or more secondary streams is easily obtained by using

$$\bar{d}_w = \bar{N}/q_2 \tag{5}$$

where

- $\bar{d}_w$ = mean waiting time in seconds;
- $\bar{N}$ = mean queue length in vehicles from Figure 6; and
- q_2 = flow in vehicles per second in the lane being considered.

Equation 5 is quite general, although of course the results suffer the same constraints as do mean queue length calculations.

The remainder of total delay, called "running delay," is caused by acceleration and retardation. In the manual, running delay is calculated for each stream in the lane as a function of speed or speed limit, radius for turning vehicles, and probability of stopping. General parameter values are used for the first two variables, while the probability of stopping for a yield sign intersection is

$$P_s = 1 - P_f(1 - P_q) \tag{6}$$

where

- P_s = probability of having to stop;
- P_f = probability of accepting the first lag (depending on total primary flow and critical gap); and
- P_q = probability of queue (equal to load factor, according to queuing theory).

The running delay is averaged for all streams in the lane and added to the mean waiting time to give mean delay per lane. From this sum, however, must be subtracted a certain correction to compensate for double counting. Retardation loss and running time along the queue are also included in waiting time during queuing conditions.

The present version of the manual does not include calculation of queue length and delay for lanes on the primary road, with the exception of reserved lanes for left-turning vehicles. But extending the manual in this respect is being considered and has in fact already been done in a computerized version.

TRAFFIC CIRCLES

Studies of traffic circles in Sweden and England (14) indicate principal operating differences between weaving sections of different lengths. Where a length exceeds 30-45 m and capacity is sufficient, there is a "true" weaving in the sense used by the 1965 HCM. For such long weaving sections the Swedish manual recommends a procedure similar to that of the 1965 HCM.

For shorter weaving sections, which are frequent in Swedish traffic circles, a procedure much like the one described above for intersections is recommended. The studies mentioned show that such traffic circles actually operate as a series of T-intersections that gives priority to the traffic inside the traffic circle.

Capacity, queue length, and delay are therefore calculated for each approach in turn, as described in Figure 3, although with different charts and parameters.

There are two major differences, however. First, for a certain secondary stream the total primary flow in the different points of conflict is totaled, as for the intersection. In a converging conflict, however, the primary flow is reduced by a factor that depends on the geometry of the weaving section and of the exit used. Second, the critical gap depends primarily on the direction of the secondary stream (right turning, straight ahead, or left turning) and on the weaving section's width-to-length ratio.

In addition, only one chart (as in Figure 5) is given for mean service time. Very little data are available on headway distributions in traffic circles, and generally the calculation model for traffic circles is felt to be less reliable than the one for intersections. No doubt this is mainly due to more degrees of freedom in road-user behavior and hence to the decreasing reliability of behavioral models. Still, the method is believed to be an improvement over existing empirical formulas, perhaps most of all with regard to consistency.

MEASUREMENT OF CRITICAL HEADWAYS

Background and Purpose

Gap-acceptance behavior has been shown to be of vital importance in modeling the performance of an unsignalized intersection, and a number of previous studies have undertaken to define and measure parameters relevant to such behavior (1, 7, 15, 16). The literature survey preceding the manual gives a fairly complete bibliography on the subject and a comparison between and an analysis of results.

The main purposes of the measurements made on unsignalized intersections were to validate these results for Swedish conditions and to further study the influence of some design variables on gap acceptance. In most intersections studied, service time, delay, and similar variables were measured alongside gap acceptance parameters in order to obtain some preliminary tests of the model.

Definitions

Several different distribution functions for accepted headways can be defined and have been used in the literature (1, 17). These definitions vary with respect to the types of choice situation studied (gaps and lags or both, one or more vehicles accepting one gap, and so forth), statistical population defined (drivers or gaps), and points of reference for studying headways (intersection limits, conflict points, and so forth).

For these measurements, gaps and lags were referred to points of conflict between the primary and secondary vehicles involved. In the final analysis, gaps and lags were pooled to produce a flow-dependent average. Critical gaps were determined by using probit analysis with correction for bias according to Ashworth (17, 18). By using these definitions, the results are believed to be compatible with the theoretical models for calculating capacity and delay.

Sample

The sample should be representative of the most frequent intersection designs. It should also, insofar as possible, admit pair-wise comparisons of the effects of the more important variables, such as type of control (yield or stop), speed limit, cross section of a major road, lane types in secondary approaches, location (urban, rural), and size of flows. In order to facilitate such comparisons, intersections with more than normal disturbances of any kind were excluded.

Eighteen intersections were selected on the basis of a nationwide inventory. Total incoming vehicle flows to these intersections ranged from 6000 to 34 000 average daily traffic (ADT). Total number of lanes in all approaches ranged from three to twelve. Nine of the intersections are situated in urban areas (speed limits 50 or 70 km/h) and three in rural areas (speed limits 50 or 90 km/h). The study considers at-grade intersections only. No ramp connections with the acceleration lanes common to grade-separated interchanges were included, nor were any weaving sections or traffic circles.

Method

An automatic data-collection technique, based on detectors and a data log, was developed and used. In the minor-road approaches, presence-type magnetic loop detectors were installed. These were combined with pneumatic tube detectors in a logic circuit to minimize errors. Two detectors were installed in each lane, so that both the first and second vehicles in a queue could be detected. Passage times in the major road lanes were usually registered by pneumatic tube detectors. Manual switches were used to indicate choice of exit for passing vehicles, lane changes, and different types of disturbances. Figure 8 shows a typical installation in a T-intersection.

The effective analyzed time of observation in each intersection ranges from 6 to 10 h divided into continuous periods of 20-30 min and distributed over different peak and off-peak flow conditions.

Results

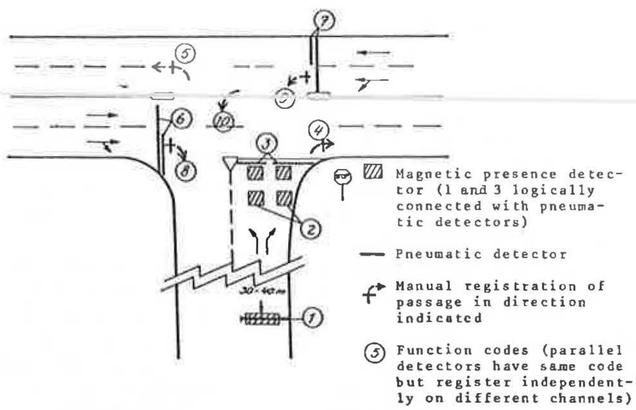
The critical headways obtained for minor-road approaches are listed in Table 2, together with values of the move-up time. Gaps and lags have been pooled in this table.

Table 2. Measurements of critical headways at 18 Swedish intersections.

Location	Control	Speed Limit (km/h)	No. of Approaches	Lanes on Major Road	Critical Headway (s) Right Turning	Straight	Left Turning	Move-Up Time (s)
Malmö	Yield	50	3	4	5.0	-	6.3	3.0
	Stop	50	4	6	5.1	6.0	6.1	3.5
Lomma	Stop	50	3	2	5.5	-	6.0	FO
Lund	Stop	50	4	2	5.6	FO	5.9	3.6
	Yield	50	4	2	5.9	5.1	FO	FO
Helsingborg	Yield	50	3	2-4	4.7	-	5.6	3.0
v15 - v23[a]	Yield	90	2	2	3.5	-	-	FO
v107 - v109[a]	Stop	90	4	2	6.0	7.0	7.4	FO
Hässleholm	Yield	50	4	2	4.6	6.0	6.0	3.1
Stockholm	Yield	50	3	4	4.0	-	4.9	2.6
	Yield	50	3	4	4.3	FO	5.1	2.9
	Yield	50	3	4	3.9	-	4.6	2.2
Eskilstuna	Stop	50	4	2	6.5	6.7	6.6	4.0
E4 - v218[a]	Stop	90	4	4	6.2	6.9	7.5	FO
Norrköping	Stop	70	4	6	3.7	5.3	5.2	3.5
	Yield	50	4	4	3.3	4.7	5.3	2.9
Västerås	Stop	70	4	4	6.7	-	7.3	4.2
	Yield	70	3	4	6.7	-	6.4	4.1

Note: FO = few observations.
[a] Road number.

Figure 8. Example of detector installation for measuring critical headways.



The confidence limits for the critical headways in the table vary between ±0.1 and ±1.0 s, and for the move-up time between ±0.1 and ±0.3 s.

The critical headways for vehicles turning left from a major road vary between 4.6 and 6.1 s for the four intersections where this parameter was studied (gaps and lags pooled). The highest value pertains to a road with a 90-km/h speed limit.

The definition of total primary flow—i.e., exactly which major road streams and lanes should define headways for a certain minor road vehicle—has a rather strong influence on critical headways. The variance of the gap-acceptance distribution was used as a criterion for choice of definition. A smaller variance would give a more relevant description of behavior.

The following main conclusions can be drawn from the measurements (see Table 2).

1. The general level of observed critical headways agrees well with previous but mainly foreign studies (1).
2. On the whole, the influence of the most important variables was also what could be expected from previous studies: the critical headways are higher at stop than at yield signs and higher with increasing speeds (speed limits) on the major roads and lower in big cities (Stockholm).
3. One point of difference may be the influence of major road width. Some previous studies (1) report substantially higher critical headways for crossing a four-lane road compared to a two-lane road. These results could not be confirmed here.
4. The differences obtained between crossing and left-turning vehicles from a minor road are somewhat less than could be expected.

The results regarding service time and delay were averaged over continous 20-min periods and compared with different theoretical models, using a sample of 55 such periods with different flow conditions and from different intersections. For mean service time, and comparing with the theoretical model finally selected and described above (with actual critical gaps), the following regression line (R^2 = 0.85) was obtained:

$$\bar{d}_{sm} = -1.11 + 0.96\,\bar{d}_{st} \qquad (7)$$

where indexes m and t denote measured and theoretical values, respectively. No systematic errors could be found. The agreement seems to be good, although of course the error in estimating critical gap remains to be added.

COMMENTS

The calculation method presented above is believed to be a valuable tool for analyzing intersection performance as well as for designing intersections themselves. Like most such tools, however, the method presumes some skill and good judgment on behalf of the user. The Swedish capacity manual, perhaps more than most such manuals, leaves it to the user to decide the context in which the method is to be used and how to apply the results.

Flexibility has also been an important objective in constructing the method. New research results and practical experiences are expected to continue flowing in, and it should be possible to accommodate these in the manual without changing its entire structure. The stepwise calculation process selected is believed to be useful in this respect.

The accuracy of the method cannot yet be judged, although some preliminary results appear promising. Much work remains to be done before the manual can be considered to be validated.

The responsiveness of the method to changes in various design parameters is perhaps as important as overall accuracy. The attempt to base the calculation process on an explanatory model for road-user behavior, rather than on a merely statistical model, ought to favor this responsiveness.

REFERENCES

5. S. Nordqvist. Gators och vägars kapacitet. IVA, TFK-medd, No. 39, 1958.
6. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
7. Capacity of At-Grade Junctions. Organization for Economic Cooperation and Development, Research Group T8, 1973.
8. G. Jessen. Ein Richtlinienvorschlag für die Behandlung der Leistungs-fahigkeit von Knotenpunkten ohne Signalregelung. Strassen Verkehrs Technik, Vol. 12, No. 7-8, 1968.
9. V. Surti. Operational Efficiency Evaluation of Selected At-Grade Intersections. HRB, Highway Research Record 321, 1970, pp. 60-73.
10. J. Tanner. A Theoretical Analysis of Delays at an Uncontrolled Intersection. Biometrika, Vol. 38, 1962, pp. 383-392.
11. J. Tanner. The Capacity of an Uncontrolled Intersection. Biometrika, Vol. 54, 1967, pp. 657-658.
12. K. Harders. Leistungsfähigkeit nicht signalregelter städtischer Strassenknoten. T. H. Hannover, thesis, 1976.
13. W. Siegloch. Die leistungsermittlung an Knotenpunkte ohne Lichtsignalsteurung. Strassenbau und Strassenverkehrstechnik, Vol. 154, 1973.
14. E. Grant. Traffic Capacity of Small Roundabouts. Traffic Engineering and Control, Vol. 10, No. 12, 1969.
15. F. Wagner. An Evaluation of Fundamental Driver Decisions and Reactions at an Intersection. HRB, Highway Research Record 118, 1966, pp. 68-84.
16. P. Solberg and J. Oppenlauder. Lag and Gap Acceptance at Stop Controlled Intersections. HRB, Highway Research Record 118, 1966, pp. 48-67.
17. A. Miller. Nine Estimators of Gap Acceptance Parameters. Proc., 5th International Symposium on the Theory of Traffic Flow and Transportation, Berkeley, 1971.
18. R. Ashworth. The Capacity of Priority Type Intersections With a Non-Uniform Distribution of Critical Acceptance Gaps. Transportation Research, Vol. 3, No. 3, 1969.

Part 3. Capacity of Signalized Intersections

Karl-Lennart Bång, Swedish Transport Research Commission,
Royal Academy of Engineering Sciences

Calculations of signal timing and capacity have been performed in Sweden by using methods based on the 1950 and 1965 Highway Capacity Manuals. These methods, however, give results that differ by as much as 50 percent from observed flows. A comprehensive development was therefore undertaken by the Swedish National Road Administration from 1971 to 1976. After extensive literature reviews, theoretical analyses, and field studies, a method was developed for calculating signal timing, capacity, queue length, proportion of stopped vehicles, and delay. The method is based on calculating saturation flows separately for each lane, which makes the method applicable to all geometric designs and phasing schemes. Adjustment factors for vehicle-actuated control are also included. Study emphasis was on the relations among approach width, lane markings, and capacity; conflicts between left-turning and opposing flows; and conflicts between turning vehicles and pedestrians. The signal-timing method is based on the minimum average delay criterion proposed by Webster.

For 20 years the calculations of signal timing and capacity done in Sweden have been based on a manual developed by Nordqvist in 1958 (5). This manual was, in turn, largely based on the 1950 Highway Capacity Manual (HCM). Some attempts have also been made to promote the use of the 1965 HCM (6).

These manuals, however, have been found to have some serious drawbacks. They do not give optimum cycle time, and capacity values often differ from observed flows by as much as 30-50 percent. Furthermore, no guidance is provided for calculating more complex geometric designs and signal phasing schemes.

To overcome these deficiencies, a comprehensive study including the development of calculation methods for roads as well as for different types of unsignalized and signalized intersections was initiated in 1971 by the Swedish National Road Administration (NRA). Chief investigator for the work was Professor Stig Nordqvist at Vattenbyggnadsbyrån (VBB); I was responsible for the part covering signalized intersections; and Arne Hansson was responsible for unsignalized intersections. The manual was completed and published in Swedish in 1977 (4).

An overview of the different phases of the work is presented as part 1 of this paper.

SCOPE AND OBJECTIVES

The manual was designed for calculations of capacity (defined as maximum flow at given conditions), queue length, proportion of stopped vehicles, and delay. These calculations serve to describe the consequences of a given set of geometric and traffic parameters. No recommendations for design standards are given, because the manual does not include all aspects, such as safety, capital costs, and external effects on the environment, that would have to be considered.

The manual does include forms and examples for practical application. The purpose of this paper is to present an overview of the proposed method and some of the material dealing with signalized intersections.

The method is based on theoretical models supported by a limited number of field surveys. Each model explains the behavior of traffic in a critical conflict or geometric design that affects the discharge rate at the stop line. By identifying the true reasons for each effect, adjustment factors such as city size and location within the city can be neglected or given only minor importance. This should greatly improve the accuracy of the results as well as the possibility of their being reproduced.

METHOD PROCEDURE

Each lane is treated individually in the method. This means that the manual can be used for practically all intersection layouts and phasing scheme designs.

The computational procedure involves the following basic steps:

1. Preparation

a. Determination of phasing scheme, lane division, and lane types.

2. Signal Timing: First Round

b. Calculation of the saturation flow (s) for each lane. Saturation flow, s, is defined as the highest stable flow in vehicles per hour of green (vphg) during existing conditions. First a base value of s is obtained. It takes into account the proportion of turning vehicles and the degree of conflict with other vehicles and pedestrians in the intersection with green in the same phase. This base value is then adjusted for conditions other than normal regarding the width, slope, length, and markings of the lane, the proportion of heavy vehicles, and so forth.

c. Distribution of the flow on different lanes. If the flow in a lane, q_i, is not known, q_i is calculated so that equal values of q_i/s_i are obtained for the adjacent lane or lanes with the same direction in the approach (Figure 9). This assumes that a driver arriving at the approach selects the lane that will minimize his or her delay before he or she crosses the stop line.

d. Identification of the critical conflict point in the intersection. The q_i/s_i ratios are calculated for each lane, and the critical conflict, defined as the point having the highest total of ratios q_i/s_i of adjoining lanes, $(q_i/s_i)_{max}$, is identified (Figure 10). The sum represents the degree of saturation of the intersection if there are no losses in effective green time.

e. Calculation of cycle, split, and green times. These calculations, carried out according to Webster (19), give the signal timing that results in minimum average delay for fixed time operation of the signals.

3. Signal Timing: Second Round

f. Renewal of steps b-e for improved accuracy. The saturation flow for lanes having turning traffic in conflict with opposing vehicle flows or having pedestrian crossings with green in the same phase is a function of the length of the green time and the cycle time. The second round includes procedures for more precise calculations of s for such lanes. The signal timing and other parameters are derived in the first round (b-e) as input. This is only necessary if consequences for individual lanes, such as queue length or delay, are to be calculated.

4. Consequences

g. Calculation of capacity, queue length, proportion of stopped vehicles, and delay. Once s, q, and the signal timing obtained from steps a-f are known, calculating the demanded consequences is a simple arithmetic procedure. Capacity is defined as the highest stable flow in vehicles per hour (vph) during existing conditions.

The manual is divided into a number of separate work moments that are numbered consecutively in the order in which they are to be performed. In Figure 11 a flowchart of the method is illustrated.

Figure 9. Method used to distribute the flow between lanes in the same direction.

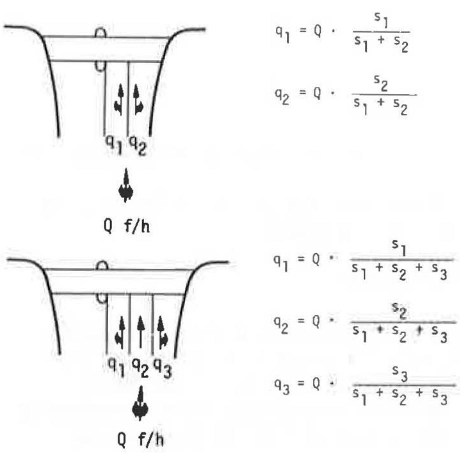
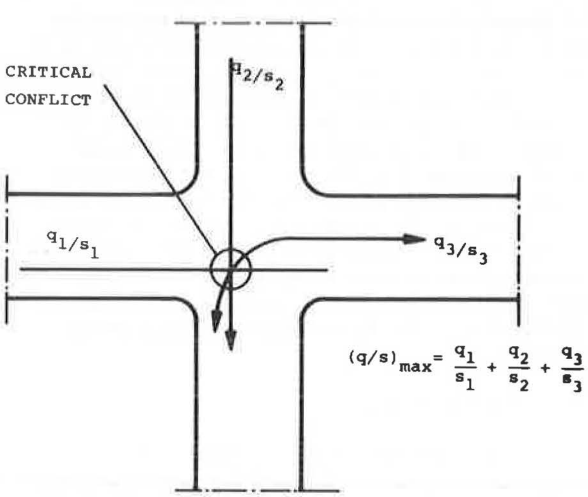


Figure 10. Definition of critical conflict.

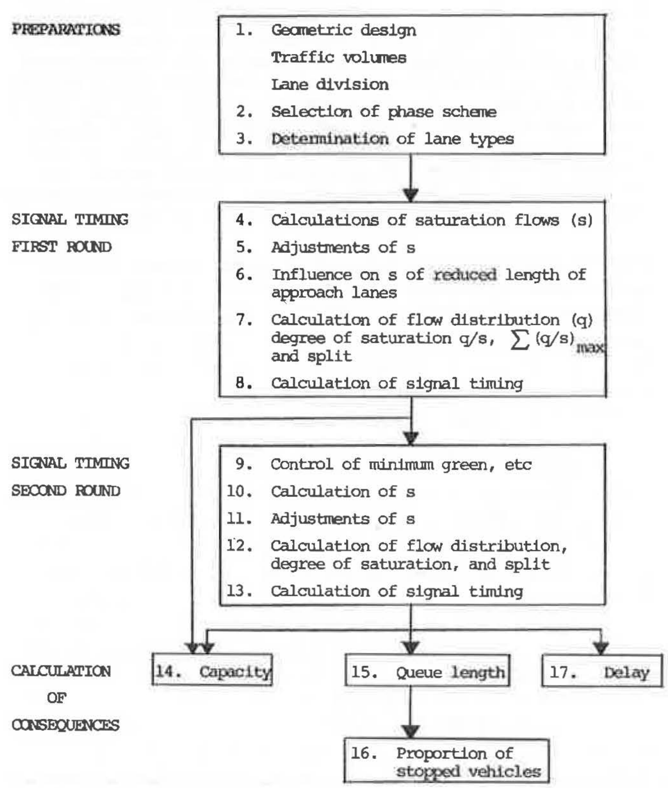


IMPORTANT FEATURES

Lane Configuration and Classification

The models that form the basis of the method assume that each lane carries a single flow of vehicles through

Figure 11. Flow chart for the manual.



the intersection. It is therefore important to make sure that actual traffic behavior corresponds to the lane markings. If this is not the case, the calculations should be performed with the number of lanes actually formed by moving traffic.

If field observation in the actual intersection is not possible, one should assume a minimum lane width of 2.5 m.

Figure 12. Relations among saturation flow, approach width, and number of marked lanes.

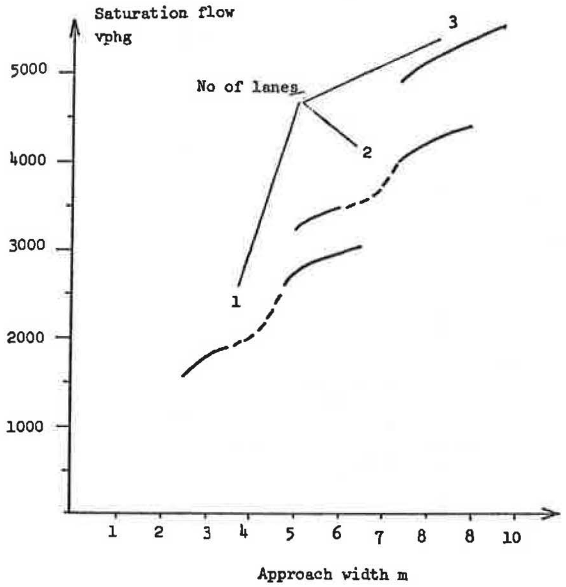
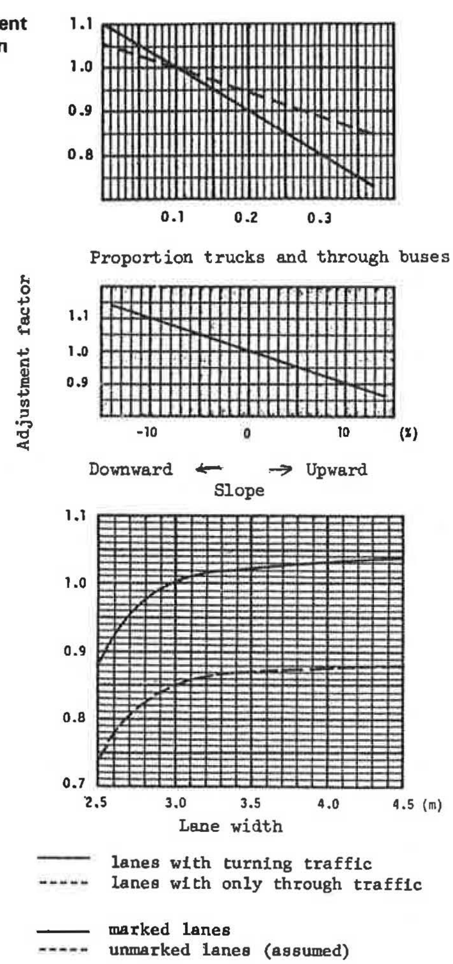


Figure 13. Adjustment factors for saturation flow.



The relations among capacity, approach width, and number of marked lanes in the approach have been evaluated by means of before-and-after field measurements at intersections with different lane arrangements. Figure 12 shows examples of results from these experiments and indicates that capacity is a function of both approach width and number of lanes. The model deals with this by making separate calculations of the capacity for each lane and by adjusting for lane widths other than 3.0 m and for absence of lane markings (Figure 13).

To simplify handling of the method, the lanes are classified as different types, depending on the presence of turning traffic and the degree of conflict experienced by the traffic in the lane. In Figure 14 seven different lane types are defined; lanes with turning traffic that conflicts with both opposing vehicle traffic and pedestrians discharged in the same phase may be assigned to more than one type, for instance to D/F or E/G.

Base Values for Saturation Flow

The base value for the saturation flow (s) for lane type A (only through traffic) is set at 1700 vphg and for type C (only turning traffic without conflict) at 1500 vphg. For type B (some turning traffic without conflict) s varies between 1700 and 1500 vphg, as a function of the percentage turning (0-100 percent).

For lane types with some degree of conflict between turning traffic and opposing flows or pedestrians (types D, E, F, G), determining s is more complicated.

Conflicts Between Left-Turning and Opposing Flows

From practical experience and limited field measurements it was concluded that a primary source of error in older manuals was inadequate handling of conflicts between left-turning vehicles and opposing vehicles on two-way streets. This is particularly important for normal, two-phase controlled intersections where all left turners have to face this conflict. In order to overcome this weakness, a thorough analysis was performed, and a model, suggested by Gordon and Miller in 1966 (20), was applied for stepwise calculations of the different stages of the conflict (Figure 15). These stages are as follows:

1. First part of the green phase, g_k, is when queue discharge from the opposite direction blocks left-turning vehicles in the lane. During g_k, only through or right-turning vehicles, N_k, can be discharged from the lane and then only until the lane becomes blocked by queueing left turners.
2. Remainder of green, $g_g = g - g_k$, is the time during which left-turning vehicles can pass when acceptable gaps occur in the opposing flow. The total discharge during this stage is noted as $N_g = g_g \times s_g$.
3. Inter-green period is that time during which vehicles stalled in the intersection can pass, N_r.

Provided that the signal timing is known from the first round of calculations, g_k can be estimated as the time required to discharge the queue that has been formed in the opposite direction during the previous red phase:

$$g_k = [q_m(c - g)]/s_m \times 1/(1 - q_m/s_m) = [q_m(c - g)]/(s_m - q_m) \quad (8)$$

Figure 14. Classification of lane types.

TYPE	DESCRIPTION	EXAMPLE
A	ONLY THROUGH TRAFFIC	
B	SOME TURNING TRAFFIC WITHOUT CONFLICT	
C	ONLY TURNING TRAFFIC	
D	SOME TURNING TRAFFIC IN CONFLICT WITH PEDESTRIANS	
E	ONLY TURNING TRAFFIC IN CONFLICT WITH PEDESTRIANS	
F	SOME TURNING TRAFFIC IN CONFLICT WITH OPPOSING TRAFFIC	
G	ONLY TURNING TRAFFIC IN CONFLICT WITH OPPOSING TRAFFIC	
D/F	SOME TURNING TRAFFIC IN CONFLICT WITH OPPOSING TRAFFIC AND PEDESTRIANS	
E/G	ONLY TURNING TRAFFIC IN CONFLICT WITH OPPOSING TRAFFIC AND PEDESTRIANS	

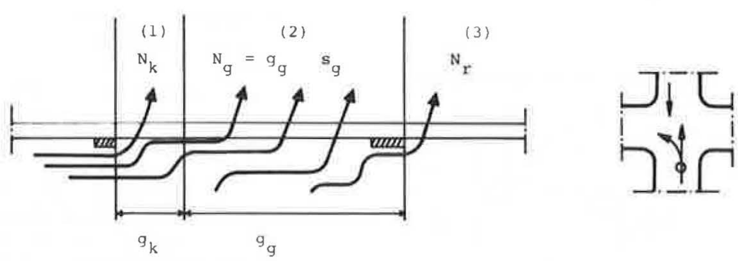


Figure 15. Calculation of saturation flow for lanes with left turning vehicles.

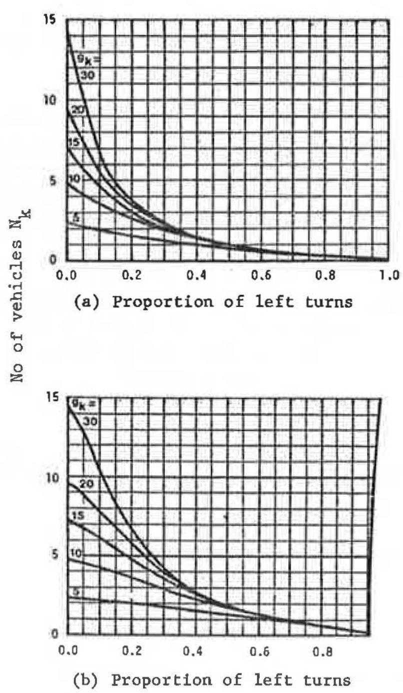


Figure 16. Number of vehicles, N_k, that can be discharged during g_k.

where

g = length of green phase in the opposite direction;
c = cycle time;
q_m = flow vph in opposite direction; and
s_m = saturation flow vphg in opposite direction.

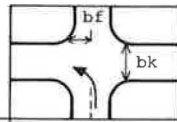
The term $1/(1 - q_m/s_m)$ accounts for the cumulative effect of vehicles arriving at the approach before the queue has been completely discharged. Numerical values for all variables are obtained from the first round of calculations for the opposing lane with the highest q/s ratio.

With g_k known, the number of vehicles, N_k, that can be discharged is a function of the proportion of left-turning vehicles, p_1, in the lane and the number of left-turners that can queue in the intersection without blocking other vehicles in the same lane. N_k can be solved by using general probability theory with the following results in Figure 16, where (a) is the case where no left-turning vehicle can queue without blocking the lane:

$$N_k = \sum_{i=1}^{N-1} [i \times p_1 \times (1 - p_1)^i] + N \times (1 - p_1)^N \tag{9}$$

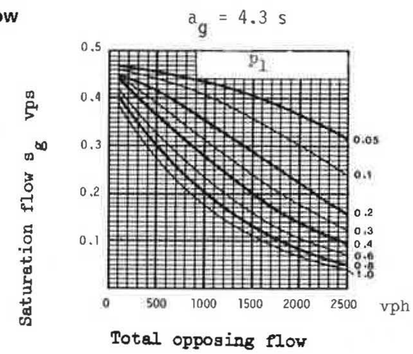
and (b) is the case where one left-turning vehicle can queue without blocking the lane:

Figure 17. Accepted critical gap, a_g, in the conflict between left-turning and opposing vehicles.



		CROSS-STREET WIDTH bk (m)	
		≤10	> 10
OPPOSITE APPROACH WIDTH bf (m)	≤ 8	4.8	5.3
	>8	5.3	5.8

Figure 18. Saturation flow s_g during g_g.



$$N_k = \sum_{i=1}^{N-2} [i(i+1) \times p_1^2 \times (1 - p_1)^i] + (N-1) \times N[p_1 (1 - p_1)^{N-1}] + N \times (1 - p_1)^N \qquad (10)$$

where N is the maximum number of vehicles that can be discharged during green ($N = g \times s$). The discharge, N_g, during g_g is a function of g_g, the proportion of left-turning traffic, p_1, and the possibilities for left-turners to cross in gaps in the opposing flow. This conflict is similar to the one that occurs in an unsignalized intersection. The capacity for left turns is in this case a function of the accepted critical gap α_g and the distribution of gaps in the primary stream.

Measurements of α_g were performed at 10 signalized intersections with a method described by Hansson in part 2; α_g was found to vary as a function of the width of the opposite approach, bf, and the cross street, bk, as presented in Figure 17. For busy intersections in large cities, the values obtained for α_g were 0.5 s lower than the values in Figure 17.

The saturation flow s_g during g_g can be expressed as

$$s_g = [q_m \times (e^{-\alpha_g \times q_m})]/[1 - (e^{-\alpha_f \times q_m})] \qquad (11)$$

where α_f is move-up time, which is 1/s.

In the field studies α_f was found to be closely correlated with α_g ($\alpha_f = 0.54\ \alpha_g$). In the manual this is presented in a separate graph for each value of α_g (Figure 18), where critical gap $\alpha_g = 4.3$ s.

The value for N is then obtained as

$$N_g = g_g \times s_g \qquad (12)$$

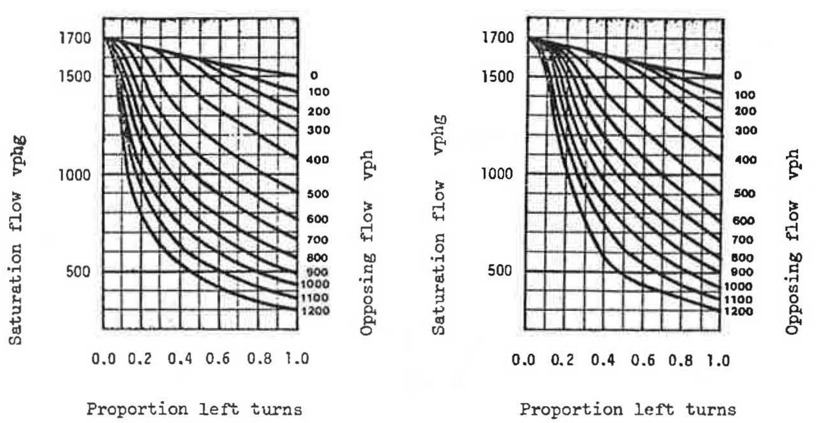
The number of vehicles discharged during inter-green, N_r, is obtained from measurements or observations of the space for queueing vehicles inside the intersection. In Sweden an average space of 8 m per vehicle is assumed. Care should also be taken to ensure that the inter-green is sufficiently long to allow these vehicles to clear the intersection before the next phase begins. Otherwise the intersection can become seriously blocked.

The resulting saturation flow for the lane in vphg is obtained as

$$s = (3600/g)\ (N_k + N_g + N_r) \qquad (13)$$

The complex and thorough procedure described above can only be carried out when signal timing and saturation flows have been estimated in the first round of calculations. The manual gives two simplified graphs for the first calculation round (Figure 19) for this purpose that are based on series of calculations with the second-round procedure for normal types of intersections. In these graphs s is determined by using the ratio of left-turning vehicles and the total opposing flow as inputs.

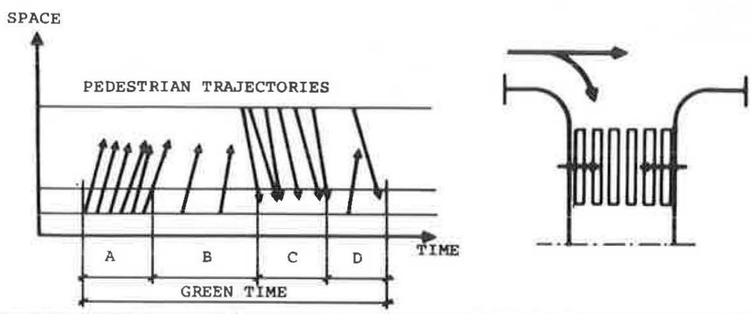
Figure 19. Saturation flow for lanes with left-turning vehicles in conflict with opposing flow.



No left-turning vehicle can queue without blocking the lane

One left-turning vehicle can queue without blocking the lane

Figure 20. Calculation of saturation flow for lanes with right-turning movements.



Conflicts Between Turning Vehicles and Pedestrian Movements

In two-phase signal control for intersections of two-way streets, right-turning as well as left-turning vehicles are usually affected by pedestrian movements in the crosswalk of the leg into which the turn is being made. This is particularly true for Sweden, where right turns on red are not permitted. An attempt was therefore made to analyze these conflicts and to develop a theoretical model as a basis for a computational procedure in the manual (3).

Collecting data with time-lapse photographs combined with inductive loops for vehicle detection was performed in four intersections with the described conflicts. A theoretical model was then developed to describe the behavior of a vehicle and a pedestrian as a function of the relative time advantage α_{by} for the party first arriving at the collision point. The reason for this was that it was felt that the right-of-way for the pedestrians during green was not fully respected by the turning vehicles. The model was therefore designed to enable tests of alternative hypotheses of this behavior.

The results from the field studies were evaluated manually, and α_{by} was derived by probit analysis. The vehicles usually required a positive time advantage of 3.2 s if they were not to give way to a pedestrian (standard deviation 4.3 s).

A capacity model based on division of the green phase into four parts was developed. Figure 20 shows A and C, where no vehicles can pass because pedestrian platoons have formed during red, and B and D, where a random arrival of pedestrians into the conflict zone is assumed. Turning vehicles can pass if their time advantage is sufficiently great.

The capacity of the turning flow can be expressed as

$$K_f = (T_D - T_C + T_B - T_A) \times \{q_p [(e^{-4.2q_p})/(1 - e^{-3q_p})]\} + N \qquad (14)$$

where

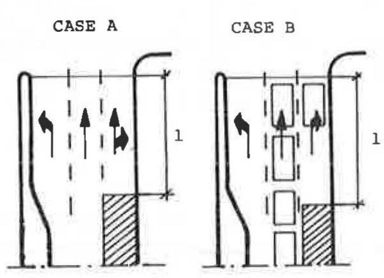
K_f = largest number of turning vehicles that can pass per green phase per lane;
$T_A - T_D$ = starting point of intervals A-D;
q_p = pedestrian flow, in number of pedestrians per second, two-way; and
N = number of turning vehicles that can pass during inter-green.

$T_A - T_D$ are functions of the length of the crosswalk and of the size and diffusion of the pedestrian platoons formed during the red phase. The equation, however, includes too many variables to allow graphic presentation. It has therefore been evaluated for a number of typical cases and the results have been put together in Table 3.

Effect of Restricted Length of Approach Lanes

In built-up areas, the length of the curb lane is often limited by parking, bus stops, or narrowing street width. If the curb lane in an approach is too short, capacity can fall. Two different cases can be distinguished (Figure 21).

Figure 21. Restricted length of approach lanes.



Case A. Curb Lane Serving Right-Turning and Through Vehicles

A reduction in saturation flow, s, of the curb lane occurs if the available lane length, l, is smaller than the space occupied by the maximum number of vehicles (8 m per vehicle) that can be discharged during the green time, g.

$$l/8 < (s \times g)/3600 \qquad (15)$$

The reduced saturation flow s' becomes

$$s' = (3600 \times l)/(8 \times g) = (450 \times l)/g \qquad (16)$$

Case B. Curb Lane Serving Only Right-Turning Vehicles

In this case the queue formed during red in the nearby lane might block the curb lane from being used to its full length. The likelihood of such blocking occurring is a function of l, g, and the ratio of right-turning vehicles in combined curb lane and adjacent lane. The number of right turns, N, that can be made during a green phase has been solved with probability theory (Figure 22).

The adjusted saturation flow s' is obtained as

$$s' = N \times (3600/g) \qquad (17)$$

If s' < s for the curb lane, s' should be used in further calculations. In this case the curb lane and the nearby lane are considered one lane with a saturation flow equal to the sum of the individual saturation flows for each lane.

The method can also be applied with minor modifications to lanes for left turners.

Signal Timing

The signal timing method the manual uses was developed by Webster (19) and is based on minimum average delay. An important factor in this method is the estimation of the loss of effective green time per cycle, F. For this purpose the manual includes methods for determining safe clearance times between phases set according to rules established by the Swedish National Traffic Safety Administration.

Traffic signals in Sweden display green plus amber at the end of green. Field measurements have shown that this interval is fully utilized by drivers and that it has therefore been included in the effective green in the manual. An inter-green interval thus consists only

Figure 22. Number of vehicles per cycle that can be discharged from a curb lane for right turns with reduced length.

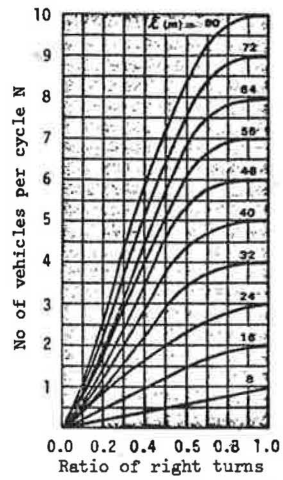


Table 3. Saturation flow of lane with turning vehicles in conflict with pedestrian movements.

		Green Time (s)								
		10			15			20		≥30
	Two-Way Pedestrian Flow	Length of Crosswalk (m)								
Turns	per Hour	7	14	21	7	14	21	7	≥14	All Widths
0	250 - ≥1500	1700	1700	1700	1700	1700	1700	1700	1700	1700
0.05	250	1580	1650	1640	1650	1670	1650	1660	1670	1680
	500	1500	1620	1600	1610	1660	1640	1640	1670	1670
	750	1390	1570	1580	1560	1570	1580	1600	1600	1630
	1000	1330	1490	1560	1510	1490	1510	1550	1540	1600
	≥1500	1290	1490	1560	1380	1120	1490	1450	1500	1540
0.10	250	1530	1620	1600	1610	1650	1640	1630	1660	1670
	500	1400	1580	1530	1540	1620	1600	1590	1640	1650
	750	1260	1480	1500	1460	1500	1490	1520	1550	1580
	1000	1150	1380	1460	1380	1370	1390	1450	1450	1520
	≥1500	1090	1380	1460	1170	1280	1350	1280	1350	1420
0.15	250	1430	1580	1560	1560	1630	1600	1590	1630	1650
	500	1270	1510	1460	1470	1580	1540	1540	1610	1620
	750	1070	1390	1400	1360	1410	1400	1440	1470	1530
	1000	960	1230	1360	1250	1230	1270	1350	1340	1450
	≥1500	890	1230	1360	1010	1420	1220	1130	1200	1300
0.20	250	1390	1560	1520	1520	1620	1580	1560	1610	1640
	500	1200	1480	1400	1410	1550	1510	1490	1590	1600
	750	1010	1320	1340	1290	1350	1340	1380	1430	1490
	1000	870	1160	1280	1160	1140	1180	1270	1260	1380
	≥1500	800	1160	1280	890	1030	1120	1020	1110	1210
0.30	250	1300	1500	1450	1440	1580	1530	1510	1580	1610
	500	1040	1390	1280	1300	1490	1420	1410	1540	1560
	750	850	1200	1210	1150	1240	1220	1270	1330	1410
	1000	700	1000	1140	1000	980	1020	1130	1120	1260
	≥1500	630	1000	1140	720	860	960	850	950	1060
0.40	250	1230	1460	1390	1380	1550	1480	1450	1550	1580
	500	930	1320	1190	1200	1430	1350	1330	1490	1510
	750	720	1100	1110	1030	1150	1130	1170	1250	1330
	1000	590	880	1030	880	860	900	1010	1000	1160
	≥1500	520	880	1030	600	740	840	730	830	940
0.50	250	1180	1410	1340	1330	1510	1450	1420	1520	1550
	500	830	1240	1100	1120	1380	1290	1260	1440	1470
	750	640	1000	1020	940	1070	1050	1090	1180	1270
	1000	500	790	930	780	770	810	920	910	1070
	≥1500	450	790	930	520	650	740	640	730	850
0.60	250	1130	1370	1300	1280	1480	1410	1380	1490	1520
	500	750	1180	1030	1050	1320	1230	1200	1400	1440
	750	570	940	950	870	990	980	1020	1120	1210
	1000	440	710	860	710	690	730	840	830	1000
	≥1500	390	710	860	450	570	670	570	660	770
0.80	250	1040	1300	1110	1200	1410	1350	1300	1450	1480
	500	640	1070	910	930	1230	1120	1090	1320	1360
	750	470	830	820	750	910	870	900	1020	1120
	1000	360	600	730	590	580	620	720	710	880
	≥1500	310	600	730	360	470	550	460	550	650
1.0	250	950	1240	1170	1120	1360	1290	1240	1410	1420
	500	550	980	820	840	1150	1030	1000	1250	1300
	750	390	740	730	660	830	780	800	930	1040
	1000	300	510	640	510	490	530	630	620	780
	≥1500	260	510	640	300	400	470	390	470	570

of all red and all red plus amber between green in two conflicting phases.

The total loss of effective green time per cycle, F, is obtained as the sum of the inter-green intervals between the conflicting phases.

The cycle time, c, is obtained as

$$c = (1.5F + 5)/[1 + \Sigma(q_i/s_i)_{max}] \quad (18)$$

where

F = sum of inter-green intervals per cycle, and

$\Sigma(q_i/s_i)_{max}$ = sum of q/s for the critical lanes, i.e., the lanes that make up the critical conflict point in the intersection.

The distribution of green times that best minimizes delay is derived from

$$g_i = (c - F) \times (q_i/s_i)_{max}/\Sigma(q_i/s_i)_{max} \quad (19)$$

The green times are checked against minimum green time requirements, and the timing is adjusted if necessary. This is done by adding the extra green time required to F and calculating a new cycle time that is distributed among the phases without disrupting the optimum distribution according to Equation 18.

The manual also includes recommendations for timing of traffic-actuated controls (minimum and maximum green, extension intervals) based on previous research (21).

Capacity, Queue Length, Proportion of Stopped Vehicles, and Delay

All the measures dealt with under this heading are first calculated on a per-lane basis. Totals and averages for an approach or for the whole intersection can also be derived according to instructions in the manual, but they are not always meaningful.

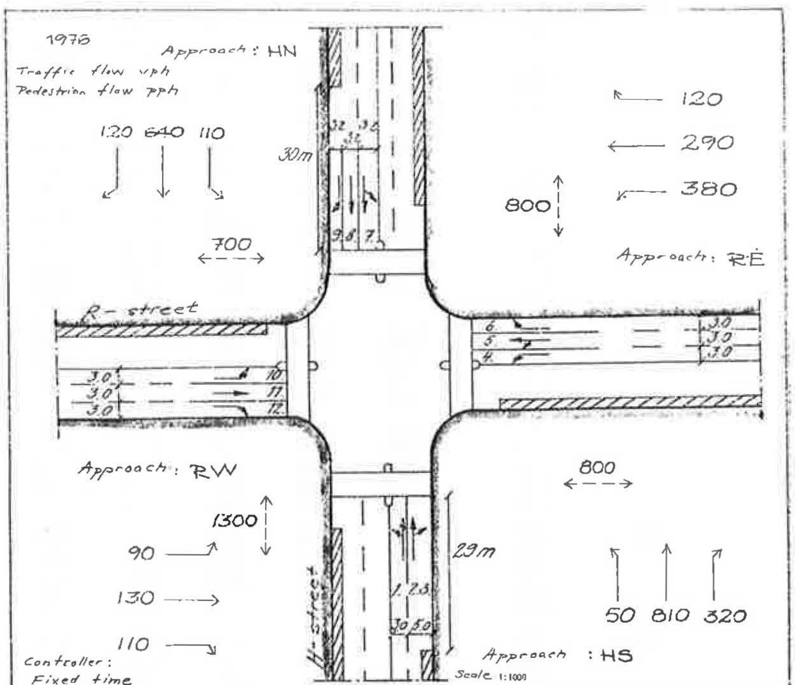
When the signal timing is established, lane capacities, K_i, in vehicles per hour are derived as

$$K_i = g_i/c \times s_i \quad (20)$$

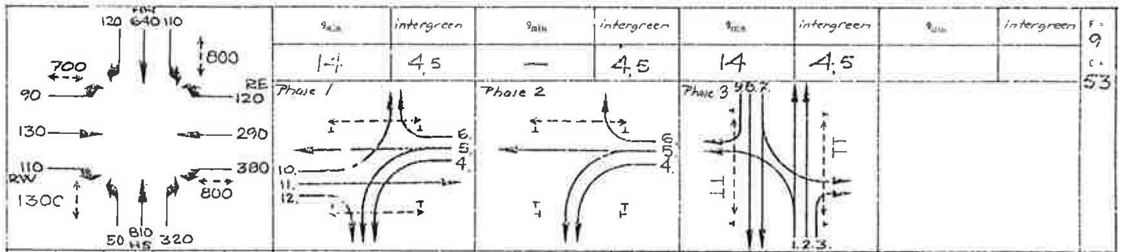
The degree of saturation, q_i/K_i, becomes highest and equal for all the critical lanes. If it exceeds 1.0, queues build up infinitely, and even at values above 0.8 substantial queueing occurs. The average number of queueing vehicles in lane i at the beginning of green, N_i, is calculated according to Miller (22):

$$N_i = N_{1i} + N_{2i} \quad (21)$$

Figure 23. Form A plan sketch, traffic volumes, and assumptions.

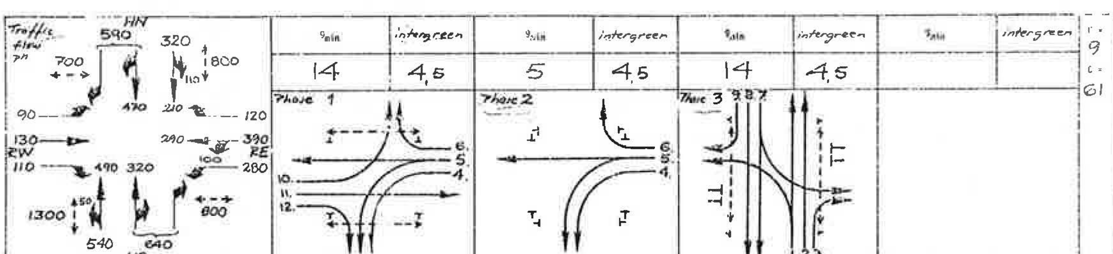


Approach	Speed limit (km/h)	Trucks and through buses %	Bicycles	Slope (%) + up - down	Min. green for pedest. (s)	Min. inter-green for pedestrians (s)	
c'dt. 1	2	3	4	5	6	7	
HS	50	10		—	9	7,0	
RE	50	10		—	9	7,0	
HN	50	10		-3	9	7,0	
RW	50	5		—	9	7,0	



Approach	Lane no i	Green in phase no	Lane type	Turns p	Saturation flow s_i: Base value vphg	Adjustments: Trucks and through buses	Slope	Lane width	Other	Adjusted value vphg	Flow q_i vph	q_i/s_i	$k_i = \frac{(q_i/s_i)_{max}}{\Sigma (q_i/s_i)_{max}}$	Green time g_i (s): Base value s	Adjusted value s	$g_{i\,max}$ s	Capacity $c_i = \frac{s_i \cdot g_i}{c}$ vph	Degree of saturation $x_i = q_i/c_i$
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
HS	1	3	D/F	0,12	1350/1637	—	—	—	—	1350	540				27			
	2	3	A	0,0	1700	—	—	0,74	—	1260	640	0,40	0,618	27	27			
	3	3	E	1,0	900	—	—	0,74	—	320*					27			
RE	4	1+2	E/G, C	1,0	900/1250 1390	—	—	—	—	1130	280	0,247	0,382	17	17			
	5	1+2	D/F, B	0,13	1490/1670,1680	—	—	—	—	1590	390				17			
	6	1+2	E, C	1,0	1000, 1500	—	—	—	—	1250	120	0,096			17			
HN	7	3	D/F	0,29	1300/870	—	—	1,02	—	890	280				27			
	8	3	A	0,0	1700	—	—	1,02	—	1730	590	0,31		27	27			
	9	3	E	1,0	530	—	—	1,02	—	160*					27			
RW	10	1	E/G	1,0	1000/1233	1,05	1,03	—	—	1080	90	0,083			14			
	11	1	A	0,0	1700	1,03	1,03	—	—	1800	130	0,072			14			
	12	1	E	1,0	900	1,05	1,03	—	—	970	110	0,113	0,175	8	14			
Sum												0,648						

Figure 24. Form C, calculation of signal timing, first round.



Approach	Lane no i	Green in phase no	Lane type	Turns %	Saturation flow s_i vpgh: Base value vphg	Adjustments: Trucks and through buses	Slope	Lane width	Other	Adjusted value vphg	Flow q_i vph	q_i/s_i	$k_i = \frac{(q_i/s_i)_{max}}{\Sigma (q_i/s_i)_{max}}$	Green time g_i (s): Base value s	Adjusted value s	$g_{i\,max}$ s	Capacity $c_i = \frac{s_i \cdot g_i}{c}$ (i/h) vph	Degree of saturation $x_i = q_i/c_i$
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19
HS	1	3	D/F	0,09	1420/1680	—	—	—	—	1420	550				29		670	0,82
"	2	3	A	0,0	1700	—	—	0,74	—	1260	630	0,387	0,558	29	29		770	0,82
"	3	3	E	1,0	1040	—	—	0,74	—	350*					29			
RE	4	1+2	E/G, C	1,0	780/1160,1350	—	—	—	—	880	270	0,307	0,442	23	23		330	0,82
"	5	1+2	D/F, B	0,26	1220/1650,1650	—	—	—	—	1300	400				23		490	0,82
"	6	1+2	E, C	1,0	780, 1500	—	—	—	—	910	120	0,132			23		340	0,35
HN	7	3	D/F	0,39	1330/890	—	—	1,02	—	910	290				29		430	0,67
"	8	3	A	0,0	1700	—	—	1,02	—	1730	590	0,300			29		880	0,66
"	9	3	E	1,0	570	—	—	1,02	—	120*					29			
RW	10	1	E/G	1,0	780/710	1,05	1,03	—	—	770	90	0,117			14		180	0,50
"	11	1	A	0,0	1700	1,03	1,03	—	—	1800	130	0,072			14		410	0,32
"	12	1	E	1,0	780	1,05	1,03	—	—	840	110	0,131	0,189	10	14		190	0,58
Sum												0,694						

Figure 25. Form C, calculation of signal timing and capacity, second round.

Figure 26. Form D, calculation of queue length, proportion of stopped vehicles, and delay.

Approach	Lane no. i	Flow q_i vps			Degree of saturation	No of queuing vehicles				Queue length (m)	Proportion of stopped vehicles						Average delay, fixed time (s)	Average delay, vehicle actuated (s)	Total delay vs
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
HS	1	0.153	0,387	0,475	0,82	1,8	4,9	6,7	9,3	75	0,63	0,096	0,22	13,4	1,87	12,2	23,0		3,5
"	2+3	0,175		0,475	0,82	1,8	5,6	7,4	10,0	80	0,64	0,112	0,22	13,4	1,87	10,7	21,7		3,8
RE	4	0,075	0,307	0,377	0,82	1,8	2,9	4,7	7,0	55	0,69	0,052	0,27	16,5	1,87	24,9	37,3		2,8
"	5	0,111		0,377	0,82	1,8	4,2	6,0	8,4	67	0,71	0,079	0,27	16,5	1,87	16,8	30,0		3,3
"	6	0,033	0,132	0,377	0,35	–	1,3	1,3	2,5	21	0,64	0,021	0,22	13,4	0,094	2,8	14,6		0,5
HN	7	0,081	0,319	0,475	0,67	0,5	2,6	3,3	5,4	41	0,57	0,046	0,22	13,4	0,480	8,4	19,6		1,6
"	8+9	0,161	0,314	0,475	0,66	0,5	5,2	5,7	7,8	61	0,60	0,097	0,22	13,4	0,641	4,0	15,7		2,5
RW	10	0,025	0,117	0,230	0,50	0,1	1,2	1,3	2,6	20	0,78	0,020	0,34	20,7	0,250	10	27,8		0,7
"	11	0,036	0,072	0,230	0,32	–	1,7	1,7	3,1	22	0,79	0,028	0,33	20,1	0,075	2,1	20,0		0,7
"	12	0,031	0,131	0,230	0,59	0,2	1,5	1,7	3,1	22	0,79	0,024	0,34	20,7	0,400	12,9	30,2		0,9
Sum		0,881										0,575							19,8
Average												0,66							22

where

$$N_{1i} = (2x_i - 1)/[2(1 - x_i)] \quad ; \quad x_i = (q_i \times c)/(s_i \times g_i) > ½$$

$$N_{1i} = 0 \quad ; \quad x_i \leq ½$$

$$N_{2i} = q_i \times (c - g_i) \qquad (22)$$

N_1 represents the number of vehicles remaing from the previous green phase and N_2 the number of vehicles arriving during red. N is then adjusted upward by means of Poisson's curves for required probability of overload.

Queue length, L_i, is derived by multiplying the number of queueing vehicles by the average space per vehicle. The manual also includes adjustments of L_i as a function of the proportion of trucks and buses in the lane.

The proportion of stopped vehicles, p_{si}, in lane i is derived from

$$P_{si} = (r_i + N_i \times t_i)/c \qquad (23)$$

where $N_i \times t_i$ represents the time from beginning of green until the N_ith vehicle in the queue has started to move. Measurements indicate that t_i normally equals 1.0.

Average delay, d_i, is calculated according to Webster (6):

$$d_i = 0.9\,[c(1 - \lambda_i)/2(1 - \lambda_i \times x_i)] + \{x_i^2/[2q_i \times (1 - x_i)]\} \qquad (24)$$

where λ_i equals g_i/c. For x_i, see Equation 22 above. Adjustment factors for queue length and delay in traffic-actuated controls based on the corresponding values at fixed time controls are also given in the manual.

APPLICATION

In the manual (4) the chapter on signalized intersections occupies about a fourth of the whole book. The important relations are all presented in graphs or tables in order to simplify use of the manual. Computer programs are also being developed for the same purpose.

The layout of the manual is in the form of stepwise calculations (Figure 10), each step leading to a set of figures to be filled into the corresponding column on the form for the calculation. Four different forms are provided for signalized intersections:

1. Plan sketch: traffic volumes and assumptions,
2. Determination of inter-green,
3. Calculation of signal timing and capacity, and
4. Calculation of queue length, proportion of stopped vehicles, and delay.

Figures 23-26 illustrate the use of the forms for an intersection in the downtown area of Stockholm. Figures 23 and 24 represent the first and the second rounds of the signal timing that are filled into form C.

The time required to carry out a complete application, such as the one in the example, may vary between 3 and 6 h, depending on the skill and practice of the engineer. The most time-consuming part is the second round of signal timing, which, however, is only necessary if queue lengths or other measures for individual lanes are requested. If only signal timing and degree of saturation are asked for, the calculations should not take more than an hour.

The manual has been introduced in a number of workshops for traffic engineers in Sweden and has been greatly appreciated for its exactness and flexibility. A number of important questions still have to be solved by future research, however, for example, the effect of bicycles in the vehicle flow, the influence of different types of traffic-actuated controls, and the influence of rain, snow, and darkness.

A thorough validation of the methods is also desirable and may take place after the manual has been used for some time.

REFERENCES

19. F. Webster. Traffic Signal Settings. Her Majesty's Stationery Office, London, Road Research Technical Paper 39, 1958.
20. L. D. Gordon and A. J. Miller. Right Turn Movements at Signalized Intersections. Proc., 3rd Annual Conference of the Australian Road Research Board, Vol. 3, No. 1, 1966.
21. K.-L. Bang. Optimal Control of Isolated Traffic Signals. TRB, Transportation Research Record 597, 1976, pp. 33-35.
22. A. J. Miller. Settings for Fixed-Cycle Traffic Signals. Operational Research Quarterly, Vol. 14, No. 4, 1963.

Discussion

Robert C. Blumenthal, Robert C. Blumenthal Associates, Boston

PART 1 BY PETERSON

Mr. Peterson initially reviews the historical background leading to development of the 1977 publication on calculation of capacity, queue length, and delay on road traffic facilities. He notes that "it has been clearly shown from measurements in several countries that actual capacity is higher than that arrived at by application of the methods" in referring to previous methods.

No design standards, such as levels of service, are provided; rather, the final decision as to the design is suggested to depend on factors of road safety, cost, and environment. Thus, Mr. Peterson states that "the choice (or design) should be based on that degree of accuracy which is right both socially and economically."

Major additions to the literature are calculations for unsignalized intersections, pedestrians (midblock), bicycle traffic (bike paths), and short weaving sections of 40-60 m.

The manual itself is a fine publication, well laid out with color for emphasis and the use of examples and forms to assist the analyst. I am generally favorably impressed by the Swedish manual and Mr. Peterson's paper. It is, indeed, a forward step in the capacity field.

I do take some minor exception to Mr. Peterson's statement that "it has been clearly shown" that actual capacity is higher than that arrived at by different methods. In my discussion of the other two papers, I shall indicate why I differ. Also, I believe that providing design guides or levels such as levels of service is an advantage because it sets standards. "The choice of that degree of accuracy which is right both socially and economically" can be a controversial choice if it is not based on engineering fact.

PART 2 BY HANSSON

Mr. Hansson's paper presents a method of calculating capacity, queue length, and delay at intersections controlled by yield or stop signs. The method is based on a queuing model (M/G/1) whose most important input parameter is the critical headway determined from field observations at 18 intersections in Sweden (Table 2 in the paper).

Data gathered for the 10 yield and 8 stop approaches provided critical headway (right, straight, and left) and move-up times. The paper notes that "values... are based on these measurements for Swedish conditions, as well as on previous studies elsewhere." Table 1 outlines the values.

I have reviewed the information in Tables 1 and 2 and cannot find any direct relations. Table 1 values are the basis for calculating intersection operations, so it would be interesting to know how they are derived.

A breakdown of Table 2 reveals 12 locations (8 yield and 4 stop) where speeds were 50 km/h, 3 (1 yield and 2 stop) locations at 70 km/h, and 3 locations (1 yield and 2 stop) at 90 km/h. I wonder what level of statistical confidence can be placed on values derived from one or two locations.

I suspect the additional data came from previous studies. In 1974, the Organization for Economic Cooperation and Development (OECD) published Capacity of At-Grade Junctions (7). Table 4 of that publication (here Table 4) summarizes critical time gap values.

Table 4. Summary of values for the critical time gap found by various researchers or used in the national design manuals.

	Through		Right Turning		Left Turning	
Source	Passenger Cars	Trucks	Passengers Cars	Trucks	Passengers Cars	Trucks
Grabe, Wörner	4.1-5.2	6.0-6.5				
Krell	6.4 (second vehicle: 1.82 t; third 2.53 t; etc.)					
Ashworth (median value)	6.5					
Hofwegen (value at a major road flow of 1000 v/h)	5.4		5.0		6.0	
Kell (median value)	5.8		5.4		6.3	
Owens (median value)	6.3					
Thomasson (median value)	8.6					
Knoflacher	6					
French design standards[b]	6	8				
German guidelines for traffic signals						
V = 90 km/h	7-8		6-7		8-9	
V = 50 km/h	6-7		5-6		6.5-7	
English design standards	4-8		8-12			
Dutch design standards	6 (+2 s reaction time)				(7)	6 (9)
Swedish manual						
V = 90 km/h	7.0		7.2		7.5	
V = 70 km/h	6-6.5		6-6.5		6.2-6.8	
V = 50 km/h	5.2-5.8		4.8-5.5		5.3-6.0	

[a] In some cases the time needed to complete the maneuver for one carriageway with two lanes is given.
[b] Sight distances are calculated taking into account the time required for merging.

I have added the Swedish manual values to the bottom of this table, and find they generally compare well. (Note, however, that the values in Table 4 are simplistic. The original texts generally classify time gaps by specific configurations.)

There are no sample calculations in the paper. However, there are two examples in the manual (attached). One is a T-junction, where both roads have speeds of 50 km/h (Figure 27). Traffic flow from the two-lane, one-way secondary road is 200 vph turning left, and 200 vph turning right. The two-lane, two-way primary road carries 500 vph each way. Capacity is calculated as 570 vph for the right turn and 300 vph for the left turn.

For comparison, I have another OECD exhibit (Figure 12 from page 35, attached here as Figure 28). This figure shows three locations in Britain where comparisons were made between observed flows and those predicted using Tanner's formula (10, 11). I note in site 3, where the critical gap time is identical to that chosen by Hansson, that the results are a perfect match. Sites 1 and 2 have lower critical gap times, resulting in higher capacity for the right-turn movement.

The next figure (Figure 11 from the OECD, here given as Figure 29) compares capacities predicted by five methods, with critical gaps of 5 and 7 s for a straight movement in conflict with two-way primary movements. As may be seen, there is a great deal of variance in predicted capacities. The OECD review concludes with the remark, "The foregoing review of methods of estimating the capacity of major/minor priority junctions has shown that the topic is a very complex one, and that despite the considerable amount of research which has already been completed, a great deal still remains unknown."

Mr. Hansson's paper and procedures add a method for calculating unsignalized intersections that appears to be simpler than others used in some areas.

I would like to see an example or two in the paper so

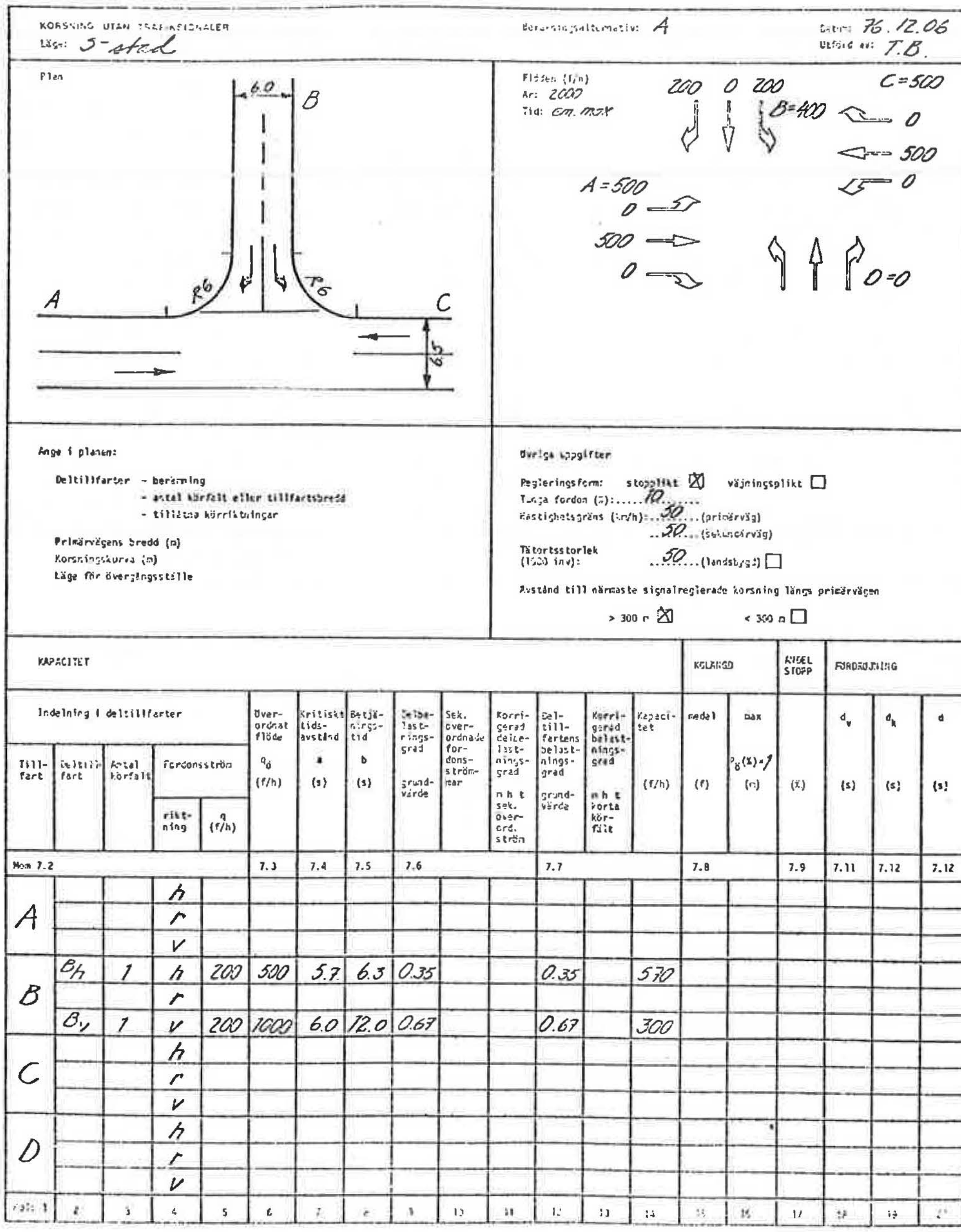
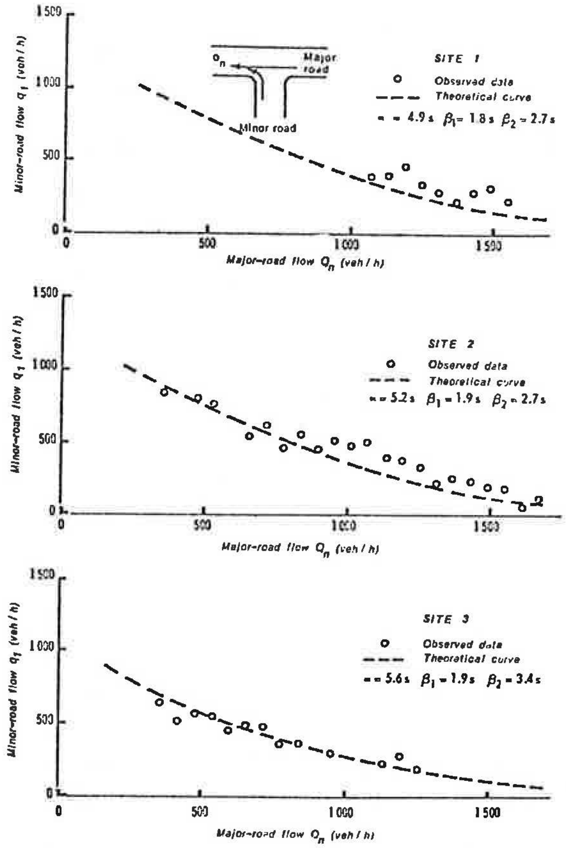


Figure 27. Example 1 from the Swedish capacity manual.

the method could be used by traffic engineers. Thus, comparisons could be made. At present, the paper provides a great deal of background, which has value but does not help the reader make practical use of what is obviously a well-developed method. I hope the Swedish manual can be translated to fill the missing elements.

Figure 28. Comparison of predicted and observed traffic flows (7).



PART 3 BY BÅNG

Mr. Bång has presented a well-written paper outlining the Swedish capacity manual method for calculating, not only capacity, but also queue length, delay, approach grades effect, proportion of stopped vehicles, and signal timing. The method makes use of previous findings by Webster of Great Britain on signal timing and Miller and Gordon of Australia on conflicts between left turns and opposing flow and signal timing, as well as previous Swedish work, especially by Arne Hansson on conflicts between vehicles and pedestrians.

Original procedures are shown for calculating

1. Effect of conflicts between left turns and opposing traffic on two-way streets where (a) a waiting left turn will block the lane and (b) a waiting left turn will not block the lane;
2. Reductions in the capacity of right turns by pedestrians in the crosswalk; and
3. Effects of restricted length of curb lanes where (a) through and right-turning traffic are in the lane and (b) the curb lane serves only right turns.

The paper contains an example of the procedures for a four-way intersection, in downtown Stockholm, with three-phase signal control. There are three examples in the manual: two-phase, three-phase—the one in this paper (Figures 30 and 31)—and four-phase.

For comparison, I calculated capacities for two of the three intersections using the HCM and then compared the results (Tables 5 and 6). Tables 5 and 6 show the comparisons. My calculations were rough, but are adequate for comparative purposes. Overall intersection capacity appears to correlate fairly well. Location A, M-Stad, shows an overall variance of 6 percent, or 220 vehicles. In this example, the HCM would predict more overall capacity than the Swedish method. Location B, Stockholm, shows an overall variance of 4 percent. In this instance, the Swedish manual predicts more capacity than the HCM.

The individual approaches do not compare as well, with variances of 2-50 percent between the methods. One of the major reasons for variations on the high end is signal timing. Table 7 reviews the level of ser-

Figure 29. Comparison of capacities (7).

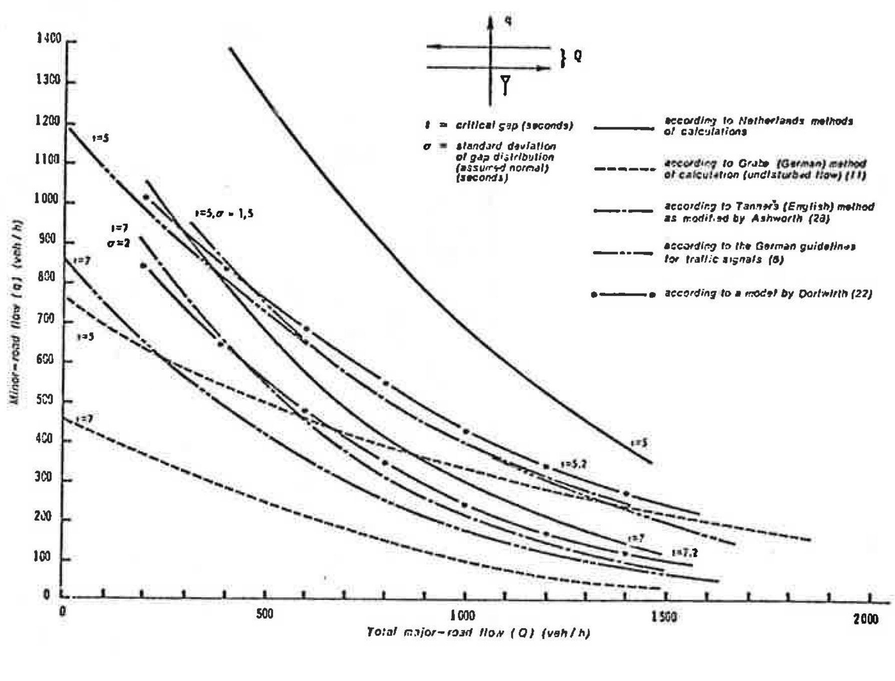


Table 5. Capacity comparison for location A: four-way two-phase intersection with all two-way streets.

Approach	Width (m)	Volume L	Volume S	Volume R	HMC Level E	Capacity*	Difference
BN	5.5	100	200	200	750	590	-160 (-0.21)
BS	5.5	50	150	150	940	470	-470 (-0.50)
AV	6.5	110	760	150	1000	1200	+200 (0.20)
AO	6.5	80	850	170	1060	1270	+210 (0.20)
Total					3750	3530	-220 (-6)

*Pedestrian volume: N-S = 500-700, E-W = 200-150.

Table 6. Capacity comparison for location B: four-way three-phase intersection with all two-way streets.

Approach	Width (m)	Volume L	Volume S	Volume R	HMC Level E	Capacity*	Difference
HN	9.6	110	640	120	1340	1310	-30 (-0.02)
HS	8.0	50	810	320	1260	1340	+80 (0.06)
RV	9.0	90	130	110	700	780	+80 (0.06)
RO	9.0	380	290	120	1090	1150	+66 (0.11)
Total					4390	4580	+190 (+4)

*Pedestrian volume: N-S = 1300-800, E-W = 700-800.

Figure 30. Location A from the Swedish capacity manual.

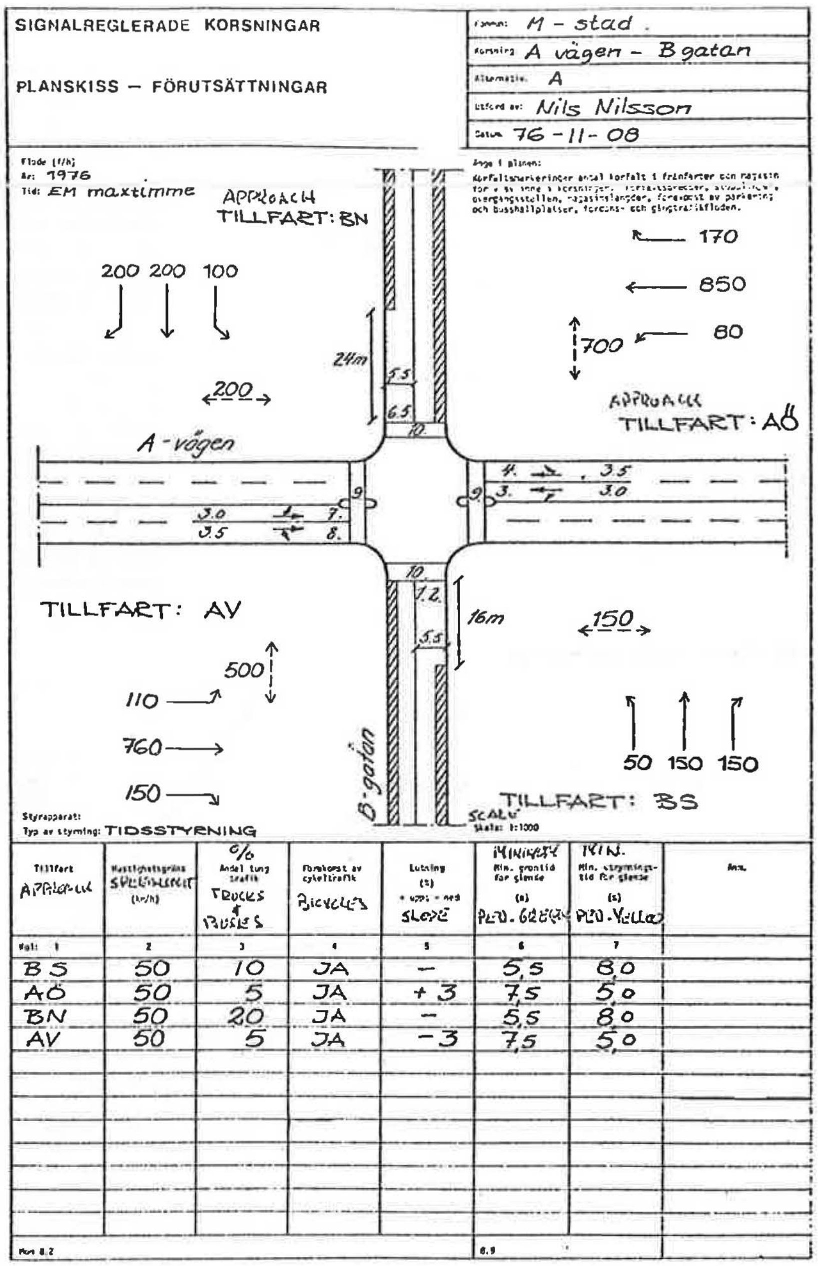


Table 7. Level of service comparison for locations A, B, and C.

Location	Approach	W.	Volume	G/C	HCM LOS "E"	HCM V/C	HCM LOS	Degree of Saturation
A	BN	18	500	0.49	750	0.53	C/D	0.86
	BS	18	350		940	0.37	C	0.74
	AV	21.3	1020	0.40	1000	1.02	E	0.86
	AO	21.3	1100		1060	1.04	E	0.87
Total			2870		3750	0.76	D	
B	HN	31.5	870	0.36	1340	0.65	D	0.67
	HS	26.2	1180		1260	0.93	E	0.82
	RV	29.5	330	0.18	700	0.47	C	0.58
	RO	29.5	790	0.29	1090	0.72	D	0.82
Total			3170		4390	0.72	D	
C	MN	21.3	550		530	1.04	E	
	MS	14.8	280		380	0.74	D	
	IY	19.7	1080		1100	0.98	E	
	IO	19.7	1270		1050	1.15	E	
Total			3180		3060	1.04	E	

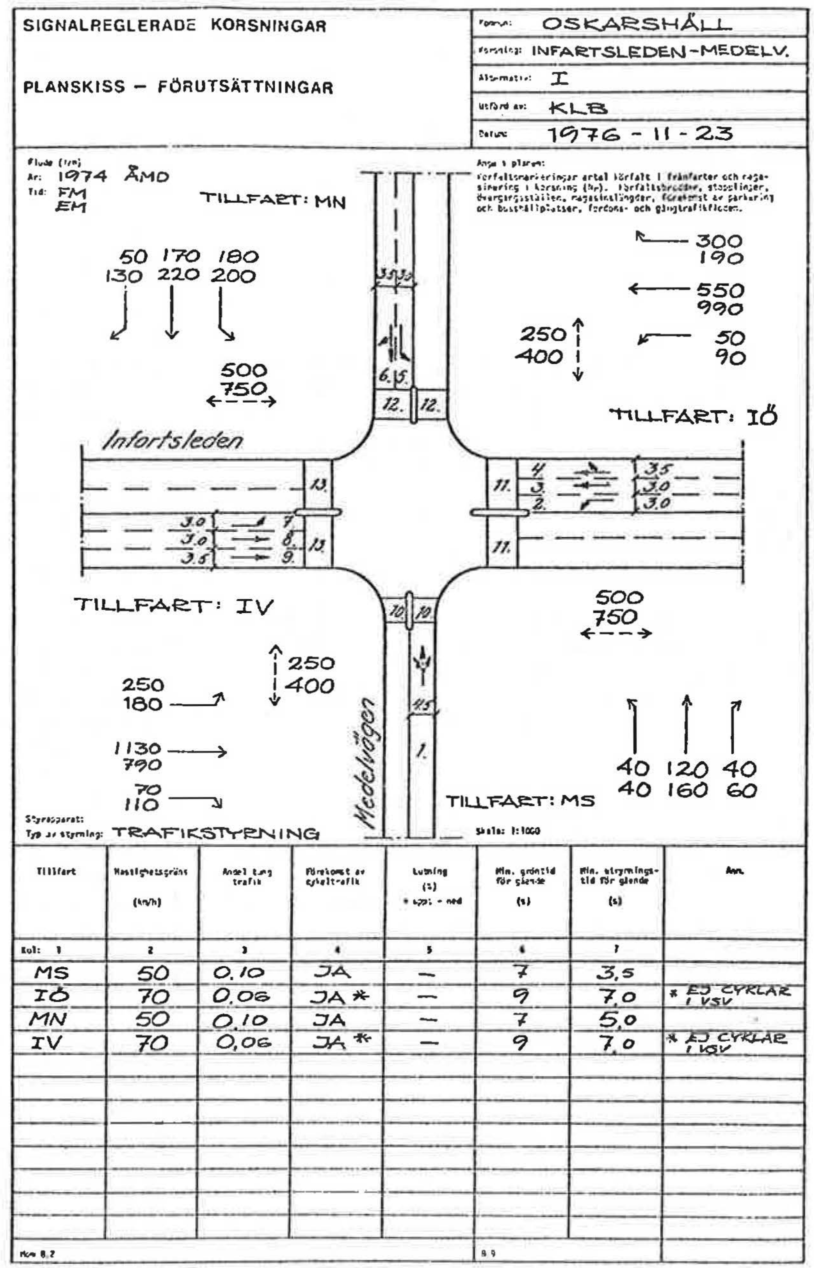


Figure 31. Location C from the Swedish capacity manual.

vice expected with the signal timing shown in the examples. Location A would have been timed differently, based on HCM analysis, and would have evened out the capacity differences. Delay was not calculated for location A. Therefore, there was no opportunity for comparison of that item.

I should note that neither capacity nor delay was calculated for location C (Oskarshall), the four-phase, actuated intersection.

I wish to congratulate Mr. Bång on a well-written paper and on his contribution of a methodology for calculating effects of conflicts between left-turning and opposing flows, conflicts between turning vehicles and pedestrians, delay, and queue lengths.

Analysis by lane and movement should give the traffic engineer an opportunity for specific insights into an intersection's operations. It will obviously not be a favored method for planners or those who seek a quick, simplistic method.

Although the comparison of approach capacities appears to indicate substantial variance between the Swedish manual and the HCM, two factors in the examples could have significance in reducing the variance: (a) signal timing, as previously noted, and (b) use of approach width versus lane. Please note the single lane approaches at location A of 5.5 m. The HCM (and others) treats this width as two lanes (although not specifically), thus providing greater capacity.

Jack A. Hutter, Traffic Institute, Northwestern University, Evanston, Illinois

Because of time and space limitations, my discussion will be directed only to the paper presented by Mr. Bång. This informative paper provides a detailed overview of an analysis procedure that offers some original and unique techniques for evaluating the performance capabilities of signalized intersections.

The primary object of this discussion paper is to highlight those attributes of the Swedish manual that may be applied to the development of an improved U.S. HCM.

In discussing Mr. Bång's paper, I shall carefully avoid the term "intersection capacity analysis," because the Swedish manual provides potential capabilities for evaluations beyond the determination of signalized intersection capacity. Additional analysis techniques include

1. Calculation of optimum signal timing,
2. Determination of queue length,
3. Calculation of the proportion of stopped vehicles,
4. Measures of delay, and
5. Determination of the degree of saturation, similar to a volume-to-capacity ratio that provides a measure of the level of operation of the critical lanes or approaches of an intersection.

Another significant feature of the Swedish manual is the important mathematical relations that are presented in graphic or tabular form for ease of use by the practicing engineer. In addition, the layout of the descriptive procedure is in the form of stepwise calculations, supplemented by clearly labeled work forms, which assist the user in following a fairly detailed and complex analysis procedure.

There are several calculation routines in the Swedish intersection analysis procedure that appear to have superior sensitivity and consistency of application than do the procedures in the HCM. For instance, Swedish calculations are based on the analysis of the individual vehicle streams entering an intersection, rather than aggregated intersection approaches. This enables the user to identify the critical lanes as well as the critical approaches of the intersection. It also provides the analyst with an overview of the total performance potential of the intersection, through the identification of the degree of saturation.

In currently practiced "critical lane analysis" procedures, the user must have access to lane volume data or make assumptions regarding the distribution of the approach lane volumes, which can produce serious inconsistencies in the results of the analysis. The Swedish manual provides a unique lane-volume distribution technique whereby the lane flows are distributed in proportion to the saturation flow values of the adjacent lanes. This technique allows all users to perform this calculation in a consistent manner.

The Swedish manual also provides a technique for quantifying the effect of the conflicts between left-turning vehicles and opposing flows and the conflicts between turning vehicles and pedestrian movements on the performance capabilities of the intersection. This simplified lane classification system enables the user to identify the type and degree of vehicle and pedestrian conflicts and to select corresponding values of saturation flow for each lane type in a consistent manner.

Most intersection analysis techniques are unable to calibrate the utilization or capacity of a curb lane if parking is not prohibited for a specified distance from the intersection approach or exit. The Swedish manual provides procedures for calculating the reduction in saturation flow values for curb lanes of restricted lengths for the two cases where the curb lane serves right-turning and through vehicles or the curb lane serves only right-turning vehicles.

The second object of this discussion is to raise questions regarding various aspects of the Swedish intersection analysis procedures, questions brought about by my lack of practical experience in applying these new techniques to real-world situations, and give to the author an opportunity to share his insights and experiences regarding the use and application of the manual by practicing traffic engineers in Sweden.

The lane-flow distribution technique is designed to produce consistent results by all users. The question is whether this technique has demonstrated that it will simulate actual lane-volume distributions under a wide range of geometric and traffic flow conditions. Also, it is stated that the method is based on theoretical models supported by a limited amount of field survey. One must ask if the experiences of the practicing engineers in Sweden in the application of the techniques supported the validity of the models.

There is an increasing need for American engineers to have models that provide quantitative measures of queue length, stopped vehicles, and vehicle delay for use in environmental and fuel conservation studies. Here the question is whether the techniques in the Swedish manual provide for reasonably accurate and consistent simulation of these variables, insofar as previous research would tend to indicate that flow distributions and traffic performance by lane are highly variable and erratic quantities when measured on a cycle-by-cycle basis.

It was reported that calculations require from 1 to 6 h per intersection, depending on the number of measures desired and the skill and experience of the user. Therefore, one must know the level of acceptance and use of the procedure by practicing engineers in Sweden.

Arthur A. Carter, Office of Traffic Operations, Federal Highway Administration

The authors should be commended for undertaking research into highway capacity in their own country, rather than depending on earlier, increasingly obsolete American criteria, or criteria from other countries. Given the performance differences that we see just from locality to locality even within the United States, I suspect that only concepts, not absolute values themselves, are freely transferable from country to country.

Having been technical editor of the American 1965 Highway Capacity Manual, I shall discuss in general terms how these Swedish efforts relate to the American state of the art.

PART 1 BY PETERSON

First, I was not surprised by Mr. Peterson's observations that capacity flows in several countries appear greater than indicated by U.S. criteria. Typically, I find that foreign flows exceed ours, possibly because of smaller vehicles and different driver behavior. Therefore, I am not particularly concerned about findings that disagree with either our 1950 or 1965 HCM, in terms of absolute numbers.

I am more interested in concepts, trends, and procedures. In this regard, the literature review and analysis of some 900 references, accomplished in 1973, seems particularly impressive. It would appear to be superior, in both numbers and scope, to anything yet done in the United States. Translation into English might be recommended.

Although Mr. Peterson indicates that a complete manual has been developed, his emphasis appears to be on urban intersection operations, the subject of the two other reports presented here, with briefer mention of urban networks, bicycle flows, and pedestrian traffic. He does not specifically cover the basic area of rural uninterrupted flow, either two-lane or multilane, or basic freeway flows. I assume that these fundamentals are included in the manual, presumably in the road sections and street sections categories of the first table. I would be interested in learning more about their handling.

I am confused by Mr. Peterson's comments regarding pedestrian crossings, which seem to stress the between-intersection case at the expense of the at-intersection case. This appears to be in conflict with Mr. Bång's paper, which emphasizes the at-intersection situation. [The authors of this paper revised the draft on which these discussions were based. Therefore, some of the points in this discussion are no longer relevant.]

Mr. Peterson describes the use of models based on queue theory for much of the work as an important innovation, with only limited use of empirical evidence. Although I am probably biased, I continue to feel that in this capacity field, practice frequently varies too widely from theory for us to make empirical evidence secondary. I strongly support validation with empirical data.

The "description of consequences" concept appears to be parallel to our level-of-service concept, but extended to better cover congested conditions, where delays, stoppages, and queueing occur. Its orientation differs somewhat from our procedures in that, in its straightforward form, design and demand volume are both known and performance is unknown. Apparently, application to design problems or to volume determinations is a trial-and-error procedure. This is in contrast with our current procedures, which are most directly oriented to flows obtainable with a given design. There probably is no one answer to which is best; it depends on the particular user's needs.

In this connection, I sympathize with Mr. Peterson concerning his quandary regarding the right degree of accuracy. Here in the United States, we have the same problem. I suspect that the final answer one day may be three entirely separate manuals, at different levels of precision and complexity, for planners, designers, and operators.

PART 2 BY HANSSON

Mr. Hansson's work on the performance of unsignalized intersections is much needed. Relatively little has been done in this field, and that which has been done is mainly on a case-history basis. I assume that four-way stops are not included.

This chapter apparently uses the term "capacity" in its general sense to cover all levels of service, rather than its specific saturation flow sense, because cautions are given against its application where the minor flow is at a load factor (which, as they define it, is similar to our volume-to-capacity ratio, rather than in the U.S. sense) exceeding 0.8 or 0.9. The observation in the report that such conditions normally would warrant traffic signals tends to substantiate the 1965 HCM statement that the true capacities of unsignalized intersections are usually of academic interest only.

The reference to M/G/1 queueing system and M/M/1 queueing model may well be meaningless to many readers of this paper. It would have been helpful to describe these models in general terms.

I am somewhat concerned by the rather large number of assumptions and deliberate omissions, in contrast to the overall considerable detail involved in the method. We have no answers regarding the influence of such factors as primary flow headway distributions, the combining of flows, move-up time, conflicts between flows, relative sizes of flows, and load factor as they define it, but I feel that the related assumptions and omissions may have a significant effect in this unsignalized case. This points up the need for validation, as suggested by the author.

Of greatest concern to me is the author's decision to omit consideration of platooning produced by nearby traffic signals, on the basis that the effect, like that of several other factors, cannot be quantified. In this country, one of the principal factors governing the performance of a particular unsignalized intersection is the proximity of nearby signalized intersections and the relative phasing of the nearest ones on either side of that under consideration. While I agree that the effect is nearly impossible to quantify, I submit that it is a principal reason why some of the work in this country has ended up quoting specific cases rather than general criteria.

With all other conditions being identical, I see one specific location operating well, with periodic breaks in the major flow occurring simultaneously in both directions to permit free crossings, while another never is free of at least one direction of the major flow. Possibly, adjacent signalized intersections are less common in Sweden than in the United States, but I question that this factor could be overlooked in practice in the United States. I agree, therefore, with the author's conclusion that primary flow headways may be the weakest link in the method.

Also of great importance in this country is the nature of the primary road—two-lane two-way, multilane undivided, or multilane divided. This element is considered in the method, and we are interested in their inability in the validation studies to confirm higher critical headways for crossing four-lane roads than for two. Possibly, the influence of adjacent signals, more likely to be found on four-lane roads than on two, is involved.

PART 3 BY BÅNG

Mr. Bång's work on signalized intersections is a valuable contribution to the state of the art, particularly with respect to its consideration of turning movements related to pedestrians. This is one area where the public expects and seems to assume that we have much more information than we actually do.

I was surprised to note that the principal Swedish intersection capacity criteria, prior to the subject work, were based on the original American 1950 HCM rather than the 1965 edition, which is the principal reference in use today. Apparently this is because Sweden was in the forefront in the 1950's and developed its own procedures soon after ours were published.

It would be useful to learn which evaluations were made in Sweden after the 1965 manual was issued that caused them to decide to start over rather than try to adapt its procedures. Such information would be particularly timely as we begin the initial steps of preparation of a new edition of the American manual.

The Swedish procedure appears to emphasize a different range of operational levels than does our manual. While they indicate that their work omits design levels and centers on saturation flow or capacity, which is our level of service E, they suggest that 0.8 times saturation flow be used in practice as a feasible maximum. In effect, then, they closely relate to our level of service D, which of necessity has become our design level for many urban applications even though level C is considered more desirable.

The Swedish procedures then appear to extend from our level D into level F, with a description of consequences throughout including level F, as compared to our levels of service A through E, with only general reference to broad breakdown of level F. This is rather characteristic of foreign signalized intersection capacity investigations and procedures as compared to current U.S. procedures; they typically concentrate on saturation flow to a greater extent than do U.S. procedures.

I agree that any new American method must include consideration of stops, queueing, and delay, which are the elements of intersection performance most visible to users. Our load factor (that is, percentage of fully utilized green intervals) is inadequate where most needed.

It is not clear why the Swedish measures developed on a by-lane basis are "not always meaningful" when expanded to a complete approach or whole intersection; this concerns me.

The conclusion that both width and number of lanes are significant is interesting. This point is controversial, not only within the United States but internationally, where viewpoints range from the 1965 HCM's overall approach width criterion to Australia's basic number-of-lanes procedure.

Regarding the pedestrian effect, I find the tabulation form of presentation quite good and a step toward what I would like to have, to answer citizens' questions. I wonder, though, whether such national criteria could ever be established for the United States as a whole, given the widely differing degrees of respect shown each other by drivers and pedestrians from one city to another. It would seem that city factors would have to be developed, much as both we and our Swedish counterparts dislike them.

It is interesting to note that several of the same factors that have escaped easy solution in this country also defy solution in Sweden, including bicycles, actuation of signals, and weather.

The problem solutions described conform with Mr. Peterson's introductory comments; they relate to situations where the traffic volumes and geometrics are known, and the nature of the resulting operation is desired. It is not indicated whether or not the method can be used effectively for other situations, where either volumes or geometrics are the unknown. (The sample problems that exist in the manual undoubtedly assist greatly in providing an understanding of the procedures.)

At first glance, the typical problem solution time, 3-6 h, looks long. However, given that this involves a detailed solution including queue lengths and so on, something much beyond our current procedures, and that a basic signal timing solution is possible in an hour, it is probably reasonable.

Finally, it is unfortunate that this work, like nearly all other work in this field in recent years, must be tagged as "needing validation." I hope users will soon do sufficient testing and evaluation so that validity can be more firmly established.

Publication of this paper sponsored by Committee on Highway Capacity and Quality of Service.

Weighing Vehicles in Motion

A. T. Bergan, G. A. Sparks, and G. Dyck, College of Engineering, University of Saskatchewan, Saskatoon

A scale for weighing vehicles in motion was developed at the University of Saskatchewan. This scale has been successfully operated unmanned at two locations in Saskatchewan for the past 2 years. An expanded evaluation program is currently under way in which the scale will be installed and evaluated in Ontario, Quebec, and New Brunswick by a project committee of the Roads and Transportation Association of Canada.

The relationship between vehicle and axle loads and the structural requirements of bridges and roadways has been and continues to be an area of particular interest to those concerned with the provision and maintenance of bridge and roadway facilities. The need for comprehen-

Figure 1. Design of the scale.

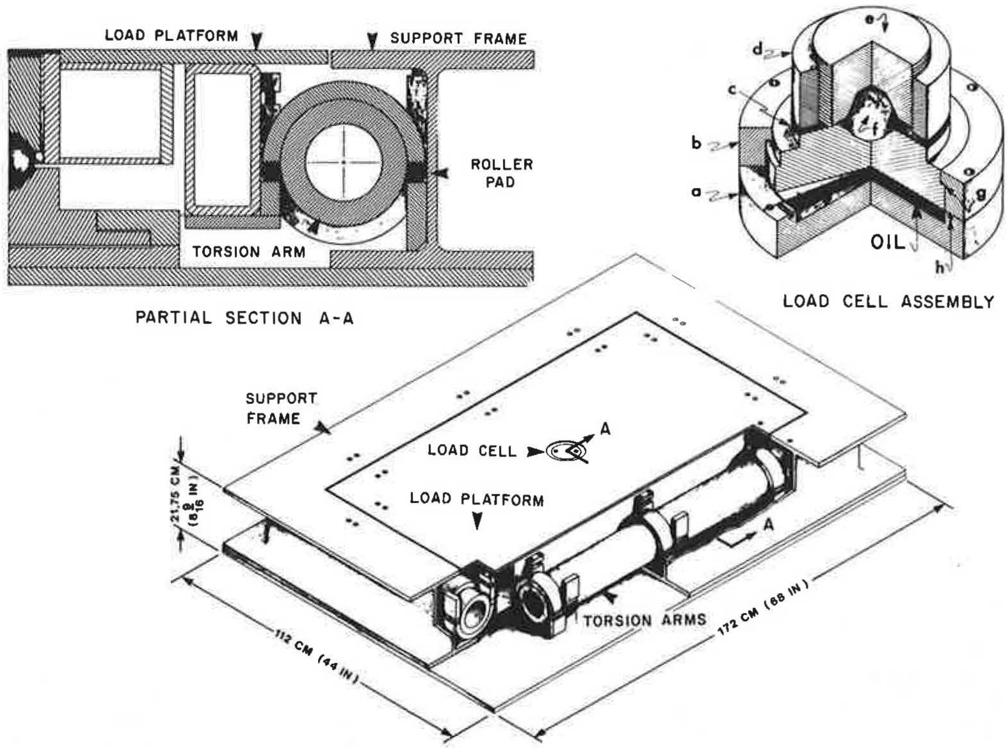
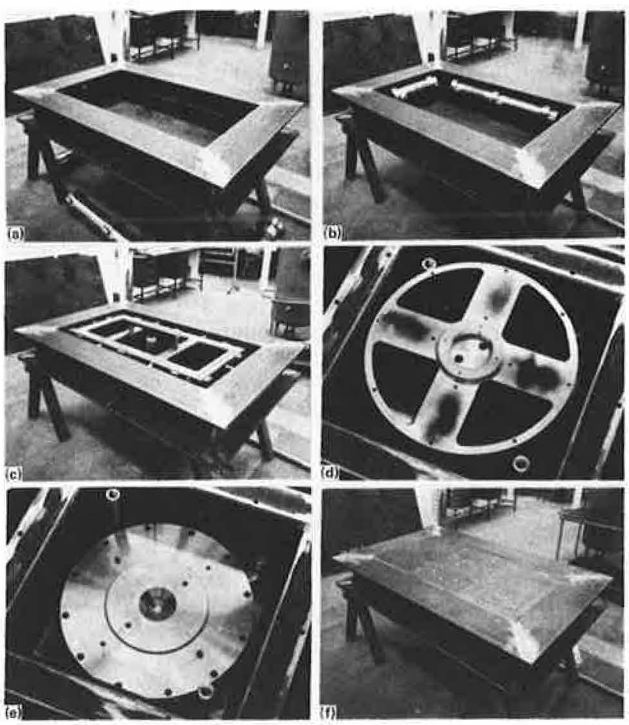


Figure 2. Steps in assembling the scale.



sive information on the nature and characteristics of roadway traffic as related to the design and maintenance of bridge and roadway structures is well documented (1, 2, 3).

In an attempt to fulfill this information requirement, a number of organizations have in recent years worked on the development of scales for weighing vehicles in motion. Such a scale has been developed at the University of Saskatchewan (4, 5, 6).

The efforts at the University of Saskatchewan were initiated primarily because existing scales for weighing vehicles in motion were incapable of operating in a continuous unmanned manner in the harsh Canadian environment.

The scale for weighing vehicles in motion that was developed at the University of Saskatchewan utilizes the hydraulic pressure principle. Loads applied to any point on the load platform are transmitted evenly around the perimeter of the platform by four torsion arms. (Figures 1 and 2 illustrate the scale assembly.)

The load platform can move only vertically as a rigid unit. This vertical motion is extremely small [i.e., of the order of 0.015 cm (0.006 in) at 4500 kg (10 000 lb)] and is nearly frictionless due to the roller pad contacts between the load platform, the torsion bars, and the support frame. The entire load is then supported by a single, centrally located load cell, which is an oil-filled piston cylinder arrangement with a strain gauge transducer.

DEVELOPMENT PROGRAM

Until the fall of 1976, the major effort was devoted to the development and evaluation of the scale unit itself. The first prototype of the scale was constructed during the summer of 1974 and was installed in an abandoned section of highway. The results of the series of tests using several vehicle weights and speeds were very encouraging. The observed differences between the actual static weight and the observed dynamic weight typically compared within 10 percent.

A second scale incorporating several small design changes but utilizing the same principles was constructed during the winter of 1974-1975 and was installed in an in-service section of roadway in the spring of 1975. Results of tests on this scale were even more encouraging; however, two problems were encountered.

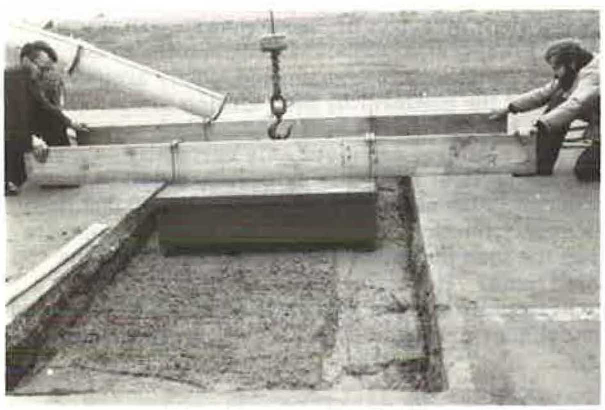
The first was the failure of a seal that allowed moisture to enter the interior of the scale and cause a corrosion problem. The second was associated with the testing program. The scale was located on a section of highway far removed from any highway weigh scale, which made calibrating and testing of the scale inconvenient.

The third-generation prototype was constructed and installed during the summer of 1975 in a section of highway 5 km (3 miles) from a government weigh scale near Clavet, Saskatchewan. This location greatly facilitated the testing procedure, and the design modifications included in this generation of the scale successfully overcame the moisture problem previously encountered.

When the third-generation scale was installed, special attention was given to the pavement surface leading to the scale. Since any irregularities in the road surface would be expected to cause transient perturbations of the vehicle suspension, an infinitely smooth section of roadway would be desirable. In attempting to approach this ideal condition, the highway was resurfaced for a distance of 60 m (200 ft) in front of and 15 m (50 ft) beyond the scale.

Figure 3. Installing the scale.

Excavation concrete pad



Installing dummy frame

Dummy frame installed in roadway

Scale installation methods have been modified with each generation of the scale. The installation methods used with the third-generation scale are illustrated in Figure 3. They involved making the appropriate excavation in the roadway, pouring the required concrete base, and installing the dummy frame. The scale units themselves were then placed in the dummy frames, which give the scale units a degree of portability.

The data acquisition equipment was housed in a temperature-controlled trailer adjacent to the scale site. Axle-load information was recorded on a 24-h/d basis. In addition to the weigh scale, two magnetic loop detectors were placed in the roadway adjacent to the scale. These loop detectors turn on and shut off the magnetic recorder that is used to record the signal from the weigh-scale load cells and can also determine vehicle speeds. Approximately once a month the magnetic tape containing the recorded weigh-scale load cell information was picked up. This tape containing the analog output was digitized and analyzed using computer facilities located at the University of Saskatchewan. The information obtained included total traffic counts (cars included), detailed speed information, individual axle weights and axle spacing, vehicle types or classifications, and time of day associated with each of the above.

In the fall of 1975, a scale was installed on a major pulp haul road in Northern Saskatchewan near Montreal Lake on Highway 2 (Figure 4). Five-axle trucks involved in the pulp haul are permitted to carry 25 000 kg (55 000 lb) per tandem on this roadway. The data-acquisition equipment used at the site was identical to that used at the Clavet site.

Data have been collected on an unmanned basis at the Clavet and Montreal Lake sites since the fall of 1975. The scales have not required any maintenance or adjustments over the nearly 2-year period.

Digitizing and processing the analog tape obtained from the field data acquisition system has proved to be a time-consuming and costly procedure when utilizing the relatively archaic methods initially developed. Recent efforts have been devoted to developing more efficient methods of handling the data-processing requirements.

Figure 4. Montreal Lake installation.

These new methods were expected to be operational in the fall of 1977. The details of these new data acquisition methods are discussed in a following section.

A number of trial run series have been undertaken over the last 2 years to evaluate the performance of the scale unit. Figure 5 illustrates the results of three of these test series for the Montreal Lake installation.

It is apparent from Figure 5 that, while the average axle load measured by the scale is relatively insensitive to speed, the variations about the average are observed to increase with speed (as might be expected because of vehicle dynamics). Further, it is apparent from the figure that the observed variations for the October and December tests are significantly greater than those for the May tests. This can be attributed to roadway roughness.

The scale was installed during the early fall of 1975. Prior to testing in October, there was considerable settlement in the vicinity of the scale that resulted in a rough approach. The approach was improved with minor patching but again, before testing in December, differential frost movement caused deterioration that resulted in a rough approach for the December tests. Prior to the testing in May, minor surface improvements within 3 m (10 ft) of the scale were made to improve the smoothness of the approach wheel paths. This improvement in riding quality resulted in the considerably improved results for the May test.

The preliminary testing and evaluation work undertaken during the first 2 years of the development program have resulted in the following conclusions: (a) the scale developed at the University of Saskatchewan is capable of weighing vehicles in motion with sufficient accuracy to meet the information requirements of pavement and bridge engineers, and (b) the scale is rugged and reliable enough to be operated on a continuous unmanned basis in the harsh Canadian environment.

CURRENT PROGRAM

As a result of the widespread Canadian interest in developing a capability to weigh vehicles in motion, the Roads and Transportation Association of Canada formed a project committee to monitor vehicle and axle weights in 1976. The primary task of the project committee is to make recommendations for equipment and procedures for monitoring programs by which to determine the vehicle axle and gross load data required in the assessment of impacts of pavement and bridge structures.

As part of the activities of this project committee, three types of weigh-in-motion scales are being tested and evaluated at various locations in Canada during 1977 and 1978. The Viatec axle weight analyzer is to be evaluated at various locations in Alberta and Ontario by Alberta Transportation and the Ontario Ministry of Transportation and Communications respectively. The Texas weigh-in-motion system is to be evaluated in New Brunswick by the New Brunswick Department of Transportation. The University of Saskatchewan scale was installed in the fall of 1977 in Saskatchewan, Ontario, Quebec, and New Brunswick and the evaluation undertaken by the Saskatchewan Department of Highways and Transportation, the Ontario Ministry of Transportation and Communications, the Quebec Ministère des Transports, and the New Brunswick Department of Transportation respectively.

The University of Saskatchewan scale units to be installed will be as previously illustrated. The data acquisition system, which has been under development since the fall of 1976, will be a substantially improved system. The two alternative data acquisition systems under development will be installed as illustrated in Figures 6 and 7. The primary difference between the 11/03 and the 11/04 systems is that the summary reports of traffic will be produced in the field by the 11/03 system, whereas such reports will be produced by remote software analysis programs for the 11/04 system. No permanent individual vehicle data records will be maintained with the 11/03 system, but the 11/04 system will produce a permanent record of the vehicle data. This permanent data record could be available for analysis of historical traffic volumes and processing or both, as may be required. Similar flexibility regarding the analysis of historical data will not exist with the 11/03 system.

Figure 5. Montreal Lake installation test results.

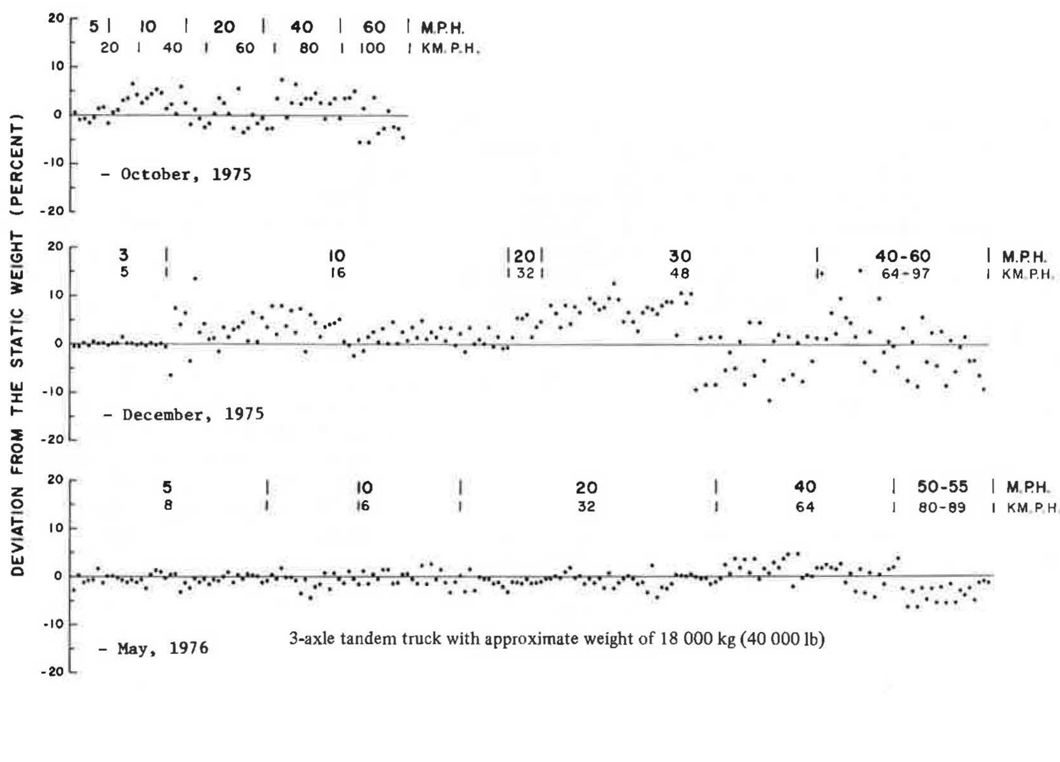


Figure 6. The 11/03 data acquisition system.

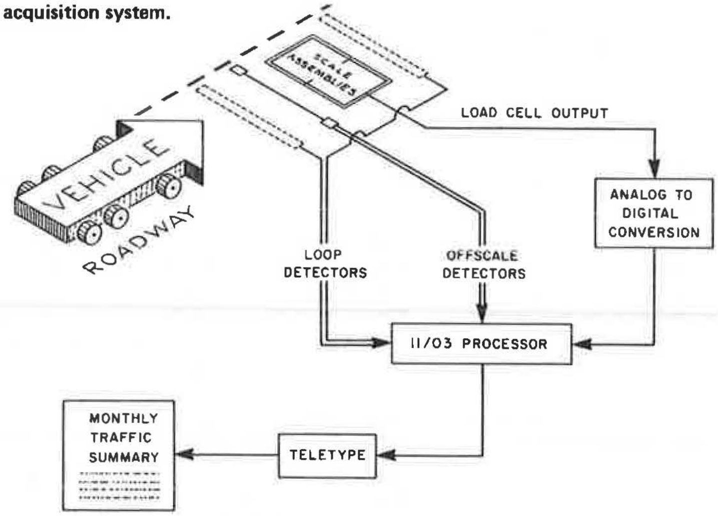
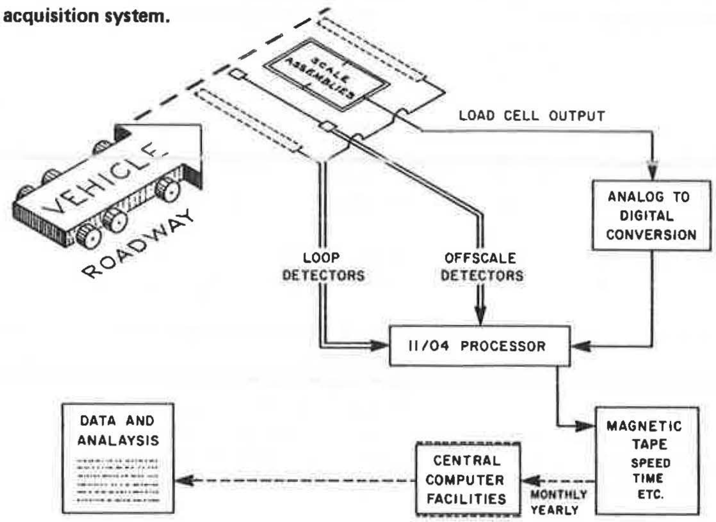


Figure 7. The 11/04 data acquisition system.



Upon completion of the evaluation of the various scales by the provincial organizations, a report giving a comparative assessment of the weigh-in-motion devices tested in the program will be prepared. The Roads and Transportation Association of Canada project committee will make recommendations regarding the suitability of the devices tested for various purposes. The report will include data on capital and operating costs and technical information on the quality of the data and the reliability of the equipment. Details concerning equipment, installation, recommended procedures for data collection, handling, and transformation will be provided.

ACKNOWLEDGMENT

The authors would like to acknowledge the support received from the National Research Council of Canada, the Saskatchewan Department of Highways and Transportation, Transport Canada, the provinces of Ontario, Quebec, Alberta, and New Brunswick, the University of Saskatchewan, and the Roads and Transportation Association of Canada.

REFERENCES

1. K. T. Solomon and B. H. Tritt. The Development of Weighing in Motion Equipment. Proc., 7th Conference of the Australian Road Research Board, Vol. 7, pt.3, paper A87, 1974.
2. H. Kalisch. Installation and Operation of 146 Axle Load and Axle Counting Units for the Research Project "Long-Term Observations of Roads." Strasse und Autobahn, Vol. 27, No. 11, 1976.
3. O. K. Normann and R. C. Hopkins. Weighing Vehicles in Motion. HRB, Bulletin 50, 1952.
4. A. T. Bergan, G. A. Sparks, and G. Dyck. Development of an Automatic Highway Scale. Canadian Technical Asphalt Association Annual Proceedings, 1975.
5. A. T. Bergan, G. A. Sparks, and G. Dyck. Development of a Transport Data Acquisition System. Roads and Transportation Association of Canada Annual Proceedings, Ottawa, 1975.
6. A. T. Bergan, G. A. Sparks, and G. Dyck. Evaluation of the Highway Scale. Department of Civil Engineering, Univ. of Saskatchewan, Saskatoon, Internal report, 1976.

Discussion

J. H. Havens, Kentucky Department of Transportation

Some agencies that have invested heavily in the development of devices to weigh highway vehicles on-the-run can fully appreciate the futility of almost achieving success. The team of authors have, here, asserted their success matter-of-factly. The paper does not reveal the pitfalls they have avoided or escaped. Indeed, two years of operating experience in an unmanned mode, with only monthly harvesting of data, is an impressive accomplishment.

Kentucky has two of the Texas scales (7); one site has produced 53 manned days of data out of 763; the other has produced 3 out of 333. Neither is operative at the present time. Downtime of the platforms, however, has been only about 20 percent.

Before succumbing to the Texas system, we abandoned a very sophisticated platform and data system developed for the Department of Highways by the University of Kentucky during 1961 and 1971 (8, 9). Of several designs investigated, a "broken-back" platform—that is, two simple spans with abutting ends supported commonly on load cells—was judiciously selected. It was modeled after one developed at the Otto-Graf Institute, Stuttgart, Germany, in 1958 (10).

This type of platform produces a triangular output signal from the load cells as a load (axle) traverses the platform. The apex of peak of the triangular signal from the load cells is calibrated in weight units. The unique feature of this type of design is that the base leg of the triangle represents the span length; the addition of an internal timing signal permits speed of traverse to be calculated. Then, by presetting a practical time gap between vehicles, it is possible to determine the number of axles per vehicle (classification) and to sum the several axle loads, which yields a gross load for each vehicle. Thus, the digitized output capabilities of the system are: load impulses of individual axles, vehicle speed, gross load, and vehicle classification by number of axles. Various statistical analyses may be programmed to determine specific characteristics of the traffic stream.

The axle loads sensed by this system are not necessarily equivalent to static weights. Vehicles in motion tend to undulate or bounce as they travel; there is a random probability or likelihood that a vehicle (or axle) will be on an "upswing" or "downswing" when it crosses the platform. The most unlikely events would be to catch an axle at either extreme or at its null (equivalent static) state; however, there is a greater probability that an axle will be closer to a null condition than to an extreme as it crosses the platform. The standard error of estimate is judged to be of the order of ±5 percent of the static weight (11). Statistically speaking, the errors tend to cancel, and so the use of the system for survey purposes is not impaired.

Whereas the scale system is capable of measuring the force exerted by a set of wheels moving at high speeds, the force impressed on the platform is simply not the static weight force of the axle. The ratio of the peak downward forces to the static weight force defines impact factor. This explanation merely emphasizes the fact that the weighing platform senses only the instantaneous, dynamic force of each transient axle.

Despite overwhelming hardware failures that beset the development of an automatic, in-stream, vehicle-weighing system—which we then became convinced we must abandon—significant measures of success were achieved. In other words, we created an automation that almost worked. The decision to abandon the prototype installation arose from pilot operations and proof testing. The basic defect was in the weighing platform in the pavement. Unfortunately, it was a design defect. Tie rods anchoring the platform in the pit induce a purposeful preload on the load-sensing elements. These tie rods change the preload as the temperature fluctuates. Thus, the balance or null point drifts. The noticeable effect was a triggering of the counting and weighing circuits when there was no live load on the platform. Since this load was not transient—but sustained—the circuitry "locked in" on the excess preload. The preload and tie rods were intended to keep the platform in firm bearing on the load-sensing units and to eliminate resonances and friction. Conceivably, it would have been possible to control the temperature in the pit, but other factors were equally dissuasive.

Whereas the cargo box or principal mass of a heavy vehicle may be on the downswing or about to "bottom out" as it passes over a weighing platform, the most abrupt change (reversal) in direction at this point induces the greatest force on the platform. The acceleration imparted tends to cause the mass to rise higher on the springs and to soar or dwell longer on the upswing. If the axle were to bounce off the pavement, the upward acceleration would necessarily have been greater than 1 g. On the other hand, the maximum downward acceleration may never exceed 1 g. Thus the "impact factor" at the end of the downward excursion is the greater. In other words, the centroid of the points exceeding the static weight will lie farther above the static-weight line than the centroid of the points showing less than the static weights. The number of points should be about equally divided—that is, half should be above and half should be below the line. These are prerequisites in the performance criteria of a weighing-in-motion device—regardless of the speed of vehicles. Perhaps the authors should comment further in regard to Figure 5 in their report and perhaps define for us what they mean by "average axle load."

REFERENCES

7. C. E. Lee and N. I. Al-Rashid. A Portable Electronic Scale for Weighing Vehicles in Motion. Center for Highway Research, Univ. of Texas, Austin, Research Rept. 54-1P, 1968.
8. Weighing Vehicles in Motion. Univ. of Kentucky Research Foundation and Kentucky Department of Highways and Bureau of Public Roads, Research Rept. 284, 1969.
9. B. W. Carr, Jr. Development of an Electronic Means of Weighing Vehicles in Motion. Division of Research, Kentucky Department of Highways, Rept. 307, 1971.
10. W. Schwaderer and W. Reimund. Die Automatische Achlastwaage bei Grunback (Remstal). Strasse und Autobahn, 1959.
11. Weighing in Motion. Department of Civil Engineering, Univ. of Kentucky and Kentucky Department of Highways and Bureau of Public Roads, Research Rept. 212, 1964.

Stephan Fregger, Bureau of Planning, Florida Department of Transportation

The paper presented by Bergan, Sparks, and Dyck is quite well prepared and clearly written. It adds an important new chapter to the knowledge of dynamic weighing. There is every reason to believe that the University of Saskatchewan scale system can accurately weigh, record, and analyze heavy vehicles with the torsion arm load transfer. Apparently it can also perform satisfactorily under very cold conditions for an extended period of time (2 years).

Two types of questions come to mind. The first relates to details that would aid in the potentially wide application of the research to operational use. For example, one wonders how many scales are installed at the weigh site. The use of the word "scale" in the singular and references to axle weights (not wheel weights) would tend to imply a single unit. Yet dimensional measurements of the scale indicate a size only big enough to accommodate a single wheel path (or indicate very narrow Canadian trucks).

Reference is made to a "degree of portability" of the scale. Does this mean only that the scale can be shop-constructed and then delivered to the weigh site, or does it imply portability in the sense of convenience for periodic relocation from site to site?

It would be helpful to know the expected order of magnitude of cost of a typical installation, the approximate cost of a scale unit, site preparation, field processor, and so forth.

In addition to the several applications questions, there is a second type of question of even greater relevance. This is directed at the implied premise of the paper, that it is desirable to obtain an in-motion weighing scale capable of operating in a continuous unmanned manner in the harsh Canadian environment.

The authors are correct in noting the need for comprehensive roadway traffic data to provide information for highway maintenance and design. Those data have traditionally been obtained from the three-tiered counts-classification-truck weight program. Traffic counts are obtained from a large sample of sites representing the range of road systems and geographic locales; vehicle classifications are obtained from a sample of the count stations to determine the percentage of trucks in the traffic stream; weigh stations are established at a sample of the classification stations in order to determine trends in truck weights, configurations, and dimensions. The weight trends are factored up through classifications and counts to predict the load replications essential to design and maintenance.

The concept of weight trend is crucial. What is needed are weight data from a representative sample of trucks in their principal uses, across a broad geographic coverage, and over a long term. The trend or time series analysis of truck weights is employed because change in fleet and deployment of trucks is generally quite slow. As a matter of fact, because the annual change is generally so slight, many states are now considering conducting weight surveys only on alternate years.

The question, then, is whether the Saskatchewan scale, with its 100 percent sample of trucks weighed at a continuously operating site for almost 2 years, is an appropriate step forward in dynamic weighing. Unfortunately, the answer is not yet clear, since it will depend upon the responses to the earlier questions of size, portability, and cost. I suspect, however, that the answer is negative.

I also suspect that the rugged design required to permit the unattended, continuous usage may have sacrificed practical portability and precluded inexpensive fabrication and installation. If such is the case, then the Saskatchewan scale may be a regression from the successful Texas weigh-in-motion system.

That system, as adapted by the Florida Department of Transportation, has been satisfactorily operated in Florida for several years. Using a single pair of transducers and operated by a three-man crew (in order to obtain 24-h coverage) and a single climatized mobile trailer that houses the field computer, Florida covered 15 weigh-in-motion survey sites in 1977 and weighed more than 425 000 vehicles, including approximately 60 000 trucks. The sites were geographically distributed throughout the state. The data collected appear to be statistically stable and satisfactory for our needs. In 1978 we plan to expand to 20 sites to improve our weigh-in-motion coverage.

Authors' Closure

Some additional comments may clarify some of the issues raised by Mr. Havens and Mr. Freeger.

A typical installation using the University of Saskatchewan scale includes two weighing platforms, one in each wheel path. The output signals from each of the these weighing platforms are then summed to yield an axle weight.

Concerning portability of the scale units, procedures have been developed that facilitate the movement of a scale from one location to another. These involve the preparation of a particular site with the installation of frames and dummy units. These dummy units can be lifted out of the frames and weighing platforms installed. This procedure takes approximately 3 h to complete. The portability feature of the University of Saskatchewan scale would permit the use of a pair of weighing platforms at a number of different sites.

The University of Saskatchewan scale is only in the development stage, and thus it is impossible to comment on the costs apart from saying that the computer equipment required for the acquisition of data makes up a substantial portion of the total cost of an installation. Further, the cost of the computer equipment is highly variable and depends on the degree of sophistication desired. Details regarding costs will be available by the fall of 1978 and will be one of the topics covered in the final report for the evaluation project.

We agree with Mr. Fregger regarding the question on the desirability of obtaining continuous data from a single location. In looking at the application of the University of Saskatchewan scale for the collection of vehicle and axle weight data, there would appear to be considerable merit in investigating the appropriateness of a small number of permanent installations within a province or state with a larger number of sites with frames and dummies, which would allow periodic sampling throughout the highway network.

Finally, regarding Mr. Havens' comment on the large number of pitfalls, he can be assured that, in the course of developing the scale, we have not escaped all of the pitfalls. We may have been able to avoid some of the more serious ones because of having the advantage of others' experiences.

Publication of this paper sponsored by Committee on Highway Capacity and Quality of Service.

Estimation of Left-Turn Saturation Flows

Panos G. Michalopoulos*, University of Minnesota, Minneapolis
Jerome O'Connor and Sergio M. Novoa*, Rensselaer Polytechnic Institute, Troy, New York

This paper addresses the problem of estimating left-turn saturation flows at both signalized and unsignalized intersections. The best known methods for estimating this traffic measure were tested for reliability against field data. A new approach to the problem is also presented. The results are applicable to left-turning traffic flowing through gaps of suitable size in the opposing traffic without the protection of a special signal phase. Both one and two lanes of opposing traffic were considered, as were unsignalized intersections. The gap-acceptance functions that best represent the behavioral patterns of left turners at these intersections are also presented. Time-lapse photography was used to collect data at five different intersections in upstate New York. Approximately 4000 completed left-turning movements were observed. Using these field observations, the ability of several existing methods to estimate left-turn saturation flow was tested by standard statistical analysis techniques. Adjustments were made for any divergence from actual conditions. Most of the original models do not reflect real-world conditions, so a new model is proposed for each type of intersection. The results indicate that the gap-acceptance characteristics of left-turners can be accurately described by a uniform cumulative density function.

Although several methods for estimating left-turn capacity at intersections have been proposed, none has been widely accepted by practicing traffic engineers as being truly representative of real-world conditions. May (1) reports that research on left-turning movements rated second in priority over twenty other items of interest related to intersection capacity, in a 1974 survey.

This paper will be concerned primarily with the left-turn capacity of an intersection, which is easily computed from saturation flow. The saturated condition under consideration is illustrated at unsignalized intersections by a stream of left-turning vehicles moving continuously and restricted only by the presence of the opposing through movement. Pedestrian traffic on the cross street that might interfere with vehicles attempting to turn is assumed to be negligible.

Since the flow at signalized intersections is controlled by the amount of green time allotted, the left-turn saturation flow under these conditions is defined as the flow rate of left-turning vehicles that would be obtained if there were a continuous queue of vehicles given 100 percent green time (2). Left-turn capacity is then given by the actual possible number of left turns in one hour, considering the effects of the signal.

Reliable estimates of left-turn saturation flows have several applications in traffic management and design. Such applications include

1. Decisions concerning the installation of a traffic signal at unsignalized intersections,
2. Determination of optimum signal timing,
3. Determination of optimum signal-phasing arrangements,
4. Estimation of the average queue length used in the design of left-turn bays, and
5. Estimation of the average and maximum delays for left-turning vehicles.

In an effort to obtain estimates that represent actual conditions, several approaches have been taken, resulting in theoretical or semi-empirical solutions to the problem. The most widely known were considered in this study and are summarized in Table 1,

where

S_l = left-turn saturation flow in vehicles per hour,
Q_o = opposing flow in vehicles per hour,
q_o = opposing flow in vehicles per second,
τ = critical gap in seconds,
h_o = mean minimum opposing headway in seconds, and
h_l = mean minimum left-turn headway in seconds.

Tanner's model (3) was initially derived for a single lane of opposing vehicles, but it was proposed that the condition of multilane opposing could be approximated by regarding the opposing vehicles as a single stream with an arrival rate twice that of a single lane and one-half the minimum headway. Webster and Cobbe (2) used these theoretical equations in developing curves for estimating left-turn saturation flows. Estimations of the model parameters such as critical gaps and minimum headways were based on data collected in the field and on the test track. However, validation of these curves with field data is still lacking (4).

Drew (5) derived an equation based on the analysis of gap-acceptance behavior for drivers merging at freeway ramps. Since this model resembles the condition under question, it was later proposed for estimating the left-turn capacity (6). Using observations to obtain appropriate values for two of these three independent variables in Drew's equation, Fambro, Messer, and Andersen (6) suggested a simplified form of the model, which is also presented in Table 1.

Table 1. Left-turn saturation flow formulations.

Model	No. of Opposing Lanes	Equation
Tanner	One	$S_l = [Q_o(1 - h_o q_o)]/\exp[q_o(\tau - h_o)]\,[1 - \exp(-h_l q_o)]$
Tanner	Two or more	$S_l = [2Q_o(1 - h_o q_o)]/\exp[2q_o(\tau - \frac{1}{2}h_o)]\,[1 - \exp(-2h_l q_o)]$
Webster	One	$S_l = \{Q_o[1 - (3)q_o]\}/\exp\{q_o[(5) - (3)]\}\,\{1 - \exp[-(2.5)q_o]\}$
Webster	Two or more	$S_l = \{2Q_o[1 - (1)q_o]\}/\exp[2q_o(6) - \frac{1}{2}(1)]\,\{1 - \exp[-2(2.5)q_o]\}$
Drew	Any	$S_l = Q_o\{[\exp(-q_o\tau)]/[1 - \exp(-q_o h_l)]\}$
Fambro, Messer, Andersen	Any	$S_l = Q_o\{\{\exp[-q_o(4.5)]\}/\{1 - \exp[-q_o(2.5)]\}\}$
HCM	Any	$S_l = 1200 - Q_o$
Australian Road Capacity Guide	Any	$S_l = 1200f$, where f is given by: Q_o: 0, 200, 400, 600, 800; f: 1.0, 0.81, 0.65, 0.54, 0.45

The formulations proposed by the Highway Capacity Manual (HCM) (7) and the Australian Road Capacity Guide (8) are brief and easy to follow. The maximum left-turn flow rate in vehicles per hour (vph) of 1200, which corresponds to a minimum headway of 3 s/vehicle and no opposing traffic, is diminished in proportion to the opposing flow. In the case of the HCM, the left-turn saturation flow is equal to the difference between 1200 vph and the opposing flow, with theoretically zero left turns at opposing flows of 1200 vph. However, in capacity calculations, it is stipulated that the number of turns will not be less than two vehicles per signal cycle, regardless of the opposing flow. The Australian method decreases the left-turn flow rate in a nonlinear manner as the opposing volume increases. However, this method is not applicable to opposing flows greater than 800 vph.

The most obvious advantages and disadvantages of the various methods are summarized in the following table.

Model	Advantages	Disadvantages
Tanner	Distinguishes between one and two opposing lanes Closed-form solution, easy to apply	Lacks sufficient validation Does not adequately agree with field data in any of the cases studied
Webster	Distinguishes between one and two opposing lanes Solution easily obtained from graphs	Lacks sufficient validation Overestimates saturation flow at the lower range of opposing flows Does not adequately agree with field data in any of the cases studied
Drew	Closed-form solution, easy to apply Relative agreement with field data in cases with two opposing lanes	Was primarily developed for estimating merging capacity at entrance ramps Oversimplifies assumptions in deriving the solution Does not distinguish between one or two opposing lanes
Fambro, Messer, Anderson	Validated with field data Simple, and easy to apply	An extension of Drew's solution and has similar disadvantages Does not adequately agree with field data in any of the cases studied
HCM	Relatively accurate for opposing flows <600 vph Easy to apply	Assumes left-turn saturation flow is zero for opposing flow > 1200 vph Underestimates left-turn saturation flow for opposing flows between 600 and 1200 vph Does not distinguish between one or two opposing lanes
Australian Road Capacity Guide	Based on a semi-empirical gap-acceptance behavioral model Close agreement in cases with two opposing lanes	Valid only for opposing flows <800 vph Does not distinguish between one or two opposing lanes

Observations made after the evaluation of the models are also indicated in the table. Since none of the methods presented has been unanimously accepted by practicing traffic engineers, field data were collected and a comparative analysis was performed to determine the models best representing actual conditions (9).

Traffic was observed at both signalized and unsignalized intersections having exclusive left-turn lanes including the effects of opposing traffic in one or two lanes. Although existing methods were found to be fairly realistic in some instances, their application is limited to only a small number of cases or to a certain range of volumes. The general disagreement of the field data with the theoretical results led to the development of the alternate method presented in this paper. The new method allows a choice among the most suitable of the existing models for a particular case or statistical models developed from the collected data.

Data were also examined microscopically to determine the distribution of gap sizes accepted by left-turners confronted with opposing traffic. Thus, gap-acceptance functions were derived by allowing the estimation of the percentage of drivers accepting a gap of a particular size.

DATA COLLECTION AND ANALYSIS

By using a time-lapse camera, approximately 4000 completed left-turn movements were observed at five different upstate New York intersections. The test sites were located in a typical suburban environment near central shopping areas. A film exposure of one frame per second was selected as desirable for recording intersection data of this type. A total of 11 h of selected data were collected. This was considered a sufficient sample size for making statistical inferences concerning the general behavior of left-turning vehicles.

Although we initially intended to investigate all possible left-turn conditions, personnel and time constraints precluded the inclusion of left-turn movements from optional through and left-turn lanes. Thus, intersections with the following characteristics were selected:

1. With or without signalization,
2. With opposing traffic moving in one or two lanes,
3. With an exclusive left-turn lane, and
4. Without a separate signal phase for left turns.

In addition to these traffic control conditions, certain other factors were considered in the selection of test sites:

1. Intersection isolation to ensure random arrivals,
2. Sufficiently high left-turn demands to allow continuous left-turn queues during most of the observation periods,
3. Full range of opposing traffic,
4. Grades (a zero grade was sought on all approaches),
5. Clear visibility,
6. Parking (no parking allowed on any approach),
7. Approach speed [an average free-flow speed of 48-56 km/h (30-35 mph) was considered in all the cases], and
8. Good pavement conditions and satisfactory pavement markings.

After the selection of test intersections representative of the conditions stated, the most influential traffic variables affecting left-turn saturation flows were identified according to observed findings and an exhaustive literature review. Thus, it was concluded that the dominant independent traffic variables are flow rate of the opposing through movements, critical gap, minimum left-turn headway, and minimum opposing headway.

Among other traffic variables considered were average approach speed and percentage of trucks, buses, and motorcycles. It was found, however, that these variables either have secondary importance or they are indirectly included in the above four.

Time-lapse photography was selected as the most efficient method of data collection because of its clearly superior advantages over other alternatives. The advantages of this system include: time savings, easy operation, low operating costs, reliability of data (precise and complete records), permanent record of data, technical dependability, and minimum personnel requirements. The benefits of the time-lapse photography ensured an exact quantitative account of all traffic variables in question for subsequent reconstruction of the test conditions.

Observations were made at intersections where a queue of left-turning vehicles had formed because of opposing through traffic. The maximum flow rate of the turning vehicles was then recorded for the time period when a queue existed and converted to vehicles per hour of green. Since left-turning drivers at signalized intersections use both the initial green period and the yellow at the end of the phase, there is little lost time. For this reason, opposing flow was measured for the entire time available for making turns while left-turn demands were present. This time included at most the entire green and the yellow clearance intervals.

With the necessary information available and in usable form, a quantitative comparison of existing methods for estimating left-turn saturation flows was carried out. Classical statistical analysis techniques were employed in testing the reliability of the models. The measures used to compare the models were the coefficient of determination (R^2) and the standard error of the estimate (S_ε). The F-test, t-test, and chi-square test were also employed for qualitative testing in the analysis.

Using the available data, the existing models were subsequently modified to obtain a higher degree of correlation with the observations. Regression analyses were performed on each of the equations of Table 1 to adjust each model to the data. (Because the Australian model is given by tabulated values, rather than an equation, it was not adjusted.) The general form of the adjusted models is

$$S_L = b_0 + b_1 X \tag{1}$$

where

S_L = left-turn saturation flow,
X = original form of the model, and
b_0 and b_1 = regression coefficients.

In the statistical analysis, the BMDP2R computer program package developed by the University of California (10) was used. This package has been extensively tested in the past and has been widely accepted for statistical analyses of a similar nature.

Because many of the models were too inaccurate, even after the adjustment, a new model was developed for each condition under consideration, using multiple regression to achieve the closest fit to the observed data. By assigning qualitative variables to represent the number of opposing lanes of traffic and the presence of a signal, a composite model was also devised, using the data collected from all the test sites. This general model allows the traffic engineer to easily arrive at a reasonably accurate estimate of the left-turn saturation flow for any combination of roadway characteristics previously described.

In the gap-acceptance study, the probability density functions most widely employed in studies of similar nature were tested with standard statistical tests, and the most appropriate for each case were singled out.

RESULTS

Based on the results of the statistical analyses performed for the original and adjusted models, the models were ranked according to their ability to match the observed data. A summary of the results is given in Table 2 with the standard error of the estimate (S_ε) and the coefficient of determination (R^2) given for a quantitative comparison. A case-by-case study was performed to detect the significance of signalization and the number of opposing

Table 2. Summary of results.

									Regression Models		
Case	Ranking	Model	S_ε	R^2	Ranking	Adjusted Model	S_ε	R^2	Form	S_ε	R^2
1	1	Drew	142	0.73	1	Fambro	141	0.74	Polynomial[b]	139	0.76
	2	Australian[a]	169	0.54	2	Drew	142	0.73			
	3	Fambro	175	0.59	3	Tanner	150	0.70			
	4	HCM	303	0.00	4	HCM	162	0.65			
	5	Tanner	324	0.00	5	Webster	168	0.62			
	6	Webster	366	0.00							
2	1	Australian[a]	168	0.50	1	Fambro	157	0.57	Polynomial[b]	148	0.62
	2	HCM	208	0.23	2	Drew	165	0.53			
	3	Webster	221	0.13	3	Webster	165	0.53			
	4	Tanner	245	0.00	4	Tanner	170	0.50			
	5	Fambro	269	0.00	5	HCM	176	0.46			
	6	Drew	341	0.00							
3	1	Australian[a]	107	0.83	1	Tanner	94	0.85	Polynomial[b]	92	0.86
	2	Drew	136	0.69	2	Webster	104	0.82			
	3	HCM	195	0.34	3	Fambro	105	0.81			
	4	Fambro	248	0.00	4	Drew	106	0.80			
	5	Webster	269	0.00	5	HCM	136	0.68			
	6	Tanner	301	0.00							
4	1	Tanner	194	0.25	1	Fambro	122	0.70	Polynomial[b]	114	0.74
	2	HCM	214	0.05	2	Webster	133	0.64			
	3	Webster	247	0.00	3	Tanner	137	0.61			
	4	Australian[a]	263	0.00	4	Drew	138	0.58			
	5	Drew	286	0.00	5	HCM	143	0.57			
	6	Fambro	427	0.00							
1, 2, 3, and 4 combined	1	Australian[a]	206	0.48	1	Fambro	169	0.55	Composite[b]	137	0.71
	2	Drew	237	0.11	2	Drew	172	0.54			
	3	HCM	247	0.03	3	HCM	197	0.40			
	4	Fambro	280	0.00							
2 and 4 combined	1	Tanner	217	0.16	1	Tanner	164	0.54	–	–	–
	2	Webster	232	0.03	2	Webster	178	0.45			
1 and 3 combined	1	Tanner	300	0.00	1	Tanner	136	0.73	–	–	–
	2	Webster	342	0.00	2	Webster	264	0.65			

[a] Tested for opposing flows <800 vph. [b] See Table 3 or the regression models.

lanes in the selection of the best model. However, since most of the methods do not make any such distinction, an overall comparison was also made.

The six tested models are listed according to increasing standard error of estimate and the corresponding decrease in the coefficient of determination for each case. R^2 values close to zero indicate that the model is unrealistic in estimating left-turn saturation flows. According to these results, Drew's original formulation is very close to the best fit of the data for signalized intersections with two lanes of opposing traffic. However, when all cases are combined, the standard error of the estimate and the coefficient of multiple determination are 237 and 0.11 vph, respectively, suggesting failure of the model.

The Australian method presented in Table 1 can be used only for opposing flows less than 800 vph, and therefore it was tested for only this range of opposing flows. Observation of the results presented in Table 2 leads to the conclusion that this is the best of the existing methods for two of the four cases studied, assuming the restricted volume range indicated earlier.

For unsignalized intersections with one opposing lane, none of the existing models adequately estimates the observed values. The low coefficients of determination in most of the unadjusted models clearly indicate that it is important to consider more complex formulations for each particular case. Thus, as a first step the existing models were adjusted according to Equation 1, and the results are presented in Table 3.

In this table, values of b_0 and b_1 close to 0 and 1 respectively indicate that the original model is realistic. Each model has individual characteristics, so the ranking of the adjusted models is not identical to that of the unadjusted models. Inspection of Table 2 reveals that most of the models are substantially improved by the adjustment. The most dramatic improvement is observed from the Fambro, Messer, and Andersen model, which is ranked as the best in three of the four cases studied and also in the combined category.

It can also be seen from Table 2 that the results obtained from the best adjusted models are fairly close to those of the polynomial regression models (where S_L is left-turn saturation flow, Q_o is flow in the opposing arm, and T is critical gap), which are based on a 99 percent confidence level. The polynomial regression models, however, are closer to the observed data for all cases as expected. The regression equations are presented below.

Polynomial Model Case	Intersection Type	Equation
1	Signalized, two opposing lanes	$S_L = -0.875Q_o + 0.000\,012\,Q_o^2T + 1145$
2	Signalized, one opposing lane	$S_L = -1.245Q_o + 0.000\,014\,Q_o^2T + 1165$
3	Unsignalized, two opposing lanes	$S_L = -0.277Q_oT + 0.000\,012Q_o^2T^2 + 1172$
4	Unsignalized, one opposing lane	$S_L = -0.324Q_oT + 0.000\,012Q_o^2T^2 + 1142$

where

S_L = left-turn saturation flow in vehicles per hour,
Q_o = flow in opposing arm in vehicles per hour, and
T = critical gap in seconds.

For all cases combined, the composite model equation is

$$S_L = -0.233Q_oT + 0.000\,015Q_o^2T^2 + 126\,L + 103\,S + 995 \quad (2)$$

where

L = 0 if there is one opposing lane and
L = 1 if there are two opposing lanes; and
S = 0 if the intersection is unsignalized and
S = 1 if the intersection is signalized.

As mentioned earlier, in addition to the independent variables and the terms which appear in the final form of the equations, others were also considered but they were found to be statistically insignificant.

It is evident from the form of the equations that signalization affects left-turn saturation flow. Further, it is observed that both signalized cases (1 and 2) are consistent in form, as are unsignalized cases (3 and 4). The relative ease with which the independent variables are obtained allows direct field application for all practical purposes.

In order to further simplify the application of the polynomial models, the results were combined to form the single comprehensive model presented above (Equation 2) as the composite model. The effects of signalization and number of opposing lanes of traffic were considered by adding qualitative (dummy) variables to the general form used in the development of the individual models.

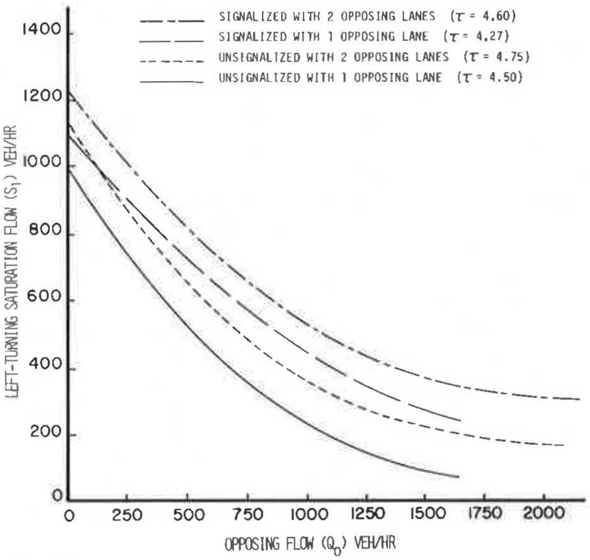
Figure 1 illustrates the results obtained from the composite model representing the four cases studied. The four distinct curves verify that there is a significant difference between the left-turn saturation flows at signalized and unsignalized intersections with one or two opposing lanes. It should be noted that the composite model yields values that are very close to those of the first four equations of the polynomial model.

As depicted in Figure 1, the curves corresponding to signalized intersections (cases 1 and 3) are shifted up by approximately 100 vph in comparison with the unsignalized pair of curves (cases 2 and 4). This phenomenon is accounted for by the change in driver behavior caused by psychological stress when turning at a signalized junction. Since the driver who is waiting to turn faces the possibility of being delayed for a full cycle if the turn is not completed before the red begins, every attempt will be made to pass through the intersection before the amber period. Thus, the flow of turning vehicles is hastened and the maximum output is increased. It is possible that the driver's critical gap will be decreased in the process, but this is not entirely necessary.

Table 3. Calibration coefficients of adjusted models.

		Case 1		Case 2		Case 3		Case 4		Cases 1, 2, 3, and 4 Combined		Cases 2 and 4 Combined		Cases 1 and 3 Combined	
Model	Equation	b_0	b_1	b_0	b_1	b_0	b_1	b_0	b_1	b_0	b_1	b_0	b_1	b_0	b_1
Tanner	$S_L = b_0 + b_1$	306	0.794	246	0.459	316	0.776	39	0.702	—	—	160	0.544	307	0.787
Webster	$S_L = b_0 + b_1$	365	0.658	303	0.487	292	0.666	65	0.852	—	—	223	0.475	347	0.656
Drew	$S_L = b_0 + b_1$	-41	0.926	-264	0.862	-115	1.070	-256	0.958	28	0.715	—	—	—	—
Fambro	$S_L = b_0 + b_1$	-44	0.777	-75	0.812	-106	0.827	-370	0.954	6	0.684	—	—	—	—
HCM	$S_L = b_0 + b_1$	459	0.414	310	0.502	292	0.520	61	0.630	367	0.345	—	—	—	—

Figure 1. Saturation flows given by composite regression model.



The left-turn saturation flow was observed to be higher with two opposing lanes of traffic than with one, for the same type of intersection (signalized or unsignalized) for any given value of opposing flow. This is due to the ability of the opposing traffic to move simultaneously on two lanes rather than one and results in a larger number of acceptable gaps for the same value of opposing flows. It should be noted that with two opposing lanes it is possible to observe gaps close to zero, but for a given opposing flow, there is a higher probability of larger gaps, so that overall there is more opportunity to turn.

As can be seen in Table 2, the composite model has a fairly high coefficient of determination and a standard error of 137 vph, which is substantially lower in comparison with the existing left-turn saturation flow models. These statistics indicate that the model is reliable for the estimation of the desired values, although the particular polynomial models are slightly more accurate on a case-by-case basis.

The effectiveness of the composite and regression models can also be visualized in Figure 2, in which they are plotted along with the observed data, the best unadjusted, and the worst unadjusted models for comparison purposes. The data correspond to the case of a signalized intersection with one opposing lane. However, it should be noted that similar results were obtained for the remaining cases. Tanner's model is one of the least accurate unadjusted models, and the figure indicates that it generally overestimates saturation flows when opposing traffic is less than 900 vph, while above this value left-turn saturation flows are underestimated. The figure also illustrates that the polynomial models are best for the entire opposing flow range, while the Australian method yields results similar to these models for opposing volumes less than 800 vph. Incidentally, it should be pointed out that the Australian method is not adjusted, since it presents results in a tabular form rather than in a closed-form expression.

Knowledge of left-turn saturation flows allows estimation of the left-turn capacity that can be expected at the intersection. In the case of an unsignalized intersection, the left-turn capacity is simply equal to the saturation flow, since a continuous queue of drivers is capable of turning left without the interruption of a signal. However, at a signalized intersection, the actual number of cars that can turn left in 1 h is affected by the traffic signal. Thus, considering the discharge time required for the queue formed at the beginning of green, the following equation is applicable (8)

$$C_L = S_L [S_o g - q_o C/C(S_o - q_o)] + 3600(K/C) \quad (3)$$

where

C_L = left-turn capacity in vehicles per hour,
S_L = left-turn saturation flow in vehicles per hour of green,
S_o = saturation flow of opposing traffic,
g = effective green time,
C = cycle length, and
K = average maximum number of turns per phase change.

Caution should be exercised when the above equation is used in conjunction with the proposed models, since the effects of cars turning ahead of time at the beginning of green or during the yellow interval should be taken into account. Furthermore, Equation 3 suggests that left-turn movements can discharge at rates to saturation flow only after the queues of the opposing traffic, formed at the beginning of green, are dispersed.

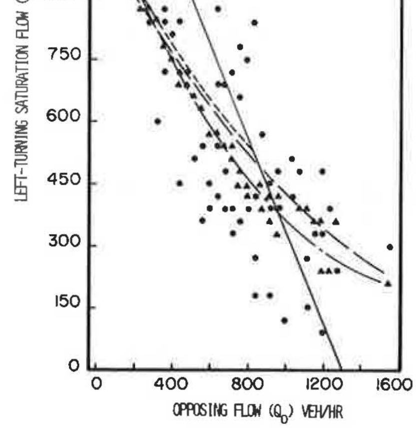
GAP-ACCEPTANCE CHARACTERISTICS

A secondary objective of this study was to determine the probability distributions that best describe the behavior of left-turning drivers at signalized and unsignalized intersections with one or two opposing lanes. Knowledge of gap-acceptance characteristics is needed in traffic simulation, the design of control systems, the computation of delays resulting from left-turning traffic, and so on.

The distributions most widely employed to describe gap-acceptance functions are uniform (trapezoidal) distribution, shifted negative exponential distribution, Erlang distribution, and log-normal distribution (5).

Using time-lapse films, an analysis was performed in order to determine the gap size (T) accepted by each

Figure 2. Left-turn saturation flow versus opposing flow at signalized intersection with one opposing lane.



driver. From these data a plot was obtained showing the cumulative number of gaps accepted (expressed in percentages) for gap intervals ranging from 1 to 10 s in 1-s increments. This cumulative function was then compared with the cumulative density functions of the above distributions.

The uniform distribution was found to be a very good description of actual driver behavior in all four cases. The resulting probability density functions are presented in Table 4, and a representative cumulative function is shown in Figure 3.

Chi-square tests were also performed on all equations to test goodness of fit, which was found acceptable at the 99 percent confidence level. The coefficients of determination (R^2), which are very close to 1, indicate that very little error is encountered when describing these data with a linear relationship. The computed standard errors of the estimate revealed deviations within less than 1 percent of the regression line.

It can be safely concluded, therefore, that for this set of intersections the uniform distribution realistically represents the actual gap-acceptance characteristics. Theoretically, then, an equal number of drivers will accept one gap as will accept any other gap between the upper and lower bounds of the equation.

The data also suggested that the first case, signalized intersection with two opposing lanes, has a longer range of acceptable gaps. The span is approximately 10 s long as compared to 9 and 8 s for the other cases, primarily because the larger gaps are required for completion of the turning movement across two traffic lanes. However, when the road junction is unsignalized, the number of opposing lanes of traffic appears to have little effect on the range of acceptable gaps. For instance, the gap accepted by 100 percent of the drivers (c_1) and the minimum gap size (c) are almost identical for cases 3 and 4.

CONCLUSION

The most widely known methods for estimating left-turn saturation flow were analyzed and found, in general, to be unsatisfactory for realistically predicting the saturated conditions observed in the field. Depending on the model and type of intersection tested, the estimates deviate from the actual values by as much as 400 vph. In a few instances, however, the existing models were found to be essentially as good as the regression models derived from the data.

For example, Drew's equation, when used at signalized intersections with two opposing lanes, provides a very reliable estimate of left-turn saturation flows. The Australian method is fairly satisfactory for signalized intersections with one opposing lane and for unsignalized intersections with two opposing lanes. For the case of an unsignalized intersection with one opposing lane, none of the models was found satisfactory.

Since most of the existing methods failed to match all data satisfactorily, they were subsequently modified to better reflect real-world conditions. The adjusted forms were found to result in substantial improvements over the original models, but even with this improvement estimation of left-turn saturation flows for some of the cases studied was still unsatisfactory.

For this reason, a new model was developed for each type of intersection studied. From the extensive statistical analysis of the factors affecting left-turn saturation flows, it was concluded that the dominant independent variables are the opposing flow, the critical gap, signalization, and the number of opposing lanes. Other variables used in theoretical derivations performed by earlier researchers, such as the minimum opposing and left-turn headways, did not prove to be statistically significant.

Naturally, since the statistical models fit the observed data more closely, one would be inclined to recommend their use over the others. However, in using these models, caution must be exercised, since it can be argued that the collected data represent only a limited number of intersections. It is primarily for this reason that the existing models were not ignored in Table 2.

Use of this table is recommended as a guideline in selecting the appropriate model for a particular situation. The close agreement of some of the existing models with the statistical ones suggests that there should be a reasonable degree of confidence to the collected data. It should be mentioned, however, that further validation of the proposed models is desirable. Validation of the composite model should be of particular interest due to its ability to represent all the cases combined.

Finally, it must be pointed out that the case of multilane opposing traffic can be taken into account in a manner similar to Tanner's (3), i.e., by considering the opposing vehicles as a single stream with arrival rate three times that of a single lane (for the case of three opposing lanes for example) and increased critical gap.

From the gap-acceptance study, it is concluded that cumulative accepted gaps are uniformly distributed over the range of permissible sizes. Inspection of the gap-acceptance functions (Table 4) leads to the conclusion that there appears to be insignificant difference in gap-acceptance characteristics when the opposing traffic moves in one or two lanes at unsignalized intersections. Finally, critical gaps were found to be shorter at signalized intersections, as expected.

Figure 3. Gap-acceptance distribution at signalized intersection with one opposing lane.

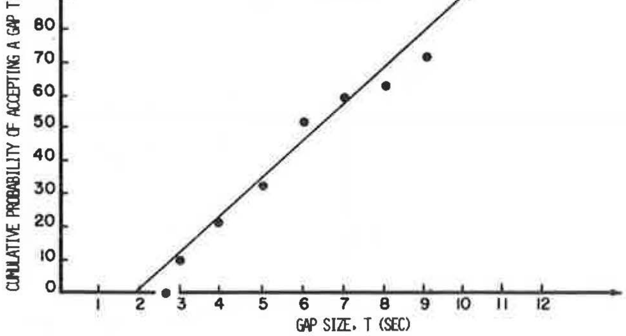


Table 4. Gap-acceptance distributions.

Case	Equation	Conditions	R^2	S_E
1	0 P(T) = (T - 2.33)/(12.37 - 2.33) 1	if T < 2.33 if 2.33 ≤ T ≤ 12.37 if T > 12.37	0.98	0.040
2	0 P(T) = (T - 1.91)/(10.91 - 1.91) 1	if T < 1.91 if 1.91 ≤ T ≤ 10.91 if T > 10.91	0.94	0.080
3	0 P(T) = (T - 2.70)/(10.80 - 2.70) 1	if T < 2.70 if 2.70 ≤ T ≤ 10.80 if T > 10.80	0.96	0.072
4	0 P(T) = (T - 2.73)/(10.80 - 2.73) 1	if T < 2.73 if 2.73 ≤ T ≤ 10.80 if T > 10.80	0.97	0.070

REFERENCES

1. A. D. May. Intersection Capacity 1974: An Overview. TRB, Special Rept. 153, 1975, pp. 50-59.
2. F. V. Webster and B. M. Cobbe. Traffic Signals. BRLL, Road Research Technical Paper 56, London, 1966, pp. 38-45.
3. J. C. Tanner. A Theoretical Analysis of Delays at an Uncontrolled Intersection. Biometrika, Vol. 49, Nos. 1, 2, 1962, pp. 163-170.
4. Notes from the Intersection Capacity Workshop. TRB, Transportation Research Circular 161, Dec. 1974, pp. 1-28.
5. D. R. Drew. Traffic Flow Theory and Control. McGraw-Hill, New York, 1968.
6. D. B. Fambro, C. J. Messer, and D. A. Andersen. Estimation of Unprotected Left-Turn Capacity at Signalized Intersections. TRB, Transportation Research Record 644, 1978, pp. 113-119.
7. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
8. A. J. Miller. On the Australian Road Capacity Guide. HRB, Highway Research Record 289, 1969, pp. 1-12.
9. S. M. Novoa and J. S. O'Connor. Estimation of Left-Turning Saturation Flow at Intersections. Rensselaer Polytechnic Institute, master's project, May 1977.

Publication of this paper sponsored by Committee on Highway Capacity and Quality of Service.

**At the time this research was done Mr. Michalopoulos was with Rensselaer Polytechnic Institute and Mr. Novoa was with Catholic University of Valparaiso.*

Signal Cycle Length and Fuel Consumption and Emissions

Stephen L. Cohen and Gary Euler, Office of Research, Federal Highway Administration

A microscopic network simulation model (NETSIM, formerly UTCS-1) was used to evaluate the relationship between fuel consumption and signal cycle length. A single intersection was simulated for three scenarios having different traffic characteristics. It was found that the cycle length that minimizes delay also minimizes fuel consumption and hydrocarbon and carbon monoxide emissions. A regression analysis showed that fuel consumption and these emissions are strongly correlated with vehicle average speed but that the relationship is not linear. Differences between the results in this work and previous results are discussed.

Since the passage of the Clean Air Act of 1970 and the oil embargo crisis of 1973, the issues of automobile fuel consumption and emissions have greatly increased in importance. Thus, it has been proposed that more emphasis be placed on the measures of effectiveness (MOEs) for fuel consumption and emissions and on such traditional measures as speed, stops, and delay. Thus, various types of policies affecting traffic flow would be evaluated as to their effect on the fuel-emission MOEs, speed, and so forth.

In recent years a number of authors (1, 2, 3, 4, 5, 6, 7) have addressed themselves to the issue of fuel consumption in urban traffic. Bauer (1) and Courage and Parapar (2) investigated the relationship between signal cycle length and fuel consumption. Lieberman and Cohen (3) and Honeywell (4) addressed the issue of finding the effects of different traffic control strategies on fuel efficiency (measured in distance traveled versus fuel consumed). Evans, Herman, and Laur (5) addressed the problem of relating fuel consumption to other traffic MOEs such as average speed, while Pattersen (6) and Cohen (7) examined the problem of estimating the concentration profile of traffic-generated carbon monoxide at signalized intersections.

Of particular interest are the findings of Bauer (1) and Courage and Parapar (2). The analysis performed by these authors showed that at an isolated intersection the cycle length at which fuel consumption is minimized is very much longer than the cycle length at which delay is minimized.

In the present study, we shall describe an analysis of this finding that was conducted using the network flow simulation [NETSIM, (8), formerly the UTCS-1] model. Our result differed from others (1, 2) in that fuel consumption and the hydrocarbon (HC) and carbon monoxide (CO) emissions were found to be minimized at approximately the same cycle length as delay. Another finding of interest was that MOE stops did not always follow Webster's expression (9, 10), which predicts that number of stops decreases as the cycle length increases.

A regression analysis was performed to examine relationships between the average speed and MOE fuel consumption and emissions. It was found that there is a strong correlation between these measures but that the relationships are not linear.

PROBLEM DESCRIPTION AND TECHNICAL APPROACH

In order to isolate the relationship between signal cycle length and average speed, stops, fuel consumption, and emissions, we confined ourselves to the analysis of single isolated intersections.

The initial configuration involved the analysis of a two-phase pretimed signal. In future work, we plan to analyze more complicated situations, in particular multiphase signals and varying geometric configurations.

Our approach to the problem was to use NETSIM as modified to compute fuel consumption and emissions (4). This approach is particularly appropriate for analyzing fuel versus emissions impacts, because it is difficult to measure the former directly in the field and impossible to measure the latter.

NETSIM is a microscopic network simulation model. Thus, individual vehicle movements are simulated according to car-following, queue discharge, and lane-changing laws. Vehicles are generated on entry links according to a shifted exponential headway law. However, since there is a spread of free-flow speeds and car-following interactions on the network links, the arrival patterns at the rear of the queue will in general be complicated.

The model has been validated in a network in Washington, D.C., and single intersections in Arlington, Virginia; Berkeley, California; and New Jersey. The model has been used on several projects for a variety of research and operations applications.

DESCRIPTION OF INTERSECTION SCENARIOS

The first intersection examined is shown in Figure 1. Each east-west approach of the intersection was assumed to have two through lanes, with the east-to-north left-turn movement served by a left-turn bay. Each north-south approach of the intersection had one through lane. This configuration provides a considerable amount of geometric variability. All approach lengths were assumed to be 305 m. The assumed free-flow speed on the east-west approach was 64 km/h and 56 km/h on the north-south approach.

Scenarios were generated by varying volume and left- and right-turn percentages on each approach of the intersection. Opposing movements on the same street (for instance, the west-to-east and east-to-west movements) always had equal volumes and turn percentages. Since the opposing volumes only interfere with left-turn movements, there is little loss of generality in imposing this restriction. On the east-west approaches, volumes ranged between 600 vehicles per hour (vph) and 2400 vph in increments of 200 vph. On the north-south approaches, volumes ranged between 300 and 1200 vph in increments of 100 vph. The left- and right-turn movements were either 0, 10, or 20 percent of the total approach volumes and varied by approach. By varying the volume and turning percentage for each approach, a total of 8100 scenarios was generated.

For each of the generated scenarios, the degree of intersection saturation was first calculated. This is the sum of all phases of the critical approach volume-to-capacity (V/C) ratios for each phase.

$$S_i = V_{i_{crit}}/C_{i_{crit}} \tag{1}$$

where

S_i = degree of saturation of the ith critical approach to the intersection,
$V_{i_{crit}}$ = critical approach volume for the ith phase, and
$C_{i_{crit}}$ = critical approach capacity for the ith phase.

The V/C ratios for both direction movements for both approaches were calculated from the following equation.

$$V/C = V/(NC \times \text{Lanes}) \tag{2}$$

where

NC = nominal capacity = (3600 s/h)/(2.4 s/vehicle) and
Lanes = number of lanes serving through traffic on the approach.

In the case of an exclusive left-turn lane, this equation was modified to

$$V/C = [V(1.0 - LT)]/(NC \times \text{Lanes}) \tag{3}$$

where LT is the fraction of left turns.

The minimum-delay cycle length for each scenario was then estimated by using Webster's formula as

$$C = [1.5(L) + 5]/(1 - S) \tag{4}$$

where

C = cycle length,
$S = \sum_i S_i$, and
L = total lost time, assumed to be 4 s, on the critical approaches.

Oversaturated intersections (S > 1) were eliminated from consideration, since these would have no minimum-delay cycle lengths. All scenarios were executed by varying the cycle lengths in 20-s intervals between 40 and 150 s. These limits were chosen to correspond to cycle lengths used in practice on two-phase signals. From these runs, the cycle length giving the lowest delay for each scenario was determined. Two further runs for each scenario were made at 10-s cycle length intervals around this value to further refine the results. The green split for each of these scenarios was calculated from Webster's demand relation as

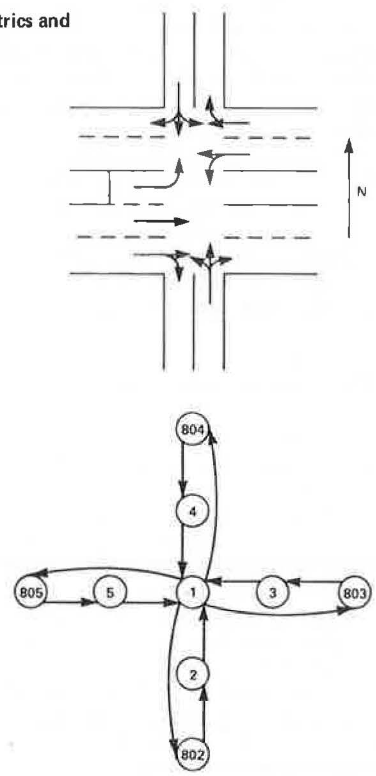
$$g_i = S_i/S \tag{5}$$

where g_i is the proportion of green time of the ith phase.

The green splits were held constant for all cycle lengths listed for a given scenario.

The research plan was to execute several scenarios that would be chosen to provide a wide range of volume and turning movement conditions. Results of runs of the

Figure 1. Intersection geometrics and associated link-node diagram.



first three of these are described in Table 1 and reported in this work.

For each scenario and cycle-length pair, ten replications of a half-hour period were simulated. The effect of varying the cycle length at the intersection can then be examined for a series of measures: average delay in seconds per vehicle, average speed in kilometers per hour, number of stops per vehicle, fuel efficiency in kilometers per liter for the intersection, and emissions of hydrocarbons, carbon monoxide, and oxides of nitrogen in grams per kilometer for the intersection.

The measures described above are computed in NETSIM as follows.

1. Average speed is calculated by dividing the total number of vehicle kilometers traveled on each link by the total number of vehicle hours spent on the link.
2. Average delay is calculated by subtracting the number of vehicle seconds that unimpeded vehicles would spend on a link from the actual number of vehicle seconds spent on a link and then dividing by the total number of vehicles discharged from the link.
3. Stops are the number of simulated vehicles that are forced to stop by traffic conditions.
4. Fuel emissions are assessed, by a table of fuel-emissions rates, for each second by each vehicle using the vehicle's speed-acceleration couplet (8).

RESULTS

Cycle Lengths

The results of the three scenarios are shown in Figures 2-7. Figure 2 is a plot of average delay versus cycle length; Figure 3 shows fuel consumption versus cycle length; Figure 4 shows stops versus cycle length; Figure 5 shows HC versus cycle length; Figure 6 shows CO versus cycle length; and Figure 7 shows NOX versus cycle length. The plotted points are averages over the ten replications run for each scenario.

It can be seen that, in all three cases, the cycle length at which minimum delay occurs is the cycle length where minimum fuel consumption and HC and CO emissions occur. This is approximately 60 s for scenario 463, 80 s for scenario 1462, and 100 s for scenario 3836. This result differs from those obtained elsewhere (1, 2). A discussion of the reasons for this follows shortly.

We also observe that stops are minimized at 60 and 80 s for scenarios 463 and 1462, respectively, while the stops curve for scenario 3836 decreased as a function of cycle length.

These findings on stops may be demand dependent, because the V/C ratios on the critical approaches are higher in scenario 3836.

Regression Analyses

The object of this exercise was to examine the correlation between average speed and the MOE's fuel consumption and emissions.

Average speed was chosen because it was a tested independent variable (4, 5). First and second order regressions were run. Since a total of 26 cycle lengths over the three scenarios was tested by executing ten replications for each cycle length and there were four intersection approaches, there was a total of 1040 data points. These points are plotted in Figures 8-11.

The results for fuel consumption and HC and CO emissions showed a very high correlation between these measures and average speed. In all three cases, the second order terms were found to be significant. In addition, HC and CO were regressed against 1.00/average speed. The regression lines were

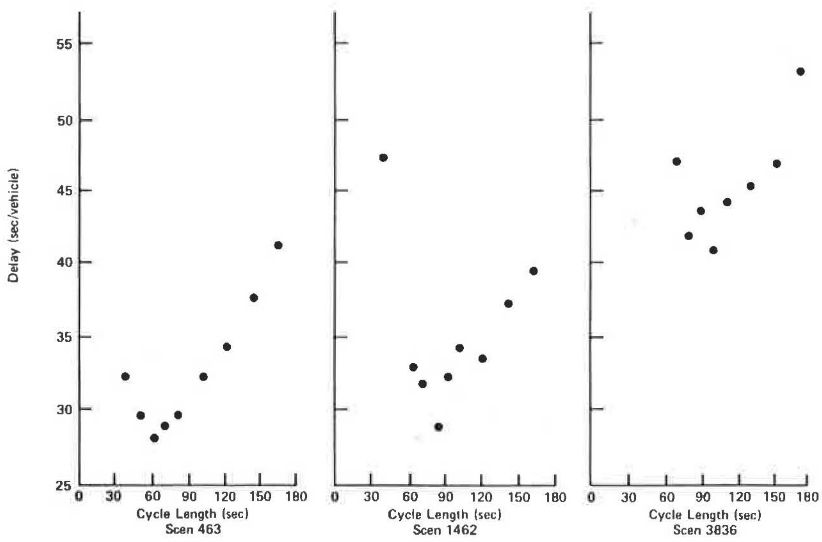
$$FC = 0.695 + 0.471*(\text{average speed}) - 0.0154*(\text{average speed})^2 \qquad (6)$$

Table 1. Scenario parameter descriptions.

Scenario No.	Phase I (east-west) Vol.	Percentage Left Turns	Percentage Right Turns	Phase II (north-south) Vol.	Percentage Left Turns	Percentage Right Turns	GS(G_1/G_2)	Minimum Delay Cycle[a]
463	1600	0	0	500	0	10	1.525	60
1462	1800	0	10	400	0	20	2.143	80
3836	1000	10	10	800	10	20	0.608	100

[a] Determined from the NETSIM simulation model.

Figure 2. Plot of delay versus cycle length for three scenarios.



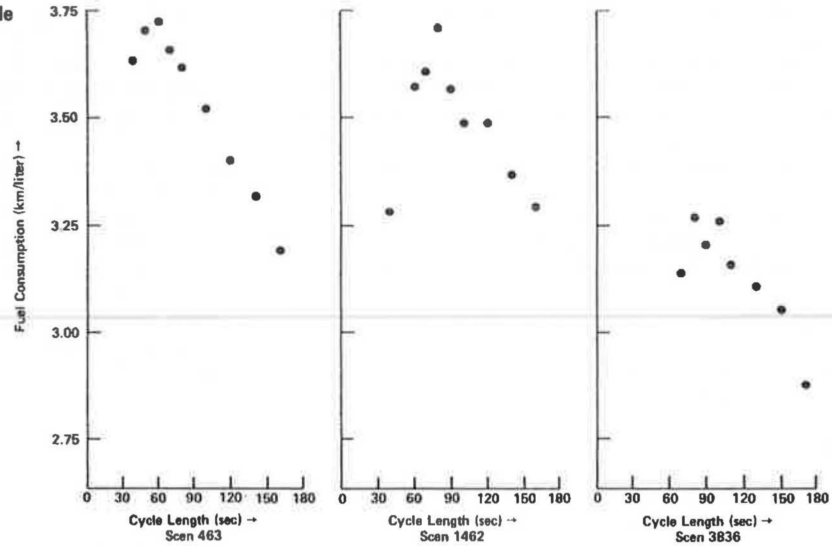


Figure 3. Plot of fuel consumption versus cycle length for three scenarios.

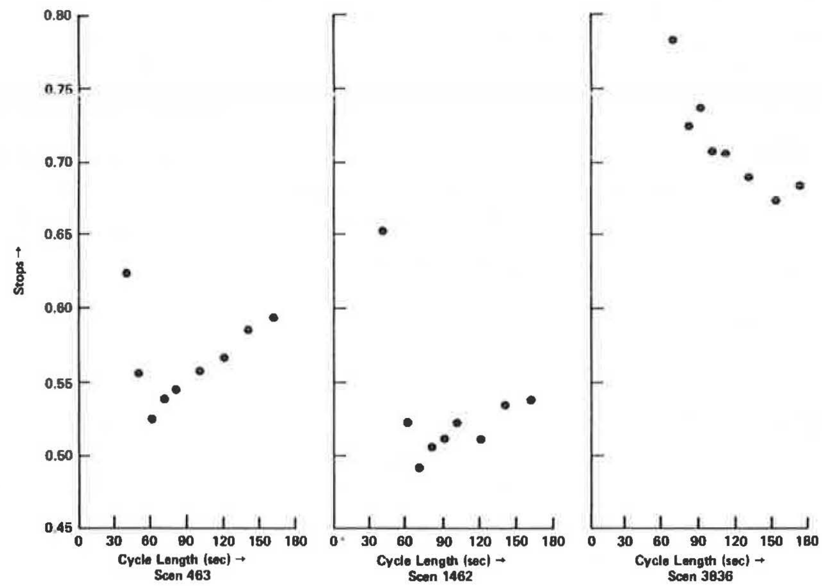


Figure 4. Plot of stops versus cycle length for three scenarios.

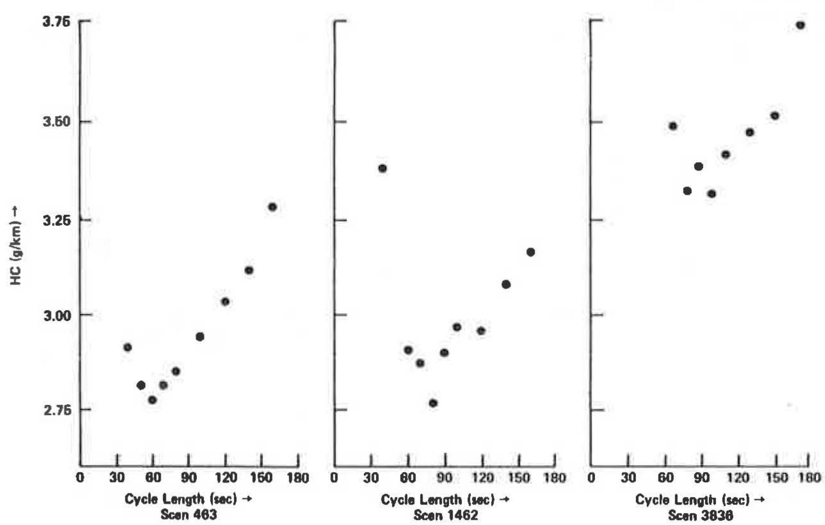


Figure 5. Plot of HC emissions versus cycle length for three scenarios.

Figure 6. Plot of CO emissions versus cycle length for three scenarios.

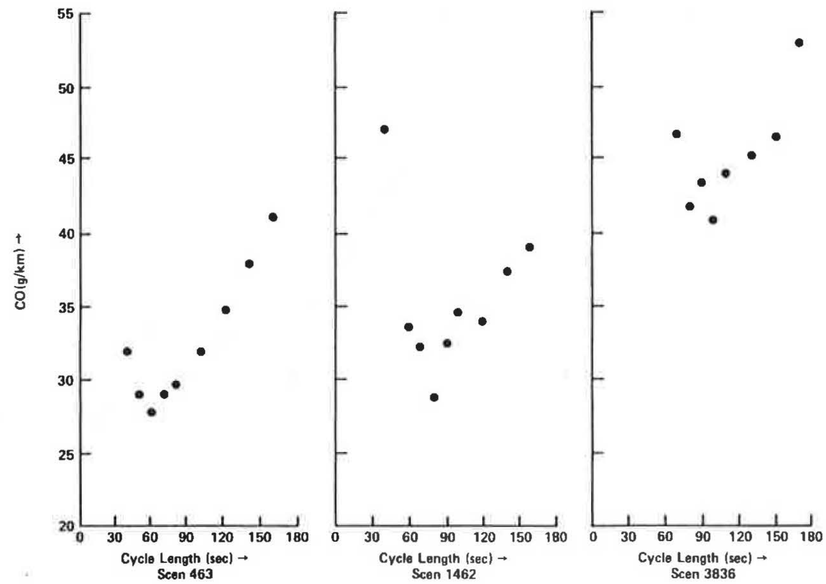


Figure 7. Plot of NOX emissions versus cycle length for three scenarios.

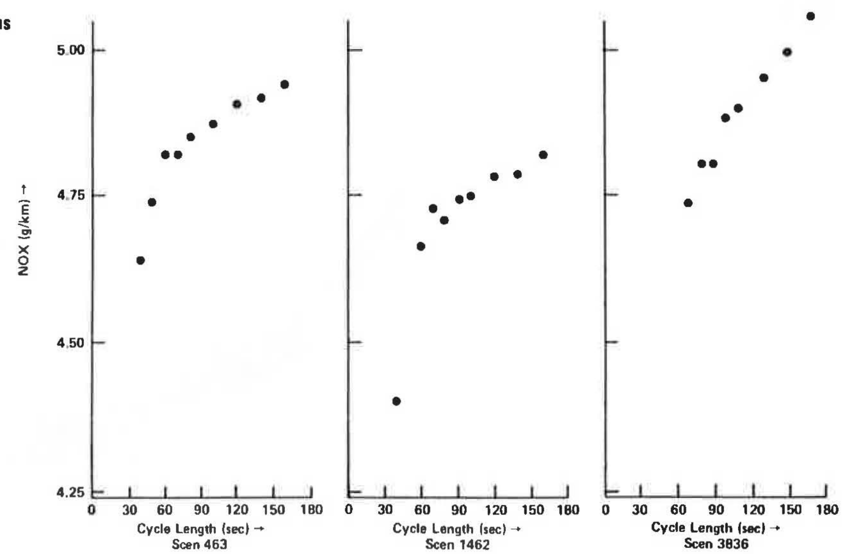


Figure 8. Plot of fuel consumption versus average speed.

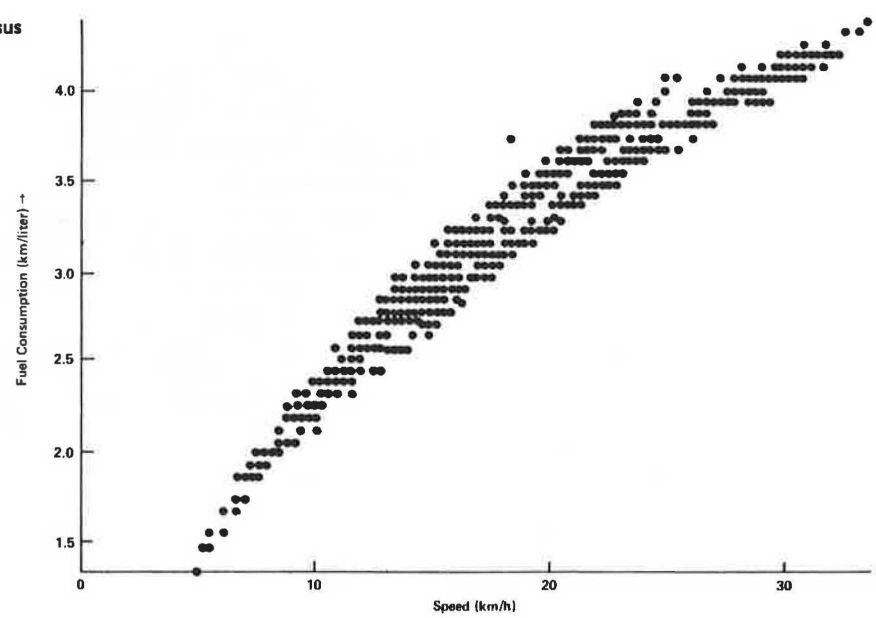


Figure 9. Plot of HC emissions versus average speed.

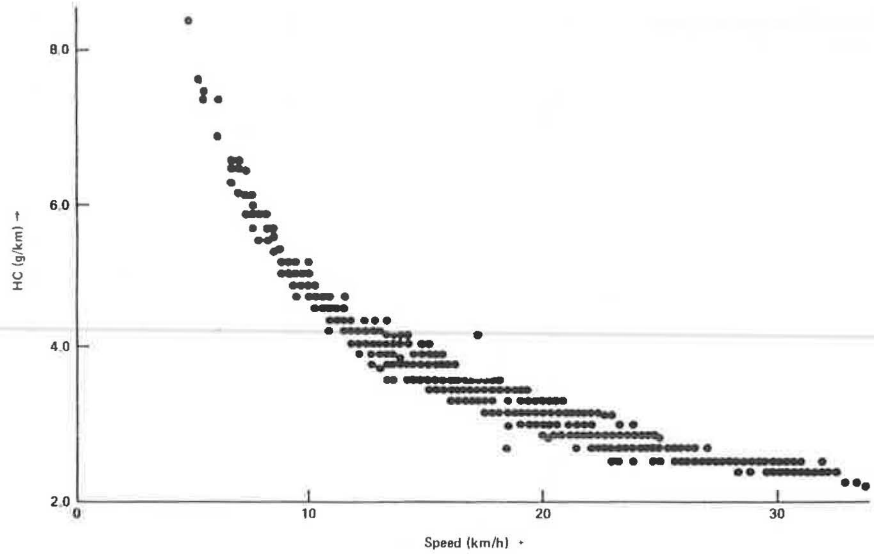


Figure 10. Plot of CO emissions versus average speed.

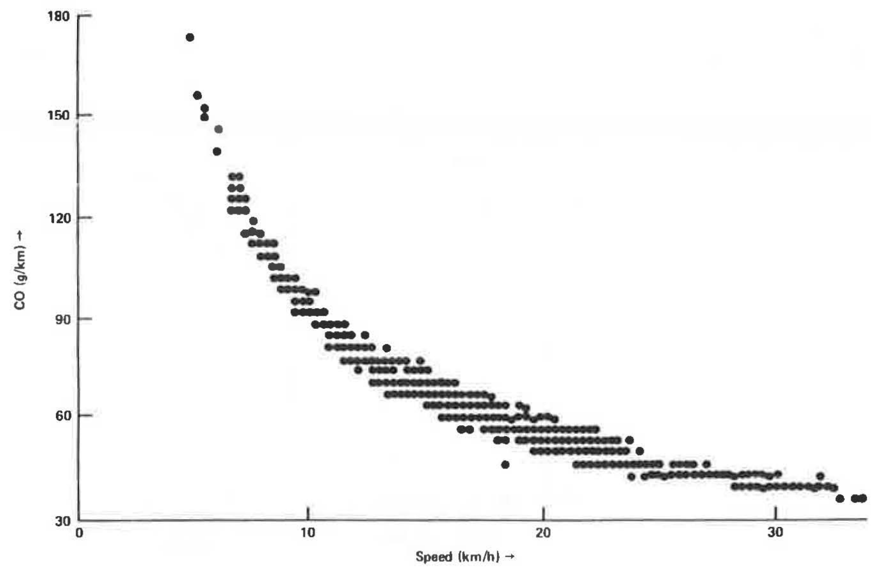
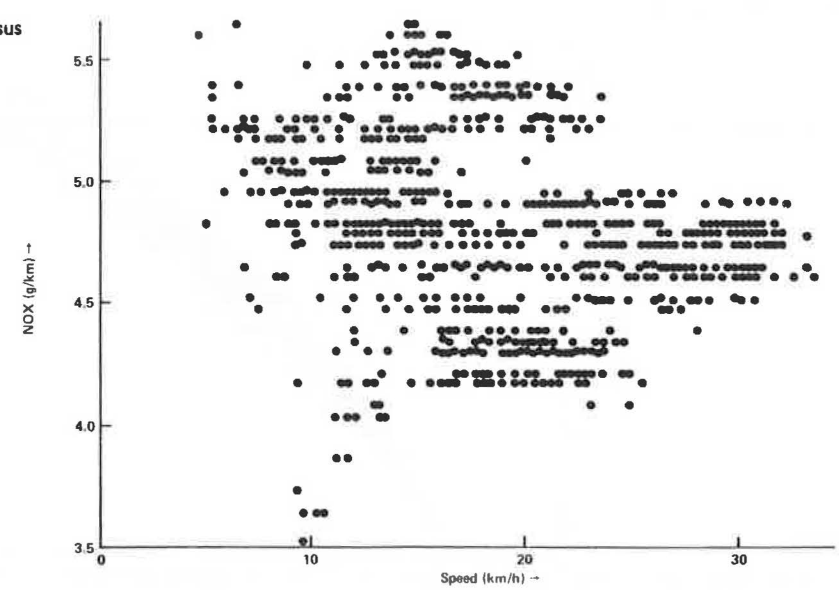


Figure 11. Plot of NOX emissions versus average speed.



for fuel consumption in kilometers per liter versus average speed in kilometers per hour ($R^2 = 0.975$);

$$HC = 8.342 - 1.065*(\text{average speed}) + 0.0483*(\text{average speed})^2 \quad (7)$$

for HC emissions in grams per kilometer ($R^2 = 0.9266$) or, alternatively,

$$HC = 1.36 + 21.46*(\text{average speed}) \quad (8)$$

($R^2 = 0.964$);

$$CO = 171.71 - 23.87*(\text{average speed}) + 0.096*(\text{average speed})^2 \quad (9)$$

for CO emissions in grams per kilometer ($R^2 = 0.934$) or, alternatively,

$$CO = (16.03 + 476.1)/(\text{average speed}) \quad (10)$$

($R^2 = 0.970$). No correlation was found between NOX and average speed.

CONCLUSIONS

The results to date of this study, which are admittedly somewhat limited, indicate that the cycle length, which minimizes delay at an isolated intersection, also minimizes fuel consumption and the emittants HC and CO.

Fuel consumption and emissions are strongly correlated with average speed, but the regression relation is not linear.

Stops appear to play no role in the analysis, probably because they are correlated closely with delay (11).

Great care must be exercised in applying regression relations such as those found in this work. The expressions we used should not be used in situations in which average speeds greater than 30 km/h occur. This was the highest speed occurring in this study, and it is unwise to extrapolate using such regressions. The fuel consumption and emission figures derived here are based on a particular vehicle mix (3). It is rather unlikely that different vehicle mixes (possibly involving trucks as well as cars) would lead to the same regression parameters. It should be noted that the regression relationships derived here are based on link-wide results aggregated over all vehicles that traversed the link. The result obtained by Evans, Herman, and Laur (5) was based on a set of floating car runs. These relationships were determined from an analysis of a single isolated intersection with a two-phase signal. The results of the regressions do not necessarily apply to other situations.

The inverse relationship between average speed and HC and CO emissions is probably better than the quadratic, as there is evidence that these emittants level out at higher speeds, and the correlation coefficients are better.

Future research in this area will be aimed at running several more scenarios on this geometric configuration and examining the effects of multiphasing and other geometric configurations and generating data at higher average speeds in order to improve the range of validity of the regression relations.

DISCUSSION

There are several reasons that probably account for the discrepancy between this work and others (1, 2). The fuel consumption increment due to cars that slow down but do not stop is not considered. Since NETSIM is microscopic, this effect is automatically included. It is assumed (1, 2) that all stop cycles are the same. This is generally not true, as the acceleration pattern from the locked wheels position is a function of queue position. Thus, the first vehicle in a queue accelerates directly up to cruising speed, while cars farther back spend considerable time traveling at speeds lower than the cruising speed while moving up to the stop line. This type of movement is usually more costly in fuel than traveling at the cruise speed. Again, the microscopic queue discharge behavior of NETSIM automatically includes this effect. The effect of multiple stops due to left turns is ignored (1, 2). Again, the microscopic logic of NETSIM automatically accounts for this effect.

The major reason why this work is at variance with Webster (9) on the issue of stops versus cycle length is probably due to the different assumptions, such as constant arrivals and departures, in his model. NETSIM has random arrivals and departures.

FIELD DATA VERIFICATION

At the present time, there are no field data that directly support either the conclusions of this work or others (1, 2). There are, however, two field studies that indirectly support our conclusions.

Our finding of a strong correlation between fuel consumption and average speed is in agreement with the finding of Evans, Herman, and Laur (5). This provides a verification of the simulation model results. The other study of relevance was one of traffic delay at signalized intersections performed for the Federal Highway Administration by JHK and Associates (11). Intersection delay and number of vehicles stopping were measured at several signalized intersections for a wide variety of volume and turning movement and geometric situations. It was found that the measures of stops and delay were strongly correlated with each other. This result supports the conclusion that stops do not enter into fuel regression relationships independently of delay.

REFERENCES

1. C. S. Bauer. Some Energy Considerations in Traffic Signal Timing. Traffic Engineering, Vol. 45, Feb. 1975, pp. 19-25.
2. K. G. Courage and S. M. Parapar. Delay and Fuel Consumption at Traffic Signals. Traffic Engineering, Vol. 45, Nov. 1975, pp. 23-27.
3. E. B. Lieberman and S. Cohen. New Technique for the Evaluation of Urban Traffic Energy Consumption and Emissions. TRB, Transportation Research Record 599, 1976, pp. 35-45.
4. Honeywell Traffic Management Center. Fuel Consumption Study. NTIS, Springfield, VA, PB 259 003, Feb. 1976.
5. L. Evans, R. Herman, and T. N. Laur. Gasoline Consumption in Urban Traffic. Paper presented at the Automotive Engineering Congress and Exposition, Feb. 23-27, 1976, SAE Paper 760048.
6. R. M. Pattersen. Traffic Flow and Air Quality. Traffic Engineering, Vol. 45, Nov. 1975, pp. 14-17.
7. S. L. Cohen. Use of Traffic Simulation in Analyses of Carbon Monoxide Pollution. TRB, Transportation Research Record 648, 1977, pp. 74-76.
8. R. D. Worrall and E. B. Lieberman. Network Flow Simulation for Urban Traffic Control Systems. NTIS, Springfield, VA, Phase II, Vols. 1-5, PB 230 760, 230 761, 230 762, 230 763, 230 764, March 1973.
9. F. V. Webster. Traffic Signal Settings. Her Majesty's Stationery Office, London, Road Research Technical Paper No. 39, Oct. 1957.

10. F. V. Webster and R. B. M. Cobbe. Traffic Signals, Her Majesty's Stationery Office, London, Road Research Technical Paper No. 56, 1966.
11. W. R. Reilly, C. C. Gardner, and D. H. Kell. A Technique for Measurement of Delay at Intersections. Vol. 1, Technical Rept., NTIS Springfield, VA.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Traffic Conflicts as a Diagnostic Tool in Highway Safety

Charles V. Zegeer and Robert C. Deen, Bureau of Highways, Kentucky Department of Transportation

Accident repeatability from one year to the next was found to be high at 60 intersections (r = 0.64) and 170 spot locations (r = 0.59). Nearly half of the 209 Kentucky locations designated as hazardous by accident criteria were found to have been so identified falsely because of random accident occurrences. Conflict counts were conducted at 5 intersections in central Kentucky to determine characteristics of conflict data. Good reliability was found between observers in simultaneous counts of conflicts and weaves with r values as high as 0.93. Traffic volumes accounted for only about 30 percent of the variation in numbers of conflicts. Reductions in conflicts and accidents that resulted from such safety improvements as installing left-turn signal phasing, raised pavement markers, and green-extension systems at numerous locations were determined. A revised procedure for collecting and utilizing conflict data was described.

Traffic conflicts are measures of accident potential and operational problems at a highway location. Many highway agencies are now using traffic conflict techniques to complement the limited accident data found in accident records. The Kentucky Department of Transportation has used various forms of conflict data since 1972 to assist in its efforts for highway improvement. While new procedures are currently under development for collection and use of conflict data in Kentucky, past experiences with conflicts have proved very encouraging.

The first formalized procedure for identifying and recording traffic conflicts at intersections was developed by Perkins and Harris of General Motors Corporation in 1967 (1). Major types of conflicts at intersections include rear-end, left-turn, cross-traffic, red-light violation, and weave conflicts. Conflict counts may be used to quickly evaluate changes in road design, signing, signalization, and environment. After a location is identified as hazardous, a study of conflict patterns can be used with accident diagrams to gain a more accurate understanding of operational deficiencies and accident causes.

Crude forms of traffic conflict counts to determine appropriate safety improvements have been made since traffic engineers first began making field observations. Formalized traffic conflict techniques give a more objective measure of observed traffic problems and allow for a permanent record of the comparative magnitude of such problems. The use of traffic conflict techniques has to date been primarily limited to intersections. However, conflict procedures for other types of locations are under development.

A more severe form of traffic conflict is an erratic maneuver, which is any sudden, unexpected movement by a vehicle that could cause an accident. An erratic maneuver usually involves only one vehicle's making an unsafe move independently of other vehicles. Such a maneuver may often result in a conflict if another vehicle is forced to brake or weave to avoid it. Poor signing and inadequate geometric design often cause erratic maneuvers.

While traffic conflict counts usually indicate the potential for accidents between two or more vehicles, erratic maneuver counts may also provide information about the potential for single-vehicle accidents.

A near-miss accident is a collision between two or more vehicles barely avoided by a last-second movement or stop. This type of accident is a very severe sort of conflict and is rarely observed at any location compared to other conflicts or erratic maneuvers.

Traffic events may be classified in terms of increasing severity from traffic volume to fatal accidents. The ordering of traffic events by severity is as follows:

1. Traffic volume,
2. Routine conflicts,
3. Moderate conflicts and erratic maneuvers,
4. Severe conflicts or near-miss accidents,
5. Minor collisions (usually not reported),
6. Property damage accidents,
7. Injury accidents, and
8. Fatal accidents.

While accident data provide only the last three levels of traffic events, traffic conflict counts provide the other five, since volume counts are usually made along with conflict counts.

NEED FOR CONFLICT DATA

Several limitations have been observed in the use of accident data alone in traffic safety studies. Accident files only contain records of reported accidents, which comprise only a fraction of the accidents that actually occur. The criteria for accident reporting vary considerably among states. For example, all traffic accidents in Colorado, Nevada, and the District of Columbia by law must be reported; only accidents with injury costs exceeding $400 damage to any one person must be reported in Connecticut. Reporting criteria in other states range between these extremes; the most common reporting criteria are $100 (23 states) and $200 (12 states including Kentucky) (2).

Because of such reporting criteria, estimates of traffic accidents actually reported range from 20 to 50 percent. The number of reported accidents at a site is, therefore, a function of local reporting laws, accident severity, and damage costs of each accident.

Another problem with using accident data alone for identifying and evaluating high-accident sites is the random fluctuations in accident data. Many accidents result from a vehicle malfunction (blowout or brake failure), an obvious driver error (speeding or drunk driving), or a weather-related problem (ice on road or heavy fog) that is unrelated to any geometric deficiency.

A study was completed in 1973 in Kentucky that illustrated the effects of random accidents on the identification of hazardous sites. Of the 208 spot locations identified by accident data as hazardous, 99 of them were wrongly identified because of random accident occurrences. These 99 sites were found by field inspections to need no improvements, and accidents decreased to normally low levels the following year. Nearly half the accident locations warranted no improvements (3).

To test the reliability of accident data for predicting future accidents at a location, an analysis of 60 intersections in central Kentucky was made. The number of accidents for a given year compared with the number of accidents the following year resulted in a correlation coefficient (r-value) of only 0.64. The 95 percent confidence level (twice the standard error) for this relationship was ±10.9 accidents per year, and the average number of accidents per year at the intersections was 11.1. This indicated that an error of almost 100 percent in either direction is possible when accident numbers from one year to the next are compared.

A similar analysis was also made for 170 rural, 480-m (0.3-mile) spots in Kentucky, and an r-value of only 0.59 was found. More than a 100 percent error was also found for this sample of locations (within the 95 percent confidence level), which illustrates the nonrepeatability of accident data.

Another problem with accident data is the waiting time needed to obtain a significant data base. A previous study in Kentucky suggested that up to 2 years of accident data are necessary to ensure reliability when selecting high-accident locations (4). After an improvement is made, it often takes several more years to determine the effectiveness of the improvement based on accident data. Also, without some other measure of safety, several accidents must occur at a site before improvements can be justified.

While accident data have many limitations, they can be quite useful when complemented by traffic conflict data. Accident histories can point out locations where conflict data should be collected. Conflict studies can then be made at these and other sites suspected of being hazardous. Conflict counts can be used to help select appropriate improvements and later to determine whether the improvements were effective in reducing the hazard to motorists.

CHARACTERISTICS OF CONFLICT DATA

An effort was made to gain a better understanding of the nature of traffic conflicts. The immediate intent was to determine consistency of conflict counts between observers, to evaluate volume and conflict relationships, and to test daily repeatability of conflicts.

Conflict and volume data were continuously collected by the General Motors (GM) procedure at each of five sites for 11 hours from 7:30 a.m. to 6:30 p.m. on Tuesday, Wednesday, or Thursday. Two days of data were collected at one site to test for conflict repeatability. Five observers alternated duties at each site to allow for breaks when needed. Some conflict counts were made simultaneously on the same approach to test observer consistency.

Conflicts were counted on the two major approaches at four intersections and one approach at the other. One observer was stationed at each approach from 30 to 90 m (100-300 ft) back from the intersection in a state-owned car wherever possible. Chairs on sidewalks were used at urban locations that had no shoulders. Volume counts were made of every movement (through, left turns, and right turns) of all intersection approaches throughout the test period.

Conflict and volume data were recorded in 15-min periods on the GM data sheets. Several new categories of conflicts and erratic maneuvers were added from observations of the specific problems at a site. Each conflict was also classified as routine, moderate, or severe.

All five intersections were located in and around Lexington, Kentucky (population 200 000), and data were collected in the spring of 1977. A summary of volume, speed, geometric, and conflict information for each intersection approach is given in Table 1. All approaches were two lanes of four-lane arterials; minor streets were all two-lane collector streets. Each was a four-way signalized intersection, except Harrodsburg Road at Larkspur Drive, which is a T-intersection with a stop sign on the minor approach.

Observer Reliability

One of the most important aspects to consider when using conflict data is the reliability of data collected by observers. There are many factors that will account for variations in conflict counts, such as alertness, experience, and different driving attitudes of the observers; location of the observer at the site; and traffic volumes. Several hours of training are routinely given to each observer before conflict data are taken alone. Typically, an experienced observer trains an inexperienced one at a site by discussing all conflicts and weaves as they occur. Periodic checks between observers are made to help ensure consistency.

The first test was conducted in June 1977 at the signalized intersection of Limestone Street and Virginia Avenue. During data collection, four observers were used, two simultaneously counting conflicts and weaves in 15-min intervals using the GM technique. A plot was made of conflicts per 15-min period for one observer versus those of another, and the overall r-value was 0.86. Numbers of conflicts per 15-min period ranged from 5 to 36, depending primarily on traffic volume. A similar plot of weaves resulted in an r-value of 0.93, and numbers of weaves varied from 0 to 24 every 15 min. A total of 25 periods were used in this analysis.

The second site was a T-intersection of Harrodsburg Road at Larkspur Drive. Again, four observers counted conflicts and weaves on the two major approaches (in July 1977). A correlation coefficient of 0.87 was found between conflict counts by observers as shown in Figure 1 for 26 periods of 15 min each. The correlation for weaves was lower than before, at 0.77. The overall reliability of observers involved in conflict counts was considered to be very good. Reevaluation of observers is made periodically, so observer reliability is expected to improve.

Volume and Conflict Relationships

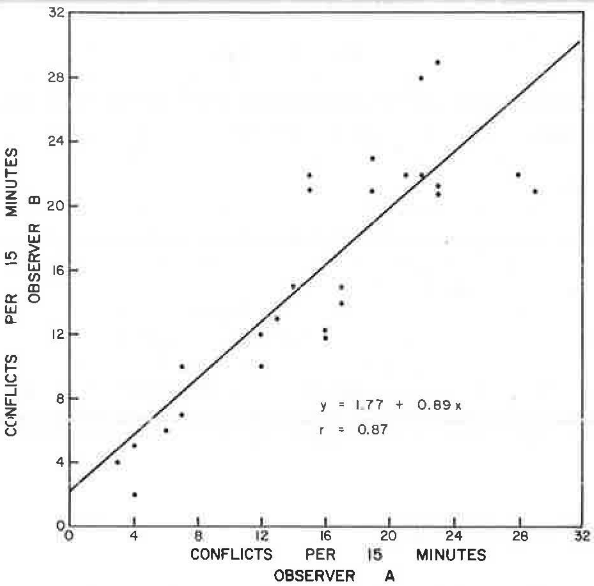
The relationship between traffic volume and conflicts was found on all intersection approaches for each day of data collection. Plots of total volume (x-axis) versus total conflicts (y-axis) were made by considering each 15-min period as one data point. A total of 44 points were plotted for each intersection approach (11 h of data with four periods per hour). The correlation coefficients varied widely from 0.24 to 0.81. Individual values of r for the approaches as ordered in Table 1 were 0.72, 0.70, 0.81, 0.35, 0.73, 0.45, 0.24, 0.51, and 0.72. Based on average r^2 of all approaches, only 37 percent of the variance in conflicts can be explained by traffic volumes.

Volume and conflict relationships were also compared on two separate days at one intersection. On the inbound approach, the r-value was 0.28 the first day and 0.35 the second day (2 weeks later). The difference was greater on the other approach where the r-values were 0.42 and 0.73 for the 2 d.

Another plot was made of conflicts per hour versus hourly volume for all approaches (11 data points), which ranged from 32 to 83; hourly volumes were between 294 and 931. The r-value was only 0.51, which indicates that only 26 percent of the conflict variation can be explained by traffic volume (r^2 = 0.26).

The previous results indicate that, while traffic volumes have some effect on number of conflicts, volume and conflict correlations vary considerably at different intersections. Also, the correlations may vary on different days at the same approach. Thus, counting conflicts is not merely another way of counting traffic volume. Most conflicts at the test sites were traced to a geometric deficiency, an inappropriate signal timing, or a capacity problem.

Figure 1. Conflict counts per 15-min period for two observers.



Conflict Repeatability

One of the questions raised concerning use of conflict data concerns the variation in conflicts from one day to the next. A large variation in conflict numbers and patterns would require several days of collection at each site to ensure reliable data. To obtain information concerning the daily repeatability of conflicts, conflict data

Table 1. Characteristics and conflict summaries of test sites.

Intersection	Direction of Approach	Date of Count	Volume: Average Hourly	Volume: Peak Hour	Other Characteristics	Conflicts: Per Hour	Conflicts: Peak Hour	Conflicts: Per 1000 Vehicles
New Circle Road at Woodhill Drive	S	5/10/77	908	1375	Three-phase signal; Left- and right-turn lanes; Separate left-turn signal phase; Left turns (9 percent); Right turns (10 percent)	81	127	89.5
Euclid Avenue at Woodland Avenue	NW, inbound	5/18/77	377	707	Two-phase signal; No separate left-turn lanes; Left turns (15 percent); Right turns (7 percent)	55	97	144.9
Euclid Avenue at Woodland Avenue	SE, outbound	5/18/77	294	530	1.2-m (4-ft) median; Left turns (10 percent); Right turns (6 percent)	33	81	111.6
Limestone Street at Virginia Avenue	N, inbound	5/26/77 6/7/77	620	1046	Two-phase signal; No separate left-turn lanes; Left turns (10 percent); Right turns (2 percent)	57	95	92.2
Limestone Street at Virginia Avenue	S, outbound	5/26/77 6/7/77	616	1162	Inadequate right-turning radius; Left turns (1 percent); Right turns (14 percent)	54	98	88.3
Main Street at Jefferson Street	SE, inbound	6/1/77	543	711	Three-phase signal; Separate left-turn lanes; Left turns (6 percent); Right turns (7 percent)	83	103	152.9
Main Street at Jefferson Street	NW, outbound	6/1/77	569	716	Minor approaches offset; Left turns (3 percent); Right turns (14 percent)	45	65	79.3
Harrodsburg Road at Larkspur Drive	NE, inbound	7/27/77	674	871	Stop sign on side street; Separate left-turn lanes; Left turns (17 percent); Right turns (0 percent)	32	52	47.1
Harrodsburg Road at Larkspur Drive	SW, outbound	7/27/77	595	985	Left turns (0 percent); Right turns (13 percent)	43	119	71.8

were collected for 11 continuous hours on each of 2 d from 7:30 a.m. until 6:30 p.m. at the intersection of Limestone Street and Virginia Avenue. Traffic volumes of each movement (through, left turns, and right turns) were taken on all approaches by one observer, while observers were stationed on each of the two major approaches.

Data were collected at the site on May 26 (Thursday) and June 7 (Tuesday), 1977, approximately 2 weeks apart. The intersection is located near the University of Kentucky, which is a strong traffic generator. The first count was scheduled to take place after the spring semester ended; the second count was conducted during the summer session. Thus, slightly higher volumes were expected on the second day, and variations in conflicts were expected to be about as high as would normally be expected from day to day at most intersections.

As expected, volumes on the inbound (northbound) approach increased by about 22 percent, from 6162 (day one) to 7514 (day two). The total number of conflicts increased from 566 to 695, a 23 percent increase. The conflict rate on this approach increased very slightly from 91.9 to 92.5 (conflicts per 1000 vehicles). Numbers of conflicts were generally higher during high-volume periods, as shown in Table 2. The highest volume (728) and number of conflicts (81) were observed between 7:30 and 8:30 a.m., the morning rush hour. All values in Table 2 are actual counts and include no adjustments.

Similar results were found on the outbound (southbound) approach. While traffic volume increased 16 percent, from 6258 to 7280, conflicts increased only 3 percent, from 586 to 604. The conflict rate was 93.6 on day one and 83.0 on day two. The highest number of hourly conflicts was 104 (4:30-5:30 p.m.) and 91 (3:30-4:30 p.m.) during afternoon peak hours. The highest hourly volumes also corresponded to these hours.

An analysis was also made to determine the variations in types of conflicts from one day to the next. The percentage of each major conflict type was calculated for each approach on each day. Rear-end conflicts were 57 and 46 percent for the 2 d on the inbound approach and 64 and 58 percent on the outbound approach. Most of these rear-end conflicts were due to traffic congestion and backups throughout most of the test period. Left-turn conflicts, 32 and 41 percent on the inbound approach, were caused by the absence of a separate left-turn lane and a high left-turn demand. On the outbound approach, the percentage of right-turn conflicts (vehicles slowing for right turners) stayed nearly constant. These conflicts were due to an inadequate right-turn radius that caused vehicles to slow drastically to complete the right-turn maneuver. Running the red and other conflicts did not change significantly on the second day.

The previous analysis was not intended to prove that conflicts repeat themselves from one day to the next at all locations. However, at this intersection, conflict numbers and types were very similar for the 2 d. Conflicts, like accidents, are produced by human reactions as well as environmental and traffic conditions. An analysis of this moderately high-volume intersection (average annual daily traffic of 24 000) was made as an initial attempt to gain a better understanding of conflict data. Similar analyses will be conducted in the future, particularly at low-volume rural intersections where greater fluctuations in conflicts are expected.

DEVELOPMENT OF A CONFLICTS PROCEDURE

The development of an effective and practical traffic conflicts procedure was sought for Kentucky. After careful review of several of the conflicts procedures in use in the United States and other countries, the GM technique was revised for use in Kentucky. Several modifications were made with respect to data-collecting procedures.

Data-Collecting Times

Using the GM technique, conflict data are normally collected for 10 h each day from 7:30 a.m. to 12:00 noon and from 12:45 p.m. to 6:15 p.m. at each site on a Tuesday, Wednesday, or Thursday. For low-volume sites, more than a day of data collection may be necessary for an adequate sample size. One observer usually records conflicts while another counts traffic volumes. After each 15 min of data collection, the following 15-min period is used to record data and to move to the opposite approach (1).

This procedure results in the use of about 20 work hours per day, excluding the lunch break (two people for 10 h each). A total of 2.5 h of data is then available for each of the two major approaches. Comparing the total work-hour requirements with the resulting quantity of data obtained from the GM technique, questions were raised as to the efficiency of this procedure. Such large allotments of time were thought to be impractical in Kentucky because of personnel limitations and the large number of locations that warrant conflict counts. Also, little or no useful information was generated from conflict counts during off-peak hours at the test sites. The adequacy of using only one 15-min conflict count to represent an hour of data also needed to be evaluated.

The GM procedure was evaluated from 11-h continuous conflict counts at nine intersection approaches. First, the 15-min count periods were removed from the data that would have been counted by the GM technique.

Table 2. Conflict reliability study.

	Inbound Approach				Outbound Approach			
	Day One		Day Two		Day One		Day Two	
Time Period	Conflicts	Volume	Conflicts	Volume	Conflicts	Volume	Conflicts	Volume
7:30-8:30 a.m.	81	728	94	1046	31	617	45	572
8:30-9:30 a.m.	49	437	38	643	48	352	46	441
9:30-10:30 a.m.	46	485	52	578	37	418	23	497
10:30-11:30 a.m.	33	577	39	566	43	519	38	480
11:30 a.m.-12:30 p.m.	26	652	77	697	71	653	63	857
12:30-1:30 p.m.	77	610	109	681	48	536	56	565
1:30-2:30 p.m.	53	475	75	661	45	444	42	607
2:30-3:30 p.m.	39	582	55	663	55	591	61	716
3:30-4:30 p.m.	60	586	63	757	91	766	91	801
4:30-5:30 p.m.	58	588	57	716	74	895	104	1162
5:30-6:30 p.m.	44	442	36	506	43	467	35	582
Total	566	6162	695	7514	586	6258	604	7280

On an inbound approach, this would correspond to 7:30-7:45 a.m., 8:30-8:45 a.m., 9:30-9:45 a.m., and so on. The outbound periods would be 8:00-8:15 a.m., 9:00-9:15 a.m., 10:00-10:15 a.m., and so on. Each 15-min conflict count was multiplied by four (to obtain an estimated hourly count) and compared to each actual hourly conflict count. A total of 121 h of data were used for this analysis.

The number and percentage of the total hours (y-axis) were plotted against the percentage of error (x-axis) in Figure 2 to summarize the results. The plot shows that an error of 10 percent or less was found in about one-third of the sample. The error is within 17 percent about half the time, and about 75 percent of the sample had an error of 32 percent or less. The difference between the total daily count (11 h) and the GM estimated count (four times the 15-min counts) ranged from 0.7 to 13.2 percent at the 11 intersection approaches. The average difference for all approaches was 4.6 percent.

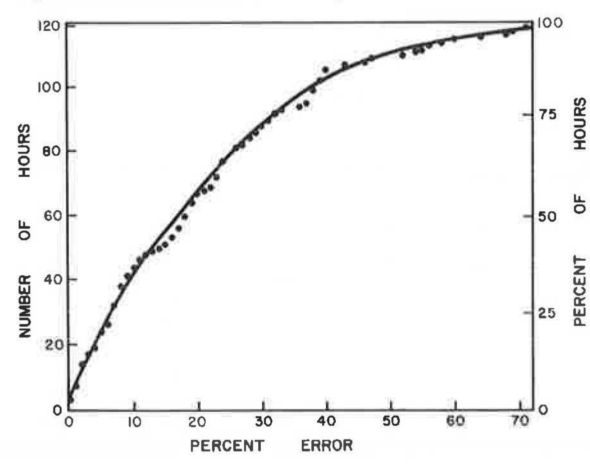
While the 15-min counts each hour proved to be reasonably close in most cases, the personnel required for each count was still a major concern. By plotting conflicts versus time of day, the highest conflict periods occurred during peak hours.

During the morning peak hour (7:30-8:30 a.m.), inbound approaches had their highest conflict numbers, while few conflicts occurred on outbound approaches. The opposite was true in the afternoon, when peak periods generally lasted from 3:30 to 5:30 p.m.

A comparison was made between the GM time periods and the three peak hours in terms of required work hours. If one observer counts conflicts on each approach and the third counts traffic volumes of all movements, only 9 work hours of observation would be required at each intersection. This would produce a total of 3 h of data. Data would represent one high-conflict hour, one low-conflict hour, and one intermediate hour for each approach. About 20 percent more minutes of data would be collected with less than half the work hours expended.

Collecting conflict data only during peak hours was found to be desirable, because off-peak hours were generally uneventful. Problems with left-turning vehicles, for example, are not usually detected until certain left-turn and opposing volumes exist. Care should be taken to avoid collecting more than 1 h of data during very congested times, when some traffic maneuvers are restricted. Data-collecting times should be when problems are suspected. These may correspond to the noon, evening, or weekend rush, or even during seasonal periods at some locations. Additional data may be needed at low-volume sites to obtain adequate samples.

Figure 2. Differences between hourly and 15-min conflict counts.



Conflict Categories

The GM conflict data sheet was revised for use in Kentucky. As currently used, there are 10 columns for counts of vehicle movements and 24 columns for counts of traffic conflicts (a total of 34 categories). Many of these columns were found to be unnecessary; they only create confusion for the observer. The cross-traffic conflicts usually pertain only to unsignalized intersections. Abrupt stops and running-the-red violations are not included on the GM conflict form. To identify left-turn problems, it is necessary in Kentucky to classify weaves, weave conflicts, running red lights, and previous conflicts.

The numbers and rates of each conflict type were summarized for 5700 conflicts observed at four signalized intersections (in the table below, conflicts per 1000 vehicles out of a total flow of 56 897).

Type of Conflict	No. of Conflicts	Conflict Rate
Congestion and traffic backup	3034	53.3
Slow for left turn	885	15.6
Slow for right turn	654	11.5
Brake for previous conflict	203	3.6
Other rear-end conflict	182	3.3
Weave conflict	172	3.0
Running red light	167	2.9
Brake for slow-moving vehicle	135	2.3
Abrupt stop	81	1.4
Opposing left turn	73	1.3
Pedestrian	50	0.9
Other conflicts and erratic maneuvers	125	2.2
Total	5761	101.3

Congestion and backup accounted for 3034 conflicts (52.6 percent), and slowing for left and right turns accounted for another 885 and 654 conflicts respectively (26 percent total). Other conflict numbers over 100 included previous conflicts (203), other rear ends (182), weave conflicts (172), running red lights (167), and braking for slow vehicles (135). Also, abrupt stops, opposing left turns, and pedestrian conflicts were 50 or more. The total conflict rate of the four intersections (all were high-accident sites) was 101.3 conflicts per 1000 vehicles.

Based on the occurrence of conflicts at the test sites, a simplified conflict data sheet was developed for signalized intersections (Figure 3). To aid in the evaluation of the left-turn problems, separate left-turn categories were included for weaves, weave conflicts, running red lights, and previous left turns. All observed conflicts should be classified as either routine, moderate, or severe. Twelve horizontal rows are provided to accommodate 3 h of 15-min counts. The form for unsignalized intersections excludes the running red lights and abrupt stopping categories. Additional categories include five types of cross-traffic conflicts as used in the GM method.

Although the conflict categories on the data sheets will account for about 98 percent of all events, there are various types of weaves, conflicts, and erratic maneuvers peculiar to certain locations. The list below was made up of all such occurrences observed at the test sites or foreseen for others.

Figure 3. Conflict data sheet for signalized intersections.

LOCATION__________ DIRECTION__________ DATE__________

STARTING TIME	WEAVES		WEAVE CONFLICTS		OPPOSING LEFT TURN	RUN RED LIGHT		ABRUPT STOP	SLOW FOR RIGHT TURN	SLOW FOR LEFT TURN	PREVIOUS LEFT TURN	OTHER PREVIOUS CONFLICT	TRAFFIC BACKUP OR CON-GESTION	SLOW-MOVING VEHICLE	PEDESTRIAN CONFLICT
	AROUND LEFT TURNER	OTHER	AROUND LEFT TURNER	OTHER		TURNING LEFT	STRAIGHT OR RIGHT								

Weaves

A Weave for stopped truck
B Weave for stalled vehicle
C Weave for stopped bus
D Weave for road maintenance or construction
E Weave to avoid pedestrian
F Weave into turn lane and back into major traffic flow

Conflicts

G Conflict from erratic maneuver
H Slow for turn out of driveway or shopping entrance
I Slow for turn into driveway or shopping entrance
J Driveway cross traffic from left
K Driveway cross traffic from right
L Slow for stopped bus
M Slow for road maintenance or construction
N Slow for stopped truck
O Weave pedestrian conflict
P Previous conflict from pedestrian (following car)
Q Right turn on red without stop
R Left-lane vehicle slow for right turner
S Slow or stop for stalled vehicle

Erratic Maneuvers

T Left turn from wrong lane
U Right turn from wrong lane
V U turn in road
W Use of shoulder for turns
X Right turner hitting curb
Y Vehicles overrunning stop bar and backing up
Z Vehicle backing from driveway across traffic lanes
AA Turn into wrong lane (opposing lane)
BB Stop in median
CC Run off road
DD Right turn on red without stopping
EE Late-entry right turn (or nonuse of turn lane)
FF Late-entry left turn (or nonuse of turn lane)
GG Vehicle unexpectedly stopping in road
HH Vehicle swerving across traffic lanes
II Vehicle backing in road
JJ Turn into turn lane and back into traffic flow
KK Vehicle on wrong side of road
LL Wide turn (encroaching into adjacent lane)
MM Multiple vehicle erratic maneuver
NN Multiple bicycle erratic maneuver
OO Bicycle on wrong side of road
PP Bicycle riding in median
QQ Illegal pedestrian crossings

This list includes 6 causes of weaves, 13 unusual conflict types, and 24 types of erratic maneuvers. Each observer should have this sheet during a conflict count and be familiar with the categories. If one of these events occurs, a corresponding letter should be put on the data sheet. If the event is repeated several times, one of the extra columns can be designated to count such events.

Volume data should be collected by an observer during all conflict-counting periods if possible. Space is provided for counting left-turning, straight, and right-turning vehicles on all intersection approaches. Most counts will take three observers: one observer per approach and one volume counter.

EVALUATION OF SAFETY IMPROVEMENTS

Shortly after completion of safety improvements at an intersection, another traffic conflict count should be made to determine the effectiveness of the improvement. The second conflict count will often identify minor adjustments, such as signal timing, which would further add to the safety of the intersection. Several evaluations of safety improvements have been completed in Kentucky in recent years in terms of both accidents and conflicts.

In one study, conflict and accident evaluations were conducted at locations where left-turn signal phasing was added. There was an 81 percent reduction in left-turn conflicts (peak hours) at three intersections. An accident study of 24 intersections with similar improvements showed an 85 percent reduction in left-turn accidents after adding exclusive left-turn phases. Based on accident and conflict relationships at 32 intersections, criteria were developed for installation of left-turn phasing. An average of ten or more left-turn conflicts in the peak hour was the conflict criterion. The recommended accident criterion was four left-turn accidents per year on an approach or six accidents in 2 years (5).

Traffic conflicts were used to evaluate the effectiveness of a green-phase extension system (GES) in another Kentucky study in 1976 (6). GES merely extends green time for through vehicles up to about 152 m (500 ft) in advance of high-speed signalized intersections. This supposedly eliminates the "dilemma-zone," which occurs during the amber phase and causes rear-end and right-angle accidents (abrupt stops and running red lights). Six types of conflicts that occur during and shortly after the amber phase were counted at two intersections. Conflict data were taken for 1 d at US-23 and Hoods Creek Pike in Ashland and 2 d at US-27 and US-150 in Stanford for each of the before-and-after periods. These conflicts were reduced by 62 percent at the two intersections after installation of the GES. Total accidents were reduced by 54 percent at three locations with similar improvements (6).

A Kentucky study of erratic maneuvers completed in 1974 tested the effectiveness of various types of raised pavement markers for traffic control at freeway lane drops. Erratic maneuvers, brake applications, and lane volumes were counted at five lane-drop locations. After installation of raised pavement markers, a statistically significant decrease in the total erratic-maneuver rate occurred in nearly all cases, particularly at night. The total reduction in erratic-maneuver rate was 27 percent. No significant change in braking rates was found. The installation of raised pavement markers at other lane-drop locations was recommended based on cost effectiveness (7).

INTERSECTION ANALYSIS

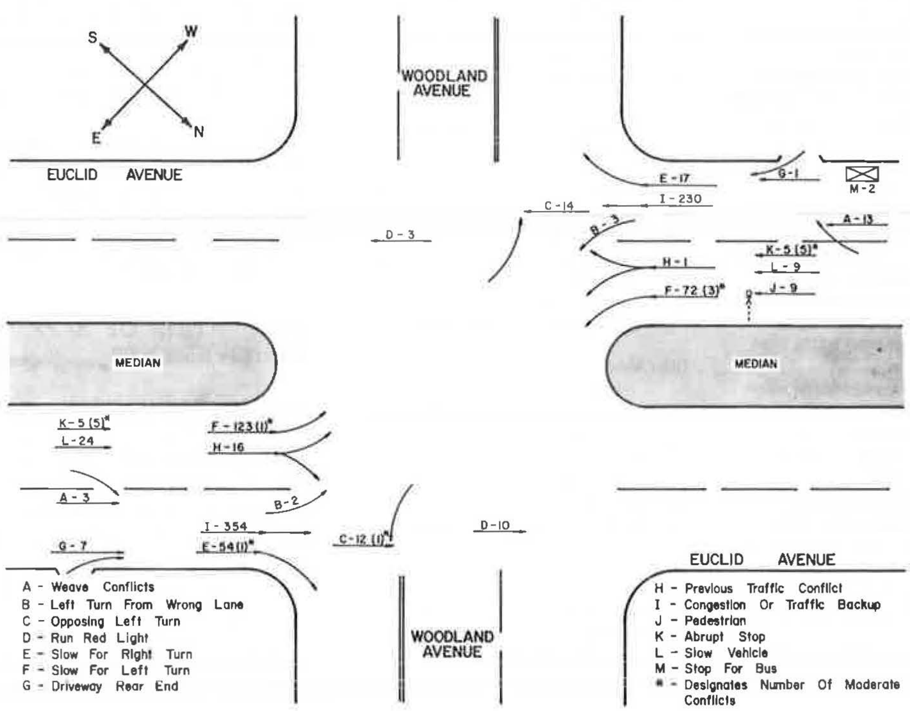
After a highway location is identified as hazardous in Kentucky, a careful analysis is made of the site. This consists of a thorough field investigation by a traffic engineer, a police officer, a local safety engineer, and sometimes other experts. A collision diagram also is used, as are data such as traffic volumes and speeds.

Because of the shortcomings in accident records mentioned earlier, collision diagrams may be of limited value in determining intersection deficiencies. To supplement collision diagrams, experiments have been done with conflict diagrams first used in Kentucky in September 1977. A conflict diagram shares many similarities with a collision diagram, and arrows are used to represent vehicle movements on each major approach. With a conflict diagram, only one set of arrows is used for each conflict type per approach, and the number of conflicts in a specified period is given.

An example of one such conflict diagram for Euclid Avenue at Woodland Avenue in Lexington, Kentucky, is given in Figure 4. The total number of conflicts is given with the number of moderate conflicts in parentheses. Erratic maneuvers and near misses may also be shown on a conflict diagram. As can be seen, the major conflict types (for an 11-h period) on the northwest approach are intersection backup and congestion (354), slowing for left turn (123), slowing for right turn (54), slow truck (24), and previous conflicts (16). Other types included opposing left-turn (12), running red light (10), driveway conflicts (7), abrupt stops (5), weave conflicts (3), and turns from wrong lane (2). The southeast approach had similar problems and also had several pedestrian conflicts and stop-for-bus conflicts (8).

Based on this conflict diagram, recommendations were made to add dual left-turn lanes on Euclid Avenue to reduce conflicts from vehicles slowing or weaving for left turners. Adjustments in signal timing were also recommended. The high incidence of backup and congestion conflicts was found to be unavoidable because of

Figure 4. Typical conflict diagram.



moderately high traffic volumes, but it was not abnormally high compared to other signalized intersections.

Another aid to intersection analysis is the use of conflict rates. The hourly conflicts, peak-hour conflicts, and conflict rates are given in Table 1 for all approaches. The highest hourly conflicts (83) and conflict rate (152.9 conflicts per 1000 vehicles) were found on the southeast approach of Main Street at Jefferson Street. Based on all available conflict data, specific problems were found on this approach, and appropriate safety improvements were recommended.

RECOMMENDATIONS

Based on the successful use of conflict and erratic maneuver data in Kentucky since 1972, increased use should be made of such data on a routine basis. A procedure for collecting and analyzing conflict data was developed and is recommended. Since 1970, a total of 904 locations have been investigated under Kentucky's spot-improvement program (about 130 per year). By routinely conducting conflict counts during such investigations, a large sample of conflict data would be available within a few years. This would provide the engineer with a systematic procedure for observing the location, and a permanent record of driver confusion and error could be generated and compared with problems at other locations. Valuable information on which to base appropriate safety improvements at the site would be obtained, and an after study of conflicts would allow for an evaluation of the improvements.

REFERENCES

1. S. R. Perkins and J. I. Harris. Traffic Conflict Characteristics—Accident Potential at Intersections. General Motors Research Publication GMR-718, Dec. 7, 1967.
2. Digest of Motor Laws. American Automobile Association, Falls Church, VA, 1974.
3. K. R. Agent. Evaluation of the High-Accident Location Spot-Improvement Program in Kentucky. Kentucky Bureau of Highways, Division of Research, Lexington, Feb. 1973.
4. C. V. Zegeer. Identification of Hazardous Locations on Rural Highways in Kentucky. Kentucky Bureau of Highways, Division of Research, Lexington, June 1974.
5. K. R. Agent. Development of Warrants for Left-Turn Phasing. Kentucky Bureau of Highways, Lexington, Division of Research, Aug. 1976.
6. C. V. Zegeer. Effectiveness of Green-Extension Systems at High-Speed Intersections. Kentucky Bureau of Highways, Lexington Division of Research, May 1977.
7. J. G. Pigman and K. R. Agent. Raised Pavement Markers as a Traffic Conflict Measure at Lane Drops. Kentucky Bureau of Highways, Division of Research, Lexington, Feb. 1974.
8. M. E. Clayton. Evaluation of Urban Intersections Using Traffic Conflicts Measures. Kentucky Bureau of Highways, Division of Research, Lexington, Aug. 1977.

Discussion

Peter J. Yauch, JHK Associates, Atlanta, Georgia
Peter S. Parsonson, Georgia Institute of Technology, Atlanta

Traffic conflict studies were performed in an attempt to determine the extent of accidents caused by the "dilemma zone" problem at high-speed, isolated, signalized intersections. The studies were made during the summer of 1977 as part of a master's degree research project.

The conflict studies were based on similar studies performed in Kentucky in 1976 that evaluated the effectiveness of green-extension systems in reducing the dilemma-zone problem. The seven categories of conflicts used were identical to those used by Kentucky except for the division of running red light into two categories based on the offending vehicle's relation to the stop line.

Five high-speed, isolated, signalized intersections having potential dilemma-zone problems were chosen for study on the recommendation of the Fulton County, Georgia, traffic engineer. All five intersections had similar approach characteristics. One intersection was operated in the semiactuated mode, with no detectors on the high-speed approaches; the other four intersections were fully actuated, but with detector setbacks far below those recommended for high-speed approaches.

Traffic conflicts for each approach were tabulated for two full hours of high traffic volume. Offending vehicles were classed by type. Volumes taken during the conflict studies were also divided into the three vehicle classes. Only non-turning vehicles were observed in the study.

Seventeen months of accident data, for the period just prior to the study, were obtained from police accident files. An attempt was made to identify the accidents caused by the dilemma-zone problem. In general, rear-end accidents, right-angle accidents caused by a high-speed vehicle running the red light, and sideswipe accidents caused by improper lane changes were classed as dilemma-zone accidents. Additional information concerning accident cause was usually not available. In addition, it is probable that not all accidents were reported, as noted in the Zegeer and Deen paper.

A comparison of the dilemma-zone accident and conflict rates showed no correlation (r = 0.16) (see Figure 5). Similar results were obtained when the total accident rate was compared to the conflict rate. This can point to two different conclusions, either that the accident data are unreliable or that conflict studies cannot be used to predict high dilemma-zone accident locations.

However, the usefulness of the dilemma-zone conflict study performed may be shown by two additional comparisons.

The clearance intervals required for each of the 10 approaches were calculated with the method described in the Traffic and Transportation Engineering Handbook. A deceleration rate of 2.7 m/s^2 (9 ft/s^2) was used instead of the suggested 4.6 m/s^2 (15 ft/s^2), the latter being obtainable only as a panic stop on dry pavement. The ratio of actual clearance-period timing (yellow plus all red) to required timing was compared to the dilemma zone conflict rate, and resulted in a good correlation (r = 0.83). The running-red-light conflict rate did not appear to be related to clearance timing. Figure 6 shows the "scattergram" and least squares analysis line for the first comparison.

By classifying the seven conflict categories into two types, in which the offending vehicle either stopped or continued on through the intersection, another comparison can be made. A regression analysis, resulting in an r-value of 0.74, shows that the ratio of nonstopping conflicts to total conflicts increases with an increase in approach volume. This agrees with the expectation that drivers are hesitant to stop quickly in heavy traffic. Figure 7 shows the results of this regression analysis.

Therefore, the conflict data taken during this study tend to agree with expected occurrences within the di-

Figure 5. Dilemma-zone accident rate versus conflict rate.

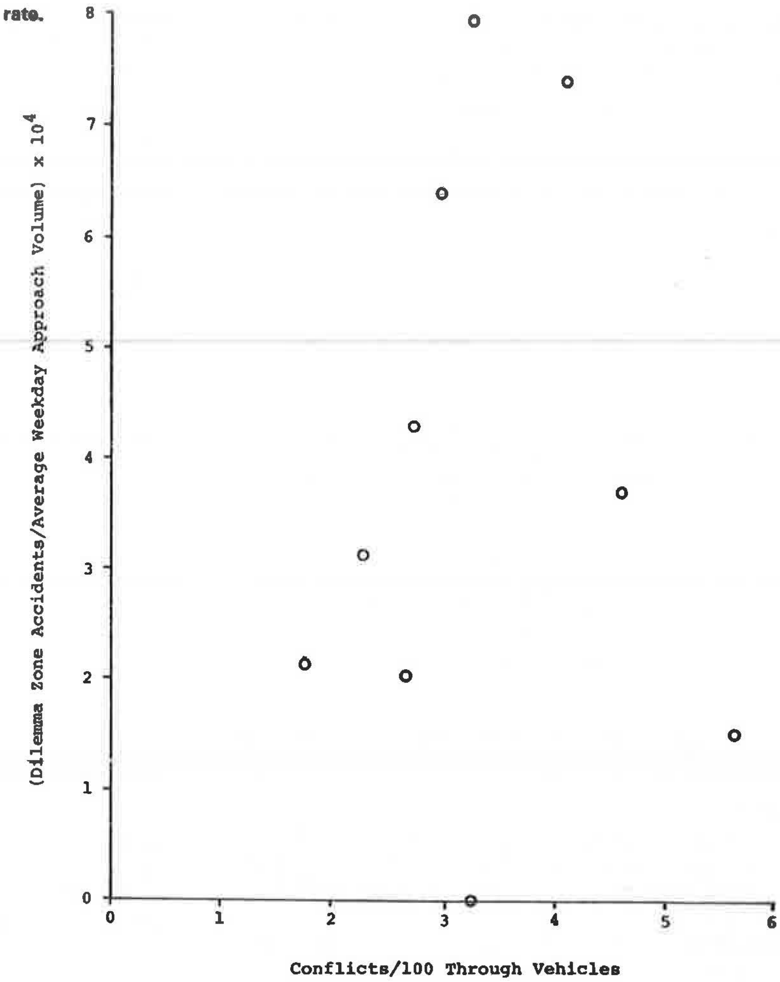


Figure 6. Conflict rate versus percentage of required clearance time.

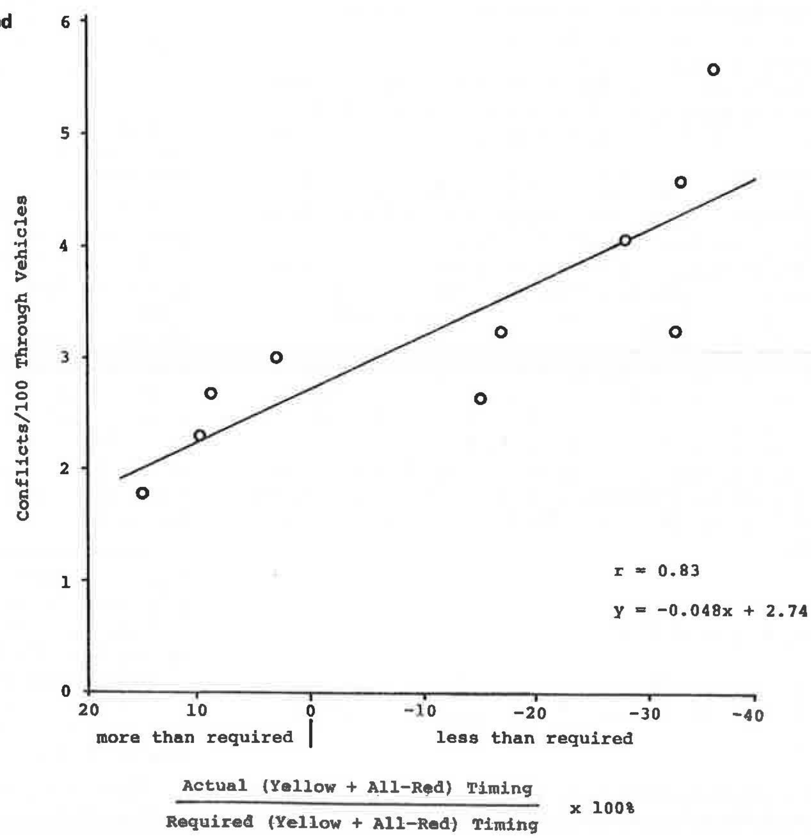
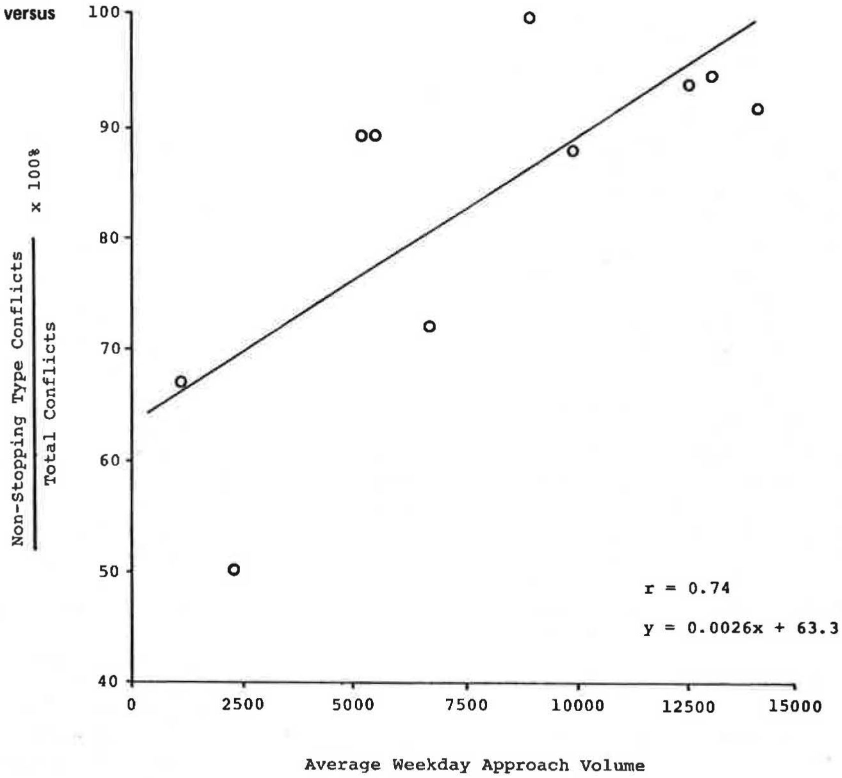


Figure 7. Percentage of nonstopping conflicts versus approach volume.



lemma zone and reinforce the usefulness of the study method.

Since no studies were made at a location where the dilemma zone was known not to be a contributor to the accident history, a comparison was not available. However, it appears that the traffic conflict study is of value in evaluating the timing of clearance intervals and stopping behavior in heavy traffic.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Design Considerations of Traffic Conflict Surveys

E. Hauer, Department of Civil Engineering, University of Toronto

The traffic conflicts technique is a device for measuring safety indirectly. It requires, at present, a field count of conflict occurrences, which gives the basis on which the rate at which conflicts occur is estimated. This report deals with the accuracy of such estimating and its dependence on the design of the field survey. Current practices in conflict-count duration are reviewed, and the relationship between count duration and estimation accuracy is examined. Using data from several sources, the daily variability of conflict counts is described. It is concluded that the expected conflict rate varies from day to day. Use of negative binomial distribution is suggested as appropriate for representing the distribution of sample means obtained from conflict studies. On this basis, confidence limits and probabilities of type I and type II errors in hypothesis testing are obtained and tabulated. Their use in study design is illustrated by numerical examples. The marginal increase in estimation accuracy diminishes rapidly as conflict-counting time increases. Thus, there is little to be gained by counting longer than 3 d. This establishes a practical limit to the accuracy with which expected daily conflict rates can be estimated.

The traffic conflicts technique is a device for indirectly measuring safety. Its early history may be traced (1, 2, 3, 4), and its recent applications have been described (5, 6, 7, 8, 9). There are also state-of-the-art surveys now available (9, 10, 11).

The traffic conflicts technique is applicable to a variety of situations. It can be used to assess changes in safety through before-and-after studies and by comparison with control sites; to investigate effectiveness of devices, layouts, design, and procedures; and to identify and diagnose hazards.

All such uses require a field study to observe and count the occurrences of conflicts and thus estimate their occurrence rate. The purpose of this paper is to examine available empirical evidence in order to provide guidance for conducting conflict surveys.

The discussion will center on conflict-rate estimation accuracy, survey duration, and sample size selection. On the basis of information from several sources, the variability of daily conflict counts will be characterized and a model suggested for this counting distribution. Various aspects of the conflict study design will be illustrated by numerical examples. Tables and graphs will be supplied for use in survey design.

Present practice in conflict-count duration is summarized in Table 1.

Glennon and others (10) recently raised grave questions about the validity of present practice. They conclude: "For all three potential uses of conflict counts, existing relationships do not allow practical sample sizes." The conclusion, if true, would have far-reaching consequences. Not only does present practice in the conduct of conflict studies seem inadequate, there also seems to be little hope that the sample sizes required by Glennon and others would leave much interest in applying the traffic conflicts technique under any conditions. As this very important conclusion has been reached on the basis of limited empirical evidence, careful reexamination is in order.

EXPECTED CONFLICT RATE

The aim of a conflict survey is to obtain satisfactory estimates of the expected conflict rate. This is not a simple concept and requires delineation. In intuitive terms, the concept of expectation is closely associated with the notion of average, in the long run. We tend to believe that just as throws of a die will, in the long run, average 3.5, so would repeated conflict counts reveal a permanent characteristic of the site. The analogy, however, is incomplete. Unlike the die, the site changes its average property. There is little reason to assume that the expected conflict rate is the same during peak and off peak, Sundays and Mondays, winter and summer. It is essential, therefore, to specify which expected rate is subject to estimation.

We will proceed on the assumption that it is the expected weekday conflict rate that is of interest. That is, the average number of conflicts occurring per unit of time during a specified period of observation characterizing any weekday during a certain season of the year. There are two reasons for this choice. First, surveys are designed in terms of team days. Thus, it makes little sense to be concerned much about, say, hourly variations. Second, the traffic conflicts technique is used principally in comparisons between sites, devices, and treatments that are usually performed in a relatively short period of time. This eliminates the need to consider seasonal variation.

COUNTING, ESTIMATION, AND ACCURACY

In Figure 1, circles represent the number of conflicts counted on 19 consecutive weekdays during 7:00-10:00 a.m. and 3:00-6:00 p.m. at the intersection of St. Clair and Keele Streets in Toronto. The bars in the figure represent the estimation of expected conflict rate obtained by averaging the first 1... 19 daily counts.

This simple graph illustrates all major features of the problem at hand.

First, the tangible evidence is the daily conflict count, which is subject to considerable variations. It is this random fluctuation that is the root source of difficulties in estimation.

Second, the fluctuating daily conflict counts are used to obtain an estimate of the expected conflict rate. In the present case, the simple average is used for estimation. Unlike the daily counts, the estimate of the expected rate is characterized by pronounced stability.

Third, the accuracy of the estimate increases with the number of daily counts. At first, every added count increases accuracy markedly. Beyond a certain point, not much accuracy is gained by counting more.

The qualitative observations made on the basis of all illustrative examples need to be quantified. This will be done, on the basis of data from several sources, with mathematical statistical methods. These allow measurement of variability in daily counts, characterization of accuracy of estimation, and so forth. However, how accurate the estimate should be cannot be determined by using statistics only. Standards of accuracy should depend on the circumstances of the survey and on the use to which its results will be put.

Figure 1. Daily conflict counts and their running averages.

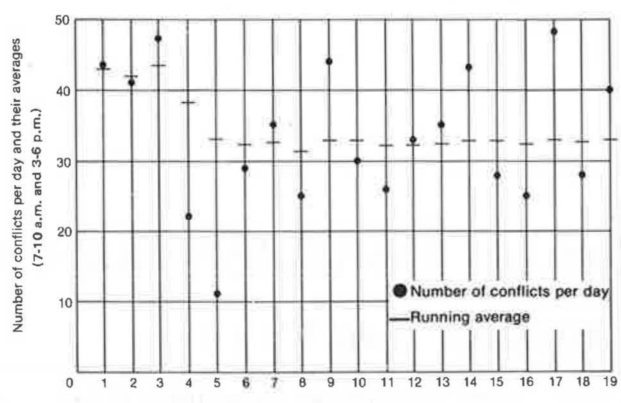


Table 1. Summary of conflict counts.

Source	Country	Intersections Surveyed	Duration	Counting Type	Team
Amundsen (9, 1974)	Norway	31 low volume, unsignalized	1 d, 7 h/d		Two- or three-person (one for conflicts)
Baker (3, 1972)	United States	392 before, 193 after modification	1 d	15-min samples on each approach with breaks between counts	Two-person
Cooper (12, 1973)	Canada	59 unsignalized, 51 urban, 5 rural	2 d, 14 h/d	Continuous	Five-person (three for conflicts)
Hydén (6, 1975)	Sweden	50 urban, varied control	2 d, 9.5 h/d	Continuous	One- or two-person
Malaterre and Muhlrad (8, 1976)	France	8 urban	1 d, 7:00 a.m.-12:00 midnight	Sample varies 13-57 percent of 1 h	Two-person
Perkins and Harris (1, 1963)	United States	30 signalized and unsignalized	3 d, 12 h/d	Alternate sampling per direction	
Spicer (4, 1973)	Great Britain	6 rural	1 d, 10 h/d	Continuous plus 16-mm time-lapse photography	Three- or four-person

This is unfortunate. One usually prefers to have firm and explicit standards for guidance. Conducting an elaborate decision analysis study in every case is impractical. There is a strong temptation, therefore, to simply adopt commonly used significance levels, because their use in medicine, sociology, quality control, psychology, and other fields lends them sufficient authority.

A word of caution is in order. It is wrong to blindly transfer levels of significance well suited for, say, concrete quality control directly to management of safety. The costs of conducting experiments are different; the implications of error are not the same. For most safety countermeasures, statistically conclusive evidence of effectiveness is unobtainable practically. Recognizing this as a fact of life, does one proceed and recommend that such things as driver licensing and vehicle inspecting be discontinued because neither can be shown effective at the 5 percent significance level? "The research community should consider carefully before doing so" (13). Uncritical use of high significance levels "can lead to the erroneous rejection of effective programs" (14). Rather, we need to adopt standards of accuracy that reflect both the benefit lost by not implementing measures that are likely but undemonstrably beneficial and the cost of implementing measures that are not effective.

DATA BASE

Investigating the distribution of daily conflict counts requires information on the number of conflicts at several sites over a relatively long period of time. Such information is not easy to come by. Daily counts are discarded or difficult to obtain, partly because conflicts are rarely counted for more than a couple of days and partly because usually only the average number is retained and archived.

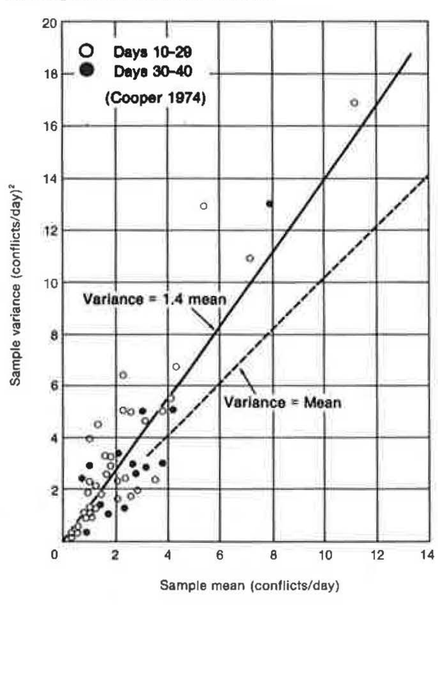
Fortunately, in the course of a study on the effectiveness of law enforcement (5), a good data base for the present purpose was generated. Conflicts were counted for 39 weekdays at each of seven urban intersections during 7:00-10:00 a.m. and 3:00-6:00 p.m.

The first 2 weeks (10 survey days) with normal police activity were followed by 4 weeks (19 survey days) with increased enforcement and another 2 weeks (10 survey days) with normal enforcement again. At each location seven conflict types were recorded. In the analysis below, the initial 10 d of the survey will not be used because "Initially, all observers tended to overcount drastically. . . stabilization of these counts did not proceed as quickly as anticipated and in most cases could not realistically have occurred by the beginning of increased enforcement" (5).

Thus, data from seven locations, seven conflict types, and two sequences of 19 and 10 d are used. It should be noted that the variations in counting will of necessity be overestimated because some will have been generated by the changes in police activity even within the phases that are analyzed separately.

To avoid reliance on one source of data, however extensive, two additional smaller sets of data were used. The first, conflict counts from 20 sites, was published by Hydén (6). In this case, 2 d of observation for each site are available. The second data set was collected by the Transport and Road Research Laboratory at four rural intersections; for three of them only 2 d of conflict counts exist, while for the remaining intersection 3 d of counts are available.

Figure 2. Sample mean and variance for homogeneous conflict classes.



VARIATIONS IN DAILY CONFLICT COUNTS

Variability is usually measured by sample variance. In Figure 2, sample variance is plotted against sample mean. Each point in Figure 2 represents the sample mean and variance of a homogeneous conflict class (cross traffic, rear end, etc.) at seven intersections for two phases of the study (5) with different enforcement levels. I have made the following observations.

First, the variance of the conflict count increases with the mean count, which is to be expected. Obviously, when the mean count is zero, so is the variance. Thus the origin is the starting point of the curve describing the relationship. In the range of mean conflict counts for which data are available, it is simplest to represent the relationship by a line through the origin. When conflicts for each homogeneous class are counted separately, the average variance-to-mean ratio is 1.4. When the sum of all conflict classes is of interest, the average variance-to-mean ratio is 2.2.

In their paper on evaluating the traffic conflicts technique (10), Glennon and others concluded that daily conflict counts are characterized by a constant variance of 530 (conflicts per day) (5) irrespective of the daily conflict rate. It is on the basis of this variance that they derive the number of days needed for conflict surveys. Our data do not confirm the assumption of constant variance; on the contrary, as in most known counting distributions, variance is found to increase with the mean. Nor can one find support for the high value of the variance used by Glennon and his coworkers.

The second observation to be made on the basis of Figure 2 pertains to the Poisson hypothesis. In the absence of empirical evidence to the contrary, it is usually assumed that rare events with a constant mean follow the Poisson distribution. Were this so, one would expect the dotted line labeled variance = mean to fit the data. It is apparent that the Poisson hypothesis does not hold. This may be so because the expected rate of conflict occurrence at intersections changes from day to day because of changes in weather, vehicle flow, pedestrian volumes, etc. In addition, some variability is introduced by the subjectivity of the observers identifying conflicts.

Third, there is no assurance that the same distribution describes the conflict-counting process irrespective of the conflict type counted, the counting procedure used, the definition of the conflict event, or the specific circumstances of the site. Hydén's results (6), for example, suggest a smaller variability than the rest of the data. When specific information about count variability is not available, use of the average values obtained in this paper is recommended.

THE MODEL

To facilitate survey design and analysis in customary statistical terms, one has to adopt a model probability distribution that is simple to use, fits the data, and represents a process that bears a reasonable resemblance to our perception of reality.

The negative binomial distribution has been used for similar purposes in the past (15). It is founded on the assumption that the daily expected conflict rate follows a gamma distribution and the actual daily conflict counts a Poisson distribution with the aforementioned daily expected conflict rate as a mean.

By adopting the negative binomial distribution, it is shown that the distribution of the sample mean ($\bar{X}$) obtained from a count over j days is given by

$$P(X = n/j) - (-1)^n \binom{-\nu j}{n} p^n q^{\nu j},\ n = 1, 2, \ldots$$

with

$$E\{\bar{X}\} = \nu/a$$
$$VAR\{\bar{X}\} = E\{\bar{X}\}(1 + a)/(aj) \tag{1}$$

where

$\bar{X}$ = sum of j daily conflict counts divided by j,
$p = 1/(1 + a)$,
$q = 1 - p$,
ν = a (expected daily conflict rate), and

$$a = \begin{cases} 2.5 \text{ for homogeneous conflict classes or} \\ 0.83 \text{ for the sum of several conflict classes.} \end{cases}$$

What follows is a calculation that shows the origin of Equation 1 and provides useful information about characteristic functions, moments, and estimation.

Let X be the number of conflicts counted on a day during a given period of time and λ be the expected number of conflicts that day. If the count of that day obeys the Poisson distribution, then

$$P(X = k|\lambda) = (\lambda^k/k!)e^{-\lambda} \tag{2}$$

Regarding λ as a continuous random variable that assumes different values on different days, the distribution of X over many days is given by

$$P(X = k) = \int_0^\infty P(X = k|\lambda) f(\lambda) d\lambda \tag{3}$$

When λ obeys the gamma distribution

$$f(\lambda) = \begin{cases} [a^\nu/\Gamma(\nu)]\lambda^{\nu-1}e^{-a\lambda} & \text{for } \lambda > 0 \\ 0 & \text{for } \lambda \leq 0 \end{cases} \tag{4}$$

with $\nu > 0$ and $a > 0$.

Substituting Equation 4 into Equation 3 and integrating, we obtain

$$P(X = k) = (-1)^k \binom{-\nu}{k} p^k q^\nu$$
$$k = 0, 1, 2, \ldots \tag{5}$$

where

$p = 1/(1 + a)$,
$q = 1 - p$,
$\binom{-\nu}{k} = (-\nu)(-\nu - 1) \ldots (-\nu - k + 1)/k!$

The probability distribution defined by Equation 5 is the negative binomial distribution. Its characteristic function is given by

$$\phi_X(t) = q^\nu(1 - pe^{it})^{-\nu} \tag{6}$$

Ordinary moments of order r are given by

$$m_r = \sum_{i=0}^{r-1} (-1)^{r-1} \binom{r-1}{i} (p/q)^{r-i} (-\nu)_{r-i} \tag{7}$$

and the two central moments by

$$\mu_1 = \nu(p/q) = \nu/a \tag{8}$$

and

$$\mu_2 = \mu_1(1 + p/q) = \nu(1 + a)/a^2 \tag{9}$$

Thus, the negative binomial distribution is complctcly specified by the two parameters ν and a. To estimate their value from Equations 8 and 9

$$a = 1/[(\mu_2/\mu_1) - 1] \tag{10}$$

Also from Equation 8 and using Equation 10 we obtain

$$\nu = \mu_1/[(\mu_2/\mu_1) - 1] \tag{11}$$

Thus, when conflicts belong to a homogeneous class, $a \cong 1/(1.4 - 1) = 2.5$, and, when the sum of all conflict classes is of interest, $a \cong 1/(2.2 - 1) = 0.83$. The variance-to-mean ratios 1.4 and 2.2 were obtained earlier.

In the final account we are interested in the distribution of the average conflict count obtained from counting j days. Denote daily counts by $x_1, x_2, \ldots, x_i, \ldots x_j$ and

$$\bar{X} = (1/j) \sum_1^j X_i$$

As counts on successive days are statistically independent, using Equation 6,

$$\phi_{\bar{X}}(t) = [\phi_X(t/j)]^j = q^{\nu j}(1 - pe^{it/j})^{-\nu j} \tag{12}$$

The characteristic function in Equation 12 belongs to the modified negative exponential distribution

$$P(\bar{X} = n/j) = (-1)^n \binom{-\nu j}{n} p^n q^{\nu j}$$
$$n = 0, 1, 2, \ldots$$

with

$$E\{\bar{X}\} = \mu_1$$
$$VAR\{\bar{X}\} = \mu_2/j \tag{13}$$

ACCURACY, ERRORS, AND DECISIONS

By using the data on daily variability in conflict counts and the suggested probability model, one can tabulate the probability distribution of conflict counts. (Tables of the probability distribution, confidence intervals, and type I and II errors for expected daily conflict rates between 2 and 20 are available from the Transport and Road Research Laboratory.) This is the basic information needed for statistical considerations of any kind.

The various uses of results obtained so far are best discussed within the framework of illustrative examples, or sample problems.

Example 1. Confidence Limits

From a 2-d survey of cross-traffic conflicts, an average daily count of 10.0 conflicts has been obtained. Find the 50 and 90 percent confidence intervals. The solution is from Figure 3, where the 50 and 90 percent confidence intervals are 4 and 11 conflicts per day.

For high confidence levels, the interval is large. Thus, one must either adopt modest standards of accuracy or invest in longer counts.

Example 2. Survey Design for Specified Confidence Limits

How many count days are needed to obtain a 90 percent confidence interval of 4 or less under the conditions of the previous example? The solutions, given in the table below for the 90 percent confidence limits, are

No. of Survey Days	90 Percent Confidence Limits
2	5.5-14.1
3	6.3-13.4
4	6.8-12.9
5	7.2-12.7
6	7.3-12.3

Figure 3. Confidence intervals for homogeneous conflict classes.

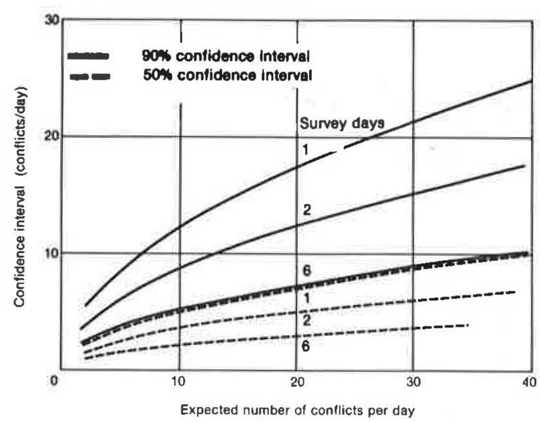
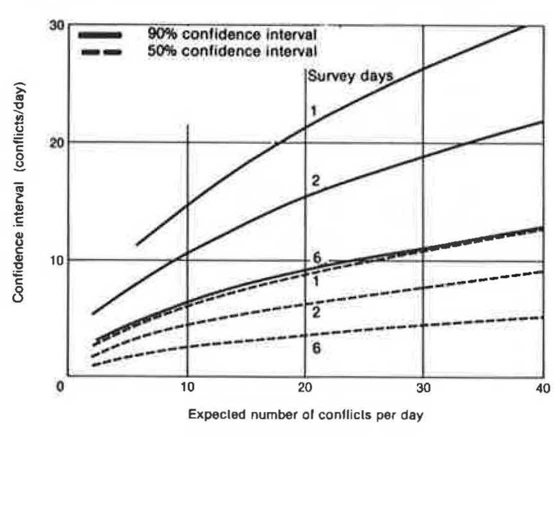


Figure 4. Confidence intervals for all conflict classes.



It appears that attainment of the specified accuracy is difficult. A count duration in excess of 10 d would be needed. As may be seen, the reduction in the confidence interval by counting longer diminishes. One should therefore weigh the increase in accuracy against the cost of prolonging the survey.

Example 3. Confidence Limits for Daily Conflict Rates Larger Than 20

Anticipating a daily conflict rate of 40 (sum of all conflict classes), how many days are needed for counting so that a 75 percent confidence interval is less than 10 conflicts per day? The answer is that, as the conflict rate exceed 20, tables of the normal probability distribution may be used. For the sum of all conflict classes, a = 0.83 (Equation 1). Also from Equation 1, VAR $\{\overline{X}\}$ = 40 × (1 + 0.83)/(0.83j). To determine the 75 percent confidence limits given below, multiply (VAR $\{\overline{X}\})^{1/2}$ by the 1.15 obtained from tables of the normal probability distribution.

No. of Survey Days	(VAR $\{\overline{X}\})^{1/2}$	75 Percent Confidence Limits
1	9.4	±10.0
2	6.6	±7.6
3	5.4	±6.2
4	4.7	±5.4
5	4.2	±4.8
6	3.8	±4.4

For the prescribed accuracy, 5 d of counting are needed. For 50, 60, 70, 80, 90, and 95 percent confidence limits, 0.67, 0.84, 1.04, 1.28, 1.64, and 1.96 are the appropriate multipliers.

Determination of confidence limits for conflict counts is relatively easy; their interpretation is straightforward. This is therefore likely to be the best basis on which to make intuitive decisions about survey duration. After all, one can readily assess the gain in accuracy from prolonging the survey one more day and the cost of doing so.

Figures 3 and 4 give the size of confidence intervals in terms of survey duration and expected conflict rate. This should prove to be an effective guide in many circumstances.

In some situations, the probability distribution of count averages and the associated confidence intervals may not be deemed sufficient for conflict survey design and survey result analysis, notably when treatment effectiveness is the main concern. This is most common in so-called before-and-after studies.

In this context, one usually wishes to ascertain whether some treatment is effective in reducing the number of conflicts. A positive answer may lead to the modification of design standards, installation of new equipment, reconstruction of inferior site features, etc.

While the question of whether the treatment is effective is simple, the answer is not. Indeed a straightforward response is not forthcoming. It is for this reason that use of confidence limits may be preferred as the device least given to misinterpretation.

To provide an answer of a sort, the problem of treatment effectiveness needs first to be recast in terms of testing a statistical hypothesis. The outcome of such a test is not a statement that the hypothesis is true or false; it is merely a statement specifying the chance of error, should one decide on the basis of conflict studies to accept or reject the hypothesis.

It is customary (for little good reason) to test the hypothesis that the treatment is not effective. If so, concluding on the basis of data that the treatment is effective when in fact it is useless constitutes a type I error. Conversely, maintaining on the basis of empirical evidence that the treatment has no effect, while in fact it is useful, is an error of type II.

These errors depend on the decision rule that determines acceptance or rejection of the hypothesis. While essential for formal decision analysis, this information is not understood intuitively and is therefore apt to be misinterpreted. It invites, therefore, use of arbitrary significance levels that are not derived from the reality of the situation at hand. As statistical hypothesis testing is deeply ingrained in present practice, the two types of error have been tabulated by the Transport and Road Research Laboratory. To guard against the possibility of misinterpretation, they are described in as clear a language as possible and their use is illustrated by numerical examples.

Example 4. Failure to Observe Improvement When Improvement Exists

Example 4 deals with the chance of failing to detect (through a reduction in the count of conflicts) a real decrease in the expected rate at which conflicts occur. The subsequent examples focus on the probability of observing a reduction in conflict counts when in fact there has been no change in the expected conflict rate.

Due to a successful treatment, the expected daily rate of cross-traffic conflicts has been reduced from 12 to 8. Thus, in the long run, the number of such conflicts is reduced by 33 percent. Determine the probability of observing no reduction in the average conflict count if 2 d of before counts are compared to 2 d of after counts. Such a result—failure to observe reduction in conflicts when improvement does exist—might lead one to conclude erroneously that the treatment had no beneficial effect.

The solution can be seen in Figure 5. For this example, the probability is 0.156. To aid interpretation, consider 100 sites at which treatment reduces the expected daily conflict rate from 12 to 8. Counting conflicts for 2 d before and after treatment, approximately 84 sites will show a reduction in the average number of conflicts; at the remaining 16 sites, no reduction will be observed.

It is natural to ask now how important it is to reduce the probability of failing to observe a reduction in the average conflict count. The answer depends on the specific objectives and circumstances of each survey. If the effect of the treatment is examined at 10 sites, a reduction in conflicts should be obtained on the average of

Figure 5. Probability of the before mean being equal to or smaller than the after mean when improvement exists: homogeneous conflict classes.

m_1	j										
2.	1	0.595									
2.	2	0.562									
2.	3	0.547									
2.	4	0.543									
2.	5	0.538									
2.	6	0.536									
4.	1	0.301	0.562								
4.	2	0.192	0.543								
4.	3	0.132	0.534								
4.	4	0.094	0.529								
4.	5	0.068	0.526								
4.	6	0.054	0.526								
6.	1	0.138	0.344	0.549							
6.	2	0.050	0.252	0.534							
6.	3	0.020	0.196	0.528							
6.	4	0.008	0.156	0.524							
6.	5	0.003	0.126	0.521							
6.	6	0.001	0.103	0.519							
8.	1	0.059	0.172	0.367	0.543						
8.	2	0.011	0.094	0.287	0.529						
8.	3	0.002	0.050	0.236	0.524						
8.	4	0.000	0.027	0.197	0.520						
8.	5	0.000	0.015	0.168	0.518						
8.	6	0.000	0.009	0.144	0.516						
10.	1	0.024	0.100	0.227	0.382	0.538					
10.	2	0.002	0.029	0.129	0.310	0.526					
10.	3	0.000	0.010	0.078	0.263	0.521					
10.	4	0.000	0.003	0.049	0.227	0.518					
10.	5	0.000	0.001	0.031	0.198	0.516					
10.	6	0.000	0.000	0.020	0.174	0.515					
12.	1	0.010	0.050	0.132	0.252	0.393	0.534				
12.	2	0.000	0.008	0.050	0.156	0.327	0.524				
12.	3	0.000	0.001	0.020	0.103	0.283	0.519				
12.	4	0.000	0.000	0.009	0.070	0.249	0.516				
12.	5	0.000	0.000	0.004	0.049	0.222	0.515				
12.	6	0.000	0.000	0.002	0.034	0.198	0.513				
14.	1	0.004	0.023	0.073	0.158	0.272	0.402	0.531			
14.	2	0.000	0.002	0.017	0.069	0.179	0.340	0.522			
14.	3	0.000	0.000	0.004	0.033	0.125	0.299	0.518			
14.	4	0.000	0.000	0.001	0.016	0.090	0.267	0.515			
14.	5	0.000	0.000	0.000	0.008	0.065	0.240	0.513			
14.	6	0.000	0.000	0.000	0.004	0.048	0.218	0.512			
16.	1	0.001	0.011	0.039	0.094	0.179	0.287	0.408	0.529		
16.	2	0.000	0.000	0.005	0.027	0.087	0.197	0.351	0.520		
16.	3	0.000	0.000	0.001	0.009	0.046	0.144	0.312	0.516		
16.	4	0.000	0.000	0.000	0.003	0.025	0.107	0.281	0.514		
16.	5	0.000	0.000	0.000	0.001	0.014	0.081	0.256	0.512		
16.	6	0.000	0.000	0.000	0.000	0.008	0.062	0.234	0.511		
18.	1	0.001	0.005	0.020	0.054	0.112	0.196	0.300	0.414	0.528	
18.	2	0.000	0.000	0.001	0.010	0.038	0.103	0.213	0.360	0.519	
18.	3	0.000	0.000	0.000	0.002	0.014	0.058	0.160	0.322	0.515	
18.	4	0.000	0.000	0.000	0.000	0.005	0.034	0.123	0.293	0.513	
18.	5	0.000	0.000	0.000	0.000	0.002	0.020	0.096	0.269	0.512	
18.	6	0.000	0.000	0.000	0.000	0.001	0.012	0.075	0.248	0.511	
20.	1	0.000	0.002	0.010	0.029	0.068	0.129	0.211	0.310	0.418	0.526
20.	2	0.000	0.000	0.000	0.003	0.015	0.049	0.118	0.227	0.367	0.518
20.	3	0.000	0.000	0.000	0.000	0.004	0.020	0.071	0.174	0.331	0.515
20.	4	0.000	0.000	0.000	0.000	0.001	0.009	0.044	0.137	0.303	0.512
20.	5	0.000	0.000	0.000	0.000	0.000	0.004	0.027	0.109	0.280	0.511
20.	6	0.000	0.000	0.000	0.000	0.000	0.002	0.017	0.088	0.260	0.510
m_2		2	4	6	8	10	12	14	16	18	20

m_1 Expected daily conflict rate 'before'
m_2 Expected daily conflict rate 'after'
j Number of survey days

eight sites by counting conflicts for 2 d. It can happen, of course, that a reduction will be observed at five sites or less. The probability of this event is less than 2 percent (using the binomial distribution with p = 0.156). In this case, a 0.156 level of significance might offer sufficient insurance against the possibility that a reduction in conflicts will not be observed at most sites.

If, however, the survey is carried out at four sites only, the probability of obtaining a reduction in conflict count at two sites or less (when counting 2 d before and 2 d after) is 12 percent. By counting 3 d in this case, one can reduce the chance of not obtaining a reduction in conflicts at the majority of sites to 6 percent.

Example 5. Failure to Observe Improvement

The expected daily conflict rates (sum of all classes) are 35 before and 30 after. What is the probability that the average after count will not be less than the average before count when counting 1, 2, . . . , 6 days? This is a repetition of example 4 but with daily conflict rates above 20 to illustrate the use of the normal approximation.

The difference between the before and after sample means is approximately normally distributed with a mean that equals 35 minus 30, or 5 conflicts per day, and a variance that equals the sum of VAR $\{\bar{X}\}$ before and after, which is $[(35 + 30)(1.83)]/0.83j$ (see Equation 1 and example 3). The standard normal variable in this case is $5/\{[(65)(1.83)]/0.83j\}^{1/2}$. The probability that the difference between the counts will be negative is listed in the table below of the normal probability distribution.

No. of Survey Days (j)	Standard Normal Variable	Probability of No Reduction in Conflict Count
1	0.42	0.34
2	0.59	0.28
3	0.72	0.24
4	0.84	0.20
5	0.93	0.17
6	1.02	0.15

The value of the standard normal variable for any two expected daily conflict rates is given by the difference between the expected daily conflict rates, or

$$[\text{sum of expected conflict rates } (1 + a)/aj]^{1/2} \qquad (14)$$

where a and j are defined by Equation 1.

Example 6. Critique of Results by Glennon and Others

Glennon and his group (10) use the probability of failure to observe a reduction in the count of conflicts as their criterion for survey duration determination. It is natural at this point, therefore, to discuss their results in the context of the following example.

Determine the duration of conflict survey needed to assure that the probability of obtaining a reduction in conflict count is 0.025, 0.005, . . . , 0.40, given that the expected daily conflict rate before treatment is 50 and that the reduction after treatment is 5, 10, . . . , 25 percent.

The solution, found by using the normal approximation as in example 5, is given in Table 2, where the results obtained by Glennon and others (10) are given in the first column. It is on this basis that they conclude that the required survey duration is not practical. The survey durations, according to our analysis, are approximately five times shorter for the same probability of failure to obtain reduction in counts (see column three). The discrepancy between the two results stems from the difference in the assumed variability of daily counts. While Glennon and others assume a constant variance of 530, our calculations are based on a variance-to-mean ratio of 2.2 as obtained from the data described above.

On the basis of Table 2, it appears that,as long as the difference between the before-and-after expected conflict rates is large, surveys of modest duration guard sufficiently against the probability of not observing a reduction in counts. When the difference between the expected conflict rates is small, even very long surveys do not offer protection against the chance that the after count will be larger than the before count.

With 4 d as a largest practical survey duration, the solid line in Table 2 is the boundary of combinations of expected conflict-rate reductions and probabilities of failure to observe a reduction in conflict counts. If the probability is to be as low as 0.05 (10), a conflict survey seems practical only when the before and after rates differ by more than 25 percent. If a 0.3 probability of not observing a reduction in conflict count is still acceptable, differences as low as 15 percent can be measured. It must be remembered that,if the effectiveness of the treatment is tested at, say, 20 sites, then, with a probability of 0.3 pertaining to each site, the chance of not obtaining a reduction at the majority of sites is less than 5 percent.

In summary, there is nothing sacred about a significance level of 0.05. In many circumstances, lower levels may be regarded as satisfactory. However, small differences between expected daily conflict rates cannot be measured even with very modest significance levels. This limitation is inherent in every estimation based on random variables with large variances.

In spite of this limitation, one needs to retain the proper perspective. At present, safety can be measured by using accident records or conflict counts. Accident records fluctuate no less than conflict counts. If a site has, on the average, 50 conflicts per day, then, for a 10 percent confidence level in a 25 percent reduction in the rates, one needs to count conflicts for 3 d. If the same site has 10 accidents per year, then, for a similar accuracy, accident records for 15 years (before and after treatment) need to be collected. Thus, the very real limitations on the conflict method of measuring safety discussed above are even more severe when accident data are used for the same purpose.

Table 2. Number of days to attain probabilities of failure to observe a reduction in conflicts (expected daily conflict rate before treatment -50).

Reduction in Expected Conflict Rate (%)	Probability of No Reduction in Conflict Count After Treatment						
	Glennon and Others (0.05)	0.025	0.05	0.10	0.20	0.30	0.40
25	26	5	4	3	2	1	1
20	41	8	6	5	3	2	1
15	72	14	12	9	6	4	2
10	162	32	27	21	14	8	4
5	650	132	110	86	56	35	17

This comparison is not quite fair, in that it disregards the question of proportionality between conflicts and accidents. It serves, however, to illustrate the main attraction of measuring safety via conflicts and the accelerated collection of information. The argument for indirect safety measurement, for instance by conflict studies, cannot be based on a claim of great estimation accuracy, which is ordinarily not attainable. It is based on the simple fact that in some circumstances indirect safety measurement is more accurate than any other method at our disposal.

So far we have been concerned about the possibility of not being able to show, through a reduction in the count of conflicts, a real reduction in the expected conflict rate. This may be thought of as the danger of not recommending a treatment that in fact is effective when implemented. The converse, of course, must be also of concern. It is quite possible (in fact it is very likely) to obtain a reduction in the count of conflicts in spite of there being no change in the expected conflict rate. This error is associated with the danger of implementing treatments, because of reduction in conflict counts, that are without effect. Such practice is wasteful of resources that could be spent more effectively elsewhere.

Example 7. Observing Improvement When No Improvement Exists

The expected daily rear-end conflict rate is 16 and remains so after treatment. What is the probability of obtaining a reduction in the average daily conflict count of 2, 4, 6, or more in a 1-d before-and-after survey?

The solution, from Figure 6, shows that the corresponding probabilities are 0.41, 0.30, and 0.20.

Note that even fairly large reductions are not unlikely, in spite of there being no real change in the rate at which conflicts occur. Out of 100 such sites to which an ineffective treatment has been applied, in a 1-d survey, some 30 will show a reduction of 4 or more in the average daily conflict rate. Conversely, if only treatments that reduce the daily conflict count in a 1-d survey by 4 or more are implemented, then 30 percent of all useless treatments under consideration will be implemented.

Also note that if the daily rate of the sum of all conflict classes is involved, Figure 7 and not Figure 6 should be used. As expected, when counting longer, the probability of obtaining a reduction exceeding a specified magnitude diminishes. This is illustrated graphically in Figures 6 and 7.

Example 8. Distribution of the Difference Between Count Averages

For large expected conflict rates, the normal approximation may be used. To illustrate, find the probability of the difference between the average counts (sum of all conflict classes) of 2-d surveys to exceed 10 if the expected daily conflict rate both before and after is 30.

The solution is that the difference is approximately normally distributed with a mean of 0 and variance that equals $[(2)(30)(1.83)]/0.83j$, or 66 (see Equation 1 and examples 3 and 5). The standard normal variate is $10/(66)^{1/2} = 1.23$. From tables of the normal distribution, the probability that the difference will exceed 10 is 0.11.

In general, the standard normal variate is given by

$$\text{Difference between conflict count averages} \div \{[2(\text{expected daily number of conflicts})(1+a)]/(aj)\}^{1/2} \qquad (15)$$

where a and j are as defined in Equation 1.

SUMMARY AND DISCUSSION

Counts of the number of conflicts occurring per day are characterized by considerable variation. After observing available data, it appears that, for homogeneous classes, the variance-to-mean ratio is 1.4; for the sum of several (seven) conflict classes, the variance-to-mean ratio is 2.2. Accordingly, conflict counts do not follow a simple Poisson distribution. It is convenient to assume that the expected conflict rate varies from day to day. The negative binomial model is invoked to account for this variation.

Using this model in conjunction with the aforementioned empirically derived variance-to-mean ratios, the probability distribution has been tabulated, and confidence intervals have been derived and probabilities of types I and II errors computed. Their use is introduced through eight illustrative examples substantiated by graphs and tables.

Interest centers on questions of result accuracy and survey duration. Result accuracy is characterized through confidence limits and probabilities or error in testing hypotheses with respect to treatment effectiveness. It is suggested that, unless coupled with formal decision analysis, the framework of hypothesis testing is given to misinterpretation. Thus, for a judgmental decision to be made on survey design and standards of accuracy, confidence limits may be preferred.

As illustrated, accuracy increases with survey duration. However, the increase in accuracy per additional survey day diminishes rapidly. In general, there is not much to be gained by counting longer than 3 d. Thus,

Figure 6. Probability of reduction in average daily conflict count to equal or exceed r when no improvement exists (homogeneous classes).

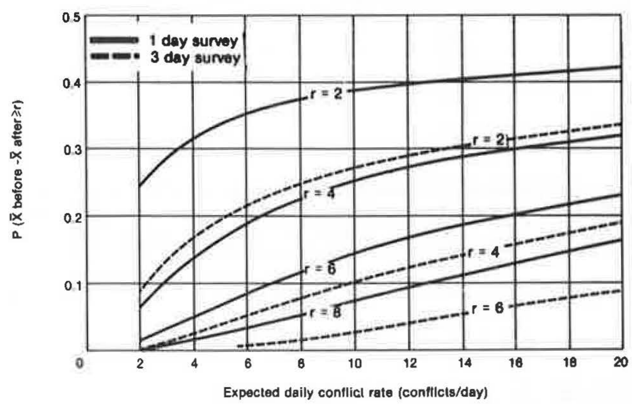
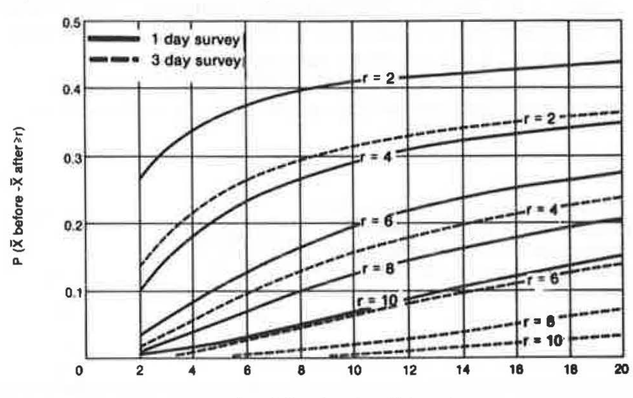


Figure 7. Probability of reduction in average daily conflict count to equal or exceed r when no improvement exists (all classes).



there is a practical limit to the accuracy with which the value of the expected daily conflict rate can be estimated. Existence of this practical limit on estimation accuracy must be considered when investigating treatment effectiveness. Conflict rate differences of 15 percent or less will prove difficult to demonstrate through conflict studies.

ACKNOWLEDGMENTS

This research was supported by a grant from the National Research Council of Canada and by the Transport and Road Research Laboratory.

REFERENCES

1. S. R. Perkins and J. I. Harris. Traffic Conflict Characteristics—Accident Potential at Intersections. HRB, Highway Research Record 225, 1968, pp. 35-43.
2. S. R. Perkins. Traffic Conflicts Technique Procedures Manual. Research Laboratories, General Motors Corp., Warren, MI, GMR-896, Aug. 1969.
3. W. T. Baker. An Evaluation of the Traffic Conflicts Technique. HRB, Highway Research Record 384, 1972, pp. 1-8.
4. B. R. Spicer. A Study of Traffic Conflicts at Six Intersections. Transport and Road Research Laboratory, Crowthorne, England, TRRL Rept. 551, 1973.
5. P. J. Cooper. Effectiveness of Traffic Law Enforcement. Traffic Canada, Road Safety, Ottawa, 1974.
6. C. Hydén. Relation Between Conflicts and Traffic Accidents. Department of Traffic Planning and Engineering, Univ. of Lund, Sweden, mimeo, 1975.
7. V. A. Guttinger and J. H. Kraay. Development of a Conflict Observation Technique. Paper contributed to the Organization for Economic Cooperation and Development, Special Research Group on Pedestrian Safety, Netherlands Institute for Preventive Medicine (TNO), Nov. 1976.
8. G. Malaterra and N. Muhlrad. Intèrêt et limite du concept de conflit. Jan. 1976 (ONSER).
9. F. H. Amundsen. Near Accidents. Development of a Recording Method and an Analysis of the Correlation With Traffic Accidents. Norges Teknisk-Naturvitenskapelige Forskningråd, Oslo, TOI-4413, May 1974.
10. J. C. Glennon, W. D. Glauz, M. C. Sharp, and B. A. Thorson. Critique of the Traffic Conflicts Technique. TRB, Transportation Research Record 630, 1977, pp. 32-38.
11. S. Oppe. Traffic Conflict Analysis, a Road Safety Research Technique. Institute for Road Safety Research (SWOV), Voorburg, The Netherlands, Dec. 1975.
12. P. J. Cooper. Predicting Intersection Accidents. Transport Canada Road Safety, Sept. 1973.
13. P. F. Waller. Challenging the Status Quo in Driver Licensing. Traffic Safety, Vol. 76, No. 10, Oct. 1976, pp. 20-21.
14. R. C. Peck. Towards a Dynamic System of Driver Improvement Program Evaluation. Human Factors, Vol. 18, No. 5, 1976, pp. 493-506.
15. M. Greenwood and G. U. Yule. An Enquiry Into the Nature of Frequency Distributions Representative of Multiple Happenings With Particular Reference to the Occurrence of Multiple Attacks of Disease or Repeated Accidents. Journal of the Statistical Society, Vol. 83, 1920, pp. 255-279.

Discussion

William D. Glauz, Midwest Research Institute, Kansas City, Missouri

A major concern with any data collection effort, and particularly when traffic safety data are involved, is how much to collect. Mr. Hauer has nicely reviewed prior work on this question as it relates to the traffic conflicts technique.

Prior assumptions regarding the statistical properties of traffic conflict counts have ranged from a presumed Poisson's distribution, which is intuitively enticing and the usual hypothesis for accident data, to the worrisome finding of Glennon and myself and others (10) that the variance may be so large as to make the technique impractical.

Mr. Hauer is to be congratulated for his careful study, which indicates that variances are more modest and within the range of practicality.

Probably the major, although not obvious, conclusion that can be reached through critical analysis of Mr. Hauer's and other studies is that the prime factor in determining the usefulness of the technique may be the experience and training of the observers. For example, the data we used were collected by several observers as a small segment of a much broader research project. There was a very large interobserver variance that can be attributed to minimal formal training. However, the Hauer paper itself furnishes a convincing argument for proper training, as I shall now demonstrate.

First of all, he states that the first 10 d of observations were not used in the analysis because of observer error. The total experiment for the Glennon data lasted less than 10 d. Thus, a certain amount of variance was arbitrarily eliminated. Even after deleting these 10 d, however, the effect of observer error is still present. This is clearly illustrated by the calculations in the table below on the effect of training on observer consistency.

No. of Days	Conflicts Mean	Variance	Variance-to-Mean Ratio
4	38.7	87.0	2.25
5	33.2	192.6	5.80
6	32.5	163.0	5.02
7	32.9	140.7	4.28
8	31.9	129.8	4.07
9	33.2	129.8	3.91
10	32.9	117.9	3.58

The data in this table are derived from Hauer's Figure 1, which shows total daily conflict counts at one specific site (after the first 10 d of observations had been discarded). Note that, although the mean count became well stabilized after 5 d or more, the variance continued to decline substantially thereafter. Within 1 week's observation, the variance-to-mean ratio was 5.80; it dropped to 3.58 after 2 weeks.

An even more vivid demonstration is to compare this ratio by using all the data given by the author for this intersection (again, discarding the first 10 d) with that obtained by deleting yet another week. The 19-d ratio is 2.9; it is only 1.8 based on the last 14 d. Clearly, traffic conflict counts made by any but the most thoroughly trained observers are, statistically, a nonstationary process.

This illustration suggests that Mr. Hauer's final results, although very encouraging as a guide to conflict counting needs, are still too pessimistic. It also suggests that the overall average variance-to-mean ratio is not as great as 2.2 but, if examined by the method of the illustration, is on the order of 1.5. In fact, I suspect that an approach to 1.0—the Poisson distribution—is theoretically possible.

He uses his statistical findings to carry out examples and concludes that "there is not much to be gained by counting longer than 3 d." The user must be cautioned that such a conclusion depends not only upon the variance-to-mean ratio (thus, observer training and experience) but also upon the definitions of traffic conflicts used (formal as well as operational interpretations) and upon the expected mean daily count.

Mr. Hauer notes that the expected conflict rate is dependent upon time of day, day of week, and season; I agree. Therefore, the user must be careful in interpreting his or her data in terms of the time frame in which they were collected. They need not be obtained as a daily count; a partial day may be more practical.

In the latter case it is important to note the particular time, if an after count is subsequently to be made. Also, for some locations, evening or weekend counts may be of more interest than weekday counts. Further, the time between before and after counts that are to be used to evaluate improvement effectiveness may well span many months in actual practice, so the possibility of seasonal variations must be considered.

Mr. Hauer states that the negative binomial distribution is appropriate for traffic conflict counts, a fact we had also noted (10). However, it is questionable whether λ, the expected number of conflicts on a day, should itself be a random variable. Moreover, based on our earlier comments, the value of a (or p) is really not yet known (nor is it likely to be a constant). Therefore, the tables and figures presented by Mr. Hauer are illustrative at best and should not be used indiscriminately.

In summary, the state of the art is that traffic conflicts and accidents are undoubtedly related (positively correlated), but the strength of the relationship is uncertain. The important need at this time is to improve and simplify the traffic conflict technique so that it can be applied easily yet uniformly by all interested agencies. In the author's words, we need to optimize the estimation of the "expected conflict rate."

Author's Closure

It is most gratifying to be congratulated for a careful study. Even more gratifying is Mr. Glauz's apparent concurrence with the main conclusions of the study. These squarely contradict their earlier findings (10), which appear to have been based on data collected by observers, as he says, with "minimal formal training" in the course of a "small segment of a much broader research project." Some points raised by Mr. Glauz warrant comment.

THE IMPORTANCE OF OBSERVER TRAINING

Mr. Glauz argues that observer training and experience are of prime importance, that observer reliability improves even after weeks of experience, and that with very experienced observers the variance-to-mean ratio may not be as large as 2.2. All these observations are most likely correct.

However, it is possibly somewhat hasty to draw conclusions about the manner in which the variance-to-mean ratio diminishes with observer training (and about its possible limiting value) on the basis of partial data obtained at a single intersection. (Complete data for all seven intersections are readily available.)

The issue cannot be resolved by using data presented for illustration and requires careful experimental design and analysis. Therefore, for the time being it seems prudent to follow the recommendation made in the paper that "when specific information about count variability is not available, use of the average values obtained in this paper is recommended." Admittedly, this may be conservative.

To keep the issue in proper perspective a glance at Figures 3 and 4 is useful. In Figure 4, the variance-to-mean ratio 2.2 has been used; for Figure 3 the ratio is 1.4. It is easy to see that the size of the corresponding confidence intervals is not all that different. Certainly, the difference is not large enough to drastically alter decisions about survey duration. It seems, therefore, that the results are not too sensitive to the variance-to-mean ratio used.

Finally, the tables available from the Transport and Road Research Laboratory can be requested specifically for the variance-to-mean ratio tables. Thus, corresponding tables will be generated.

SURVEY DURATION

Mr. Glauz takes exception to the conclusion that there is not much to be gained by counting longer than 3 d.

My entire paper is an attempt to provide the user with convenient tools for the statistical design of a conflict survey—equations, graphs, and tables. Also, the paper contains examples and illustrations demonstrating the use of such tools.

The user can, therefore, balance costs against accuracy and need not apply rules of thumb. However, the practitioner needs and appreciates useful generalizations that apply to most situations of interest.

Possibly the most useful generalization that emerges from the paper is that the increase in estimate accuracy after the third survey day is rather small. This is clear from inspection of Figures 3 and 4. The conclusion will not change for any reasonable conflict definition, expected daily conflict count, or variance-to-mean ratio. Therefore, Mr. Glauz's caution to the user is unnecessary.

PARAMETERS OF THE NEGATIVE BINOMIAL DISTRIBUTION

Mr. Glauz expresses some unspecified doubt about whether the expected number of daily conflicts (λ) should itself be a random variable. His basis for this doubt is difficult to understand. He seems to subscribe to the use of the negative binomial model, saying that its appropriateness has already been noted (10), in a note I could not find. The negative binomial model arises precisely when λ is a random variable. He also explicitly states a few lines earlier that λ "is dependent on the time of day, day of week," etc.

In summary, the first part of the discussion by Mr. Glauz raises the issue of observer training, experience, and conflict count variability. It is hoped that future research will shed more light on this problem. The second part of the discussion, unfortunately, is less constructive. It exhorts the user to be cautious, "to be careful in interpreting," to regard results as "illustrative," not to use them "indiscriminately," but without apparent reason.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Analysis of Traffic Conflicts and Collisions

Brian L. Allen and B. Tom Shin, Department of Civil Engineering, McMaster University, Hamilton, Ontario
Peter J. Cooper, Road and Motor Vehicle Safety Branch, Transport Canada, Ottawa

Parameters intrinsic to the sequence of events leading to vehicle collisions and traffic conflicts are investigated in an attempt to develop a more practical and reliable application of the traffic conflicts technique. Sequences of collisions and conflict events were videotaped and are analyzed in detail. Preliminary investigations reveal that using the common method of brake application is not adequate for describing conflict. As a result, seven methods of defining a conflict situation are introduced and evaluated. It is concluded that at least two of the proposed methods will provide a practical investigative tool that explains accident occurrence better than brake application.

Since its introduction by the General Motors Research (GMR) laboratories in 1967 (1), the traffic conflicts technique (TCT) has been employed by many traffic engineers as a method of measuring the degree of hazard at a roadway location. This permits corrective action to be taken and avoids the undesirable practice of waiting for accident information to accumulate.

The GMR procedures defined a traffic conflict in a way that included visible evasive actions taken by drivers and the occurrence of traffic violations. It was suggested that evasive actions be identified by brake lights or lane changes. Traffic violations were automatically recorded as conflicts regardless of the presence of other conflicting vehicles. This original definition of traffic conflicts was apparently well received and adopted by traffic engineers for use in numerous other studies (2, 3, 4, 5).

However, using brake application as the principal descriptor for the TCT procedure is unsatisfactory and has failed to gain global acceptance as a collision-predicting model. Alternative parameters such as time measured to collision (TMTC) and time to accident (TA) have been proposed (6, 7). These methods, too, do not appear to satisfy all the requirements of an acceptable TCT method that is capable of reliably explaining collision occurrence.

This paper briefly presents the results of a 2-year research effort conducted to develop a more reliable application of the TCT procedure (8). Fundamental weaknesses inherent in present TCT methods are identified and discussed. Since the most serious deficiencies involve the capability of properly explaining collision occurrence, a sequence of events leading to a collision is hypothesized and related to seven methods for describing a conflict situation.

The development, analysis, and evaluation are based principally on detailed investigation of actual collision and conflict events recorded by video equipment over an extended period of time. A comprehensive evaluation of the proposed TCT methods is presented, and recommendations for future development are made.

CRITIQUE OF EXISTING TCT METHODS

Using brake indications to measure traffic conflicts has several indisputable advantages. Brake applications can be easily identified and counted; subjectivity in data collection can be avoided; and brakes are applied in all categories of conflict types. This method, however, is often affected by several undesirable characteristics as well.

First, braking habits may vary from driver to driver. Some drivers are very cautious and may apply brakes on entering an intersection regardless of the hazard present, while others may not brake even when presented with a very hazardous situation. Consequently, it is possible that one could falsely identify such situations in terms of conflict.

Second, braking produces only an on-off or binary set of information that does not permit further distinction regarding the severity of a conflict situation. An abrupt brake application to avoid an imminent collision and an unnecessary precautionary application will therefore receive the same rating unless a completely subjective rating scheme is introduced.

Third, deceleration evidence by brake application is not always a wise evasive action. In some conflict cases, as witnessed in an actual collision to be discussed later, acceleration rather than deceleration might be a superior reaction to avoid a collision. Had an acceleration action taken place in such a case, it would not have been identified as a conflict by the present TCT method.

Fourth, the common procedure of observing brake application by only one of the vehicles involved in a conflict situation (by definition, the vehicle with the right-of-way), information describing the actions of the other vehicles involved is lost. Collisions are occasionally precipitated by the party with the right-of-way, such as a through vehicle that speeds toward an opposing left-turning vehicle but applies no brakes. In such cases the driver with the right-of-way may not apply his brakes, and the situation will not be considered as a conflict by the present TCT method.

There are other weaknesses. Brake lights may not be visible because of mechanical failures. The purpose of braking is also sometimes not identifiable. In some cases, application is only for the purpose of obeying traffic signal indications, while in others brakes are applied for purposes of avoiding a collision in a conflict case.

Principally because of the inability to grade the severity of a conflict situation, alternative methods have been proposed. The Transport and Road Research Laboratory (TRRL) in Great Britain introduced a modified procedure in which all conflicts were graded into five classes according to perceived severity of the event (2, 3, 4).

Hayward (6) suggested that the TMTC between two vehicles involved in a hazardous event could be employed as a reasonable scale to judge the severity of near-miss cases. The TMTC value typically varies in a dynamic mode as both drivers react to each other and reaches a minimum at the moment closest to a collision.

Hydén (7) introduced a similar term, TA, which was defined as the time that would have passed from the moment at which a driver reacts and commences a braking or turning action until the moment the collision would have occurred, if both parties had proceeded with approach speed and direction unchanged.

Although these methods can grade the severity of the conflict situation, they have critical limitations. The TRRL method would be influenced excessively by subjective decisions. The TMTC or TA are also theoretically incomplete, since the methods are based on the time measured up to the expected moment of collision. Such measurements become infinite values in the cases where collisions would have been avoided even by a fraction of a second.

For example, if two vehicles are in a situation where a collision could be avoided by only 0.1 s if no evasive action were taken, TMTC or TA would be identified as infinite simply because the expected collision time is interpreted to be nonexistent. Also, these methods require accurate speed information for interacting vehicles, which is usually difficult to obtain.

COLLISION ANALYSIS

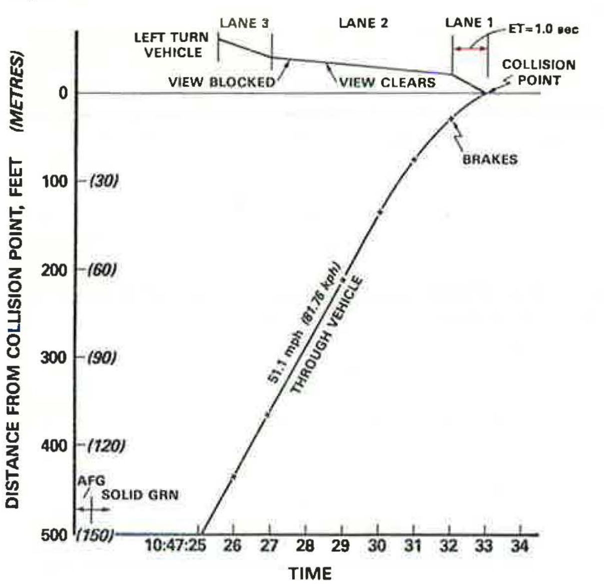
Based on the above reasoning, exploring alternative means of defining traffic conflict situations appeared warranted. This was attempted by examining actual collision and conflict scenes monitored by video equipment. Observation was concentrated on events in which left-turning vehicles conflict with opposing through vehicles. A video camera, installed permanently at a busy intersection in Toronto, recorded traffic movements continuously by time-lapse photography and collected collision scenes. The observation period extended from October 1975 to December 1976. A total of 25 collision scenes, including 9 left-turn collisions, were recorded in this way.

A typical left-turn collision is described below and is depicted in the time-space diagram in Figure 1 and in the simulation in Figure 2. In this event, the left-turning vehicle on the eastbound approach collided with a westbound through vehicle approaching at a speed of 82 km/h (51 mph). The left-turning vehicle had waited for a gap since the beginning of green and approached slowly to the middle lane of the three through lanes (lane 2). The vehicle then moved rapidly into the outer through lane (lane 1), hitting the side of the through vehicle. The through vehicle was pushed aside more than one lane width by the impact. The sight line of the left-turning driver was blocked briefly by a van in the opposite left-turn lane approximately 5 s before the collision. It was also observed that, if the through vehicle had proceeded with its original speed of 82 km/h without any deceleration, the collision might have been avoided.

Other left-turn collisions recorded by the video camera were also analyzed. Of the nine collisions recorded, five occurred during the red signal and one occurred in the third second of an amber clearance interval. One occurred as a result of the second left-turning vehicle attempting a so-called "rabbit jump" at the commencement of the green, and the last one seemed to result from an improper left turn that began from the outer third lane from the center line of the approach.

It became apparent from the above that left-turn collisions often occur (or the process begins) during the clearance interval. Drivers apparently panic and attempt to clear themselves quickly from the intersection. As seen above, six out of nine left-turn collisions occurred in an amber or a red period.

Figure 1. Time-space diagram of the left-turn collision on November 24, 1975.

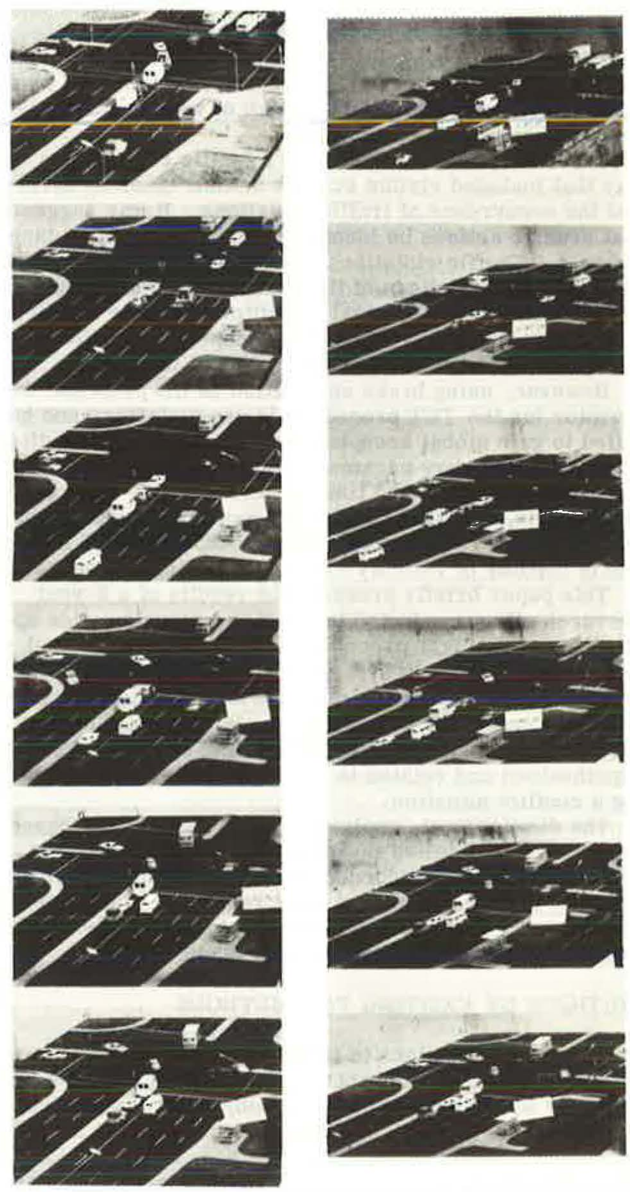


COLLISION GENERATION PROCESS

After the recorded collision scenes were analyzed, conflict situations monitored by the same equipment at the same location were examined in a similar manner.

A traffic conflict has generally been described as a situation in which the driver perceives that evasive ac-

Figure 2. Simulation of the left-turn collision on November 24, 1975.



tion is required to avoid a collision or to secure a safe maneuver. Evasive action may be decelerating or weaving or any other move that the driver considers useful and expedient. In some cases such actions may not be directly or easily observable, and in others collisions may occur without evasive action being taken.

Therefore, the currently accepted concept of traffic conflicts, which is based on observable evasive actions, may fail to include all cases that could lead to a collision. This is obviously a very undesirable situation, since conflicts should comprise a population of events within which collisions reside. Thus definitions that allow for the occurrence of a collision not preceded by a conflict are not desirable. The flow diagram in Figure 3 illustrates this point. It should be noted particularly that the logic depicted in the diagram identifies potential for a collision even when evasive action is not taken.

In an attempt to further qualify and quantify a conflict event, consider the time-space diagram of a left-turning and a through vehicle in Figure 4. The through driver would perceive the potential for a collision, take evasive action usually by reducing speed (the deceleration rate being in accordance with the perceived severity of the potential), succeed in avoiding a collision, and thus terminate the conflict situation. The driver will then attempt to recover his previous driving condition by accelerating (if he previously decelerated as an evasive action).

Figure 3. Collision generation process.

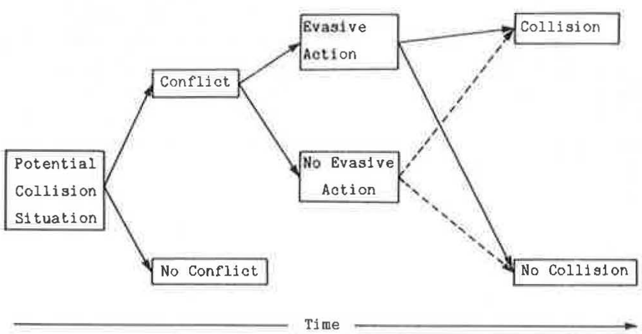


Figure 4. Time-space diagram of a typical left-turn conflict.

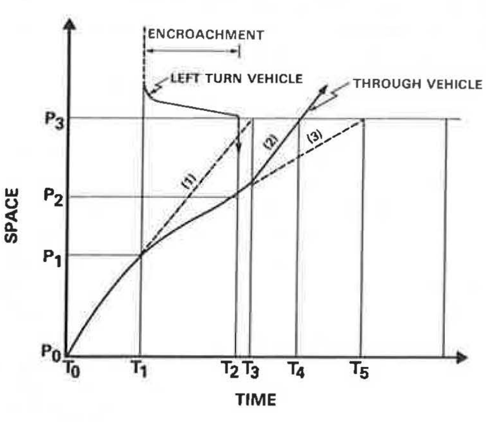
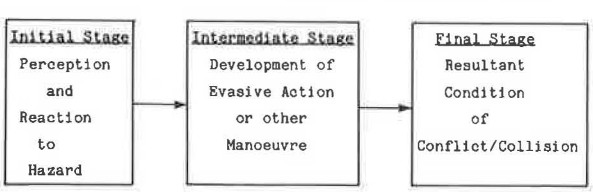


Figure 5. Sequential development of a conflict-collision event.



More specifically, the through vehicle in this example perceives a potential conflict at location P_1 at time T_1 as the left-turning vehicle encroaches on the occupied through lane. If the through vehicle were to maintain the approach speed V, it would arrive at the potential collision point at time T_3 as depicted by the dashed line 1. Assuming that the left-turning vehicle had cleared the through lane by time T_2, the through vehicle would have missed the collision by a gap time (T_3-T_2).

Obviously, if the through vehicle had traced line 1 with an insufficient gap time, a collision would have occurred as in the situation presented in Figure 1. However, in the case shown, the driver of the through vehicle considered such a gap time too short and accordingly decelerated. Upon observing that the left-turning vehicle had cleared his path and was no longer encroaching, or at least confidently predicting the end of encroachment, the driver accelerated to recover his desired speed. This behavior is traced by the solid line 2. If the vehicle had not accelerated after the encroachment, the time-distance trace would be the dashed line 3.

Observing this relatively typical example, one can easily perceive that a conflict situation should be described as a series or sequence of identifiable events, as suggested in Figure 4, rather than as the occurrence of a single event.

In the initial stage of a conflict, a specific situation is perceived by the drivers involved. This generally results from an unexpected action by one party offending the other. Encroachment on a through lane by a left-turning vehicle is a typical offending action against the through vehicle on the same lane. The degree of hazard perceived by the drivers in this initial stage will of course contribute to subsequent developments throughout the conflict situation, but it will not wholly dictate the outcome of the event sequence.

Given the circumstances that evolved during the initial stage, the drivers involved would subsequently react in an intermediate stage by attempting to further correct the remaining hazardous situation. The success of this attempt would depend on their degree of driving skill, the degree of hazard remaining from the initial stage, and the impact of any new circumstances surrounding the event.

The sequence of events that occurred during the initial and intermediate stages will result in a condition in the final stage (Figure 5). If a very hazardous situation had been faced in the first stage and the drivers did not react effectively, then a near-miss or a collision would have occurred as a result.

According to this notion of events sequentially leading to a final conflict situation, one can imagine the usefulness of measurement parameters capable of describing one or more of these stages. For example, brake application is perhaps indicative of events in the intermediate stage but does little to describe events in the final stage.

In order to specify traffic conflict parameters that adequately describe the events in various stages of the development of a collision or conflict, several measurements were proposed and investigated. A discussion of these measurements is presented in the following section.

PROPOSED TCT MEASUREMENTS

When a hazardous situation is perceived in the initial stage of conflict development, the vehicles involved will try to stop or decelerate to avoid a collision. The capability of stopping is a function of the available stopping distance, which is determined by the approach speed, the attained deceleration rate, and the distance remaining to the collision point.

Specifically, the ratio of the distance available for a driver to maneuver to the distance remaining to the projected location of collision should describe the seriousness of the given situation. This ratio is defined as the proportion of stopping distance (PSD) and is given by

$$PSD = RD/MSD \quad (1)$$

where

- PSD = proportion of stopping distance, or the ratio of RD to MSD,
- RD = remaining distance to the potential point of collision,
- MSD = acceptable minimum stopping distance = $V^2/2D$, and
- D = acceptable maximum deceleration rate.

As an example, a PSD value of 0.50 would mean that the driver would have only half the acceptable minimum stopping distance at a chosen maximum deceleration rate, whereas one needs a PSD value of 1.0 or more to stop safely before the expected collision point. The PSD values were measured at time T_1 in Figure 4, the moment when the left-turning vehicle starts to infringe upon the right-of-way of the through vehicle.

A measure of the gap time (GT) as illustrated by the vector T_3-T_2 in Figure 4 could also describe a conflict event in the initial stage of development. (The term "gap time" should be distinguished from the term commonly used in gap acceptance.) Time T_3 denotes the time at which the through vehicle was expected to arrive at the potential point of collision provided the vehicle had maintained the original approach speed and direction. Time T_2 is identified as the time at which encroachment by the left-turning vehicle on the through lane ended. By virtue of this definition, GT could assume a positive or a negative value.

It seems intuitively obvious that GT would adequately describe the potential for a collision. For example, a large positive gap time would indicate that a long time duration exists between the end of encroachment and arrival of the through vehicle at the potential point of collision, and vice versa. Therefore one could assume that the severity of the conflict or the potential for a collision was indicated by the magnitude of the GT value.

Encroachment time (ET) is defined simply as the time during which the left-turning vehicle infringes upon the right-of-way of the through vehicle. In Figure 4 this can be identified as the time T_2-T_1. Presumably, ET would be an accurate reflection of conflict severity only if all through approach speeds were uniform and the position at the commencement of each conflict identical for all vehicles. If such were the case, a relatively long ET would indicate a relatively severe conflict, since this would cause the through driver to decelerate substantially. Although such speed uniformity assumptions are not expected to hold true in reality, ET would describe driver actions in the intermediate stage of the conflict development depicted in Figure 5.

A method identified as the "zonal braking technique" was employed in an attempt to grade conflicts according to severity based on the location where the evasive action (brake application) took place. Assuming that the approach speeds of through vehicles do not vary widely, one would intuitively think that an evasive action taken farther away from the point of potential collision would constitute a less severe conflict situation than one taken closer to the collision point. Using this logic, one could conceive of zones within which a brake application would represent a conflict of specified severity.

The deceleration rate (DR) is an event that occurs during the intermediate stage of a traffic conflict and can be interpreted as indicative of the severity of the situation. Although it may vary from driver to driver, rapid deceleration will generally occur in a severe situation, whereas moderate deceleration normally implies a minor occurrence.

Post encroachment time (PET) for a conflict is identified as the time from the end of encroachment to the time that the through vehicle actually arrives at the potential point of collision (T_4-T_2 in Figure 4). This is an obvious measurement of how nearly a collision has been avoided. PET is also a suitable measurement for identifying the resulting events in the final stage of a traffic conflict. Although it directly describes neither the situation defined in the initial stage nor the action taken by the drivers in the intermediate stage, it does represent the result of the combined effects of the two earlier stages. For example, a PET value approaching zero demonstrates that a collision was avoided by only the very smallest of margins. This could result from a very severe situation perceived in the initial stage, a very poor driving maneuver during the intermediate stage, or a combination of the two.

Although PET does give a reasonable indication of the severity of a conflict, the value could be affected by the common driving habit of accelerating during the termination of a conflict. Therefore, it could become a better measurement by eliminating such early acceleration effects. This could be accomplished if one were to use the initial deceleration rate as a forecast of the initially attempted post encroachment time (IAPE) as identified by T_5-T_2 in Figure 4. Using this notion one can identify IAPE as

$$IAPE = T_5 - T_2 \quad (2)$$

$$T_5 = T_1 + (P_1P_3/V_2) \quad (3)$$

where

- T_1 = time of commencement of encroachment,
- P_1P_3 = distance from the location of the through vehicle at the beginning of encroachment to the potential point of collision, and
- V_2 = average speed of the through vehicle during the period of encroachment.

Each of the preceding conflict measures incorporates a degree of weakness by its very definition. However, it was felt that each had the potential to more adequately explain collision occurrence than the conventional brake application procedure. In particular, one would expect that those measurements that identified events in or near the final stage of the conflict-generation sequences would possess the greatest explanatory power.

To evaluate the suitability of each candidate measure, several hundred conflict events were thoroughly analyzed, measures compared, and relationship to collision history examined. A detailed discussion of that evaluation is contained in the following section.

EVALUATION OF CONFLICT MEASUREMENTS

The significance and applicability of the measures discussed above were examined by selecting a total of 347 left-turn conflict events from 2 weeks of video records. For each event all values of the parameters were obtained. Several statistical analyses were applied to

Table 1. Summary of criteria evaluation.

Criterion	Ranking 1	2	3	4	5	6
Relation to collision history	PET	GT	DR	IAPE	ET	PSD
Relations among other measurements	ET	GT	IAPE	PSD	PET	DR
Consistency over time	GT	PET	DR	IAPE	PSD	ET
Relation to brake application	GT	ET	PET	DR	IAPE	PSD
Ease of measurement	PET	ET	DR	GT	IAPE	PSD
Applicability to other conflict types	DR	PSD	GT	PET	ET	IAPE
Final rating	GT	PET	DR	ET	IAPE	PSD

evaluate the conflict measures according to the following criteria: (a) relation to collision history, (b) relations among candidate conflict measures, (c) consistency among different days' observations, and (d) relation to brake application technique.

In addition, the measures were evaluated in terms of ease of field observation and applicability as a measure of conflict types other than left turns. A summary of the evaluation according to all criteria is given in Table 1.

Collision history for westbound through vehicles and opposing left-turning vehicles at the study site was compiled for the previous 4 years. The compiled collisions and conflict measurements were ranked by time of day (three periods) and by lane position (three lanes) into a nine-cell matrix. Correlation of conflict measurements with collision history was computed by using the corresponding nine elements. The highest correlation coefficient, 0.495, went to PET and the lowest, 0.413, to PSD. Neither the overall correlations among collision and conflict measurements nor the differences among the measurements were sufficiently significant to allow any strong conclusions.

If we assume that the conflict measures introduced in this paper are all reasonable and independent, the correlation between the parameter values should give an indication of their suitability. To investigate this relationship, the same analysis technique used above was applied to the different conflict measures. On an average, ET had the highest correlation coefficient value, 0.53, to all other measurements. The lowest was DR with 0.28. However, taking the results from other analyses into account, we considered that PSD and GT were among the most highly evaluated methods, based on the relation to other measurements.

Since the TCT procedure normally utilizes 1 or 2 days' observations, consistency among different days of data collection is a desirable trait for an ideal measurement. Hourly and daily variations in observations were computed for each measurement, and the variance expressed as a percentage of the mean was computed. GT rendered the lowest variation, while ET produced the poorest result.

The average values of all measurements, classified by whether the brakes were applied or not, were compared. One intuitively thinks that the population with brake application should yield more severe parameter values than that without brake application. All measurement values agreed with this intuitive notion, with GT and PSD being respectively the most and least closely conforming measures.

While all proposed measurements could be most precisely obtained by using sophisticated observation techniques such as video recording or other time-lapse photographic methods, it appears to be completely feasible to measure at least PET, ET, and DR with simpler devices. At the simplest level, values of PET and ET can obviously be obtained by manual observation with a stopwatch. A device composed of a multichannel event recorder and specially adapted dual tape switches has been used in other research projects to obtain DR, among other data. The other measurements are not so easily obtained and would create considerable difficulties in collecting vast amounts of data.

With respect to the applicability to other conflict types, it should be noted that the development of conflict measures in this study was based on the notion of a fixed potential collision point and the time periods from beginning to end of encroachment upon the through lane. Therefore, for a conflict type for which such baseline information is not readily observable, such as a moving rear-end conflict situation, all proposed measurements are not applicable. However, it is easy to conceive of a slightly different definition of the encroachment period for right-turn and crossing conflicts. By doing so, the proposed methods could be readily adopted.

An evaluation summary of the different measurements is shown in Table 1, which lists the measurements in order of perceived performance. When the ranking was not obvious from the more quantitative information available, subjective judgments were applied to establish the ranks. The resulting rating for each measurement was determined by comparing the total numerical values of ranking obtained for each evaluation criterion. No differential weighting to either the ranks or the evaluation criterion was applied in the sum. In this way, the lowest sum of the scores was assumed the best. Although this is obviously a very crude comparative tool, it does give an overall indication of relative merit.

It can be seen that GT and PET are the most highly evaluated measurements, while PSD and IAPE are the lowest. By considering that deceleration results from brake application and by accounting for the relative ease of measurement and applicability to other conflict types, one could say that the conventional brake application technique would be in a slightly higher than average rank.

CONCLUSION

Although substantial efforts were made to objectively identify the precise sequence of events surrounding traffic conflict and collision situations in quantitative terms, many conclusions were based, in part, upon more qualitative assessments. Although this immediately implies that a degree of caution in interpretation be invoked, we feel confident that experience gained from this project has properly guided the statements that follow.

Perhaps the most controversial aspect of this study has been rejection of conventional brake application methods for the TCT. It is apparent that the assessment of alternate conflict measurements presented in the preceding section did little to firmly convince anyone that the "new" techniques should perform more satisfactorily as predictive methods.

The lack of confidence likely to be generated about this point can be attributed in large measure to the absolute size of the available collision record. Since the intersection site chosen was in a city of very large size and the collision history was the most active of all possible locations in that city, one can quickly become convinced that attempts to confidently correlate conflict measures with collision experience will never be successful. It would logically follow, then, that low correlations can always be expected and that brake-light counts are just as acceptable a technique as any other. This view is likely to be particularly appealing, since a ponderous momentum of approximately 10 years of TCT experience with brake application methods is currently in force.

However, one cannot ignore the completely sound conceptual notions presented earlier. Those notions clearly negate all arguments for continued use of conventional TCT brake application counts and will not be repeated here. It is sufficient to summarize by stating that the proposed techniques by definition can explain collision and conflict mechanisms in a more rational manner. In this regard, PET seems to hold the greatest potential as a conceptually sound descriptor while retaining considerable ease of field measurement.

As a result of this and extensive discussions presented by Allen and Shin (8), we concluded: Enumeration of brake applications is not an acceptable traffic conflict measurement technique for the TCT.

This conclusion is particularly pertinent when one considers the definition requirement that all collisions must be preceded by a conflict. A collision is in fact the most severe category of conflict possible. Clearly, brake application does not precede all collisions and is therefore an inadmissible measure.

RECOMMENDATIONS

The major task facing researchers and practitioners in the immediate future is to apply the proposed TCT under a wide variety of environmental conditions at several intersections so that acceptable estimates of the probability of collision occurrence can be derived. It is obvious that the range should include a variety of locations representing a variety of traffic, geometric, and control characteristics for a variety of conflict types.

This implies the need to standardize the measurement technique and method as quickly as possible so that several agencies may undertake the studies and still retain complete data compatibility. As noted earlier, collection of PET measurements is a relatively simple task, and it is anticipated that relatively little effort will be required to collect the vast amounts of data necessary for establishing relationships.

Finally, the applicability of PET measurements to other conflict types should be investigated in more detail. Although earlier suggestions were made in this report regarding transferability, firm procedures must be established and tested.

As a result of experience and insight gained by the study team during the project, we recommend that

1. Post encroachment time (PET) as defined in this report be thoroughly investigated as the principal traffic conflict measure for the TCT (DR is also worthy of investigation if one assumes that suitably portable tapeswitch mechanisms can be used),
2. Modest research efforts be undertaken to develop, test, recommend, and apply PET to conflict types other than left turns, and
3. Relatively major efforts be undertaken to collect large amounts of PET conflict data for a variety of conditions to establish a data base on which derivation of collision probabilities can be confidently estimated.

There is no doubt that several of the statements presented in this report will be subjected to severe criticism. However, we feel that the results of an intensive research effort conducted over the past 2 years have justified those statements. In particular, we feel that the traffic conflicts technique does indeed hold promise as a reliable predictive and evaluative tool for traffic engineers, provided that the appropriate measure of traffic conflicts be used. To this end, we suggest that post encroachment time as defined here is the appropriate measure and is capable of explaining more about collision occurrence than brake application alone can. All that remains is to use this very practical procedure expeditiously and comprehensively.

ACKNOWLEDGMENTS

We express appreciation for assistance given by the many people involved with this project. Support provided by the Toronto Branch of DeLeuw Cather Canada Ltd. was invaluable. In particular, Mr. C. J. Liew (now with the city of Calgary) deserves a large degree of credit for lending his special talents to the many work tasks and difficult development periods during the first year of the project. In addition, the extensive video monitoring of collisions would not have been possible without the dedicated assistance of Mr. J. Kimble.

Mr. Said M. Easa receives our sincere thanks for the many tireless hours he devoted during the often frustrating development work on conflict concepts in late 1975 and early 1976. His assistance was indeed valuable, timely, and very much appreciated.

A special acknowledgment is extended to Mr. E. Hauer of the University of Toronto for his continued support and valuable critiques throughout the project. His insightful assistance was certainly reflected in the direction with which project work proceeded.

This research was conducted under contract with Transport Canada. The opinions, findings, and conclusions expressed in this publication are ours and not necessarily those of Transport Canada.

REFERENCES

1. S. R. Perkins and J. I. Marris. Traffic Conflict Characteristics: Accident Potential at Intersections. General Motors Research Publication, Lansing, MI, GMR-718, Dec. 7, 1967.
2. W. T. Baker. An Evaluation of the Traffic Conflicts Technique. HRB, Highway Research Record 384, 1972, pp. 1-8.
3. R. D. Paddock and D. E. Spence. The Conflicts Technique: An Accident Prediction Method. Ohio Department of Transportation, July 1973.
4. D. E. Pugh and T. J. Halpin. Traffic Conflicts in Washington State. Washington State Department of Highways, Mar. 1974.
5. P. J. Cooper. Predicting Intersection Accidents. Canada Ministry of Transport, Ottawa, Sept. 1973.
6. J. C. Hayward. Near-Miss Determination Through Use of a Scale of Danger. HRB, Highway Research Record 384, 1972, pp. 24-34.
7. C. Hydén. Relation Between Conflicts and Traffic Accidents. Dept. of Traffic Planning and Engineering, Univ. of Lund, Sweden, mimeo, 1975.
8. B. L. Allen and B. T. Shin. Traffic Conflicts Study: A New Look at the Conceptual and Empirical Aspects of the Traffic Conflicts Technique. Department of Civil Engineering, McMaster Univ., Final Rept. 2312-77/1, Feb. 1977.

Discussion

Martin R. Parker, Jr., Virginia Highway and Transportation Research Council

Because of the unreliability of accident records and other problems associated with using accident data to estimate the relative safety of highways, other, more inexpensive, but reliable measures to describe events that may lead to traffic collisions have long been needed for the effective management of highway safety. The authors are to be congratulated for their efforts in deriving seven new descriptors for measuring safety at intersections. Their presentation of the collision generation process (Figure 3) and their sequential development of a conflict-collision event (Figure 4) are especially appealing concepts. However, their conclusions appear not to be warranted by the data presented.

In their concluding remarks, they reject the use of brake applications as an acceptable technique for measuring traffic conflicts. The study data do not adequately support this conclusion for the three reasons given below.

First, one of the authors' reasons for rejecting the brake application technique is based on data and experience gained by observing conflict-collision events on one approach at one Toronto intersection. In fact, their evaluation of the new safety measures, as well as the brake applications, was limited to only the left-turn conflict situation at that approach. The problems associated with using such an extremely limited data base to develop general conclusions are obvious. Results obtained by observing a single condition at any given site may be related to inherent characteristics of the driver-vehicle-environment system that exist at that site. If the authors had used all conflict events—cross traffic, rear end, left-turn, etc.—or the left-turn conflict at several intersections with a variety of traffic and geometrical configurations, the resulting rating (Table 1) of the measures might have led to other conclusions. Because of the limited data base, it is doubtful that any general conclusions are justified.

Second, in describing the collision generation process, the authors point out that the current method of observing evasive maneuvers may fail to include all cases that could lead to a collision. Because of this deficiency, they conclude, the conventional definitions are undesirable. Conceptually, this deficiency may be a good reason for rejecting the brake criterion. However, the authors do not test the hypothesis with data collected during their observations. For example, if their data indicated that some of their conflict-collision observations were not preceded by the application of brakes, this information would have been useful in supporting their conclusion. Without these data, one can only speculate that the deficiency is serious enough to cause rejection of the brake application technique.

Third, evaluation of the data for the study site does not provide sufficient evidence to warrant rejection of the brake application technique. Correlation coefficients computed for the new measures and accidents were low; for example, PET had an r value of 0.495. Other researchers have reported correlation values of similar magnitude between the brake-light counts and accidents (2, 3, 4, 5). Thus, as the authors mention, the new measures are not superior to brake-light indicators as accident predictors. The authors also point out that, considering the ease of measuring and applying the technique to other types of conflicts, brake applications could be interpreted as having a slightly higher than average rank when compared to the new measures.

The authors' recommendation that PET is an appropriate conflict technique capable of explaining more about collision occurrence than brake applications is also questionable. The recommendation is based on the analysis of only the left-turn maneuver and was not applied to include other major conflict types commonly found at intersections. The authors admit that the new measures are not applicable for the moving rear-end conflict situation. This discrepancy may prove to be a major fallacy in the proposed measures, because rear-end conflicts and accidents are, on the average, major events at intersections.

The study is based on the premise that the traffic conflicts technique can be used as a practical and reliable predictor of collisions. The low correlations between the new measures and accidents indicate that either the measures are not closely associated with accidents or the accident-reporting procedures are unreliable. Although 25 collisions were recorded, the authors did not give the number of accidents reported as a result of the collisions. A record of the discrepancy between observed and reported collisions would be useful in estimating the unreliability of accident records.

This discussion is not intended to discourage development of new safety measures or to support continued use of brake applications as the appropriate traffic conflicts descriptor. The authors have correctly identified problems associated with brake applications. However, it should be noted that the new measures also contain deficiencies.

Because highway accidents are often the result of complex events, no single measure can adequately explain the occurrence of collisions. Ultimately, the integration of several measures may provide a better accident descriptor. It is also possible that a more practical and reliable use of traffic conflicts measures would be evaluating highway operating and safety improvements instead of predicting accidents. These hypotheses must be thoroughly investigated before researchers and practitioners can either accept or abandon the traffic conflicts technique.

Authors' Closure

While there is no denying that collisions and the number of brake applications at a location are positively associated, it has been shown by a number of researchers that the correlation is not independent of the influence of traffic volumes. In fact, the degree of dependence is very great. This, coupled with another known phenomenon, that in a significant proportion of accidents no evidence of prior braking on the part of one or both vehicles can be found, leads logically to our decision to explore the chain of events leading up to collisions for clues as to which parameters can be employed for predictive purposes.

As such, this study was never intended to produce a definitive statistical assessment. Rather, the project was to serve as a pilot investigation into identifying suitable traffic conflicts concepts worthy of further, more detailed investigation. We anticipate that such additional work will commence this year. The standard against which success of any new technique will be judged is of course defined by the performance of established methods, such as the counting of brake applications that GMR developed.

We thank Martin Parker for his thoughtful discussion, which virtually amplifies several important issues already identified by the study team and referred to in the paper. It is particularly interesting to note that Mr. Parker has expanded upon precisely those issues of greatest concern to the authors, concerns that will be explictly accounted for in the continuing research effort mentioned previously.

However, despite the stated preliminary nature of the study, we did look in detail (8) at the specific relationships questioned by Parker. For example, less than 30 percent of the conflicts studied exhibited observable brake applications. Although this statistic surely lends itself to the premise that brake application alone is at least conceptually inadequate, it should be noted that the percentage increased as conflict severity increased, indicating a degree of positive correspondence.

Parker also raised an extremely important point concerning the adequacy and availability of police collision reports. In our study, all police reports of collisions occurring during the data collection phase were obtained. Of some 20 police reports, only eight collisions were identified by scanning the video records. This was caused by a combination of equipment failure, night recording conditions, and the fact that our records covered only one approach at the intersection. Perhaps more significantly, the police records did not contain information on 17 (68 percent) of the collisions identified by the video records. The implications of such a weak correspondence for correlation between conflict and collision occurrence are obvious.

Finally, we agree completely with Mr. Parker when he states that important hypotheses on use of the traffic conflicts technique "must be thoroughly investigated before researchers and practitioners can either accept or abandon" the procedure. Emphasis on the words "either" and "or" is extremely important, and we hope that this approach will guide future research into this very interesting and important traffic safety topic.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Abridgment

Evaluating Highway Guide Signing

Fred R. Hanscom, BioTechnology, Inc., Falls Church, Virginia
Wallace G. Berger, U.S. Senate Appropriations Committee

An operational method for the field evaluation of highway guide signs that is readily applicable by traffic engineers or researchers was developed on the basis of valid measures of effectiveness (MOEs) and sensitive, off-the-shelf data-collecting techniques.

The need for this method arises from the diversity of approaches in previous guide sign evaluations and the lack of an established, uniform, valid method for appraising results. A recent NCHRP effort (1) addressed this problem through a literature synthesis, the field development of guide-sign MOEs, and a sensitivity assessment of applicable data-collecting techniques. The product of this effort is the evaluation method reported here.

MOEs OF GUIDE SIGNS

The field development of measures established four types of vehicle behavior at interchanges as guide-sign MOEs. Each is listed and operationally defined in Figure 1, which summarizes the evaluation procedures. An analysis of over 1100 interviews that compares responses of drivers behaving in these ways with those of drivers not behaving in any of these ways revealed group differences linking driver guide-sign responses with each type as follows:

1. Gore weave and high-risk gore weave: (a) greater information-processing difficulty with all guide signs on interchange approach, (b) less certainty of action response to all guide signs on approach, (c) less time available to read and respond to intermediate exit direction guide signs, (d) lower preference rating for intermediate exit direction guide sign, and (e) less likelihood of detecting at least one guide sign.
2. Late lane change: (a) greater information-processing difficulty with at least two guide signs and (b) less certainty of action to be taken to gore-located exit direction guide sign and one advance sign.
3. Driving slowly: (a) greater information-processing difficulty with at least one guide sign and (b) lower preference rating of gore-located exit direction sign.

FIELD STUDY APPROACH

There are currently two approaches for examining the effects of a traffic control device: a study and an experiment. A study is an examination of effects only at the site where the device is installed. Generally, before-and-after observation at the site would form the basis for a judgment regarding the effectiveness of the new guide sign. An experiment, on the other hand, involves simultaneous before-and-after observation at another location (control site) not receiving the treatment. The advantage of the experiment is to permit insight into other changes in traffic behavior that are not caused by the new sign.

Data-collecting methods and guide-sign MOEs suggested here are equally applicable to a study and an experiment. Although the experiment is favored in view of the increased sensitivity of the MOEs to a signing change, a study may apply in situations where control of spurious effects is not considered necessary. The before data-collecting period must closely follow the signing change, and the after period should allow for a minimum adjustment period of 30 d. Before-and-after data-collecting periods must be matched by time of day and day of week. Sound experimental procedure dictates that these periods occur exactly 52 weeks apart. It is important that all data be gathered concurrently at the test and control sites to maintain the experimental integrity of the design.

DATA COLLECTION TECHNIQUES

Statistical reliability of off-the-shelf techniques was obtained by comparing the data with those of the traffic evaluator system (TES), a highly reliable collection method involving electronic road switch sensors. Recommendations for applicable techniques took into account the cost and general suitability of each method for use by a practicing traffic engineer. For the four guide-sign MOEs, the following reliable method factors were found.

1. Gore weaves—manual coding of vehicle weaves occurring in two directions over a gore approximately 183 m (600 ft) long was found to be 98 percent reliable using 30-min coding periods with 10-min rest intervals between each for the duration of a normal working day.

2. High-risk gore weaves—these maneuvers require tracing a vehicle's path within an interchange area; thus, time-lapse photography at an exposure rate of 2 frames/s is recommended.

3. Late lane changes—exit maneuvers before the gore can be obtained by using manual coding with equal accuracy as described above for gore weaves, given that a compatible length of highway section is monitored. Yet, for lane-changing maneuvers before the gore, which require monitoring of longer sections of highway, manual coding was only 88 percent accurate. Therefore, the recommended method for gathering data on these maneuvers is to deploy a time-lapse camera before the interchange and position it so that lane changing occurs in the foreground of the picture.

4. Speed measurements—stopwatch timing of vehicles between two inconspicuous roadway markings was found to be an inexpensive, unobtrusive, and reasonably accurate method of gathering vehicle speeds. A currently available digital-display stopwatch (cost: $100) that displays time increments to the nearest 0.01 s was found to reliably gather speed data with an accuracy of 1.8 km/h (1.1 mph) for any given vehicle. Sample means were obtainable with no statistically significant error using this type of stopwatch.

DATA-SAMPLING PROCEDURE

Field sampling procedures are discussed for each of two classes of data to be gathered.

Gore Weaves and Lane Changes

Figure 1 depicts the preferred orientation of the time-lapse camera. Although manual coding of these measures has been shown to be quite reliable, the advantage of time-lapse photography cannot be overemphasized. This technique permits a permanent, accurate record of traffic volume and weaving maneuvers. As the camera runs unattended, the operator is available to code some of the measures; the remainder can be reduced from the time-lapse film. The coder can record gore weaves and late lane changes while at the site; this leaves only volume and high-risk gore weaves for subsequent film data reduction. For gore areas longer than 305 m (1000 ft), manual coding can be used for gore weaves, and time-lapse photography for high-risk gore weaves and late lane changes. The recommended camera position, in this instance, would be overhead and before the interchange so that lane changes occur in the foreground and gore weaves remain in the field of vision.

Half-hour data collection periods with 5-min rest intervals are suggested, because one 15-m (50-ft) film cartridge will store 30 min of data. Two days are suggested as a minimum before or after study period. Sundays and holidays are recommended to sample unfamiliar motorists.

Driving Slowly

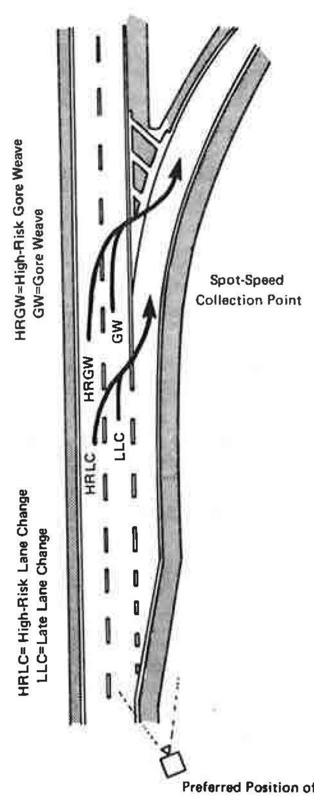
Incidences of vehicles traveling at least one standard deviation below the mean speed are obtained by two alternatively applied manual timing steps. The first procedure is to randomly sample spot speed, while the second involves timing slow vehicles. The former is discussed first.

The suggested method for determining the speed of an individual vehicle is to manually time arrivals of the

Figure 1. Guide-signing MOEs in before-and-after study or experiment.

Guide-Sign MOE:	High-Risk Gore Weave	Gore Weave	Late Lane Change	Drive Slowly
Operational Definition:	A vehicle movement into deceleration lane across painted or physical gore, in addition to crossing at least one through traffic lane.	A vehicle movement into deceleration lane across painted or physical gore.	A vehicle movement into deceleration lane across painted gore extension line.	A vehicle speed ≤ one standard deviation below mean, 240 m in advance of physical gore point.
Collection Method:	Time-lapse photography	Manual coding* or time-lapse photography *Manual coding is preferable if total weave area is 305m or longer	Time-lapse photography	Manual timing via electronic stopwatch
Collection Procedure:	Measure or count all occurrences continuously for half-hour periods simultaneously at experimental and control sites at time of day and day of week to permit matching data in before and after conditions.			Slow and mean speeds during alternate periods.
Analysis Procedure:	Pre-post design with control group: apply two-by-two factoral analysis of frequencies or proportions for each target behavior type and exit volume. Use χ^2-test for frequencies, Z-test for proportions.			
Impact of Deficiency Correction:	Decrease of 35 to 100 percent.	Decrease of 25 to 54 percent.	Decrease of 4 to 19 percent.	Decrease of 6 to 77 percent.

Note: 1 m = 3.28 ft.



vehicle (using a specific point of reference such as the front wheel) between two unobtrusive transverse pavement markings extended across all lanes and spaced 91 m (300 ft) apart. Dark green paint or tape may be used for markings. It is clearly visible to an observer, yet would probably not be detectable to the driver. One very important aspect of this procedure is to obtain a random sample of the total vehicle population. Common observer bias, for instance, too-frequent sampling of large or fast vehicles, in collecting spot data must be avoided. Approximately 60 vehicles can be sampled in a period of 30 min. Using this sample, the mean and standard deviations will be calculated as a baseline against which speeds of slow vehicles are compared.

Slowly traveling vehicles can be timed during alternate 30-min periods. Our manual coding reliability study demonstrated that vehicles traveling one standard deviation below the mean speed can be correctly estimated in 80 percent of the cases. The field procedure suggested here is to time all vehicles appearing to meet the slow-driving criterion; data on those actually traveling faster than one standard deviation below the mean speed can be discarded during the subsequent data reduction.

The measure obtained will be the proportion of exiting traffic volume meeting the slow-speed criterion. Each lane must be separately analyzed for speed variations between lanes. Since trucks, particularly large combinations, are generally driven by professional drivers, a general procedural suggestion for data collection is to observe automobiles and trucks as separate subpopulations.

ANALYSIS OF DATA

Recorded data must permit analyses of vehicle behavior as a proportion of exit volume. Comparisons of before data between test and control sites in an experiment provide a check of site configuration match. Before-and-after differences at the test site provide a gross indication of the impact of guide signing changes. Comparative before-and-after differences and the test versus the control site provide a rigorous indication of signing change impact with the time element effectively factored out and the effects of confounding variables minimized.

For each of these comparisons, it is important to use data that are collected during corresponding time periods. It is suggested that traffic volume differences be first examined for significant differences between the before-and-after condition using the chi-square test. Proper designation of before-and-after data collection periods (1-year interval) will likely result in insignificant volume differences. In this case, one should examine differences in target behavior occurrence, using the chi-square test to make the comparisons cited above. If before-and-after volumes differ, one should convert traffic behavior data to proportions of exiting traffic volume and perform the comparisons using the z-test to determine significant differences. The conversion to proportions should reduce the likelihood of spurious results caused by changes in volume.

A reduction in the frequency of the behavior types designated in Figure 1 should indicate that a measurable benefit was elicited by the signing change. The significance tests described above are the primary means for determining changes in MOE behavior.

ACKNOWLEDGMENT

The work reported in this paper was sponsored by the American Association of State Highway and Transportation Officials, in cooperation with the Federal Highway Administration, and was conducted in the National Cooperative Highway Research Program. Opinions and conclusions expressed in this paper are solely those of the authors.

REFERENCE

1. F. R. Hanscom and W. G. Berger. Motorist Response to Highway Guide Signing. Report prepared by BioTechnology, Inc., Falls Church, VA, Jan. 31, 1976, for NCHRP, Project 3-21.

Publication of this paper sponsored by Committee on Methodology for Evaluating Highway Improvements.

Abridgment

Macroscopic Simulation Models for Use in Traffic Systems Management

Stephen L. Cohen, Office of Research, Federal Highway Administration

In recent years, traffic simulation has become a powerful tool for testing alternate traffic control strategies. The NETSIM (formally UTCS-1) network simulation model (1) was developed for the Federal Highway Administration for this purpose and has found increasingly widespread application.

More recently, the favored approach to urban transportation problems has shifted from traffic control to transportation systems management (TSM). Here, too, simulation should be a powerful tool in testing alternate strategies. These strategies, however, will in general be very different from the pure control strategies developed previously in that they will involve route changes.

Unfortunately, the NETSIM model, which is so successful in testing these strategies, is inappropriate for testing many TSM strategies because of its microscopic vehicle-tracing interactions. This microscopic approach is responsible for the flexibility and accuracy of UTCS-1 but is too expensive in terms of computer time and core to be used on networks larger than 40-50 intersections. Thus, it is evident that a model that does not track individual vehicles is needed to test TSM strategies.

MACROSCOPIC NETWORK MODELS

Three existing macroscopic models, TRANS (2), TRANSYT (3), and SIGOP-II (4), were chosen for comparison. They were executed on a network in Washington, D.C., for which both input and measures of effectiveness (MOE) data were available. The MOE average speed computed by these models was then compared with field data and a NETSIM run of the same network.

TRANS Model

The TRANS model divides each link into zones of length T/S, where T is the time scan length and S is the free-flow speed. Vehicles are moved from the zone currently occupied to the next downstream zone, which is either on the same link (intralink movement) or the next downstream link (interlink movement). Traffic movements may be impeded, and vehicles enter into the queue state when (a) the zone immediately downstream is full, (b) the vehicle is in the downstream zone and faces a red signal indication, (c) the vehicle is a left turner in the downstream zone facing an unacceptable gap in oncoming traffic, and (d) the vehicle is a right turner in the downstream zone facing pedestrian interference.

In the queue state, vehicles discharge in a hit-or-miss Monte Carlo approach based on the mean queue discharge headway input for each link.

Left turns and right turns on red (RTOR) are simulated by using gap-acceptance logic. When a vehicle is discharged from the last zone on a link, the rest of the vehicles in queue are moved up in the next time step to fill the vacancy. Thus, the queue-discharge expansion wave is not modeled, and cases where spillback conditions may be expected to prevail are not properly modeled.

The version of the model used here, TRANS-IV, allows pretimed signals and midblock sinks and sources, but there is no platoon dispersion feature or any internal provision for trucks and buses or midblock rare events.

TRANSYT Model

The TRANSYT model is used as an off-line program to optimize signal settings. The evaluation portion is a simulation model that is more macroscopic than TRANS in that the detailed intersection performance is not modeled. Delay is calculated using an algorithm based on these parameters: volume-to-capacity ratio (V/C), green time, and offset, together with a platoon dispersion algorithm. For links with V/C > 1, queue buildup is computed.

The platoon dispersion algorithm, which is applied to both the primary and secondary flows on a link, is a recurrence relation based on exponential smoothing that has been validated in the field by the Transportation and Road Research Laboratory in Great Britain.

Two types of delay, uniform and random, are calculated. Uniform delay is based on the assumption that the traffic pattern is static from cycle to cycle, while random delay is based on fluctuations from uniformity and is determined by the V/C ratio. This latter term can contribute quite substantially to the total delay for intersections near saturation.

The version of the model, TRANSYT V, used in this exercise simulates pretimed signals and midblock sources but does not treat buses, trucks, turning movements, midblock rare events, midblock sinks, or RTOR.

SIGOP-II Model

The SIGOP-II model, like TRANSYT, is used as an off-line program to optimize signal settings. The evaluation portion is a simulation model similar to that of TRANSYT but with some major differences: (a) SIGOP-II assumes that all platoons are rectangular in shape; (b) a random component of delay is not calculated; (c) continuity of platoon structure beyond the intersections immediately surrounding each intersection is not maintained as rigorously as in TRANSYT; and (d) the case where V > C is not treated. If a flow occurs with V > C, it is truncated. On the other hand, a correction is made to the free-flow speed to account for acceleration and deceleration effects at the link ends. This correction has a substantial effect on simulated average speeds.

The SIGOP-II model internally handles turning movements, trucks, and buses by converting them to equivalent passenger-car units (PCUs). Thus a truck is considered 2.25 PCUs and a right turn as 1.25 PCUs; a left-turn equivalent is determined by using an algorithm developed by Fellinghauser (5). Sinks and sources and pretimed signals are simulated, but not RTOR.

The table below gives a comparative summary of the models.

Element	Model: TRANS	TRANSYT	SIGOP-II
Platoon dispersion	No	Yes	Yes
Turning impedance	Gap acceptance	None	Equivalent PCUs
Data updating	Yes	No	No
Queue discharge	Monte Carlo	None	None
Trucks and buses	No	No	Equivalent PCUs
Midblock events	No	No	No
Pedestrian blockage	Delay right-turn discharge	No	Equivalent PCUs
Computer language	IBM 7090 Assembly	FORTRAN IV	FORTRAN IV
RTOR	Yes	No	No
Free-flow speed correction	No	No	Yes

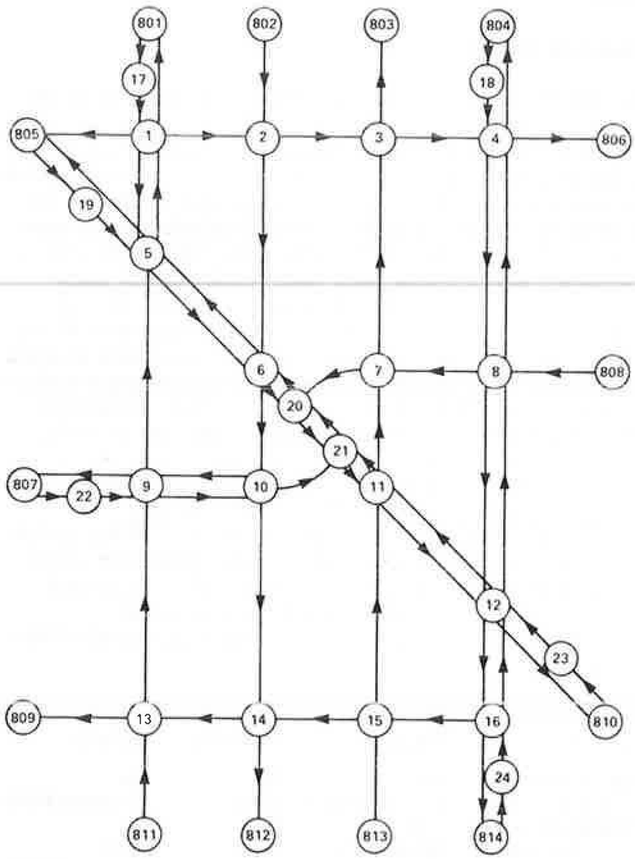
SIMULATION TEST CASE

In order to test the models for accuracy and computer time requirements, they were executed using a data set (1) gathered for the purpose of validating the UTCS-1 model. The data set used consisted of 32 min of morning peak data collected by aerial photography on a 16-intersection network (Figure 1) in downtown Washington, D.C.

This data set was chosen because it includes accurate MOE data that allow a good comparison of model accuracy. Complete information was available on volumes, turning movements, vehicle types, lane blockages, bus movements, signal settings, pedestrian volumes, and midblock sink and source volumes. When the data were reduced during the UTCS-1 validation (1), the 32-min period was split into eight 4-min intervals over which the data were aggregated. However, TRANSYT and SIGOP-II are static models and do not include a data update feature. For this reason, the eight 4-min subintervals were aggregated into one 32-min interval.

The following features that are available in some models but not in others were represented in the latter in order to make the accuracy comparison of the models as independent of these features as possible.

Figure 1. Network in downtown Washington, D.C.



1. Turning movements. SIGOP-II algorithms were used externally in TRANSYT to convert the input volumes to equivalent PCUs.

2. Source and sink volumes. TRANSYT provides for sources only. Sinks were added by inserting an extra link at each node that was the tail of a link with a sink. Sink traffic was diverted onto this exit link.

3. Stop signs. In Figure 1, nodes 9, 10, 20, and 21 are stop signs that only UTCS-1 simulates. These nodes were handled in the other three models by replacing them with sinks and sources.

4. Buses and trucks. The same factor of 2.25 PCUs used in SIGOP-II was externally introduced into the volume inputs for TRANS and TRANSYT.

5. Exclusive turning lanes. These were introduced in TRANSYT by assignment of separate links.

6. Midblock lane blockages. These are treated by UTCS-1 but not by the other three models. No easily implementable way was found to represent this effect in the other three models.

7. Free flow speed. The same link-specific free-flow speeds were chosen for all models. This was because the average running speeds that should be used in TRANS and TRANSYT were not available (the other two models take account of the link end acceleration and deceleration effects).

TEST CASE RESULTS

Using the data set described above, the four models were executed. Three replication runs were made using UTCS-1. UTCS-1, TRANSYT, and SIGOP-II were executed on the U.S. Department of Transportation's IBM-360-65, while the TRANS model was executed on an IBM-7090. Comparative computer times for the 32-min period were about 13 min for UTCS-1, about 6 min for TRANS, about 6 s for TRANSYT, and about 22 s for SIGOP-II.

The four MOEs that most stringently test model operation are travel time, average speed, stops, and total delay. The number of stops was not available in the field data. Travel time is not meaningful unless related to some distance traveled, and total delay is dependent on a free-flow speed. Average speed, on the other hand, relates travel time to distance traveled and is more independent of free-flow speed than total delay. Thus, average speed was chosen as the MOE to be used to compare the models.

The link-specific and overall results for average speed are given in Table 1. Only those network links that appear in all models are tabulated and included in the networkwide results, which indicate that, as expected, UTCS-1 performed the best, followed in order by TRANS, SIGOP-II, and TRANSYT. To determine the link-specific comparative performances, the sum of squares of the differences between each model's predicted link-specific average speed and the field value was calculated. Each term was weighted by the link volume. The results were UTCS-1: 413 027 $(km/h)^2$ (159 471 mph^2); TRANS: 1 397 663 $(km/h)^2$ (539 641 mph^2); TRANSYT: 1 449 471 $(km/h)^2$ (559 644 mph^2); SIGOP-II: 1 022 589 $(km/h)^2$ (394 824 mph^2).

DISCUSSION OF RESULTS

Several conclusions can be drawn or inferred from these results.

The SIGOP-II model performed better than TRANSYT in that the results were in closer agreement with the field data. However, this result is possibly misleading because it is probably the result of the free-flow speed correction in SIGOP-II. This correction is especially important in a network with short block spacing such as

Table 1. Comparison of simulated and field results for average speeds.

Link	Simulation Results (km/h)				Field Results (km/h)
	UTCS-1	TRANS	TRANSYT	SIGOP-II	
(5, 1)	21.50	21.1	21.50	20.6	21.13
(1, 2)	17.36	28.5	21.37	18.8	19.47
(2, 3)	16.48	19.0	18.26	21.6	20.45
(7, 3)	17.15	26.9	26.87	23.0	14.66
(3, 4)	12.45	14.0	14.25	17.0	14.85
(8, 4)	7.72	14.3	16.36	9.8	8.93
(1, 5)	20.23	9.9	20.37	28.0	7.87
(13, 5)	13.50	26.6	17.44	16.7	24.93
(6, 5)	12.47	19.1	14.27	8.0	13.14
(2, 6)	11.24	11.9	12.28	10.2	12.71
(5, 6)	20.58	23.6	21.48	20.1	17.29
(11, 6)	11.02	4.5	9.33	15.3	16.10
(11, 7)	21.59	25.3	23.80	16.7	18.66
(8, 7)	15.80	32.2	20.78	23.6	16.17
(4, 8)	13.77	11.9	18.37	17.5	9.67
(12, 8)	27.11	27.0	33.50	36.5	17.63
(6, 11)	22.53	29.3	39.92	27.8	20.77
(15, 11)	14.34	17.0	16.86	13.4	13.90
(12, 11)	23.78	25.4	26.97	20.8	22.14
(8, 12)	14.79	23.8	16.48	13.4	11.91
(11, 12)	24.73	13.4	28.19	25.9	18.52
(16,12)	6.45	7.7	6.40	6.00	6.85
(14, 13)	10.64	10.9	11.33	9.20	17.23
(6, 14)	28.22	38.0	35.93	31.85	25.60
(15, 14)	14.01	11.6	13.72	16.90	10.36
(16, 15)	18.48	19.1	18.58	34.3	13.26
(12, 16)	28.18	36.7	29.81	24.0	27.80
Network-wide	15.70	17.28	19.52	17.76	15.85

Note: 1 km/h = 0.62 mph.

was chosen for this study. Further, only one link has a V/C ratio value approaching 1, which yielded a substantial random delay contribution. Thus, it is possible that SIGOP-II might not perform better than TRANSYT in a network with more links having V/C ratios near 1 and longer block lengths (or, in fact, if a similar free-flow speed correction factor were applied to TRANSYT).

The results of TRANS and TRANSYT were mixed. TRANS was closer to the field data on a networkwide basis, but in the link-specific sum of squares test, the two models were about even. The reason can be seen from looking at the link-specific results; TRANSYT almost consistently gives a higher value for average speed than is observed in the field, while TRANS often gives a lower value. It is highly probable that the reason why TRANSYT is consistently high is the use of free-flow speed rather than average running speed, which will also be a factor in the TRANS model. In the latter, however, the hit-or-miss Monte Carlo queue-discharge mechanism is equivalent to a negative exponential headway distribution (6). This means that there will be some probability of long headways being generated. These are not observed in field data, unlike UTCS-1 in which the longest headway is 1.8 times the mean headway. This will have the effect of overestimating delay on an intersection approach in which V/C approaches 1.

REFERENCES

1. R. D. Worrall and E. B. Lieberman. Network Flow Simulation for Urban Traffic Control Systems—Phase II. Federal Highway Administration; National Technical Information Service, Springfield, VA, 1973, Vols. 1-5.
2. F. A. Wagner, F. C. Barnes and D. L. Gerlough. Refinement and Testing of Urban Arterial and Network Simulation. U.S. Department of Transportation; National Technical Information Service, Springfield, VA, 1967.
3. D. Robertson. TRANSYT—Users Manual. Federal Highway Administration; National Technical Information Service, Springfield, VA, 1974.
4. E. B. Lieberman and J. L. Woo. SIGOP-II: Program to Calculate Optimal, Cycle-Based Traffic Signal Timing Patterns. Federal Highway Administration; National Technical Information Service, Springfield, VA, 1976.
5. E. Fellinghauser and D. S. Berry. The Effects of Opposing Flow on Left-Turn Reduction Factors at Two-Phase Signalized Intersections. Paper presented at the 53rd Annual Meeting, TRB, 1974.
6. S. M. Ross. Applied Probability Models With Optimization Applications. Holden-Day, San Francisco, CA, 1970.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.

Some Properties of Freeway Density as a Continuous-Time, Stochastic Process

A. V. Gafarian, J. Pahl, and T. L. Ward, Department of Industrial and Systems Engineering, University of Southern California, Los Angeles

Density is an important macroscopic parameter of traffic flow. A number of studies have based estimations of the density on a section of roadway on speed and flow measurements at the section entrance and exit. This paper views density as a continuous-time, stochastic process and considers the characteristics of the process itself. The study relied on freeway traffic data previously obtained by sequential aerial photography. Position data were smoothed and interpolated to construct individual trajectories, which were aggregated to obtain continuous vehicle counts in roadway sections of various lengths. Autocorrelation functions and power spectra were calculated for these records. It was found that, for the traffic flow under consideration, correlation time was proportional to freeway section length. The power in the process was concentrated below a cutoff frequency that was inversely related to section length. The implications these results have for sampling real traffic processes are discussed.

Density was recognized as an important parameter early in the study of traffic flow. For example, Greenshields (1) concluded that time mean speed was a linear function of density in vehicles per kilometer. His density, the ratio of flow to the arithmetic average of the speeds of vehicles passing the measurement point, is now known to be a biased estimate of the number of cars on a given roadway section (2, 3).

A number of studies have considered the problem of basing estimations of density on a section of roadway on speed and flow measurements at the section entrance and exit (4, 5, 6, 7, 8, 9). This study views density as a continuous-time, stochastic process and considers some of the characteristics of that process.

The data for this study were originally obtained by taking sequential aerial photographs of a three-lane section of the westbound Long Island Expressway (10). The selected flow sequence had a mean concentration of 9.3 vehicles/lane-km (15 vehicles/lane-mile). This corresponds to the Highway Capacity Manual (11) level of service B. The four test sections, 91, 305, 558, and 853 m (300, 1000, 1830, and 2800 ft) long, are examined in column 1 of the table below (1 m = 3.3 ft).

The sections were nested and centered on the same point in the roadway.

Section No.	Column No. 1 Δx	2 $\hat{n}(\Delta x)$	3 $var(\hat{n})$	4 (τ_e)	5 (f_c)	6 $\frac{1}{2}(f_c)$
1	91	2.57	2.96	3.01	0.190	2.6
2	305	8.48	13.4	7.99	0.0489	10.2
3	558	15.3	25.6	13.3	0.0315	15.9
4	853	23.4	40.1	18.2	0.0256	19.5

where

Δx = the length of the three-lane test section in meters (column 1);

$\hat{n}(\Delta x) = (1/T) \int_0^T \hat{n}(t,\Delta x)dt$, with $\hat{n}(t,\Delta x)$ as the observed number of vehicles in Δx at time t and T = record length (column 2);

$var(\hat{n}) = (1/T) \int_0^T \{[\hat{n}(t,\Delta x)] - [n(\Delta x)]^2 dt\}$ (column 3);

τ_e = the time in seconds at which the autocorrelation function has decayed to 1/e (column 4);

f_c = 90 percent of the power in the process is distributed over frequencies f, $0 \leq f \leq f_c$, with f and f_c in hertz (column 5); and

$\frac{1}{2} f_c$ = the sampling interval in seconds that yields a Nyquist frequency equal to f_c (column 6).

The original data (10) had previously been reduced to a series of position measurements at 2-s intervals. These measurements were treated as the sum of a systematic part plus a random error. The errors were assumed to be an uncorrelated sequence of random variables with zero mean and known variance.

Figure 1. Number of cars versus time for section 1.

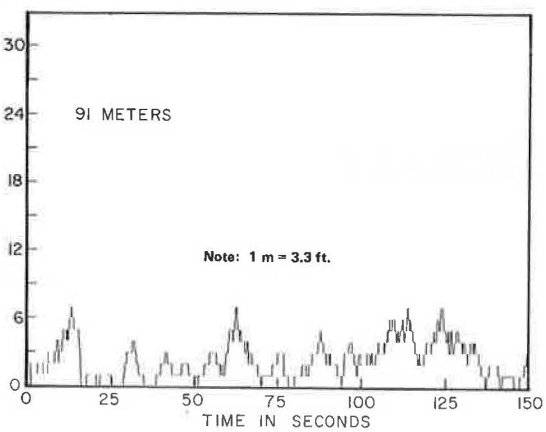
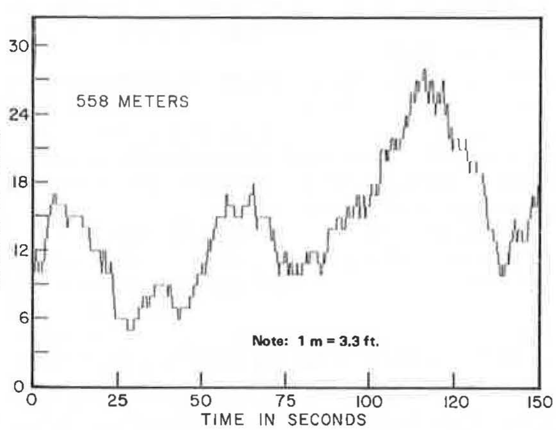


Figure 2. Number of cars versus time for section 3.



This is a good approximation to reality. The positions were smoothed, subject to the constraint that the total error was preserved (12), by minimizing the sum of squares of second differences at the measurement points. The smoothed values were then interpolated using a cubic spline procedure (12, 13, 14). The interpolation provided a point estimate of the times each vehicle entered and left the test section.

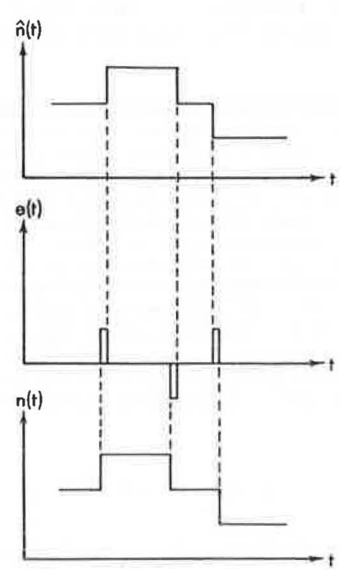
The smoothing and interpolation were evaluated by a simulation study. Trajectories were constructed to provide a known acceleration spectral density function (12). The parameters of the acceleration spectral density were estimated from the experimental data of Torres (15). Exact section entry and exit times were obtained from these trajectories. The simulated trajectories were sampled at 2-s intervals and corrupted by adding noise, which was known to have a zero mean and 0.19-m^2 (2.0-ft^2) variance. The resulting noisy trajectories were smoothed and interpolated as described above to obtain estimates of the entry and exit times. The magnitude of the differences between the exact and estimated times was found to be generally less than 0.05 s.

DENSITY PROCESS

Continuous time records of the number of cars in three test sections were developed. Figures 1 and 2 show examples of the resulting plots for sections 1 and 3. These records are essentially continuous, in that the number of cars was counted every 0.1 s. This was accomplished by first computing the absolute times of section entry and exit for each vehicle and then simply determining which vehicles were in the test section. This gave the desired point.

The error in estimating boundary-crossing times has been determined to be no larger than 0.05 s. Therefore, the location in time of a jump point in Figure 1 may be off by at most ±0.05 s. Hence, the true function—n(t) equals number of cars in the test section—is equal to the estimated function, $\hat{n}(t)$, plus an error function, e(t). That is, n(t) equals $\hat{n}(t)$ plus e(t). The error function would be a superimposition of three pulse trains—$e_i(t)$—where i = 1, 2, 3, one for each lane. Each pulse of each train would be no wider than 0.05 and would have either a positive or a negative sign. Figure 3 shows, as an

Figure 3. Example of error pulse train superimposed on car-counting process.



example, a situation in which no pulses occur simultaneously. The first two jumps in $\hat{n}(t)$ occurred 0.05 s later than they should have, while the third jump occurred 0.05 s earlier than it should have.

The average number of jump points per unit time per lane is equal to 2q/3, where q is the freeway flow (or vehicles per unit time for all three lanes). The factor 2 arises because each vehicle causes a jump on entering and leaving the test section. Thus, on the average, the total time per unit time occupied by these pulses is at most (2q) (0.05).

The flow level in this analysis was of the order of 0.75 vehicle/s. Therefore, the error pulses occupy at most 7.5 percent of the time.

To calculate the variance of e(t), we note that it is a zero mean stochastic process, since

$$E[e(t)] = E[e_1(t)] + E[e_2(t)] + E[e_3(t)] \tag{1}$$

and the errors $e_i(t)$, i = 1, 2, 3, are ±1 with equal likelihood. Therefore, $E[e_i(t)] = 0$, i = 1, 2, 3, and so also E[e(y)]. Furthermore, when i does not equal j,

$$E[e_i(t)e_j(t)] = 0 \tag{2}$$

This follows from the fact that the product $e_i(t)e_j(t)$ is also ±1 with equal probability. Thus,

$$\begin{aligned} \text{var}[e(t)] &= E[e^2(t)] \\ &= E[e_1^2(t)] + E[e_3^2(t)] \end{aligned} \tag{3}$$

Using the probabilities

$$P[e_i(t) = -1] = (1/2)[(1/3)2q\Delta t] \tag{4}$$

$$P[e_i(t) = 0] = 1 - (1/3)2q\Delta t \tag{5}$$

and

$$P[e_i(t) = 1] = 1/2[(1/3)2q\Delta] \tag{6}$$

where Δt is the pulse width of at most 0.05, it follows that

$$\begin{aligned} E[e_i^2(t)] &= (2)(1)(1/2)[(1/3)2q\Delta t] + 0[1 - (1/3)2q\Delta t] \\ &= (2/3)q\Delta t \end{aligned} \tag{7}$$

For Δt = 0.05, then

$$\text{var}[e(t)] = 2q(0.05) = q/10 \tag{8}$$

and q = 0.75 vehicle/s gives var[e(t)] = 0.075. Column 3 of the table above shows the variance of $\hat{n}$, the observed number of vehicles. Clearly, the variance of the error is small relative to the variance of even the smallest test section.

This error has a negligible effect on the characteristics being studied. For example, consider the autocorrelation function of the stochastic process [n(t)]. This is given by

$$\rho_n(\tau) = (E\{[n(t) - \mu][n(t+\tau) - \mu]\})/\{\text{var}[n(t)]\} \tag{9}$$

when

$$E[n(t)] = \mu \tag{10}$$

Replacing n(t) with $[\hat{n}(t)] + [e(t)]$ in Equation 9 and taking expectations gives four terms in the numerator, namely,

1. $E\{[\hat{n}(t) - \mu][n(t+\tau) - \mu]\}$,
2. $E\{[\hat{n}(t) - \mu] e(t+\tau)\}$,
3. $E\{[\hat{n}(t+\tau) - \mu] e(t)\}$, and
4. $E[e(t)\, e(t+\mu)]$.

But $\hat{n}(t)$ and e(t) are uncorrelated processes, so the middle two expectations are zero. Furthermore, we have

$$E[\hat{n}(t)] = E[\bar{n}(t)] = \mu \tag{11}$$

$$\text{var}[\hat{n}(t)] = \text{var}[\bar{n}(t)] + 2\,\text{cov}[\hat{n}(t), e(t)] \tag{12}$$

But

$$\text{cov}[\hat{n}(t), e(t)] = E\{[\hat{n}(t) - \mu]e(t)\} = 0 \tag{13}$$

so that

$$\text{var}[n(t)] = \text{var}[\hat{n}(t)] \tag{14}$$

Hence, Equation 1 reduces to

$$\rho_n(\tau) = \rho_{\hat{n}}(\tau) + (\{\text{var}[e(t)]\}/\{\text{var}[\hat{n}(t)]\})\,\rho_e(\tau) \tag{15}$$

where $\rho_n(\tau)$ and $\rho_{\hat{n}}(\tau)$ are the autocorrelation functions of the $\hat{n}(t)$ and e(t) processes respectively. Now $|\rho_e(\tau)| \le 1$ and

$$\{\text{var}[e(t)]\}/\{\text{var}[\hat{n}(t)]\} << 1 \tag{16}$$

so that

$$\rho_n(\tau) \doteq \rho_{\hat{n}}(\tau) \tag{17}$$

STATISTICAL CHARACTERISTICS OF THE DENSITY PROCESS

The total length of each record used was 720 s, and the sampling interval, as stated earlier, was 0.1 s. With these particular values, it would be possible to distinguish frequency peaks in the spectrum separated by $^1/_{720}$ = 0.0014 Hz, and to estimate frequencies as high as $(^1/_2)/0.1$ = 5 Hz (16).

The autocorrelation and power spectral density estimates follow Blackman and Tukey (16). Program BMD02T of the Biomedical Data Programs (17) was used in these analyses. The autocovariance of the series is first computed by using

$$R(\tau) = [1/(N-p)] \sum_{i=1}^{N-p} (n_i - \bar{n})(n_{i+p} - \bar{n}),\ p = 0, 1, \ldots, m \tag{18}$$

where

τ = pΔt,
mΔt = maximum lag considered = 180 s,
NΔt = total length of the series = 720 s,
Δt = 0.1 s, and

$$\bar{n} = (1/N)\sum_{i=1}^{N} n_i.$$

Section 4 is NΔt = 178 s, and the maximum lag considered is mΔt = 44.5 s, with Δt = 0.1 s. The number of data points and the number of lags exceeded the limitation of BMD02T. Program modifications were made to accommodate the data. The autocorrelation function is given by

$$r(\tau) = R(\tau)/R(0) \tag{19}$$

Figures 4 and 5 show the autocorrelation functions of sections 1 and 4. Figure 6 shows a plot of the correlation time versus the section length. Correlation time is defined as the time lag at which the autocorrelation function decays to the value $1/e$.

An examination of Figures 4 and 5 shows that the autocorrelation functions of the density process can be considered to be initially exponential, followed by an exponentially damped sine curve tail. The correlation time shown in Figure 6 increases almost linearly with section length, which is certainly to be expected. It follows from the fact that, as length increases, a given flow level along with its random fluctuations takes more time to change the number of vehicles in the section from its average.

The estimated one-sided spectrum is calculated in two stages, again using BMD02T (17). First, a truncated, unweighted cosine transform of the data is taken to give a raw estimate of the spectrum:

$$Q(f) = (\Delta/\pi)(r_0 + 2 \sum_{k=1}^{m} r_k \cos 2\pi fk) \tag{20}$$

calculated at $f = j/4m$ cycles/s, where $j = 0, 1, \ldots, 2m$. These estimates are then smoothed by using the weights 0.23, 0.54, and 0.23 to give the Hamming estimates

$$P(f) = 0.23Q[f - (1/4m)] + 0.54Q(f) + 0.23Q[f + (1/4m)] \tag{21}$$

at $f = j/4m$, where $j = 1, 2, \ldots, 4m - 1$. At zero frequency and at the Nyquist frequency

$$P(0) = 0.54Q(0) + 0.46Q(1/4m) \tag{22}$$

$$P(1/2) = 0.54Q(1/2) + 0.46Q[(1/2) - (1/4m)] \tag{23}$$

For a discrete process, the Nyquist frequency $1/(2\Delta t)$ Hz is the highest frequency about which we can get meaningful information from a set of data (18). Plots of the one-sided power spectra are shown in Figures 7 and 8.

An inspection of these figures shows that, as the roadway length increases, the amount of power at the higher frequencies diminishes. A numerical integration was performed to estimate the 90 percent cumulative power

Figure 4. Autocorrelation function for section 1.

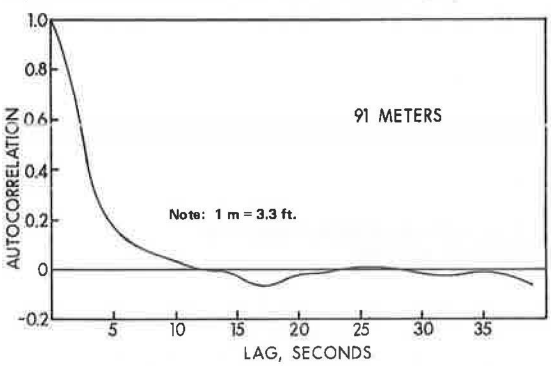


Figure 5. Autocorrelation function for section 4.

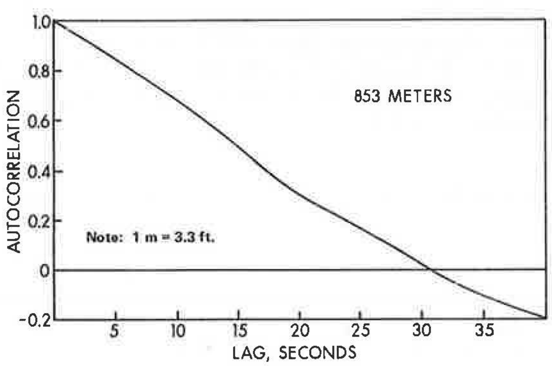


Figure 6. Correlation time versus test-section length.

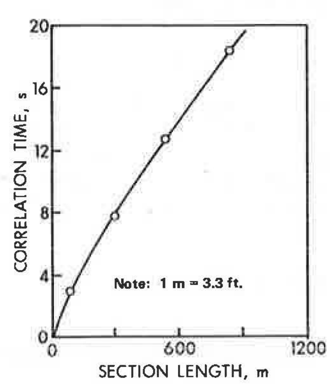


Figure 7. Power spectrum for section 1.

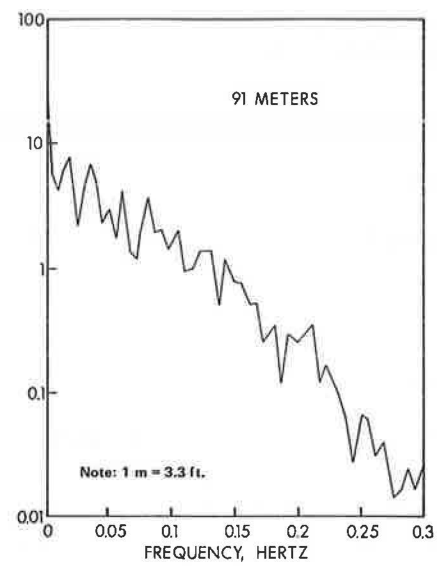
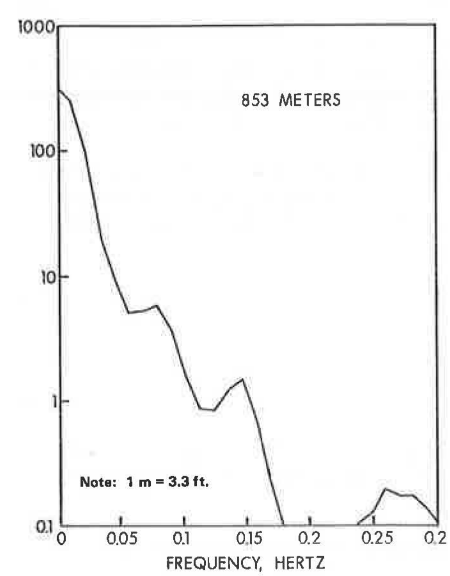


Figure 8. Power spectrum for section 4.



points. These are shown in column 5 of the table given previously.

DISCUSSION

Each of the four series considered in this study is a continuous process sampled every 0.1 s. Thus, the Nyquist frequency is 5.0 Hz. In sampling a continuous time series, an important question is how to choose the sampling interval. It is clear that sampling leads to some loss of information and that this loss gets worse as the sampling interval Δt increases. If Δt is too large, "aliasing" may occur. This is the phenomenon in which variation in the continuous process at frequencies above the Nyquist frequency will be "folded back" and will produce an effect at lower frequencies. If the continuous series contains no variation at frequencies above the Nyquist frequency, then the spectra of the continuous and sampled processes are the same. In this case, no information is lost by sampling.

From a practical point of view, aliasing will cause trouble unless the Δt chosen is so small that the spectrum of the continuous process is essentially zero for frequencies larger than $\frac{1}{2}\Delta t$. If, for a given Δt, the spectrum estimate approaches zero near and above the Nyquist frequency, then the choice of Δt is sufficiently small. Clearly, in each of the cases shown in the table, this obtains for frequencies larger than 5 Hz.

In another study (19), three of the cases reported in our table, namely, the 91-, 305-, and 558-m (300-, 1000-, and 1830-ft) records were sampled at intervals of 2, 5, and 5 s respectively.

The time-series analytical techniques of Box and Jenkins (20) were used to identify the structure of autoregressive models for each of the three. Model parameters were then estimated and statistically tested. It was found that reasonable forecasts of the density could be made for lead times comparable to the correlation times (column 4 of the table).

REFERENCES

1. B. D. Greenshields. A Study of Traffic Capacity. Proc., HRB, Vol. 14, Pt. I, 1935, pp. 448-477.
2. M. J. Lighthill and G. B. Whitman. On Kinematic Waves II. A Theory of Traffic Flow on Long Crowded Roads. Proc., Royal Society of London, Series A, Vol. 229, No. 1178, May 10, 1955, pp. 317-345.
3. L. Brieman. Space-Time Relationships in One-Way Traffic Flow. Transportation Research, Vol. 3, No. 3, Sept. 1969, pp. 365-376.
4. D. C. Gazis and R. S. Foote. Surveillance and Control of Tunnel Traffic by an On-Line Digital Computer. Transportation Science, Vol. 3, No. 3, Aug. 1969, pp. 255-275.
5. D. C. Gazis. On-Line Estimation of Traffic Densities From Time-Series of Flow and Speed Data. Transportation Science, Vol. 5, No. 3, Aug. 1971, pp. 283-301.
6. C. H. Knapp. Traffic Density for Single and Multilane Traffic. Transportation Science, Vol. 7, No. 1, Feb. 1973, pp. 75-84.
7. N. E. Nahi and A. N. Trivedi. Recursive Estimation of Traffic Variables: Section Density and Average Speed. Transportation Science, Vol. 7, No. 3, Aug. 1973, pp. 269-286.
8. D. C. Gazis and M. W. Szeto. Design of Density-Measuring Systems for Roadways. TRB, Transportation Research Record 495, 1974, pp. 44-52.
9. M. F. Chang and D. C. Gazis. Traffic Density Estimation With Consideration of Lane-Changing. Transportation Science, Vol. 9, No. 4, Nov. 1975, pp. 308-320.
10. P. K. Munjal and Y. S. Hsu. Characteristics of Freeway Traffic. Analytical Models of Multi-Lane Traffic Flow, Vol. 1, Federal Highway Administration, contract no. FH-11-7628, 1972.
11. Highway Capacity Manual. HRB, Special Rept. 87, 1965.
12. R. Agahi, A. V. Gafarian, P. Jaffer, L. T. Nguyen, and J. Pahl. Characteristics of Multilane Traffic Flow From Aerial Data. Department of Industrial and Systems Engineering, Univ. of Southern California, Los Angeles, Technical Rept. 76-2, Oct. 1976.
13. J. J. Ahlberg, E. N. Nilson, and J. L. Walsh. The Theory of Splines and Their Applications. Academic Press, New York, 1967.
14. J. J. Schoenberg, ed. Approximation With Special Emphasis on Spline Functions. Academic Press, New York, 1969.
15. J. F. Torres. Acceleration Noise, Power Spectra, and Other Statistics Derived From Instrumented Vehicle Measurements Under Freeway Driving Conditions. System Development Corporation, Santa Monica, CA, TM-3858/025/00, Sept. 30, 1969.
16. R. E. Blackman and J. W. Tukey. The Measurement of Power Spectra. Dover, New York, 1958.
17. W. J. Dixon, ed. BMD Biomedical Computer Programs, Univ. of California Press, Los Angeles, 1970.
18. G. W. Jenkins and D. G. Watt. Spectral Analysis and Its Applications. Holden-Day, San Francisco, 1968.
19. A. V. Gafarian, J. Pahl, and T. L. Ward. Discrete Time Series Models of a Freeway Density Process. Proc., 7th International Symposium on Transportation and Traffic Theory. Institute of Systems Science Research, Kyoto, 1977, pp. 387-411.
20. G. E. P. Box and G. M. Jenkins. Time Series Analysis Forecasting and Control, Holden-Day, San Francisco, rev. ed., 1976.

Publication of this paper sponsored by Committee on Traffic Flow Theory and Characteristics.