TRANSPORTATION RESEARCH

RECORD

No. 1327

Highway Operations, Capacity, and Traffic Control

Visibility, Rail-Highway Grade Crossings, and Highway Improvement Evaluation 1991

A peer-reviewed publication of the Transportation Research Board

TRANSPORTATION RESEARCH BOARD
NATIONAL RESEARCH COUNCIL
WASHINGTON, D.C. 1991

Transportation Research Record 1327
Price: $21.00

Subscriber Category
IVA highway operations, capacity, and traffic control

TRB Publications Staff
Director of Publications: Nancy A. Ackerman
Senior Editor: Naomi C. Kassabian
Associate Editor: Alison G. Tobias
Assistant Editors: Luanne Crayton, Norman Solomon
Graphics Coordinator: Diane L. Snell
Production Coordinator: Karen S. Waugh
Office Manager: Phyllis D. Barber
Production Assistant: Betty L. Hawkins

Printed in the United States of America

Library of Congress Cataloging-in-Publication Data
National Research Council. Transportation Research Board.

Visibility, rail-highway grade crossings, and highway improvement evaluation.
p. cm.—(Transportation research record ; 1327)
Includes index.
ISBN 0-309-05167-3
1. Traffic signs and signals. 2. Railroads—Crossings.
3. Roads—Visibility. I. National Research Council (U.S.).
Transportation Research Board. II. Series.
TE7.H5 no. 1327
[TE228]
388 s—dc20
[625.7'94] 92-7774
CIP

Sponsorship of Transportation Research Record 1327

GROUP 3—OPERATION, SAFETY, AND MAINTENANCE OF TRANSPORTATION FACILITIES
Chairman: H. Douglas Robertson, University of North Carolina–Charlotte

Facilities and Operations Section
Chairman: Lyle Saxton, Federal Highway Administration

Committee on Visibility
Chairman: Antanas Ketvirtis, Fenco Engineers, Inc.
Secretary: Richard L. Vincent, Lighting Research Institute, Inc.
John B. Arens, Peter G. Contos, Charles W. Craig, Eugene Farber, Mark Freedman, Fred Gronberg, S. Allen Heenan, Ronald J. Henson, Gerhart F. King, L. Ellis King, Ken F. Kobetsky, Marc B. Mandler, Richard A. Mather, Herbert A. Odle, Richard Arnold Olsen, Justin J. Rennilson, Duco A. Schreuder, Richard E. Stark, Henry L. Woltman, Robert R. Wylie, Helmut T. Zwahlen

Committee on Railroad-Highway Grade Crossings
Chairman: Charles Raymond Lewis II, West Virginia Department of Transportation
Secretary: William H. Gossard, Jr., National Transportation Safety Board
John E. Baerwald, Carl R. Ball, William D. Berg, Brian L. Bowman, Donald W. Dodson, Denis J. Dwyer, Bruce F. George, Robert J. Herstein, Sonia V. Kamel, Richard A. Mather, William J. Matthews, G. Rex Nichelson, Jr., William T. O'Brien, Paul C. Oakley, Ernie Oliphant, Eugene R. Russell, Sr., Cliff Shoemaker, Thomas D. Simpson, Otto F. Sonefeld, Betty H. Tustin, Earl C. Williams, Jr., Robert C. Winans, Thomas R. Zeinz

Committee on Parking and Terminals
Chairman: Merritt A. Neale, Merritt A. Neale & Associates
Daniel E. Benson, Wayne Berman, Kiran Bhatt, Frederic R. Caponiti, Raymond H. Ellis, James M. Hunnicutt, John Douglas Hunt, Herbert S. Levinson, Gabe Mullan, Stephen G. Petersen, James B. Saag, Andrew P. Schlaefi, William J. Scully, Steven A. Smith, Alice Snow-Robinson, James A. Thompson, W. G. Van Gelder, Robert A. Weant, Karen P. Wilson, Marie E. Witmer, Christopher C. Wright, William Young

Committee on Methodology for Evaluating Highway Improvements
Chairman: Forrest M. Council, University of North Carolina
Secretary: Charles Philip Brinkman, United States Department of Transportation
Karen K. Ajluni, William T. Baker, William D. Berg, Brian L. Bowman, Jaisung Choi, C. W. Colson, Jr., William D. Glauz, Alfred-Shalom Hakkert, Jerome W. Hall, Fred R. Hanscom, Ezra Hauer, David L. Helman, Julia L. Higle, Warren E. Hughes, Ruediger Lamm, Sidney J. Louick, King K. Mak, Thomas L. Maleck, Ernst Meyer, Olga Pendleton, Bhagwant N. Persaud, Jerry G. Pigman, Ramey O. Rogness, Harold T. Thompson, James M. Witkowski

Richard A. Cunard, Transportation Research Board staff

Sponsorship is indicated by a footnote at the end of each paper. The organizational units, officers, and members are as of December 31, 1990.

Transportation Research Record 1327

Contents

Foreword

The papers in this Record concern illumination issues, rail-highway grade crossings, parking, and methodologies for highway improvements.

Readers with a specific interest in illumination and visibility issues can begin with Zwahlen and Yu, who determine the distances at which the color and shape of reflectorized targets can be identified under low-beam headlight illumination. Schreuder relates the visibility conditions as determined in laboratory experiments with actual visibility conditions for long tunnels, and Adrian and Topalova compare recommendations for the luminance transition in tunnels.

For those interested in rail-highway grade crossings, Ryan develops a methodology for priority ranking improvements to rail-highway grade crossings, and Eck and Kang develop computer software to simulate the movement of trucks over hump-like grade crossings and determine when low-clearance vehicles will not be able to clear a crossing.

Methodologies for evaluating highway improvements is the subject of the next group of papers. McCoy et al. evaluate the safety effects of converting on-street parking from parallel to angle. Al-Bakri et al. relate accidents at traffic circles to geometric and flow characteristics, and Persaud introduces a methodology for estimating the accident potential of road sections. The results of improved signing at traffic circles to reduce traffic conflicts is examined by Sadegh et al., and the effectiveness of traffic system management measures in relieving traffic congestion after the collapse of Oakland, California's Cypress Viaduct is evaluated by Meyer and Rajappan. Lee and Smith report on the capability of traffic signal interconnection and timing optimization to reduce fuel consumption, and Kikuchi and Chakroborty determine the traffic volume warrants for the justification of adding a left-turn lane at unsignalized T-intersections.

Color and Shape Recognition of Reflectorized Targets Under Automobile Low-Beam Illumination at Night

HELMUT T. ZWAHLEN AND JING YU

Two independent studies were conducted to determine the distances at which the color and the outside shape of reflectorized targets were recognized at night under automobile low-beam illumination. The targets were flat plates presented in one of three outside shapes and covered with retroreflective sheeting using one of six colors. All flat plates were prepared in such a way that they had the same area and close to the same specific intensity per unit area. In Study 1, the color recognition distances and the shape recognition distances were determined. In Study 2, the only color recognition distances investigated were for square targets that had the same area and specific intensity per unit area as used in Study 1. Summary measures for recognition distances and confusion matrices were obtained to examine the effects of color and shape. The results of Study 1 indicate that the recognition distances for highly saturated colors are about twice the recognition distances for outside shapes. The statistical analysis results of both studies show that the recognition distances for highly saturated colors are significantly different and that some colors have longer recognition distances than others. In addition, on the basis of a statistical analysis, the results of Study 1 indicate that the recognition distances for the shapes are not significantly different. It may be concluded that the highly saturated colors are superior stimuli when earliest possible recognition of a reflectorized target under automobile low-beam illumination at night is important.

Color and outside shape have always been regarded as important stimulus dimensions in considering the proper recognition of reflectorized targets such as retroreflective traffic signs during daytime and nighttime. As stated in a FHWA manual (*1*): "standardized colors and shapes are specified so that the several classes of traffic signs can be promptly recognized. Simplicity and uniformity in design, position and application are important." In addition, the manual stipulates design criteria for signs using retroreflective materials and mandates that reflectorized signs show the same shape and color by both day and night.

Research to determine the detection distance of retroreflective targets at night has been carried out by Zwahlen (*2–4*). Research to determine the nighttime recognition of white reflectorized warning plates as a function of outside shape and target brightness, as well as a function of full area reflectorization and borders-only reflectorization for different target brightness levels, has been conducted by Zwahlen et al. (*5,6*). The results of the study by Zwahlen et al. (*6*) indicate that increasing target brightness has either no effect or only a small detrimental effect on the correct outside-shape recognition distances for the full reflectorization and the borders-only reflectorization. Targets with only the borders reflectorized were recognized from farther away than targets that were fully reflectorized. As Kantowitz and Sorkin discuss (*7*), detection requires only a go/no go decision—a stimulus is either present or absent. In contrast, recognition requires that the subject identifies not only whether a signal is present or absent, but, if it is present, what signal occurred. Studies investigating how well outside shape, symbols, and color can communicate abstract concepts or messages have been reported in literature by authors including Jones (*8*) and Saenz and Riche (*9*). For example, Jones studied the symbolic representation of two abstract concepts used in road signs—type of message and prohibition—and found that the use of two coding variables, color and outside shape, was unnecessary in most cases; the shape of the sign alone proved to be enough to convey these concepts. Jones stated,

> The effect of removing the usual colour cues from the road sign did not affect their capacity to communicate the two abstract concepts of message type and prohibition to any great extent. The shape of the sign alone, in the case of message containing orders and warnings, enables them to be clearly differentiated. Only in the case of messages intended to convey information did removal of the usual colour cue (blue) lead to a significant loss of interpretability. To distinguish information type messages from others, therefore, some additional cue (although not necessarily colour) appears to be necessary.

In these types of studies, the researchers were primarily interested in how well an abstract concept or message can be communicated by the outside shape, symbol, or color against a selected background condition. The researchers used experimental conditions in which the stimulus dimensions—such as the color and outside shape—were well within the recognition capabilities of the subjects (suprathreshold levels).

The present studies did not investigate how well the meaning of a particular shape, color, or shape-color combination was communicated; they were designed to assess whether one of the two stimuli (color or outside shape) can be recognized from farther away than the other and thus be superior in terms of recognition distance. The overall objective was to determine the distances at which the color of a reflectorized flat plate can be recognized. Another major objective of the first study was to determine the distances at which the outside shape of a reflectorized flat plate of a specific color can be recognized.

Department of Industrial and Systems Engineering, Ohio University, Athens, Ohio 45701.

METHOD

Subjects

In Study 1, a group of seven subjects was used. The average age of the subjects was 21 years, and the visual acuity for distance ranged from 20/17 to 20/22. In Study 2, a group of six subjects was used. The average age of the subjects was 23.3 years, and the visual acuity for distance ranged from 20/17 to 20/25. No major color deficiencies were found in any of the subjects. Visual acuity and color capability were tested by using the Bausch & Lomb Vision Tester. All subjects showed normal contrast sensitivity over the examined spatial frequency range as determined by the Vistech Contrast Sensitivity Test.

Experimental Design, Site, and Apparatus

In Study 1, the independent variables were color and outside shape. The colors were red, orange, yellow, green, blue, and white (six levels), and the outside shapes were a circle, a square, and a diamond (three levels; see Figure 1). The dependent variables were the distance at which the subject made a color recognition decision and the correctness of that decision and the distance at which the subject made an outside-shape recognition decision and the correctness of that decision. The randomized block design was used in the experiment. There were two kinds of blocks: color recognition blocks and shape recognition blocks. Each block had 18 conditions: 6 colors × 3 shapes. Within each block each color-shape recognition condition appeared in a unique random order exactly once. For each subject there were two color recognition blocks and two shape recognition blocks, and the number of observations was 72: 6 colors × 3 shapes × 2 recognition types × 2 replications. The order of the color and shape recognition blocks was randomized and approximately balanced for the seven subjects.

In Study 2, the independent variable was the color of square targets, which had the same six levels as were used in Study 1. The dependent variable was the distance at which the subject made a color recognition decision and the correctness of that decision. The targets used in Study 2 were the same six square targets as used in Study 1. The randomized block design was also used in Study 2. The six colors were randomized within each block, and each color appeared in a unique random order exactly once. For each subject there were 10 blocks and 60 observations.

All reflectorized color targets used in the two studies had the same area (36 in.^{2}) and close to the same specific intensity per unit area (28.2 cd/fc/ft^2 at 0.2 degrees observation angle and −4 degrees entrance angle). All the targets used 3M high-intensity encapsulated lens retroreflective sheeting material, which provides highly saturated colors (desirable) under nighttime automobile beam illumination and also meets the daytime requirements for colors used on traffic signs as specified by FHWA (*1*).

The daytime 45/0.2-degree OBS Y,x,y tristimulus values for ILL D_{65} were 3.35, 0.6645, 0.3169 for red; 18.03, 0.5498, 0.4027 for orange; 16.39, 0.5332, 0.4626 for yellow; 6.46, 0.1340, 0.4503 for green; 3.24, 0.1413, 0.1391 for blue; and 31.41, 0.3106, 0.3311 for white. The nighttime 0.33-degree OBS, Beta 1 = −5 degrees, Beta 2 = 0 degrees, Rot Ang. = 0 degrees, for CIE 1931 2-degree std. obs., std. Illuminant A, the X, Y, Z tristimulus values were 106.09, 50.33, 0.32 (chromaticity, CIE 1976, $u' = 0.4923$, $v' = 0.5255$) for red; 140.99, 93.41, 1.91 ($u' = 0.3643$, $v' = 0.5431$) for orange; 201.75, 160.69, 1.06 ($u' = 0.3086$, $v' = 0.5530$) for yellow; 11.91, 43.33, 16.78 ($u' = 0.0669$, $v' = 0.5476$) for green; 11.43, 22.77, 38.57 ($u' = 0.0975$, $v' = 0.4372$) for blue; and 253.81, 237.77, 62.66 ($u' = 0.2533$, $v' = 0.5339$) for white. The same specific intensity per unit area for the different color targets was obtained by using the blue target, which had the lowest specific intensity per unit area as the standard (e.g., 28.2 cd/fc/ft^2). All other color targets were then covered with thin, black, equally spaced self-adhesive stripes in such a way that the overall specific intensity per unit area of a color target was close to 28.2 cd/fc/ft^2. Figure 2 illustrates how the square targets were covered and provides dimensions and photometric values. The grating pattern was made fine enough so that it produced a spatial frequency (greater than 50 cycles/degree) that was beyond the recognition capability of the human visual system for any contrast value for the recognition distance range of interest.

Both studies were conducted using the same unused concrete airport runway (75 ft wide and 1,500 ft long) at the outskirts of Athens, Ohio. Figure 3 illustrates the setup of the site. A two-lane state highway with moderate traffic runs parallel to the runway about 200 ft away. A number of luminaires, a few illuminated advertising signs, and other light sources were within the subjects' field of view, mainly in the left half of the visual field. The dark background surrounding the target had a luminance value range from 0.02 to 0.05 fL. In Study 1, a 1987 Pontiac Grand Am with properly aimed 4652 low beams was used. In Study 2, a 1979 Chrysler New Yorker with properly aimed 4652 low beams was used. During the experiment, the car engines were kept idling.

Experimental Procedure

In both studies the stationary car was positioned so that the center of the front of the car was exactly above the centerline of the runway and the longitudinal centerline of the car formed a 3-degree angle to the left of the runway centerline. A darkly clothed experimenter rode a dark bicycle toward the car at approximately 10 mph, 6.25 ft to the right of the runway centerline (from the subject's point of view) (see Figure 3). The reflectorized target was attached to the front of the bicycle in such a way that the target surface was vertical and the center of the target was 25 in. above the ground. The subject would fixate on the target approaching the car with

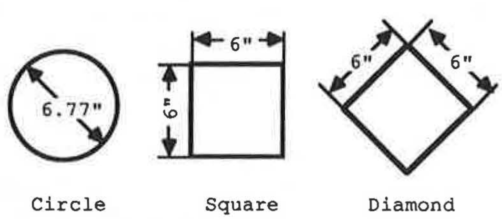
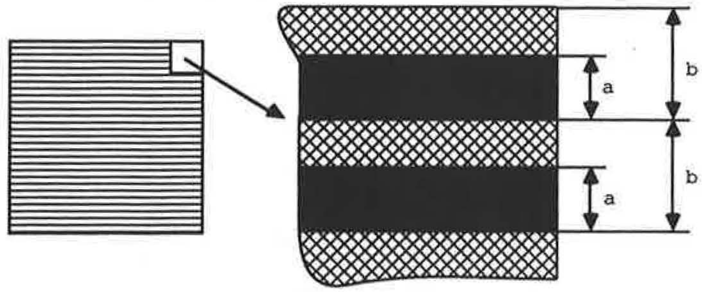


FIGURE 1 Dimensions of reflective targets.



Color	Avg. SIA (cd/fc/sq.ft)	Width (in.)		Cycles/deg. at 200'dist.
		a	b	
White	346.6	0.5	0.544	77
Yellow	256.2	0.5	0.562	75
Orange	118.5	0.25	0.328	128
Red	72.8	0.125	0.204	205
Green	54.5	0.125	0.259	162

Targets were covered by black self-adhesive stripes to obtain an equivalent overall average SIA of 28.2 cd/fc/sq.ft as was measured for the blue color

FIGURE 2 Target covering patterns, dimensions, and photometric values (SIA = specific intensity per unit area).

the low beams on. When the subject recognized the color or the outside shape of the approaching target, he or she would turn on the high beams of the car temporarily. The experimenter riding the bicycle would then drop a small sandbag. The distance from the front of the car to the sandbag was recorded as the recognition distance. The subject would also call out the color or shape of the target to an experimenter in the car, and the subject's response would be recorded. The bicycle rider would return to the starting position with the sandbag and be outfitted with another target for the next run as soon as the distance was measured.

RESULTS

Figure 4 shows the recognition distances for the six colors and the three outside shapes obtained from Study 1. From Figure 4 it can be observed that for all the color and shape conditions, the average color recognition distances consistently are considerably longer than the average outside-shape recognition distances. An analysis of variance (ANOVA) test using a .05 significance level showed that the color recognition distances were significantly different (longer) than the outside-shape recognition distances. This appears to be true for the standard deviations also. The overall average color recognition distance is 719.1 ft, but the overall average outside-shape recognition distance is only 356.4 ft (color recognition distance 2.02 times longer than outside-shape recognition distance, or 3.4 min of visual angle versus 6.8 min of visual angle for a 8.5 in. target dimension).

Figure 5 shows an overall comparison between color and outside-shape recognition distances obtained in Study 1. The average color recognition distance appears to be about twice the average shape recognition distance. The overall average standard deviation of the color recognition distances (215.8 ft) was 1.84 times larger than the overall average of standard deviation of the shape recognition distances (117.2 ft). The ANOVA tests (using a .05 significance level) further indicated that in Study 1 there was a significant difference in the color recognition distances. A Newman-Kuels test (using a .05 significance level) showed that the color recognition distances for red are significantly different when compared with the color recognition distances for greens; this was the only statistically significant difference among all color pair comparisons. The ANOVA for the shape recognition distances obtained in Study 1 showed that neither the color nor the outside shape produced statistically significant differences in the outside-shape recognition distances. The confusion matrices for the color recognition and for the outside-shape recognition tasks obtained in Study 1 are given in Tables 1 and 2. The confusion matrices and the values for the information transmitted show that the percentage of stimulus information transmitted in the color recognition task (73.3 percent) is higher than that for the shape recognition task (58.4 percent).

Figure 6 shows the recognition distances for the correct color recognition decisions obtained in Study 2 (square target) along with the averaged (all three outside shapes combined) color recognition distances obtained in Study 1. An ANOVA test using a .05 significance level showed that the recognition distances for different colors were also significantly different

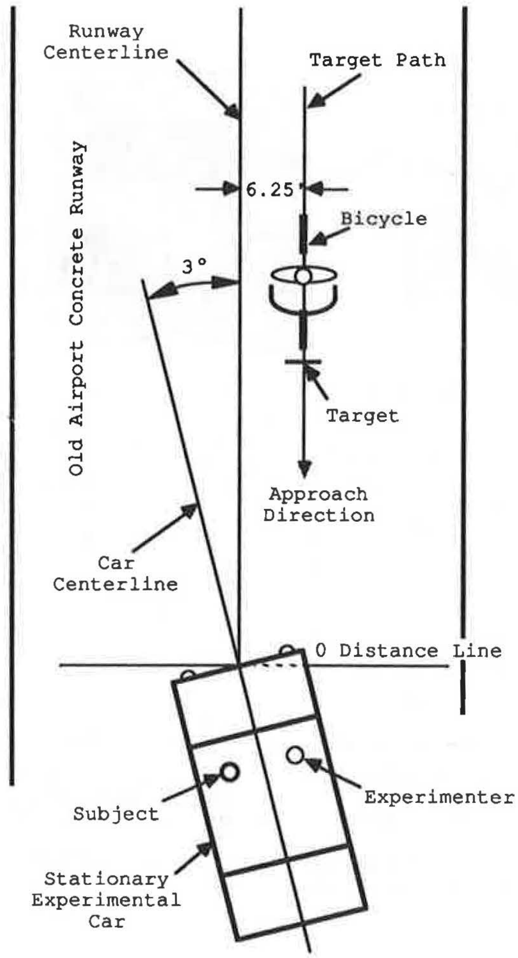


FIGURE 3 Experimental site.

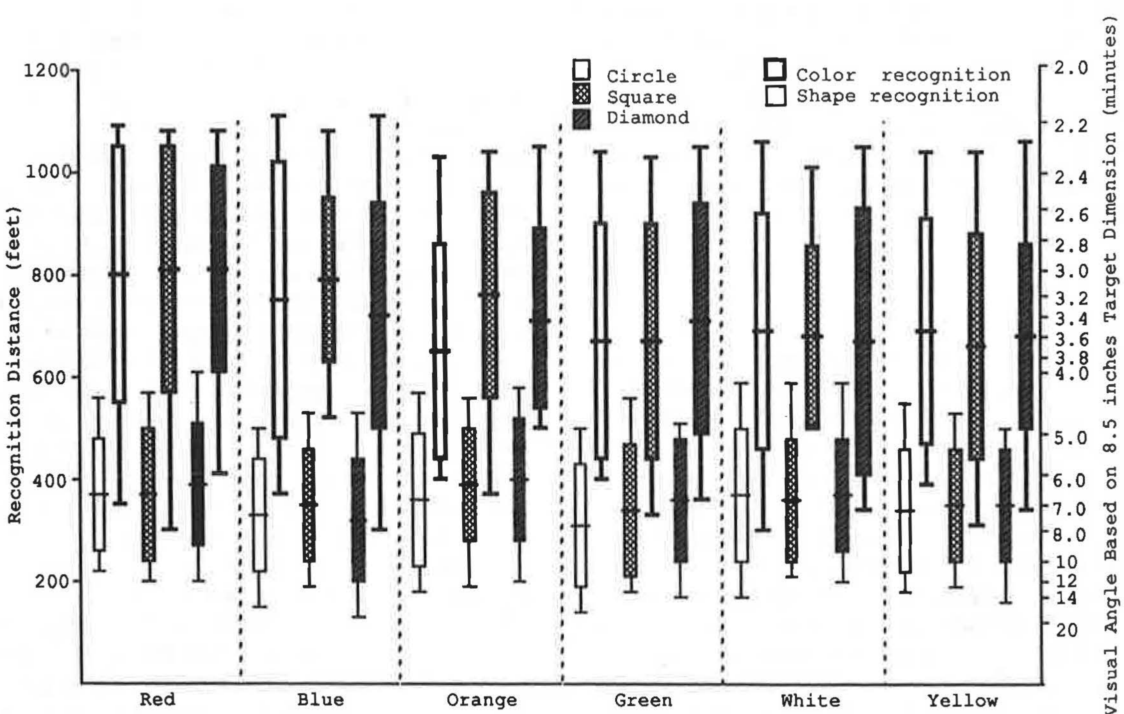


FIGURE 4 Averages, standard deviations, and maximum and minimum values of color and shape recognition distances, Study 1 (n = 14).

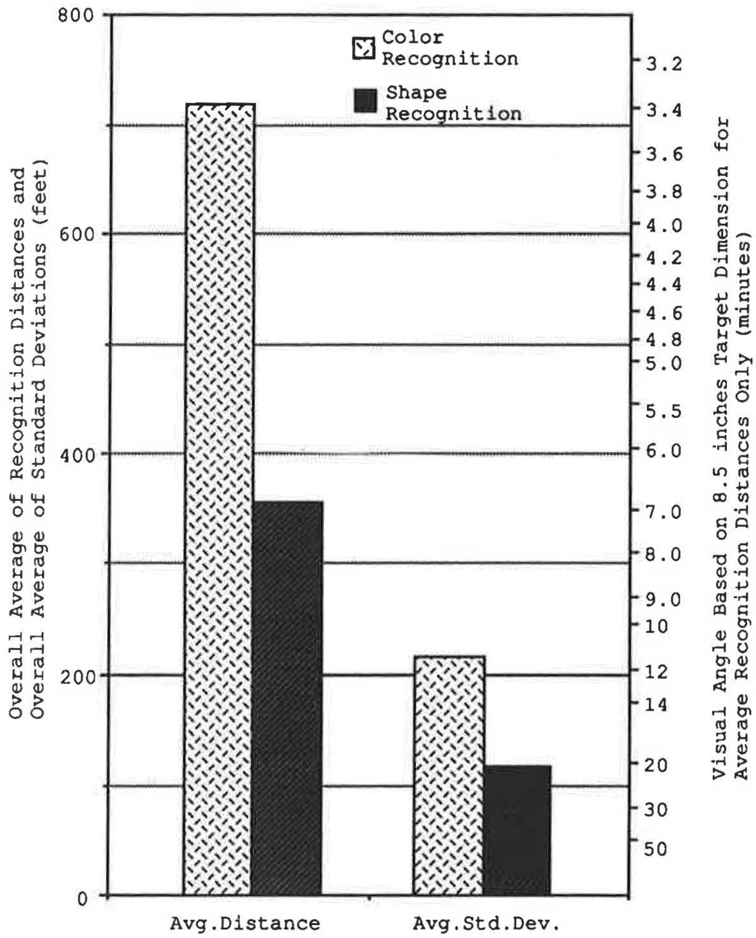


FIGURE 5 Comparison between color and shape recognition distances (overall average and overall standard deviation), Study 1.

in Study 2. The overall average recognition distance (using distances associated with correct recognition decisions only) was 659.5 ft, which is slightly shorter (8.3 percent) than the average color recognition distance obtained in Study 1 (719.2 ft). The difference between the averages obtained in the two studies could be partly because two different groups of subjects and two different cars were used, or because of slight differences in car heading angles and eye-headlamp dimensions, or because the average color recognition distances in Study 1 included some recognition distances for which incorrect recognition decisions were made. The results of the Newman-Kuels test for Study 2 are shown in Table 3 and indicate that the recognition distances for red and orange are significantly different (longer) than those for all the other colors. A confusion matrix based on the results of Study 2 for correct and incorrect responses is given in Table 4. The percentage of the color stimulus information transmitted in Study 2 is 84.5 percent.

TABLE 1 CONFUSION MATRIX FOR SHAPE RECOGNITION, STUDY 1

	Responses			
Stimuli	Circle	Square	Diamond	Σ
Circle	74	3	7	82
Square	6	73	5	82
Diamond	9	2	73	82
Σ	89	78	85	252

H(S) = 1.585 bits *H(R) = 1.583 bits*

H(S,R) = 2.242 bits *T(S:R) = 0.926 bits*

Percentage of stimulus information transmitted = ***58.4%***

DISCUSSION OF RESULTS AND CONCLUSIONS

The results suggest that, for subjects with normal color vision capabilities, the highly saturated colors used in this study are better stimuli than outside shapes when earliest possible recognition of a reflectorized target under automobile low-beam illumination at night is important (near threshold conditions). Furthermore, the results show that one cannot automatically count on simultaneous color and shape stimulus redundancy, especially when long recognition distances are involved (near

TABLE 2 CONFUSION MATRIX FOR COLOR RECOGNITION, STUDY 1

	Responses						
Stimuli	Red	White	Orange	Blue	Green	Yellow	Σ
Red	39		3				42
White		39	1			2	42
Orange	2		36			4	42
Blue				34	8		42
Green				9	33		42
Yellow		8	4			30	42
Σ	41	47	44	43	41	36	252

H(S) = 2.585 bits *H(R) = 2.58 bits*

H(S,R) = 3.27 bits *T(S:R) = 1.896 bits*

Percentage of stimulus information transmitted = 73.3%.

TABLE 3 RESULTS OF NEWMAN-KUELS TEST, STUDY 2 (0.05 SIGNIFICANCE LEVEL)

	Blue	Green	Yellow	White	Orange
Red	S	S	S	S	S
Orange	S	S	S	S	
White	NS	NS	NS		
Yellow	NS	NS			
Green	NS				

S = Significantly different.

NS = Not Significantly different.

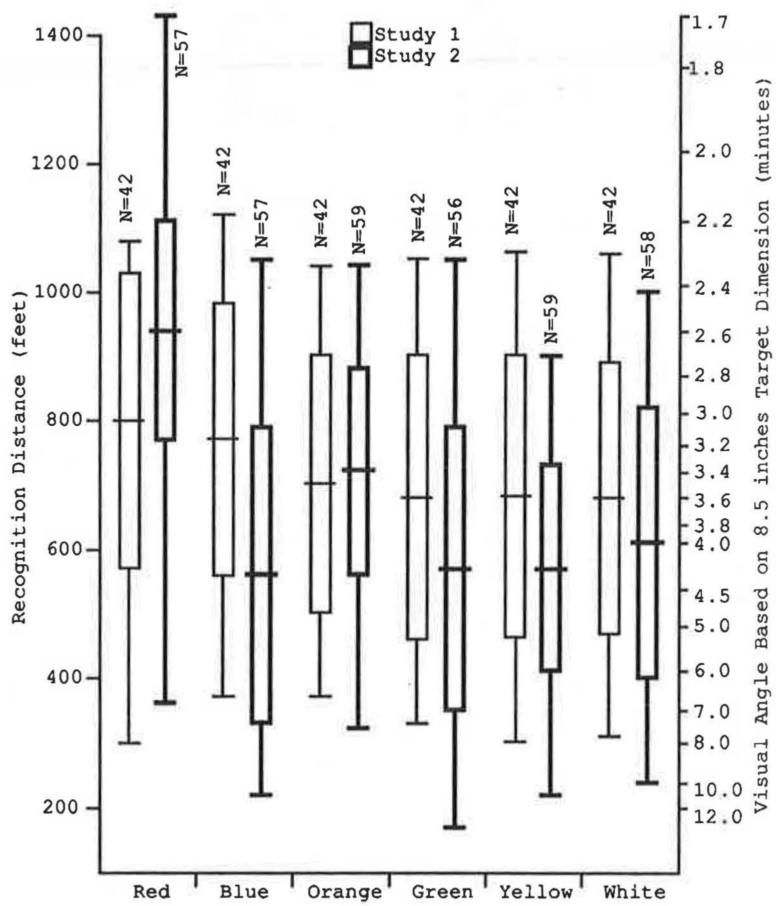


FIGURE 6 Averages, standard deviations, and maximum and minimum values of correct color recognition distances, Study 2, and averages, standard deviations, and maximum and minimum values for averaged color recognition distances, Study 1.

TABLE 4 CONFUSION MATRIX FOR COLOR RECOGNITION, STUDY 2

	Responses						
Stimuli	Red	Blue	Yellow	White	Orange	Green	Σ
Red	58						58
Blue		53				7	60
Yellow			55	4	1		60
White			3	57			60
Orange	2				58		60
Green		17				45	62
Σ	60	70	58	61	59	52	360

H(S) = 2.585 bits *H(R) = 2.58 bits*

H(S,R) = 2.979 bits *T(S:R) = 2.185 bits*

Percentage of stimulus information transmitted = ***84.5%.***

threshold conditions). A highly saturated color, in addition to the outside-shape stimulus of a retroreflective target, appears to increase the average recognition distance by a factor of 2. To maximize color recognition and minimize confusion for individuals with normal color vision capabilities, a highly saturated red color of the retroreflective target is recommended.

REFERENCES

1. *Manual on Uniform Traffic Control Devices for Street and Highways.* FHWA, U. S. Department of Transportation, 1988, pp. 2A-4–2A-7.
2. H. T. Zwahlen. Nighttime Detection of Bicycles. *Transportation Research Circular 229.* TRB, National Research Council, Washington, D.C., 1981, pp. 38–49.
3. H. T. Zwahlen. Detection of Reflectorized License Plates. *Transportation Research Record 1072,* TRB, National Research Council, Washington, D.C., 1985, pp. 63–71.
4. H. T. Zwahlen. Peripheral Detection of Reflectorized License Plates. *Proc., 30th Annual Meeting of the Human Factors Society,* 1986, pp. 408–412.
5. H. T. Zwahlen, D. J. Gardner, C. C. Adams, and M. E. Miller. Nighttime Recognition of Reflectorized Warning Plates as a Function of Shape and Target Brightness. *Proc., 32nd Annual Meeting of the Human Factors Society,* Vol. 2, 1988, pp. 971–975.
6. H. T. Zwahlen, J. Yu, S. Xiong, Q. Li, and J. W. Rice. Nighttime Shape Recognition of Reflectorized Warning Plates as a Function of Full Reflectorization, Border Only Reflectorization and Target Brightness. *Proc., 33rd Annual Meeting of the Human Factors Society,* Vol. 2, 1989, pp. 970–974.
7. B. H. Kantowitz and R. D. Sorkin. *Human Factors: Understanding People-System Relationships.* John Wiley & Sons, New York, N.Y., 1983, pp. 90–91.
8. S. Jones. Symbolic Representation of Abstract Concepts. *Ergonomics,* Vol. 21, No. 4, 1978, pp. 573–577.
9. N. E. Saenz and C. V. Riche, Jr. Shape and Color as Dimensions of a Visual Redundant Code. *Human Factors,* Vol. 16, No. 3, 1974, pp. 308–313.

Publication of this paper sponsored by Committee on Visibility.

Practical Determination of Tunnel Entrance Lighting Needs

D. A. SCHREUDER

The determination of the field factor that relates the visibility conditions in the laboratory to the visibility conditions for a car driver approaching the entrance of a long tunnel is discussed. Experimental results are given from research in the Netherlands. Suggestions are given for incorporation of these results in design methods for the lighting of tunnels.

During the day, the visual systems of car drivers and cyclists are adapted to the very bright daylight. When they enter a tunnel, their visual systems must adapt to the low luminance in the tunnel interior. The adaptation is usually disturbed by two factors: (a) the bright surroundings of the tunnel entrance restricts the adaptation, and (b) the adaptation to a relatively low luminance level may take considerable time.

The entrance at day is the major visual problem of tunnels for motor traffic. When the tunnel entrance (the threshold zone) is not adequately lit, the entrance is a "black hole," where no details can be seen (*1,2*). Usually the most crucial part of tunnel lighting recommendations is the entrance lighting during the day (*3–6*).

The black-hole effect results from several characteristics of the human visual system. First, it takes time—often a considerable amount—for the visual system to adapt from one level of brightness to another. Second, the perception in a dark part of the field of view is hindered when the dark area is surrounded by bright areas. These bright parts act as "glare sources," causing a light veil over the field of view. The effect of this veil can be expressed in terms of its luminance.

When considering the daytime entrance lighting, one must take into account one of the peculiarities of the visual system. When the visual system is adapted in a steady state to luminance values between 30 and 3000 cd/m^2, adaptation to another value in this range takes only a short time. Usually, it appears instantaneous. This may not always be the case, as was suggested also by Bourdy et al. (*7*,p.35;*8*). When, however, the steady-state adaptation level is higher than 3000 cd/m^2, it is well established that the adaptation may take considerable time; for high values (8000 cd/m^2 or more), it may take up to half a minute. This peculiarity leads to two distinct theoretical frameworks and to two distinct systems of tunnel lighting.

The first theoretical framework is the steady-state theory. The principle is that the steady state of the visual adaptation that builds up on the open road when approaching the tunnel is the determining factor for assessing visual problems at tunnels, as well as for solving them. This framework was developed in the 1960s by Schreuder and Narisada, more or less independent of each other; the discrepancies that appeared to exist proved to be no more than differences in the parameters. The steady-state theory was the basis for the original International Commission on Illumination (CIE) recommendations (*5*) and for many national codes in different countries.

The second theoretical framework is the stray-light theory. The basic idea is that near a tunnel entrance the visual system adapts immediately and that the only restriction in visual observation is the veil that extends over the field of view. The veil is thought to result from the light that is scattered in the visual system (eye lens and eye fluids). The idea was described in the late 1930s, and it was developed into a tunnel lighting system by Adrian. The revised CIE tunnel lighting document (*6*) is based on this principle, as are several national codes. The system is based on one of the different glare formulas currently in use. Adrian uses the Stiles-Holaday formula of the 1930s, and so do the French recommendations (*9*). Measurements are usually based on the almost equally old Fry formula, whereas the new recommendations of the Netherlands are based on the recent Vos formula (*4*). It should be stressed that as a source of stray light, the atmosphere and the car windshield often are more important than the ocular media. These aspects have been studied in considerable detail in France (*10,11*) and in the Netherlands (*12,13*).

The steady-state and stray-light theories are often considered to conflict; they are, however, conjoint and they will be described the following chapters.

STEADY-STATE THEORY

When the visual system of a car driver who approaches a tunnel is adapted in a steady-state mode to a very high level of luminance (L_1, e.g., 8000 cd/m^2 or more, corresponding to full summer sun on cement concrete or to sun on snow), for many seconds the adaptation is almost unchanged when the driver enters the tunnel. To ensure that the driver can look into the tunnel while still outside (to avoid the black-hole effect), the luminance in the threshold zone (L_2) must be high as well. Experiments made by Schreuder (*1*) and reconfirmed by Narisada (*14,15*) indicate that L_1/L_2 should be lower than 10 in high-speed tunnels and lower than 15 in other important tunnels. These values are taken as the basis for the original CIE recommendations (*5*).

There appears to be a conflict between the results of the experiments of Schreuder and Narisada. A precise analysis shows, however, that the differences in the results—and, even

SWOV Institute for Road Safety Research, P. O. Box 170, 2260 AD Leidschendam, The Netherlands.

more so, the differences between the recommendations based on these experiments (the CIE and the Japanese recommendations)—are mainly differences in the selection of the parameters. As Schreuder has shown (*16*), the actual research results are almost identical when they are normalized for the time of observation, the preadaptation time, and the size and contrast of the object. A difference in the parameters relates to the conditions for which the research results are used. CIE focuses on tunnels in a flat, open country, where the adaptation to the dark entrance can begin at only a very short distance in front of the tunnel; the Japanese studies refer primarily to tunnels in mountainous areas, where the adaptation may begin at a much longer distance.

STRAY-LIGHT THEORY

The second theoretical framework is the stray-light theory. The influence of stray light on perception is a well-established fact. The first to point out the importance of stray light for tunnel lighting was Adrian (*17,18*). The lighting design system based on it is a.o. described in the new CIE guide (*6*).

When the luminance in the field of view of a driver approaching a tunnel is between about 30 and 3000 cd/m², the visual system adapts very rapidly—almost instantaneously—to other luminances within that range. When the driver is close enough to the tunnel portal so that he or she can fix the entrance opening (at a distance of 50 to 100 m), the visual system adapts to the luminance in the tunnel entrance: the threshold zone luminance (L_2). The value of L_2 should be selected in such a way that the appropriate observations can be made, taking into account the fact that the driver has a driving task to fulfill and that the time for observation of objects is limited.

L_2 can be assessed when the threshold of visibility is known and when the field factors that allow for the influence of driving and of the restricted observation time are known as well.

If one would install a lighting scheme with a value of L_2 assessed in this way, the visibility in the tunnel entrance would be unacceptably low. The most important factor has not yet been considered: the stray light that originates from the surroundings outside the tunnel. That light is scattered, and it forms a veil over the complete field of view. The veil increases all luminance values with the same amount (the equivalent veiling luminance, L_{seq}). All contrasts between objects and backgrounds decrease; consequently, objects are more difficult to see.

The veil consists of three important parts:

- The light scattered in the eye (the entopic stray light),
- The light scattered in the atmosphere, and
- The light scattered in the windshield of the vehicle.

All three parts are highly variable: the entopic stray light depends heavily on the angle between the source of the scattered light and the line of sight, the conditions of the eyes of the observer, and on the age of the observer. The atmospheric stray light depends heavily on the transmission of the atmosphere, on the meteorological visibility, and on the type of particles floating in the atmosphere (the aerosol). The windshield scatter depends heavily on the maintenance condition of the vehicle, on the windshield itself, and on the windshield wipers and washers—and of course on the driver's willingness to use them. In all three cases a variation of a factor of 10 can easily be found under circumstances that are otherwise perfectly normal.

In bright weather the luminous veil over the field of view poses the heaviest requirements on the lighting of tunnels. Such a veil reduces all contrasts in the field of view, including the objects that may be in the tunnel, "behind" the veil. The only way to ensure adequate visibility is to increase the threshold zone luminance (L_2) in the tunnel entrance. This counteracts the contrast reduction caused by the veil.

The veil and its causes, assessment, and composition are described in detail elsewhere (*19*). A brief summary follows.

Contrast is defined as

$$C = \frac{L_2 - L_3}{L_2}$$

Here, C is the intrinsic contrast, and L_2 and L_3 are the luminances of the background and of the object to be perceived.

When a veil with a luminance L_d is present, all luminances are increased by L_d. The "visible" (proximal) contrast C' becomes

$$C' = \frac{(L_2 + L_d) - (L_3 + L_d)}{L_2 + L_d} = \frac{L_2 - L_3}{L_2 + L_d}$$

so

$$C' = \left(\frac{L_2}{L_2 + L_d}\right)C$$

Because $L_d > 0$, $C' < C$. In real traffic, an object can be seen only when C' is larger than the practical threshold of visibility. This practical threshold is larger than the "real" threshold (C''), such as is measured in the laboratory. The relationship is usually described by means of a field factor (f). The object is visible when $C' > fC''$. The minimal value of L_2 can be found as follows:

$$L_2 = \frac{L_d f C''}{C - fC''}$$

This formula will be called the basic formula. Obviously, L_2 can be established when the different values in this formula are known. Practice indicates that particularly the field factor (f) is important to know.

C is the intrinsic contrast of the objects that must be visible. C is chosen arbitrarily, usually 0.2 or 0.3. Such contrasts may represent such traffic obstacles as stones, boxes, and exhausts. Retroreflective devices and vehicle marker lights show much larger contrasts (apart from the negative sign).

Schreuder and Oud (*19*) show that the veiling luminance is composed from several parts:

$$L_d = L_{adef} + L_{eye} + L_{atm} + L_{glass}$$

where

L_{adef} = adaptation deficiency, that is, the degree to which the adaptation lags when the luminance in the field of view changes;
L_{eye} = ocular stray light;
L_{atm} = stray-light components of the atmosphere; and
L_{glass} = stray-light components of the vehicle windshield.

As indicated earlier, it is customary in flat countries to consider L_{adef} equal to 0. Details of the other components of L_d and their measurements are given in the literature (*12,13,20*).

The basic formula given for L_2 could not be used in practice because the field factor was not known accurately enough in the design stage of the tunnel. Some researchers used values that were based loosely on physiological studies that were irrelevant to tunnel lighting conditions. In fact, most recommendations are based on the selection of a number of standard or "nominal" conditions. The selection is somewhat arbitrary and reflects a policy decision as to what is the borderline between acceptable and unacceptable risk. Nevertheless, this procedure led to many outstanding tunnel lighting schemes.

FIELD FACTOR

At present, the field factor seems to be the major unknown factor in tunnel lighting design. Attempts by several researchers to assess the field factor either on the basis of physiological research (*21*) or on practical measurements (*10*) did not yield results that were accurate enough for design purposes. To fill this gap in the knowledge, special experiments were made in the Netherlands and in Japan, using a similar experimental setup.

When the formula for L_2 is reconsidered, it is clear that all of its elements can be either measured at a tunnel or assessed on the basis of the tunnel lighting design. Thus, the luminance in the threshold zone of the tunnel, the value needed for the tunnel lighting design, can be determined. The formula runs

$$L_2 = \frac{L_d f C''}{C - fC''}$$

As indicated, all items in the formula are known or can be found, apart from the field factor. The experiments described here are designed to find the field factor. For the determination of f, the basic formula is rewritten as

$$f = \frac{L_2 C}{C''(L_d + L_2)}$$

In the experiments to establish f, C represents the value of the intrinsic contrast of the object that can just be seen.

EXPERIMENTAL SETUP

Experiments were recently conducted in the Netherlands and in Japan. The Dutch experiments, which had the characteristics of a field test, will be described here. The experiments were done in 1989 in the tunnel in Dordrecht, a medium-sized town in the western part of the Netherlands. The experiments and the results are described by Schreuder (*22*), and a further analysis of the results is also given by Schreuder (*23*).

At the outset it was required that the observer be simultaneously the car driver. The measurements were made in normal traffic to ascertain that the driving task would be measured. The setup consisted of two cars that drove through the tunnel at a fixed speed and a fixed interdistance. In the front car was the object to be observed (a set of 0.25-m-high matte-gray numbers from 1 to 6 with decreasing contrast from 0.78 to 0.06); in the second car were the driver-observer and registration equipment. The numbers were visible continuously; every 1.5 sec the driver had to indicate the number with the lowest contrast that he or she could just perceive. All relevant data were recorded on one video recorder either directly or by radio link; data included the position and interdistance of both cars, their speeds, the daylight level, the lighting mode in the tunnel, the traffic intensity, the visibility, and the serial number of the test.

This setup differs in two respects from normal traffic. First, the observer is aware of taking part in an experiment, and second, the object is continuously visible. Studies on the attention and vigilance of drivers near tunnels—including eye-marker studies—do suggest, however, that in normal traffic, drivers near tunnels always focus their attention, and their gaze, at the tunnel (*14*). It is questionable whether this setup could be applied on normal open roads.

The experiments included a number of lighting measurements in and near the tunnel and the separate measurement under laboratory conditions of the contrast sensitivity threshold (C'') for all observers.

EXPERIMENTS

The field tests were made in the Drecht Tunnel in Dordrecht on July 4 and 5, 1989, between 6:00 a.m. and 8:00 p.m.—that is, from dawn to dusk. The weather was exceptionally bright and clear: the highest illumination in the open was more than 110,000 lux—almost a record; the meteorological visibility was well over 40 km most of the time. The Drecht Tunnel is a 2- by 4-lane tunnel 400 m long with pronounced horizontal and vertical curvatures. The lighting consists of several continuous rows of fluorescent tubes reinforced with high-pressure sodium lamps near the entrance. The lighting has been described by Foucart (*24*). It should be noted that since then the lighting has been renovated and the lighting levels have been reduced.

More than 150 experimental runs were made, each consisting of two passes through the tunnel. The total was 275 good measurements. The driving speed was 90 km/hr; the distance between the cars was between 60 and 80 m. During the measurements, the lighting levels outside and inside the tunnel were changed—the outside level as a result of the changes in natural daylight, the inside by selecting one of the available six lighting (switching) modes.

RESULTS

The results of the lighting measurements are given in Tables 1 and 2.

The veiling luminance (L_{seq}) is measured with a Pritchard luminance meter with a Fry glare-lens attachment. The average result was L_{seq} = 320 cd/m^2 (reduced by E = 100,000 lux). L_{seq} is also approximated with the standard equipment installed at the tunnel for the light control. Here, the average luminance is assessed in a measuring field of 7 by 14 degrees. This luminance is called L_p. L_p is measured continuously, and it is noted down in (arbitrary) scale units. It is therefore necessary to calibrate L_p in terms of L_{seq}. The measured known value of L_{seq} (320 cd/m^2) corresponds with 74 scale units of the control equipment. Using some "smoothing," the L_p-values given by the device could be calibrated in values of L_{seq} (Table 3).

TABLE 1 VALUES OF ILLUMINATION IN DRECHT TUNNEL (lx)

mode	distance 0-60m	60-100m	100-160m	160-400m
1	120	120	120	120
2	220	220	220	220
3	427	427	220	220
4	1288	1288	427	220
5	2303	1288	427	220
6	4427	2303	821	220

TABLE 2 LUMINANCE AND ILLUMINANCE VALUES IN THRESHOLD ZONE OF DRECHT TUNNEL

mode	left wall	right wall	road	average	E	E/L
1	5,5	5,5	5	5,3	120	22,6
2	13	10	12			
				12,3	220	17,9
3	14	14	11			
4	33	23	27	28	427	15,3
5	68	54	44			
				58	821	14,2
6	71	59	54			

Average of tubes for North- and South bound traffic. E: from Table I.

TABLE 3 CALIBRATION OF MEASURED L_p-VALUES IN TERMS OF L_{seq}-VALUES

light level outside		L_p (scale values) (interpolated)	L_{seq} (cd/m^2)
transition	0 >> 1	2,5	11,29
level 1		8,2	37,04
trans.	1 >> 2	15,0	67,76
level 2		21,6	97,57
trans.	2 >> 3	28,3	127,83
level 3		36,4	164,42
trans.	3 >> 4	45,3	204,62
level 4		59,3	267,86
trans.	4 >> 5	75,3	340,13
level 5		92,2	416,47
trans.	5 >> 6	113,4	512,23

The results of the measurements of the contrast in the tunnel are given in Table 4. Values are the average of all measurements for each value of the outside and the inside level. Not all combinations are represented: the lowest inside levels are missing at high outside levels. For reasons of road safety, these have been excluded from the measurements, because the measurements were made in normal traffic.

To determine f, the contrast as measured in the threshold must be established. The first two measuring points were excluded in view of the penetrating daylight. C is determined as the average for the points 3 to 9.

The threshold of the contrast sensitivity was assessed separately for all subjects that took part in the measurements, using a sinusoidal grid. In the first analysis the average value of C'' was used. In the final analysis each observer's own value of C'' was used.

In the tentative analysis (*22*), several shortcuts were taken. As a preliminary result, a value of f = 4,5 was given. A more precise analysis, in which the influence of a number of experimental and environmental factors was taken into account, did produce the value of f = 6.044, or—rounded off—f = 6.

The experiments in Japan produced a similar value for f. However, further work is needed to make a more accurate comparison between the data from Japan and those from the Netherlands. The preliminary results of the Japanese measurements are given by Yoshikawa (*25*).

CONCLUSIONS

It can be concluded from the experiments that the value of the field factor is f = 6.

This seems to be a reasonable representation of the experimental data. When applied to the Drecht Tunnel, the result is that the actual tunnel lighting scheme (where the luminance in the threshold zone is coupled automatically with the luminance on the open road) allows that a contrast close to 0,2 is usually visible. The results are not reconfirmed yet by studies of accidents or driving behavior; the latter is part of a plan for experiments.

The field factor is, as indicated earlier, an essential element for applying the stray-light theory to the design of tunnel entrance lighting installations. Further experiments are required, because it may be assumed that the field factor depends on the prevailing traffic conditions and on the design characteristics of the tunnel entrance.

In a proposal submitted to CIE, it is suggested to set up a more systematic international experiment (*26*). The main reasons for an international approach are the facts that in most countries the variations of climate and geography (morphology) are limited and that the tunnel lighting equipment follows national codes; in short, it is not possible to have enough experimental variability within one country to arrive at truly general results.

Furthermore, the results of the experiments in the Netherlands suggest that f, contrary to the assumption made at the outset of the study, is not a constant but depends to a certain degree on the outside luminance.

TABLE 4 RESULTS OF CONTRAST MEASUREMENTS

measuring point					1	2	3	4	5	6	7	8	9					
distance					4	8	12	16	20	24	28	32	36	40.. etc. meter				
	..contrasts (in %)..																	
	..outside..				........entrance........									.transition				
light																		number
1-2	18	18	18	18	.	.	.	.	.	.	35	.	.	30	35	35	35	n=2
1-3	15	18	18	24	.	.	24	35	.	.	35	.	35	32	37	32	32	n=4
2-0	12	12	12	.	.	.	.	53	.	.	.	.	.	53	53	53	53	n=1
2-1	14	14	14	12	.	35	.	.	35	.	.	.	.	39	39	40	35	n=5
2-2	12	12	12	24	.	.	.	.	.	.	.	35	.	35	35	35	53	n=1
2-4	24	24	24	.	.	.	35	.	.	.	.	.	.	24	35	35	35	n=1
3-1	14	17	19	24	.	12	12	.	.	.	.	53	44	44	42	44	44	n=5
3-2	12	12	12	15	.	.	24	24	.	.	35	35	35	35	33	30	33	n=6
3-3	12	14	14	18	24	12	.	.	.	.	35	.	35	26	28	28	40	n=5
3-4	12	12	12	12	.	.	.	24	.	.	.	24	.	20	24	24	28	n=3
3-5	12	12	12	18	.	24	.	.	.	.	12	.	24	20	20	28	31	n=3
4-2	12	12	12	14	12	14	12	19	44	.	33	35	39	38	35	34	33	n=33
4-3	12	12	12	14	17	14	15	19	20	.	36	20	33	34	32	32	33	n=41
4-4	11	11	11	12	11	15	12	15	16	.	21	24	21	20	21	26	28	n=47
4-5	12	12	13	13	12	16	15	13	12	.	17	26	18	18	20	24	29	n=46
4-6	11	11	12	12	11	12	12	14	.	.	12	12	16	15	16	21	24	n=39
5-3	9	12	12	14	.	12	18	12	.	.	44	12	45	36	35	31	29	n=13
5-4	9	10	10	12	.	12	12	12	.	.	.	.	24	22	22	24	26	n=6
5-5	9	9	9	9	12	.	12	12	.	.	18	12	24	17	17	17	17	n=7
5-6	7	9	9	12	.	9	12	12	.	.	.	.	10	10	10	13	17	n=7

The average of the threshold contrasts for different lighting combinations (3-2 means level 3 outside, level 2 inside; values taken from Table I and II). Distances from the tunnel portal.

Also, the experimental equipment allows for a rapid assessment of the visibility in tunnel entrances. In this respect, the equipment can be used to test the performance of finished lighting installations. In the Netherlands the possibilities of setting up a commercial measuring and testing establishment are studied.

ACKNOWLEDGMENTS

The experiments described in this paper were made by SWOV in the summer of 1989 under a contract with the Netherlands Ministry of Transport, Division of Locks and Weirs. The experiments were prepared by Biesheuvel and Planken of the Technical College at Rotterdam and carried out by Stam & Van Vuren Consultants. The light measurements were made by Philips, Eindhoven and the contrast threshold measurements by the Eye Hospital of the University of Rotterdam. The authorities of the Drecht Tunnel in Dordrecht contributed in important ways to the experiments. The experiments are described in detail elsewhere (*22,23,27*).

REFERENCES

1. D. A. Schreuder. The Lighting of Vehicular Traffic Tunnels. Centrex, Eindhoven, the Netherlands, 1964.
2. D. A. Schreuder. De verlichting van tunnelingangen. R-81-26 I + II. SWOV, Leidschendam, the Netherlands, 1981.
3. NSvV Aanbevelingen voor tunnelverlichting. *Electrotechniek*, Vol. 41, 1963, pp. 23–32, 46–53.
4. *Aanbevelingen voor de verlichting van lange tunnels voor het gemotoriseerde verkeer; Ontwerp*. NSvV, 1989.
5. *International Recommendations for Tunnel Lighting*. Publication 26. International Commission on Illumination, Paris, France, 1973.
6. *Guide for the Lighting of Road Tunnels and Underpasses*. Publication 26/2. International Commission on Illumination, Vienna, Austria, 1990.
7. C. Bourdy, A. Chiron, F. Cottin, and A. Monot. Visibility at a Tunnel Entrance: Effect of Temporal Adaptation. In *Lighting Research and Technology*, Vol. 19, 1987.
8. C. Bourdy, A. Chiron, F. Cottin, and A. Monot. Visibility at a Tunnel Entrance: Effect of Temporal Luminance Variation. In *Lighting Research and Technology*, Vol. 20, 1988, pp. 199–200.
9. *Guide de l'éclairage des tunnels routiers*. CETU, Bron, 1985.
10. F. Novellas and J. Perrier. New Lighting Method for Road Tunnels. In *CIE Journal*, Vol. 4, No. 2, 1985, pp. 58–70.
11. M. Tesson and B. Monié. Road Tunnel Lighting: Simplification. In *Lighting Research and Technology*, Vol. 21, 1989, pp. 171–179.
12. P. Padmos and J. W. A. M. Alferdinck. *Verblinding bij tunnelingangen II: De invloed van atmosferisch strooilicht.* Rapport IZF 1983 C-9. IZF-TNO, Soesterberg, the Netherlands, 1983.
13. P. Padmos and J. W. A. M. Alferdinck. *Verblinding bij tunnelingangen III: De invloed van strooilicht van de autovoorruit.* Rapport IZF 1983 C-10. IZF-TNO, Soesterberg, the Netherlands, 1983.
14. K. Narisada and K. Yoshikawa. Tunnel Entrance Lighting—Effect of Fixation Point and Other Factors on the Determination

of Requirements. In *Lighting Research and Technology*, Vol. 6, 1974, pp. 9–18.
15. K. Narisada and Y. Yoshimura. Adaptation Luminance of Driver's Eyes at the Entrance of Tunnels. In International Commission on Illumination, 1977.
16. D. A. Schreuder. Tunnel Entrance Lighting—A Comparison of Recommended Practice. *Lighting Research and Technology*, Vol. 3, 1971, pp. 274–278.
17. W. Adrian. Method of Calculating the Required Luminances in Tunnel Entrances. In *Lighting Research and Technology*, Vol. 8, 1976, pp. 103–106.
18. W. Adrian. Visibility of Targets: Model for Calculation. In *Lighting Research and Technology*, Vol. 21, 1989, pp. 181–188.
19. D. A. Schreuder and H. J. C. Oud. *The Predetermination of the Luminance in Tunnel Entrances at Day*. R-88-13. SWOV, Leidschendam, the Netherlands, 1988.
20. J. J. Vos. Verblinding bij tunnelingangen I: De invloed van strooilicht in het oog. Rapport IZF 1983 C-8. IZF-TNO, Soesterberg, the Netherlands, 1983.
21. W. Adrian. Investigations on the Required Luminance in Tunnel Entrances. Presented at the National Conference, CIBS, Cambridge, 1980.
22. D. A. Schreuder. De veldfactor bij de bepaling van de verlichtingsniveaus bij tunnelingangen; Verslag van experimenteel onderzoek. R-90-10. SWOV, Leidschendam, the Netherlands, 1990.
23. D. A. Schreuder. De veldfactor bij de bepaling van de verlichtingsniveaus bij tunnelingangen; Verslag van een nadere analyse van het experimentele onderzoek. SWOV, Leidschendam, the Netherlands, in preparation.
24. A. Foucart. *Verkeersregeling en automatisering van de technische diensten in de Drechttunnel te Dordrecht (Nederland)*. Ateliers de Constructions Electriques de Charleroi, Belgium, 1979.
25. K. Yoshikawa. *Short Report on the Experiments of the Field Factor for Lighting Level at the Tunnel Entrance Zone*. Preliminary draft report. Matsushita Electric, Osaka, Japan, 1991.
26. D. A. Schreuder. Lighting Requirements in the Entrance of Tunnels in Traffic Conditions. Presented at the International Commission on Illumination Session, Melbourne, Australia, 1991.
27. D. A. Schreuder. The Field Factor for the Determination of Tunnel Entrance Luminance Levels. Presented at SLG/International Commission on Illumination Symposium on New Developments in Tunnel Lighting, Lugano, Switzerland, Oct. 1989.

Publication of this paper sponsored by Committee on Visibility.

Visibility Under Transient Adaptation

WERNER K. ADRIAN AND RADOSVETA V. TOPALOVA

The Illuminating Engineering Society (IES) has adopted recommendations for the luminance transition in a tunnel that deviate greatly from the curves proposed by the International Commission on Illumination (CIE) and contained in the DIN (German Standards Institute) standard. IES has based its guidelines neither on practical considerations nor on scientific or experimental foundations. To clarify the discrepancy, the physiological processes of adaptation of the eye during a change in luminance have been modeled, and their impact on the required luminance in the transition necessary to ensure visibility has been derived. Using Fry's model for the kinetics of the eye's response and Adrian's ΔL model, the course of the luminance transition has been calculated. The results are compared with the IES and CIE standards. A comparison of the resulting curve with the curve suggested by CIE reveals only small differences. In general, the comparison indicates that the experimentally determined CIE curve agrees with results derived from the fundamentals of dark adaptation. The IES suggestion, however, falls short. The eye requires about twice as long to adapt to the luminance transition as the IES proposal allows.

Recently, the Illuminating Engineering Society (IES) subcommittee on tunnel lighting released new guidelines on the lighting of tunnels and underpasses. These guidelines have been passed by various institutions and have become the basis for public practice.

The guidelines differ substantially in three points from the recommendations worked out by the International Commission on Illumination (CIE) (*1*), revealing a contradiction in basic concepts that cannot remain unresolved. One such concept relates to the transient adaptation taking place when a tunnel is entered in daytime and the subsequent adaptation to the decreasing luminance. The recommended length of the luminance transition zone is, according to the CIE guidelines, three times longer than the length suggested by IES. To reconcile this fundamental difference, scientific investigation of the transient adaptation and the underlying physiological processes appeared necessary. The aim of this study is to provide indisputable scientific evidence as the basis for critical discussion of the problem and to allow for a solution.

BACKGROUND

The CIE recommendations are based on a modified curve worked out in 1962 by Schreuder and DeBoer (*2*). A similar curve is contained in the DIN (German Standards Institute) standards on tunnel lighting. The course of the luminance transition was found during tunnel simulation experiments with observers and is based on their 75 percent acceptance level of the luminance reductions after entering the tunnel. The results show that from high levels prevailing at the entrance of the tunnel, the eye needs approximately 12 sec to adapt to the interior luminance level of 10 cd/m^2. For a speed of 80 km/hr (50 mph), the length of the transition zone must therefore be greater than two stopping distances (accepted safe stopping distance on dry pavement at 80 km/hr is 130 m). The IES guidelines, in contrast, recommend one stopping distance only. This recommendation resulted from assumptions and practical aspects and has neither experimental nor scientific foundation.

The length of the transition zone depends on the rapidity with which the eye can change its adaptation and has therefore been determined on the basis of visual physiology. Two mechanisms underlying the temporal course of adaptation are distinguished. One is the fast phase, called network or neuronal adaptation; the second is the slower phase, which is due to photochemical processes.

KINETICS OF VISUAL ADAPTATION

In 1946, Jahn (*3*) attempted to describe the kinetics of dark adaptation mathematically using Wald's (*4*) visual cycle of rhodopsin. Fry (*5*) has further developed this model and implemented equations for the reaction kinetics of various intermediate products between the unbleached photopigment and the bleached stage in which we have vitamin A_1-aldehyde and opsin involved in the process of adaptation. His interpretation of the substances s, m, and h as indicated in Figure 1 fell short because the phosphoresterase of photopigments was not yet known, but the formulas for the reaction kinetics can generally be applied to describe the actual processes. Figure 1 is adopted from Fry's 1973 publication (*5*) and displays the mechanisms of excitation following the change in stimulus. The variables in Figure 1 are defined as

- s—Concentration of visual pigment (rhodopsin in Jahn's model, but generally any photopigment).
- O=R—Opsin and retinene that recombine to form the photopigment.
- m—Visual white into which part of R is transformed. The m is a substance of a secondary process that is decomposed by the energy released by the primary process that is proportional to the retinal illumination in trolands (E) and the concentration of s ($E \cdot s$).
- n—Decomposed substance m.

Department of Systems Design Engineering Optometry, University of Waterloo, Waterloo, Ontario N2L 3G1 Canada.

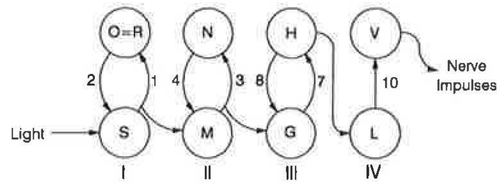


FIGURE 1 Mechanisms subserving the retinal response to a flash of light.

- g—Substance in the tertiary mechanism that responds to the energy given off by the decomposition of m; this energy is proportional to Esm.
- h—Substance to which g changes. For generating a response in the quarternary mechanism, it is the concentration of h that counts: it constitutes a catalyst for initiating the nerve impulses and therefore visual exitation.

The primary process I indicates the bleaching and regeneration of the photopigment. The rate of release of energy in the dissociating photopigment is directly proportional to $k_1 \cdot E \cdot s$, in which k_1 is a constant. It may be assumed that the concentration of free opsin o is $o = s - 1$, where s is the photopigment concentration. So $1 - s$ constitutes the reformation rate. The kinetics of the maintenance of photopigment can be described as follows:

$$\frac{ds}{dt} = k_2(1 - s) - k_1Es \tag{1}$$

In a similar way, the equations for the following processes have been found to be

$$\frac{dm}{dt} = k_4(1 - m) - k_3Esm \tag{2}$$

For the purpose of visibility consideration during the transitional adaptation, it is more appropriate to use h because its concentration is proportional to the frequency of the nerve impulses and therefore proportional to the brightness sensation.

The necessary luminance difference ΔL of an object to be perceptible during the transition of adaptation is proportional to

$$\frac{dE}{dh} = \Phi_{\text{abs}} \qquad \Phi_{\text{rel}} = \frac{dE/dh}{dE/dh_{t\to\infty}} \approx \frac{\Delta L}{\Delta L_o} \tag{3}$$

where ΔL_o is the luminance difference threshold in steady-state condition at the end of the transition.

Time Course of Processes 1 to 4

The transient processes start and end at equilibrium values. For example, if the fovea is exposed to a steady stimulus, the concentrations of s, m, and g reach those values. These concentrations can be calculated with Equation 1 in which ds/dt becomes equal to zero. Thus

$$\frac{ds}{dt} = 0 = k_2(1 - s) - k_1E \cdot s$$

$$k_2 = (k_2 + k_1E) \cdot s$$

$$\frac{1}{s} = 1 + \frac{k_1}{k_2} \cdot E$$

or

$$s = \frac{k_2}{k_2 + k_1 \cdot E} \tag{4}$$

The concentrations in steady-state conditions for m and h can be calculated in the same way.

$$\frac{1}{m} = 1 + \frac{k_3}{k_4} sE \tag{5}$$

$$\frac{1}{h} = 1 + \frac{k_8}{k_7 \cdot smE} \tag{6}$$

The graphs in Figures 2 and 3 show the time course of m and h after a sudden luminance reduction from 2000 to 8 cd/m^2.

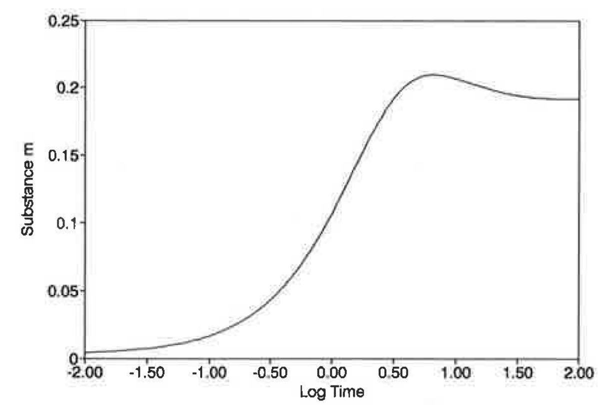


FIGURE 2 Calculated time course of concentration *m* after luminance suddenly changes from 2000 to 8 cd/m^2.

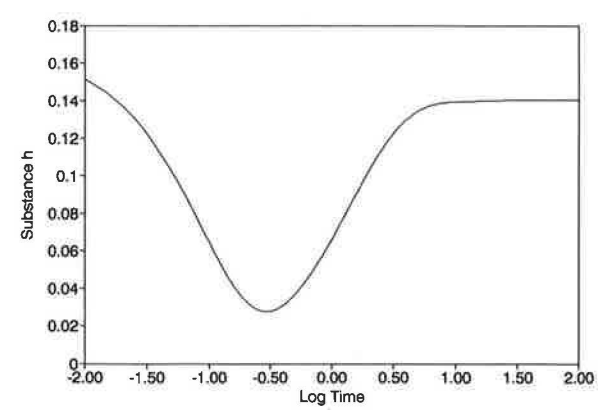


FIGURE 3 Calculated time course of substance *h* after luminance suddenly changes from 2000 to 8 cd/m^2.

The concentrations before and after the transition is completed would follow from Equations 5 and 6 and indicate the equilibrial conditions. Calculation reveals that for high levels of E—for example, $E > 10^4$ trolands—the product smE becomes constant and leads to a concentration of $h = 0.652$.

Calculation of Time Course of *s*, *m*, and *h* Using Euler's Method

According to Euler's method, a numerical solution for differential equations can be calculated, providing that initial values for the variables involved are supplied. Using the steady-state equations given in Equations 4, 5, and 6, the starting values for s, m, and h were obtained. These values were then used in Euler's difference equations.

$$s_{t+\Delta t} = s_t + \Delta t[k_2(1 - s_t) - k_1Es_t] \tag{7}$$

$$m_{t+\Delta t} = m_t + \Delta t[k_4(1 - m_t) - k_3Es_tm_t] \tag{8}$$

$$h_{t+\Delta t} = h_t + \Delta t[k_7Es_tm_t(1 - h_t) - k_8h_t] \tag{9}$$

where Δt is the time increment.

Once this model was set up in the computer spreadsheet, many trials were run to achieve results that fit the experimental data as found in the literature. Measurements of re-adaptation courses were made as early as 1936 by Lossagk (*6*) and more recently by Greule (*7*). Greule used more modern equipment that allowed precise control of the stimulus and presentation time.

In his investigation, Fry examined the required brightness difference threshold necessary for a response to take place, and he went on to study the effect of the stimulus duration on the threshold intensity required. Because h is the catalyst for the nerve impulse reaction, Fry proposed that dE/dh is proportional to the temporal threshold elevation over the steady state.

For clarity, Fry's deviation of dE/dh, as approximated by $\Delta E/\Delta h$, will be repeated here.

$$\frac{\Delta E}{\Delta h} = \frac{k_8}{k_7sm(1 - h_{on})^2} \tag{10}$$

The $\Delta E/\Delta h$ is proportional to the luminance threshold value ΔL necessary for perception. Basing its value on the ΔL_o that is achieved in the equilibrium (steady state) at the end of the transition, we obtain the threshold multiplier Φ that follows from

$$\phi = \frac{\frac{dE}{dh}(t)}{\frac{dE}{dh}(t \rightarrow \infty)} \quad \textit{from} \quad L_{\text{high}} \text{ to } L_{\text{low}} \tag{11}$$

Using this formula for Φ, we attempted to apply Fry's model, using Equation 10 for long flashes that would simulate sudden luminance changes to approximate the transition in a tunnel. Figure 4 depicts the course of the Φ of the perception threshold that is necessary to keep the target visible during the transition from 2000 to 8 cd/m². Φ indicates the multiple of the ΔL threshold at the steady-state level of 8 cd/m². As can be read from the curve, the steady-state threshold is reached approximately 3 sec ($\log t = 0.5$) after the luminance drop.

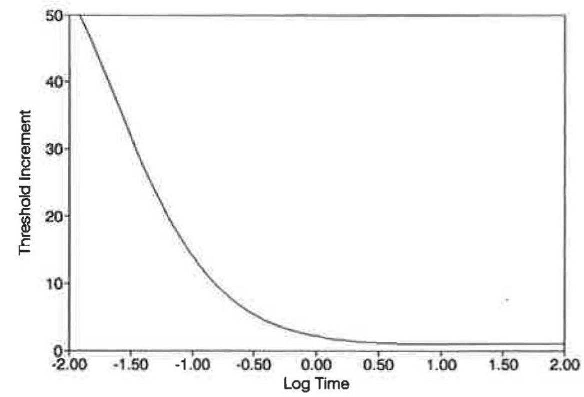


FIGURE 4 Calculated time course of threshold increment after luminance suddenly changes from 2000 to 8 cd/m².

Figure 5 shows the threshold increase Φ for a sudden drop of the adaptation luminance from 2000 to 8 cd/m² as well as Φ found for an increase from 8 to 2000 cd/m². The solid lines are calculated according to the model using the constants found from Table 1. The data measured by Greule are also plotted; they are in good agreement with the calculated values. This seems to justify the use of the modified Fry's equations to describe the transient adaptation processes even though they were originally based on inappropriate assumptions.

The constants were found by the best match with the experimental data in Figure 5. It is interesting to note that only k_3 and k_4 had to be changed for up and down luminance jumps, because this appears to indicate different intermediate chemical processes as described before.

In Figure 6 early data by Lossagk are reproduced that were obtained with an apparatus that did not allow precise timing. Still, the data show reasonable agreement with the calculated curves.

Determination of Required Luminance Transition in a Tunnel

When entering a tunnel during daylight conditions, the eye must be able to adapt to the luminance inside the tunnel. The visual task is to detect objects at a much lower luminance level at any time during the entrance in order to ensure traffic safety. Therefore, it is extremely important that adequate lighting be provided, on the basis of the requirements of the course of dark adaptation. Luminance levels should also meet the subjective demands for safety that were found to parallel the lighting conditions (*8*).

Procedure

Whenever the adaptation luminance suddenly changed, from higher to lower levels or vice versa, the luminance difference

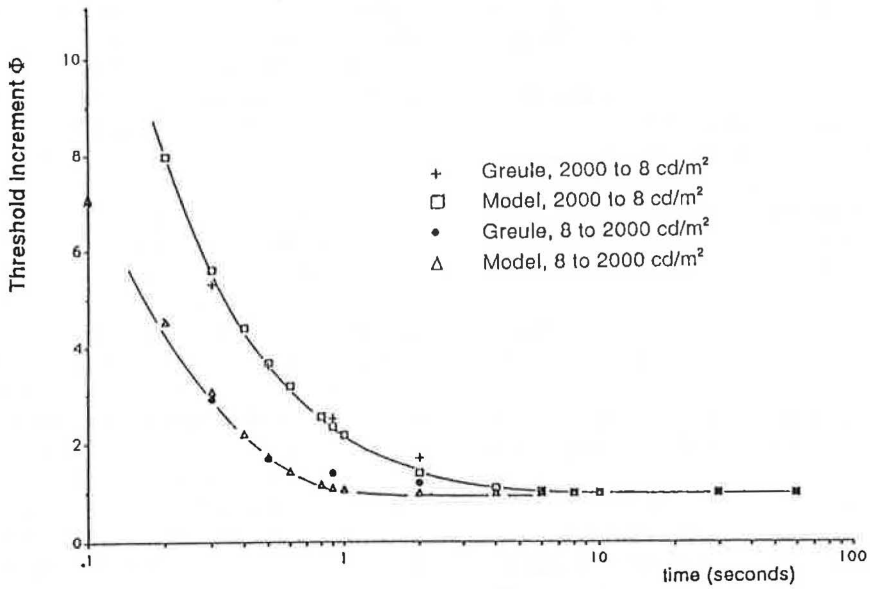


FIGURE 5 Calculated threshold increments ϕ for luminance changes from 2000 to 8 cd/m^2 (upper curve) and from 8 to 2000 cd/m^2 (lower curve) compared with Greule's data (*7*).

TABLE 1 CONSTANTS FOR INCREASE IN STIMULUS (a) AND DECREASE IN STIMULUS (b)

Constant	Fry		Adrian/Fleming	
	a	b	a	b
k_1	$2 \cdot 10^{-7}$	$2 \cdot 10^{-7}$	$2 \cdot 10^{-6}$	$2 \cdot 10^{-6}$
k_2	1/130	1/130	1/13	1/13
k_3	$2.56 \cdot 10^{-5}$	$2.56 \cdot 10^{-5}$	$7.0032 \cdot 10^{-3}$	$7.0032 \cdot 10^{-4}$
k_4	0.00481	0.00481	0.1415	0.1415
k_7	0.1	0.1	0.1	0.1
k_8	10	10	10	10

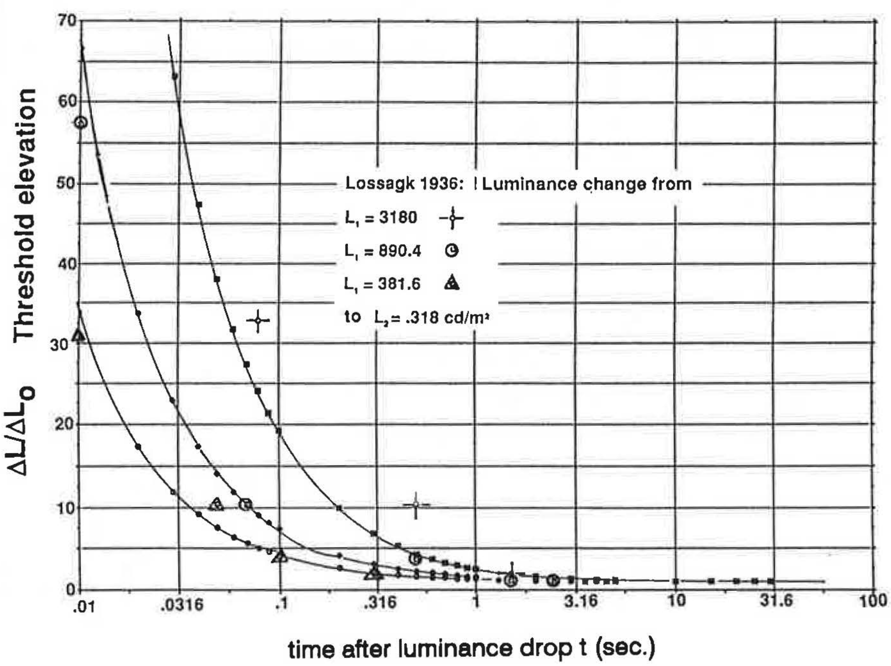


FIGURE 6 Lossagk's data compared with calculation.

threshold ΔL increased until it reached—after a transitional period—its steady-state value ΔL_o determined by the final L-level.

With Φ known as a function of time after entering the tunnel, following from the model, ΔL can be calculated.

Adrian (*9*) developed a method to calculate the ΔL threshold for various adaptation luminances, targets, and exposure times. On the basis of his method, the required background luminance for a critical target size and practical observation time can be calculated for every ΔL. As the necessary ΔL is known from Equation 3 as a function of time in the transition, that can be transformed into the corresponding background luminance $L_B = f(t)$. This procedure was applied to all the graphs in Figures 7–10.

The actual required luminance levels L_B, following from the ΔL in the transition, were computed and are shown in Figures 7–10.

The first calculation was carried out assuming that an observer is adapted to 500 cd/m² and enters a tunnel that produces (without any transition) a sharp drop to the inside luminance level of 3 cd/m². The assumption of 500 cd/m² follows from practical considerations. Adrian (*8*) suggested a method to obtain the actual adaptation luminance (L_A) when approaching a tunnel. According to this method, L_A is composed of the average luminance in the foveal field (~2 degrees) on which the stray light, created by the bright tunnel environment, is superimposed.

$$L_A = L_{2^\circ} + L_{seq} \tag{12}$$

where L_{2° equals luminance in the central 2-degree field and L_{seq} equals equivalent veiling luminance.

Practical measurements on tunnel sites revealed that a luminance in the entrance zone of 200 cd/m² prevails in reasonably lit tunnels. In bright daylight at about 47 to 52 degrees northern latitude, equivalent veiling luminance values between 150 to 450 cd/m² generally occur. With this in mind, an average of L_{seq} = 300 cd/m² was chosen that resulted in L_A = 500 cd/m², as in the example.

For the $L_B = f(\Delta L)$ calculations, a target size of 10 min arc and an observation time of 0.2 sec were used throughout. A target of that size is used internationally to express the visual task of drivers. It is the critical size that must be seen if a collision is to be avoided. It relates to an object 22 cm in square located ~85 m in front of the car. This is close to the safe stopping distance on an even and dry road for a speed of ~65 km/hr (40 mph). Studies on eye movements and fixation times have revealed that drivers devote in daytime an average 0.2 sec looking at locations in front of the car (*10*), which is the rationale for adopting that observation time.

Figure 7 shows the minimal luminance as a function of time after the abrupt drop assumed to take place at $t = 0$ to the interior level of 3 cd/m². The luminance course recommended by the CIE and that according to the DIN standard are also depicted for comparison. The CIE curve is supposed to start from 100 percent level of the entrance luminance, which again had to be assumed to be 200 cd/m². The DIN curve, which was somehow derived from that of CIE, shows in the beginning a lesser slope than that recommended by CIE but drops slightly faster after t ~ 12 sec. The calculated function shows a steeper slope in the initial phase of the transition, but levels off to approach the CIE curve after t ~ 16 sec.

The steeper slope of the calculated L-function suggests that the luminance in the initial phase of the transition could be

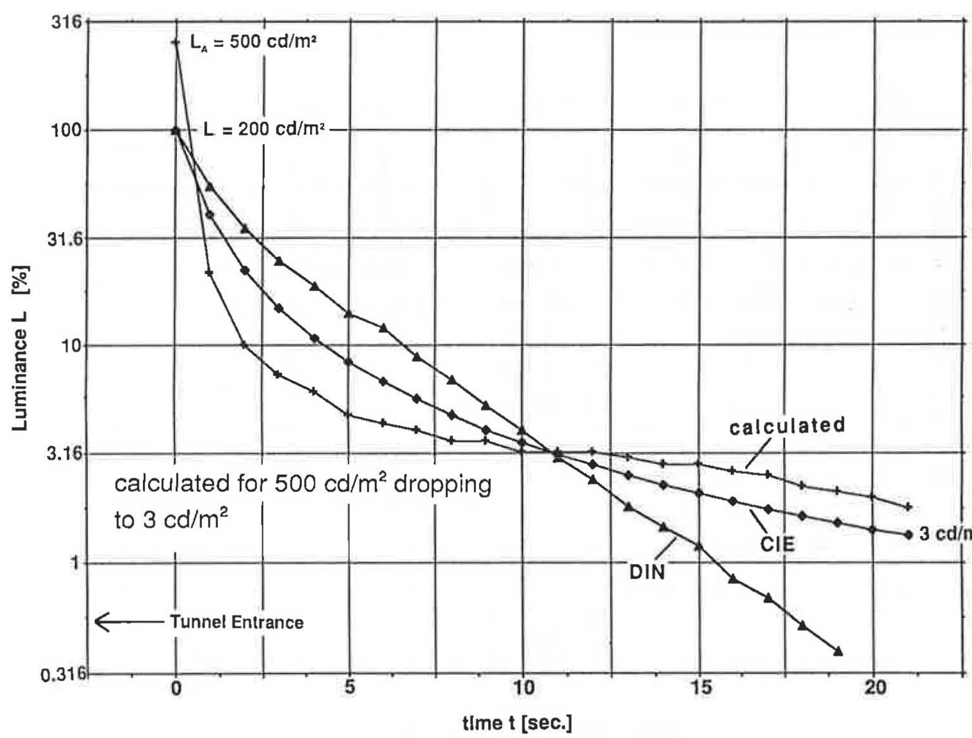


FIGURE 7 Comparison between CIE, DIN, and calculated transitions in one step.

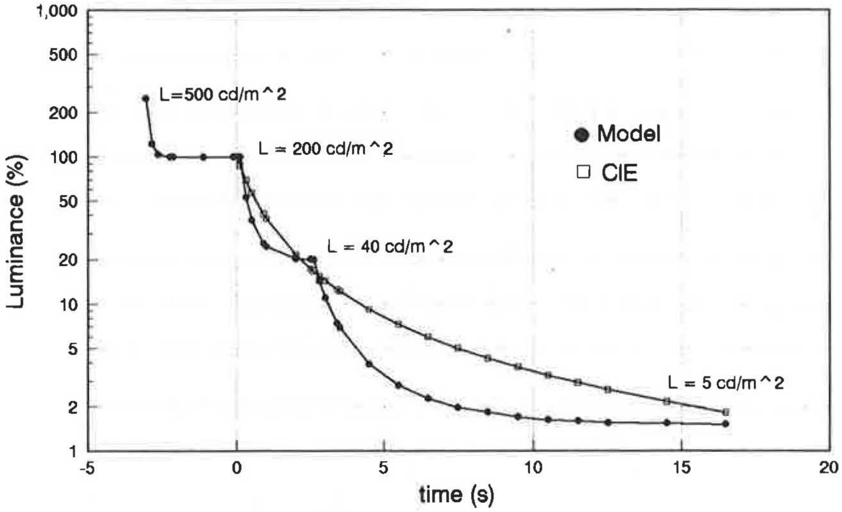


FIGURE 8 CIE transition and calculated functions following a two-step transition.

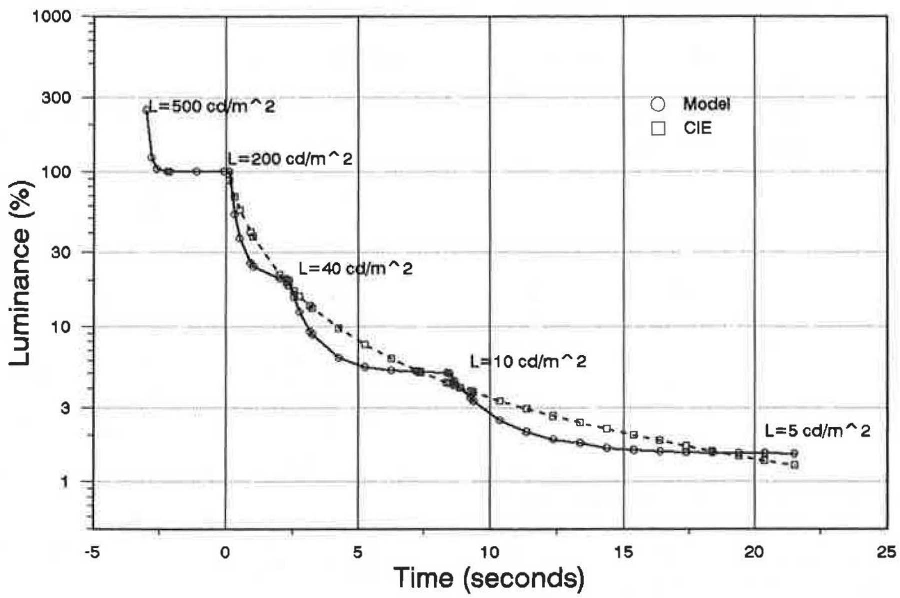


FIGURE 9 CIE transition and calculated functions following a three-step transition.

lower than suggested by CIE or DIN. This might be because an abrupt L-transition from the adaptation luminance of 500 to 3 cd/m^2 was assumed to occur at the end of the entrance zone, the length of which must measure one stopping distance. Over that distance a constant luminance level must be provided. Such an abrupt change, however, never occurs in practice. The continuously decreasing L-level must be approximated stepwise instead.

A first attempt is made in Figure 8 using two steps from the level in the entrance zone to the interior level. The first part shows the adaptation transition from L_A (500 cd/m^2) to the level in the entrance zone (200 cd/m^2). At the beginning of the transition zone, the luminance level is assumed to drop to 40 cd/m^2 and then to 3 cd/m^2. The section of the curve (200 to 40 cd/m^2) was truncated at the time the $\Delta L/\Delta L_o$ ratio reached 1.02 of its steady-state value, or when the increment was only 2 percent.

In Figure 9 the same calculation was carried out approximating the transition in three luminance steps, from 200 to 40, 40 to 10, and 10 to 3 cd/m^2. As can be observed from the graph, the calculated course of the luminance comes very close to that of CIE, which was derived from experiments in models (*2*).

Figure 10 allows us to compare the transition steps contained in the North American Standard Practice RP22 and the calculated transition using the same luminance level of each step. According to the recommendation, the interior level of 5 cd/m^2 can be reached in 6 sec (165 m, 99 km/hr) after the end of the entrance zone having a luminance of 330 cd/m^2. It is obvious that the adaptation transition takes longer

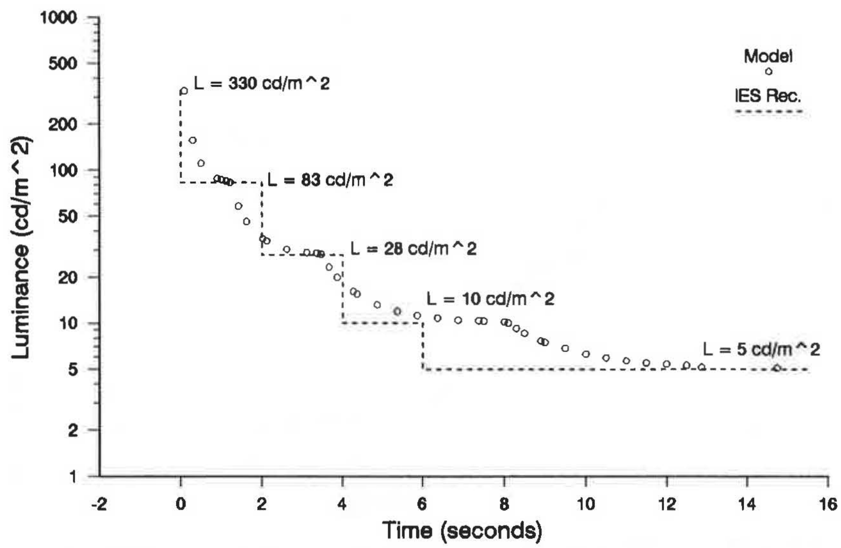


FIGURE 10 Calculated function for transient adaptation and steps proposed in IESNA RP22.

and that an observer entering a tunnel with such specifics would not be able to follow the rapid changes of interior luminance. The required distance to accommodate such a transition must be approximately twice as long.

CONCLUSION

The purpose has been to investigate the adaptation of the eye following changes in luminance of the visual field. The criterion used for the adaptation in the transition from one luminance level to the other was the perceptibility of a defined, internationally used target to describe the visual task of a driver.

The calculation of the time required for transient adaptation was based on equations used by Fry (*5*) describing the kinetics of the chemical and physiological processes in the retina.

In conclusion, it appears reasonable to adopt the CIE suggestions in respect to the course of the luminance in the transition zone, provided that the starting point is 100 percent at the end of the entrance zone and not—as CIE allows—truncated function to start at 40 percent of the entrance *L*-level.

REFERENCES

1. *Guide for the Lighting of Road Tunnels and Underpasses.* Publication 26, A-130. International Commission on Illumination, Vienna, Austria.
2. J. B. DeBoer. *Lichttechnik*, Vol. 15, 1963, pp. 124–127.
3. T. L. Jahn. *Journal of Optical Society of America*, Vol. 36, 1946, pp. 659–665.
4. G. Wald. *Journal of Optical Society of America*, Vol. 19, 1935, pp. 351–371.
5. G. A. Fry. *American Journal of Optometry and Archives of American Academy of Optometry*, Vol. 50, 1973, pp. 355–375.
6. H. Lossagk. *Das Licht*, Vol. 5, 1936, pp. 126–131.
7. R. Greule. Dissertation. Technische Hochschule, University of Karlsruhe, Germany, May 1986.
8. W. K. Adrian. *Lighting Research and Technology*, Vol. 19, 1987, pp. 73–79.
9. W. K. Adrian. *Lighting Research and Technology*, Vol. 17, 1989, pp. 181–189.
10. H. T. Zwahlen. *Driver Eye Scanning Behavior at Tunnel Approaches*, Vol. 1. Franklin Institute Research Laboratory, Philadelphia, Pa., 1979.

Publication of this paper sponsored by Committee on Visibility.

Priority Programming Methodology for Rail-Highway Grade Crossings

TIMOTHY A. RYAN

The objective was to develop a comprehensive methodology to assist in setting priorities for improvements to rail-highway grade crossings. The objectives of the methodology were as follows: to compute all existing costs at a RHGC, changes in costs that would result from each of a number of improvements for that crossing, and costs of implementing those improvements; to select the set of projects that would maximize expected net benefits, subject to a budget constraint; and to be of practical instead of theoretical use to decision makers. The methodology developed and presented fulfills these objectives, subject to the constraints of existing data bases. Accident costs, delay costs, diversion costs, and costs of delay to emergency vehicles are considered, and up to five improvement projects per crossing are evaluated. The benefit and cost computations are made in a Fortran computer program developed in the research. The methodology was applied successfully to 1985 conditions at all RHGCs in a particular jurisdiction. The results of the application indicate that, for the same total budget, expected net benefits could have been approximately $7 million higher if the methodology's projects had been implemented.

A rail-highway grade crossing is an at-grade intersection of one or more railroad tracks and a roadway. At such a crossing, railroad vehicles and roadway vehicles must share the right-of-way.

COSTS OF RHGCS

Many costs are associated with operating a RHGC; the most obvious one—particularly to the public and to elected officials—is the cost of accidents. Other costs, however, should be recognized as well.

The most obvious cost of a RHGC accident is that of damage and injury to persons and property. Other costs include the rerouting of railroad and roadway traffic while accident investigation and cleanup are under way.

When a train blocks a RHGC, roadway traffic is delayed. Such delays can be minor, as they are when a short train crosses a little-used roadway during the middle of the night. But they can also be major, as they are when a long train traveling slowly crosses a major arterial roadway during the afternoon peak hour. Delays have obvious costs, such as the time of the delayed motorists and additional fuel consumption, vehicle operation, and air pollution.

Even without accidents or blockages, RHGCs impose costs on society. Because a crossing often disrupts the grade of the roadway, vehicles must reduce speed to traverse the crossing or risk mechanical damage. In some cases the train must stop completely to allow a flagger to disembark and warn roadway traffic. Such reductions in roadway and railroad speed use up time and energy. In addition, there are expenses associated with maintaining the crossing surface and traffic-control devices.

A special type of cost is incurred when an emergency vehicle is delayed by a blockage at a crossing. Delays to emergency vehicles can, in the most extreme cases, result in deaths. In less severe cases, these delays can cause additional property damage—if, for example, fire apparatus is held up in reaching a fire. These costs are not obvious, and they frequently go unnoticed until incurred.

PRIORITY RANKING IMPROVEMENTS

Because it would be impractical to eliminate all RHGCs, and because of limited funding for safety improvements for crossings that cannot be eliminated, a methodology that efficiently allocates available resources for crossing improvements is needed.

One of the earliest efforts to priority rank improvement projects for rail-highway grade crossings was made by Richards and Hooks in 1970 (*1*). In that effort, cost-benefit analyses were proposed to be used to examine the upgrading of traffic-control devices. Only direct accident costs were included. A similar approach was taken by Richards and Lamkin, also in 1970 (*2*). Cost-benefit analyses were also used by Schulte (*3*) to determine, on a single-crossing basis, which type of upgraded traffic-control device to install. Again, only direct accident costs were considered.

A generalized approach to priority programming highway projects was presented by Harness and Sinha (*4*). The approach, called the successive subsetting technique, allows for one set of criteria to be used in establishing a short list of candidate projects and for a second set of criteria to be used in establishing a shorter list from the short list. This approach is continued until a subset is developed that allows a fixed budget constraint to be met. The criteria for each iteration may be subjective. This technique suffers from the deficiency that subjective evaluations (such as a decision without data analysis that delay is more important than safety) may be used throughout and could greatly influence the end result.

In an effort to improve on the cost-effectiveness of improvement decisions, the U.S. Department of Transportation (DOT) formed the Rail-Highway Crossing Resource Allocation Procedure in 1982 (*5*). The measure of effectiveness used in the procedure is the efficiency ratio: the percentage

JHK & Associates, Chester Building, 8600 LaSalle Road, Suite 313, Baltimore, Md. 21204.

reduction in accidents that can be expected as a result of a proposed improvement.

The procedure is certainly a helpful tool, but it suffers from two of the weaknesses of the priority ranking systems: first, it considers only safety (as measured by number of accidents and severity of accidents) to be an important parameter; second, it does not consider all the costs of accidents in its approach.

The inclusion of factors other than direct accident costs has been suggested, in general terms, as being desirable. The *Railroad-Highway Grade Crossing Handbook* (*6*), for example, identifies improvements in operating efficiency in addition to improvements in safety as fundamental objectives of an improvement program.

Tidwell and Humphreys (*7*) suggested that stopping sight distance and highway vehicle speed should be considered at crossings with passive traffic-control devices, but they did not suggest a method of numerically including these factors with a hazard index/accident prediction formula to develop a priority list.

The Colorado Department of Transportation (*8*) has developed a formula for grade-separating RHGCs that considers 12 factors: average daily traffic (ADT), daily train volume, project cost, roadway speed limit, maximum train speed, crossing geometrics, delay, alternate route availability, accident history, hazardous material, people factors, and emergency-vehicle access. Although this technique goes beyond the usual limitation of considering only direct accident costs, it has three major weaknesses: several factors, such as the activity levels of pedestrians and emergency vehicles, need to be determined subjectively; all factors are heuristically weighted; and individual site visits are required to obtain many of the factors.

Ryan and Erdman (*9*) suggested including, in addition to accident costs, costs of delay to roadway vehicles and potential emergency-access problems. The three factors are subjectively weighted, and each RHGC is given a score based on a summation of the values of the three weighted factors. The crossings are then ranked on the basis of their total scores. Though this procedure does broaden the basis of the ranking to include factors other than hazard, it has weaknesses: the weightings are subjectively assigned, the procedure ranks crossings instead of improvement projects, and costs and benefits are not explicitly considered.

NEED FOR RESEARCH

Ideally, a priority programming methodology should determine all costs associated with current conditions at a candidate RHGC, anticipated costs of conditions at the crossing if an improvement is made, costs of making the improvement, and the expected net benefit of the improvement. If such a methodology were used, potential projects could be ranked in terms of their expected net benefits, and only those projects with the greatest expected net benefits would be implemented (herein the terms "net benefits," "expected net benefits," and "expected benefits" are used interchangeably).

As far as it is known, however, such a comprehensive methodology is not used—nor does it even exist. Most of the methodologies in use consider only the direct costs of accidents. Thus, a clear need exists for an improved methodology, one based on sound economic principles, so that crossing improvements can be programmed efficiently and scarce resources can be used optimally.

DEVELOPMENT OF PROPOSED METHODOLOGY

In developing a methodology that can be used for efficiently allocating limited resources among RHGC improvement projects, all costs associated with the operation of a RHGC are identified and quantified insofar as possible; in this manner, a monetary value is determined for current conditions for any crossing chosen for analysis. These costs include safety, delay, and emergency access.

The proposed methodology makes use of the costs thus developed in assessing the effectiveness of potential improvements to a given crossing. Upgrading traffic-control devices, providing grade separations, providing additional travel lanes, and closing RHGCs are all options for improvement; the cost reductions resulting from such improvements are identified and quantified. The costs of implementing such improvements are also identified and quantified. The net benefit (the cost of operation with the proposed improvement minus the cost of operation with existing conditions minus the cost of implementation) is computed for each improvement option, and a list of projects with net benefits greater than zero is compiled. Finally, through the use of zero-one integer programming, the set of improvements that optimize net benefits subject to a budget constraint is selected (zero-one integer programming is a specialized form of linear programming; it is an optimatization technique that yields a go/no go decision for each of a set of discrete options).

A methodology that is theoretically flawless but difficult to use would be of little practical value. When a government agency must choose between using an outstanding and complete methodology that requires data collection and using an inefficient but widely accepted technique that requires no data collection (and thus is faster and less expensive), the agency is likely to use the inefficient technique. For this research, it was decided to address this conflict between perfection in theory and practicality in implementation by developing a methodology as complete and theoretically correct as possible while using only those data known to be easily accessible to most agencies responsible for RHGCs.

The data known to be on file with (or readily available to) those agencies come from only two sources: the DOT Rail-Highway Grade Crossing Inventory (to be referred to as "the Inventory"), which consists of data provided by the individual states, and the FRA accident/incident files (FRAIRS). Because of the expense involved in data collection, responsible agencies generally have no information for most of their RHGCs, except that which is provided by the Inventory and FRAIRS. Thus, to have any chance of being practicable to a responsible agency, a methodology for RHGCs should use only information from these two sources or other easily collectible data, or it must generate additional information itself.

Development of Parameters

A detailed description of the portions of the methodology in which the costs of accidents, diversion, and delays to emergency vehicles are computed is presented elsewhere (*10*); it is not included here for the sake of brevity.

Briefly, the methodology computes the direct costs of accidents through the use of an accident prediction formula developed by the Transportation Systems Center (unpublished data), a severity index developed by the same institution (*5*), and NHTSA cost data. Costs of diverting roadway traffic around a blockage are computed through application of limited field data about bypass routes and cost data produced by FHWA (*11*).

Delay costs are computed through application of a deterministic delay model, limited field data, and the cost data described earlier (*11*). Costs of normal operations and maintenance were omitted from the methodology because of a lack of available data. Emergency-access delay costs were developed through application of historical data and unit costs obtained from the literature (*12–16*).

Potential Improvements

A range of options is available for improving each RHGC. In terms of cost, at one extreme would be the installation of crossbucks signs at a completely unprotected RHGC; at the other, the construction of a grade separation. On the basis of the procedures developed or chosen for use in this methodology, the potential improvements in Table 1 was analyzed for each grade crossing. A K-value for each improvement is given in Table 1 as well; the significance of this K-value is explained later. Crossbucks, flashing-light signals, and automatic gates are taken to be as defined in the *Manual on Uniform Traffic Control Devices for Streets and Highways* (*17*).

The list of potential improvements is rather short and does not include several traditional improvements, such as increasing sight distance, improving warning signing, adding flashing-light signals for greater conspicuity, installing flashing-light signals with larger lenses, and improving the crossing surface. Though these and other improvements are certainly worthwhile, the methodology has no mechanism for considering them explicitly. The primary reason for this lack of mechanism is the fact that almost all of these traditional improvements are safety-related and would thus need to be considered in the methodology's accident prediction formula to be included in the methodology. However, none of these possible improvements would affect the value of the selected accident prediction formula. The art of predicting accidents at RHGCs is simply not developed enough to allow consideration of such factors, so there is no way to consider them explicitly in the methodology.

TABLE 1 IMPROVEMENT OPTIONS

K	Improvement Option
1	Existing Conditions
2	Flashing Light Signals
3	Automatic Gates
4	Add 2 lanes to approach roadways (one in each direction)
5	Close roadway
6	Grade Separation in Place

PRIPROG

A computer program executing the cost assessment portion of the desired methodology was developed. This program, which is written in Fortran, is called PRIPROG.FOR. It is up and running on the VAX computer system at the University of Maryland, Baltimore County (UMBC). As external input, PRIPROG.FOR requires the Inventory data and FRAIRS data for the RHGCs in question, and it also requires the user to specify the year for which the analyses are to be performed. All other necessary information is included in the program and based on the preceding analyses and discussions. The program follows the simplified flowchart in Figure 1. The iterations based on the value of K allow for recomputations of costs assuming that a specific set of improvements is made.

Of course, not all six iterations of PRIPROG.FOR are necessarily run for each RHGC: one already equipped with automatic gates, for example, will skip the second and third iterations.

PRIPROG.FOR generates two output files. The first of these files, PRIPROG.DAT, contains the following information for each iteration for each RHGC: Inventory identification number; start and end dates used for accident history analyses; annual ADT (AADT); total number of trains per day; predicted number of accidents per year; accident costs per year; excess vehicle operating costs per year due to delays; excess gallons of fuel consumed per year due to delays; excess pounds of carbon dioxide, hydrocarbons, and nitrous oxides

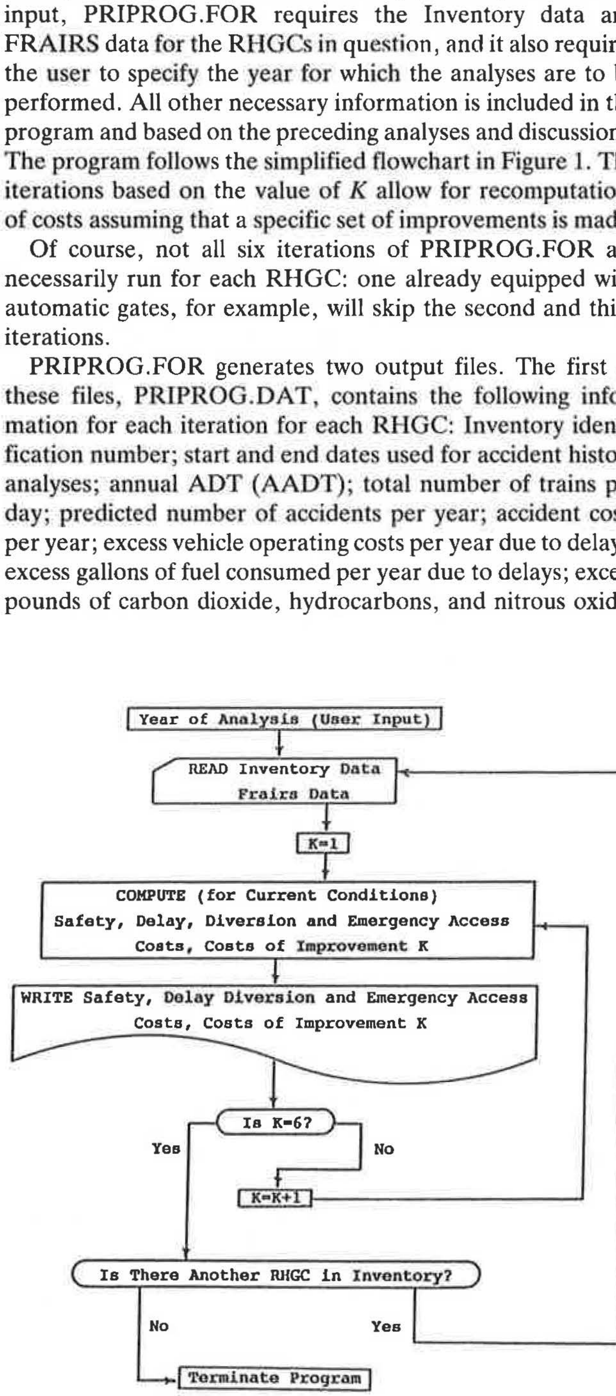


FIGURE 1 PRIPROG.FOR flowchart.

produced per year due to delays; annual delay costs, excess vehicle operating costs, and gallons of fuel consumed on bypass roadways used for diverting traffic; annual pounds of carbon dioxide, hydrocarbons, and nitrous oxides produced on bypass roadways used for diverting traffic; annual costs due to delays to fire equipment; and annual costs due to delays to emergency medical equipment. The reporting of all of this information allows the user to observe the effects of proposed improvements on all key parameters and the relative importance of each parameter to the total costs at each crossing.

The second output file, IMPASS.DAT, contains a summary of the data in PRIPROG.DAT and additional information. For each iteration for each crossing, the following is produced: Inventory identification number; iteration number; total of the monetary costs provided in PRIPROG.DAT (accident, delay, excess operating due to delay, delay on bypass routes, excess operating on bypass routes, and delay to emergency vehicles); the current worth of those costs, assuming an interest rate of 10 percent and a study period of 20 years; installation costs of the proposed improvements; and totals of the fuel consumption and carbon dioxide, hydrocarbon, and nitrous oxides production information in PRIPROG.DAT.

The interest rate of 10 percent appears reasonable in light of current economic conditions (the user can, of course, modify the interest rate as desired). The study period of 20 years was chosen for several reasons, among them that highway projects are typically assumed to have a useful life of about 20 years and that data gathered by FRA indicate that the life expectancy of active warning devices is about 20 years. The signs and barricades that would make up the major portion of the costs of closing a RHGC generally have a useful life of 7 to 10 years. For the purposes of this paper, it was assumed that all signs and barricades would need to be replaced at 10 years. Examination of IMPASS.DAT allows the user to assess, in a six-line format, the conditions at each RHGC and the effects of each potential improvement on that RHGC.

LINDO

LINDO (Linear, Interactive, and Discrete Optimizer) is a software package available on UMBC's VAX system. It is used in this methodology to perform the zero-one integer programming functions described earlier as being desirable for matching potential improvements with particular RHGCs. Of course, other software systems can perform the same functions and can be used in place of LINDO; the decision to use LINDO was based simply on its availability.

Before executing LINDO, some simple arithmetic functions are performed on the IMPASS.DAT file. These functions are performed in PRIPROG.FOR, and the results are written to a file named BENCOMP.DAT. The costs of each RHGC under each option are compared with existing costs; only those options that reduce the costs are analyzed further. The present worth of the reduction in costs is then compared with the implementation cost of the option, and only those options for which the reduction in cost exceeds the implementation cost are written to BENCOMP.DAT. Next, the information in BENCOMP.DAT is written in a format that, with minor adjustments by the user, is directly usable by LINDO. This final file generated by PRIPROG.FOR is called LINDIN.DAT.

LINDO then performs the zero-one integer programming analyses and yields the optimal set of improvements for the RHGCs under study. Strictly speaking, the formulation of the zero-one integer programming problem solved by LINDO in this research is as follows:

Maximize

$$\sum_{i=1}^{n} \quad \sum_{k=1}^{6} x_{ik} AB_{ik}$$

subject to

$$\sum_{i=1}^{n} \quad \sum_{k=1}^{6} x_{ik} C_{ik} \leq B$$

$$X_{ik} = 0 \text{ or } 1 \quad \text{for} \quad k = 1, 2, \ldots, 6$$
$$\text{and} \quad i = 1, 2, \ldots, n$$

$$\sum_{k=1}^{6} x_{ik} \leq 1$$

where

i = crossing under consideration,
n = total number of crossings,
k = improvement option under consideration,
AB_{ik} = expected benefit of option k at crossing i,
C_{ik} = implementation cost of option k at crossing i,
B = budget, and
X_{ik} = go/no go variable (limited to the values 0 or 1).

The final constraint ensures that the methodology selects no more than one option per crossing. This is necessary to avoid mathematically correct but physically meaningless selections, such as installing flashing-light signals at a RHGC that is to be closed.

APPLICATION OF METHODOLOGY

It was believed that it would be useful to compare the improvement projects yielded by the methodology with those projects actually implemented by an agency. Data for one jurisdiction's improvement projects for 1985 were obtained, and the methodology was executed, using all the crossings in the jurisdiction and the same budget ($1,232,500).

Data on the agency's projects, as computed by the methodology, are given in Table 2. Examination of Table 2 reveals that 17 improvements projects were implemented and the total expected net benefit was approximately −$880,000.

These negative benefits are caused by two factors. First, the installation of active traffic-control devices can actually increase delays at crossings, thus increasing delay costs. Second, particularly when AADTs on the intersecting roadways are low, the accident prediction formula used in the methodology may forecast an increase in accident rate when active traffic-control devices are installed.

TABLE 2 AGENCY PROJECTS

RHGC	Improvement (See Table 1)	Expected Benefits (1985 Dollars)
A	2	-72,177.13
B	2	-65,312.70
C	2	-68,112.31
D	3	-10,738.97
E	3	-30,052.14
F	2	-75,075.18
G	3	-36,757.87
H	2	-53,689.75
I	2	-72,048.76
J	3	76,752.35
K	2	-76,576.92
L	2	-77,944.89
M	2	-81,154.63
N	2	-79,322.33
O	3	-33,780.83
P	3	-2,800.01
Q	3	-122,560.81

NOTE: RHGC is rail-highway grade crossing.

Data on the methodology-chosen projects are given in Table 3. Examination of Table 3 reveals that 24 improvement projects would be implemented, having a total expected net benefit of approximately $6,600,000. Thus, in terms of costs and benefits as computed by the proposed methodology, the proposed methodology is a major improvement over the technique the agency now uses.

AVAILABILITY AND USAGE

Copies of PRIPROG.FOR are available from the author. Potential users of the program are cautioned that a number of assumptions, which are based on knowledge of local conditions, are incorporated into the program. These assumptions should be reviewed for compatibility with a user's local conditions.

SUMMARY

The objective of this research was to develop a comprehensive methodology to assist in the priority ranking of improvements to rail-highway grade crossings. The methodology developed and presented herein fulfills the stated objectives, subject to the constraints of existing data bases. Accident costs, delay costs, diversion costs, and costs to emergency vehicles are considered, and up to five improvement projects per crossing are evaluated. The benefit and cost computations are made in a Fortran computer program developed in this research; the program requires two common data bases (the Inventory and FRAIRS) for each crossing under consideration. Several sets of intermediate output can be obtained from the program to give the user detailed information about the options under consideration; the final output can be input directly to a preexisting linear programming computer package, after minor clerical adjustments are made.

The methodology was applied successfully to 1985 conditions at all RHGCs in a particular jurisdiction. The results of this application indicate that, for the same total budget, expected net benefits could have been approximately $7 million higher if the methodology's projects had been implemented instead.

In summary, the methodology developed in this research is a major improvement over at least one current technique for priority ranking. The explicit inclusion of delay costs, diversion costs, and costs of delays to emergency vehicles within such a technique is a significant step forward. In addition, the introduction of zero-one integer programming directly into the process greatly improves the cost-effectiveness of the selected projects and thus of the improvements program as a whole.

TABLE 3 METHODOLOGY PROJECTS

RHGC	Improvement (See Table 1)	Expected Benefits (1985 dollars)
R	3	265,407.63
S	3	285,193.88
T	3	213,710.34
U	5	155,979.75
V	3	291,027.88
W	3	286,110.00
X	3	268,465.13
Y	3	268,846.38
Z	3	276,767.19
AA	3	274,237.63
BB	3	293,343.88
CC	3	289,030.63
DD	3	264,602.88
EE	3	281,623.50
FF	4	140,561.56
GG	3	270,139.00
HH	3	265,888.31

TABLE 3 (*continued on next page*)

TABLE 3 (*continued*)

RHGC	Improvement	Expected Benefits (1985 dollars)
II	3	209,220.50
JJ	2	424,242.75
KK	4	251,307.00
LL	4	116,992.13
MM	4	722,403.00
NN	4	346,201.25
OO	4	108,796.38

NOTE: RHGC is rail-highway grade crossing.

REFERENCES

1. H. A. Richards and D. L. Hooks. Establishing Priorities for the Installation of Traffic Control Devices: The Rail-Highway Intersection Example. In *Highway Research Board Special Report 107*, HRB, National Research Council, Washington, D.C., 1970, pp. 70–80.
2. H. A. Richards and J. T. Lamkin. *Statistical and Economic Aspects of Rail-Highway Grade Crossing Safety Improvement Programs in Texas*. Report 111-2. Texas Transportation Institute, College Station, Nov. 1970.
3. W. R. Schulte. Effectiveness of Automatic Warning Devices in Reducing Accidents at Grade Crossings. In *Transportation Research Record 611*, TRB, National Research Council, Washington, D.C., 1976, pp. 49–57.
4. M. D. Harness and K. C. Sinha. Setting Priorities of Highway Projects by Successive Subsetting Technique. In *Transportation Research Record 955*, TRB, National Research Council, Washington, D.C., 1984, pp. 35–40.
5. J. Hitz. *Rail-Highway Crossing Resource Allocation Procedure—Users Guide*, 2nd ed. DOT-TSC-FHWA-85-2. FHWA, U. S. Department of Transportation, July 1986.
6. *Railroad-Highway Grade Crossing Handbook*. FHWA-TS-78-214. FHWA, U. S. Department of Transportation, 1979.
7. J. E. Tidwell, Jr. and J. B. Humphreys. Improving Safety at Passive Crossings with Restricted Sight Distance. In *Transportation Research Record 841*, TRB, National Research Council, Washington, D.C., 1982, pp. 29–36.
8. *Colorado State Rail Plan 81-82 Update*. Colorado Department of Highways, Denver, 1982.
9. T. A. Ryan and J. W. Erdman. A Procedure for Priority Ranking System for Rail-Highway Grade Crossings. In *Transportation Research Record 1010*, TRB, National Research Council, Washington, D.C., 1985, pp. 117–120.
10. T. A. Ryan. *A Priority Programming Methodology for Rail-Highway Grade Crossings*. Ph.D. dissertation. Department of Civil Engineering, University of Maryland, College Park, Dec. 1988.
11. C. W. Dale. *Procedure for Estimating Highway User Costs, Fuel Consumption and Air Pollution*. FHWA, U. S. Department of Transportation, 1980.
12. J. D. Mayer. Emergency Medical Service: Delays, Response Time and Survival. *Medical Care*, Vol. 17, No. 8, Aug. 1979, pp. 818–827.
13. W. D. Weaver et al. Factors Influencing Survival After Out-of-Hospital Cardiac Arrest. *Journal of the American College of Cardiology*, Vol. 7, No. 4, April 1986, pp. 752–757.
14. W. D. Weaver et al. Considerations for Improving Survival from Out-of-Hospital Cardiac Arrest. *Annals of Emergency Medicine*, Vol. 15, No. 10, Oct. 1986, pp. 1,181–1,186.
15. E. Ignall et al. *Fire Severity and Response Distance: Initial Findings*. Rand Institute, New York, N.Y., Aug. 1978.
16. J. Hogg. *Losses in Relation to the Fire Brigade's Attendance Times*. Report 5/73. Scientific Advisory Branch, Home Office, London, England, 1973.
17. *Manual on Uniform Traffic Control Devices for Streets and Highways*. FHWA, U. S. Department of Transportation, 1978.

Publication of this paper sponsored by Committee on Rail-Highway Grade Crossings.

Low-Clearance Vehicles at Rail-Highway Grade Crossings: An Overview of the Problem and Potential Solutions

RONALD W. ECK AND S. K. KANG

There are no readily available highway design standards aimed at providing adequate ground clearance at rail-highway grade crossings that have humplike profiles. As a result, low-clearance vehicles—such as low-bed equipment trailers, automobile transporters, and car- and truck-trailer combinations—can become lodged or hung up at a crossing. A number of accidents of this type have been reported in recent years, but a literature search indicated that there is very little quantitative data on the magnitude of the problem. Traffic count data from West Virginia indicate that low-clearance vehicles make up about 2 percent of the traffic stream. Such vehicles are highly variable in their physical dimensions: ground clearances as low as 2 in. for a variety of wheelbases have been reported. A literature review is presented summarizing approaches to this problem that have been taken. These include specifying crossing physical characteristics and developing advance warning signs. In response to an identified need, the researchers developed microcomputer software that incorporates graphics and animation capabilities to simulate the movement of trucks over grade crossings and to predict where hang-ups will occur for a given crossing geometry. The software is described and its use demonstrated in a sample application.

Safety at the more than 340,000 public and private rail-highway grade crossings in the United States has long been a concern of highway agencies, railroads, and local communities. Crossing safety has been addressed from several standpoints, including the development and use of formal techniques for identifying hazardous crossings and the subsequent upgrading of warning-device crossing protection at many of these crossings. The top of traffic-control devices for grade crossings has been studied extensively, resulting in increased use of active warning devices, improvements in track circuitry and control logic, and the installation of advance warning signs and pavement markings.

According to the *Rail-Highway Crossings Study* (*1*), the total numbers of accidents, fatalities, and injuries at rail-highway crossings have all declined significantly since 1970, despite a steady increase in highway vehicle miles traveled. The study noted that although the total number of crossing fatalities is decreasing along with the total number of accidents, the average number of fatalities per accident has increased. This indicates that crossing accidents, already the most severe kinds of highway accident, are becoming even worse.

In spite of these substantial accomplishments, certain types of crossing accidents still occur with disturbing frequency. One type involves long truck trailers with low ground clearances becoming lodged or hung up on a crossing at which the grades of the crossing and its approaches are not adequate. These vehicles—which include low-bed equipment trailers, car carriers, mobile homes, and car- and truck-trailer combinations—are relatively common in the traffic stream. Because of the size of the vehicle involved or the load being carried, such accidents usually result in substantial property damage and loss of life. Accidents of this type have been reported in recent years in the news media. The problem can be expected to increase in the future, given the current diversification of the truck fleet.

The National Transportation Safety Board (NTSB) has investigated several of these accidents (*2,3*). NTSB found that the American Railway Engineering Association's recommended practice on the profile and alignment of crossings and approaches is not generally available and thus not usually followed. A literature review (to be discussed later in more detail) revealed there are no highway design standards aimed at providing adequate ground clearance on highways or at grade crossings having humplike profiles. There is a need to develop a standard roadway profile design for rail-highway grade crossings. Furthermore, for those grade crossings identified as having hazardous surface profiles, a standard warning sign should be developed to alert truck drivers to the problem.

STUDY OBJECTIVES

To address these problems, a research effort is under way that has the following objectives: (a) to determine the extent of the hang-up and accident problem involving low-clearance vehicles at grade crossings, (b) to identify specific classes or categories of vehicles with low ground clearance, (c) to develop a computer model for checking whether a given class of vehicle can safely negotiate a particular grade crossing profile and for redesigning a hazardous crossing profile, and (d) to develop, on the basis of the results of these computer runs, specific design criteria for crossing profile alignment to accommodate the classes of vehicles identified as having problems at grade crossings.

This paper presents results of the research effort so far. Each of the objectives identified will be addressed, with the exception of the design criteria for grade crossing profiles. This work has just been initiated and will be reported on its completion.

Department of Civil Engineering, P.O. Box 6101, West Virginia University, Morgantown, W.Va. 26506.

PREVIOUS WORK

The first task of the research was to estimate the magnitude of the hang-up problem and to identify any previous work on solving the problem. The effort took the form of a literature search and telephone and personal contacts with personnel from highway and railroad agencies and the trucking industry.

Magnitude of Problem

It became clear early on that there is very limited information available on the hang-up problem. Certain severe accidents have been publicized by the media and investigated by NTSB. In telephone conversations, personnel from the Public Utility Commission of Oregon indicated that Oregon averages about one accident a year in which a low-clearance vehicle is hung up on the tracks and struck by a train. However, from accident data in general, it is difficult to identify which accidents are the result of low-clearance vehicles' becoming lodged on the tracks.

Perhaps a more significant, but difficult to ascertain, quantity is the number of vehicles that experienced hang-ups but did not have an accident because the truck was freed before a train arrived or because the train stopped before it passed through the crossing. None of the agencies contacted by the researchers was aware of data of this type. If such data exist, it would most likely be in isolated instances in which a local law enforcement agency or towing operator maintained such statistics.

The researchers identified certain indicators that suggest the problem of low-clearance vehicles at grade crossings is a significant one. One of these is the fact that NTSB has investigated several serious accidents of this type and made a series of recommendations relative to the problem (*4*). Discussions with local and state highway agency personnel and trucking company officials indicate that even though hard data are lacking, the problem is thought to be significant. For example, the Mid-Atlantic region safety director for a trucking company that transports automobiles noted that his fleet experiences 50 to 60 hang-up incidents per month. He could not break this down by type of hang-up—for example, railroad crossings, parking lot entrances, or pavement crowns—but the numbers are significant nonetheless and demonstrate the need for geometric design guidelines for ground clearance. A third indicator is that the Office of Motor Carriers of FHWA recently issued an "On Guard" advisory to truck drivers advising them to be aware of the problems posed by low-ground-clearance vehicles at grade crossings.

Approaches to Problem

As early as 1958 (*5*) there is mention in the highway engineering literature of vehicle ground-clearance problems. Bauer (*6*) described the problem of insufficient ground clearance for automobiles traversing driveway entrances in suburban or residential areas. He proposed the use of a 2-in. safety margin for vehicle ground clearance to accommodate the downward thrust that cars experience when traversing the varying profile grades of driveways and when braking. Because this was during the precomputer era, he recommended a manual procedure for checking designs: it involved cutting out a model car using a piece of paper at the same scale as the profile. The model could be slid along the profile to find any trouble spots and the profile adjusted as necessary.

After its investigation of a 1983 accident in North Carolina, NTSB (*4*) warned that crossing profiles with humplike vertical curves can impede the operation of a vehicle if the distance between any two axles of a vehicle spans the hump and the height of the hump exceeds the vehicle's ground clearance. The report recommended that grade crossings that have a potentially hazardous roadway profile should be identified, and that, subsequently, improvements should be made.

The American Railway Engineering Association's *Manual for Railway Engineering* (*7*) states that it is desirable that the surface of the highway be neither more than 3 in. higher nor more than 6 in. lower than the top of nearest rail at a point 30 ft from the rail, measured at right angle thereto, unless track superelevation dictates otherwise. There is no evidence that these guidelines were used in any of the accidents investigated by NTSB, most likely because the document is not generally available to highway designers. Similar guidelines could not be found in publications of AASHTO (*8*) or FHWA. Note that these guidelines do appear in the railroad grade crossing section of the 1990 AASHTO policy on geometric design.

The State of Alabama Highway Department (*9*) standardized a safe treatment at crossings. The agency developed an advance warning sign for low-clearance trucks and trailers that includes an arrow indicating a detour. Halfway from this sign to the crossing is a pictorial sign showing a truck stuck on a railroad track; below it reads the message Low-Clearance Railroad Crossing. The guidelines include a figure that presents allowable grades for various approach configurations.

As a result of several hang-up accidents and the danger of various types of trailers striking or getting stuck at railroad crossings, the North Carolina Department of Transportation has developed a new symbol-type warning sign to address the situation. The sign, shown in Figure 1, has been submitted for approval to the National Committee on Uniform Traffic Control Devices. It would be a standard 30- × 30-in. diamond-shaped warning sign used at grade crossings with insufficient clearance for low-bed trailers. The sign would be supplemented with the word message Trailers May Drag.

The NTSB investigations and findings pertaining to crossing accidents prompted the Florida Department of Transportation (*10*) to develop and implement a program to eliminate the hazard at crossings with high-profile surfaces in that state. The program included profile criteria, encouraged coordination between government and the rail industry in building and maintaining grade crossings, and developed measures to identify crossings that would not accommodate low-ground-clearance vehicles. In addition, the department designed an advance warning sign that depicts a truck lodged on a grade crossing. As high-profile crossings were identified, signs were installed on the approaches.

On the basis of the Florida effort, NTSB (*4*) issued a safety recommendation to AASHTO. The board recommended the adoption of Florida's program or a comparable program developed by an appropriate AASHTO committee. NTSB also issued a safety recommendation for advance warning signs

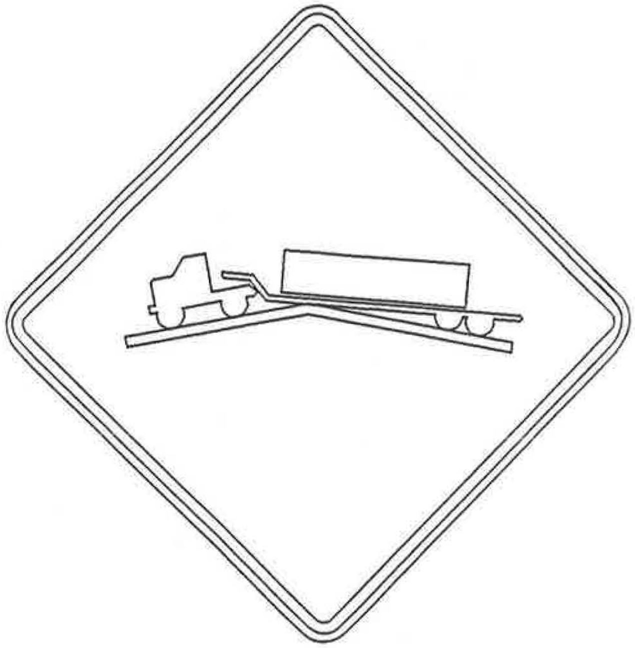


FIGURE 1 Low-clearance-vehicle warning sign proposed by North Carolina Department of Transportation.

for high-profile surfaces, because the *Manual on Uniform Traffic Control Devices (11)* does not elaborate on this issue. In response to the North Carolina request, the National Committee on Uniform Traffic Control Devices is currently researching the development of a sign for this situation.

Several European countries have geometric design guidelines for rail-highway grade crossings (*12*). England provides a circular curve roadway-crossing profile. There are three categories of radii, depending on traffic volumes and traffic moment (the product of vehicular and rail traffic). For example, the profile requirement for Category 1 is to provide a 1,250-ft nominal radius profile. This profile has a 1.75-in. elevation over a 38-ft wheelbase to which must be added a 3-in. construction and maintenance tolerance so that a vehicle with a 6-in. ground clearance can pass over the crossing without lodging or grounding. The French provide a sign—Low Vehicles Proceed with Caution—on high-profile crossings if the crossing has a rising gradient exceeding 6 percent or an elevated (hogged) radius of curvature of less than 164 ft.

RESEARCH APPROACH

Collection of Truck Data

Because of the lack of available data about the magnitude of the low-clearance vehicle incident problem and about low-clearance vehicle characteristics, the researchers decided to collect data on such vehicles. One part of this effort involved performing a vehicle classification count on Interstate 79, a major north-south route in West Virginia. The intent was to estimate at least the potential magnitude of the problem by determining the types and quantities of low-clearance vehicles on particular highways. The count was made along I-79 just south of the Pennsylvania border in May 1990. Northbound and southbound traffic were counted. General classes of vehicles were categorized, and a detailed classification of trucks was made. For example, trucks were categorized as van semitrailers, dry-bulk semitrailers, liquid tankers, acid tankers, pole trucks, ready-mix concrete trucks, and flatbed semitrailers, among others. Low-clearance trucks included double-drop low-bed equipment trailers, boat transporters, automobile transporters, and double-drop livestock trailers. Counts were made 4 hr a day for 7 days.

Literature from low-bed truck trailer manufacturers was also acquired. These data illustrated the wide variety in dimensions of low-clearance vehicles. The lowest published ground clearances were on the order of 8 in., but data collected at a port-of-entry on I-5 in Oregon showed ground clearances as low as 2 in. To explore these characteristics further, it was decided to collect additional wheelbase and ground-clearance data in West Virginia.

With the assistance of the weight enforcement unit of the West Virginia Division of Highways (WVDOH), truck wheelbase and ground-clearance data were collected at two locations. Low-clearance vehicles (defined here as any vehicle with a ground clearance of 9 in. or less) using the East Fairmont weigh station on I-79 in northern West Virginia were examined. Vehicle wheelbase (defined as the distance from the center of the rear axle of the tractor to the center of the front axle of the trailer) and ground clearance (defined as the vertical distance from the ground to the lowest point on the vehicle between the rear axle of the tractor and the front axle of the trailer) were measured. Whenever possible, the field crew asked the drivers whether they had ever been hung up or knew of colleagues who had. Additional truck characteristics data were also collected on US-48, a major east-west route through West Virginia and Maryland.

Development of Microcomputer Software

The objective of this task was to incorporate appropriate graphics and animation capabilities to develop microcomputer software for simulating the movement of trucks over grade crossings. The roadway profile geometry required for each crossing should be such that it could be acquired economically in the field and readily input to a crossing geometry data base. It would also be desirable if the user could easily modify the input data to permit evaluation of different design alternatives. Part of this task involved selecting an appropriate programming language or software package and specifying hardware requirements.

Because none of the existing software packages reviewed met the researchers' criteria, a program was written. The package developed, referred to as "HANGUP," is intended to run on the IBM-PC or compatible computers. It was programmed using the Microsoft QuickBASIC compiler Version 4.5 under MS-DOS Version 3.3. The executable HANGUP software occupies about 120 kilobytes of random-access memory and requires 300 kilobytes of total free memory. Hardware requirements include a VGA screen, a printer, and a plotter.

A simplified flowchart for the package is shown in Figure 2. HANGUP is completely menu-driven and designed to be user-friendly. The main menu, shown in Figure 3, illustrates the capabilities available to users. Note that unfamiliar users

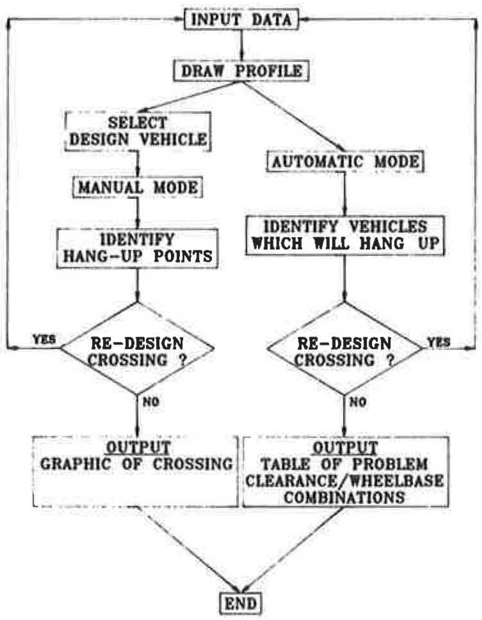


FIGURE 2 Simplified flowchart for HANGUP.

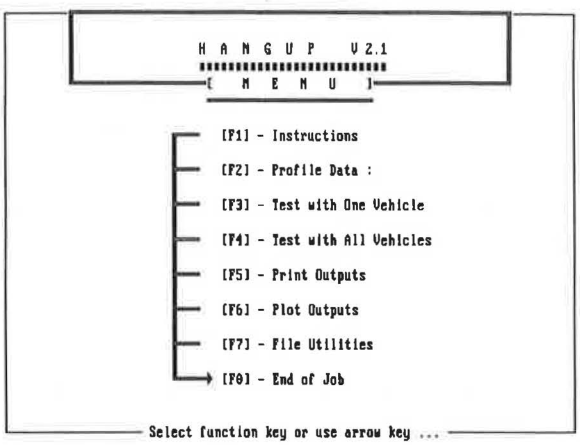


FIGURE 3 Main menu indicating capabilities of HANGUP.

can retrieve help or an instruction routine simply by pressing the F1 function key.

The program plots user-entered roadway profile data and graphically presents vehicle movement over the roadway, either on screen or on plotter. The program can use any data file from a predefined DOS subdirectory as the input or output file. The files do not have to be on the same logged directory, which may be A, C, or any other logical device. The plots allow highway and traffic engineers to determine sections of the roadway where hang-up problems will occur because of low ground clearance or vehicle overhang.

RESULTS

Characteristics of Low-Clearance Vehicles

Results of the vehicle classification counts on I-79 indicated that during the period studied, trucks made up slightly more than 13 percent of the traffic stream. Low-clearance trucks accounted for 0.8 percent of the traffic stream (or about 5.7 percent of all trucks). Obviously, this varied with the day of the week. The highest percentage of trucks (17 percent) occurred on Tuesday; the lowest percentage of trucks occurred on Sunday (6 percent). Low-clearance trucks followed an identical pattern, making up 1.2 and 0.2 percent of the traffic stream on Tuesday and Sunday, respectively.

Table 1 gives a more detailed breakdown of the classification data to provide a comparison of low-clearance vehicles with other types of vehicles on the roadway. Note that the table presents selected vehicles that account for just less than 96 percent of the total volume.

A few of the vehicle types listed in Table 1 deserve additional explanation. Double-drop vehicles refer to low-bed, also known as low-boy, vehicles designed to haul heavy equipment and oversized items. Single-drop vehicles are also low-bed vehicles, but they are not as low as the double-drop rigs; examples include warehouse, furniture, and electronics vans. Single-unit trucks with trailers refer to straight trucks, such as dump bodies or flatbeds, towing a trailer (e.g., a piece of construction equipment) by a hitch mechanism rather than a fifth wheel connection.

When taken individually, each type of low-clearance truck makes up a negligible portion of the traffic stream. However, in the aggregate, the numbers are worthy of attention.

As expected for I-79, low-bed equipment trailers were the predominant type of low-clearance truck, followed by automobile transporters. Obviously, the relative frequency will vary with the facility because certain roadways, because of their proximity to industrial or agricultural centers, may handle more (or less) of a particular type of low-clearance vehicle.

As this paper has suggested, there is at least anecdotal evidence that car- and pickup-truck-trailer combinations are also susceptible to hanging up at grade crossings. Table 1 indicates that such combinations accounted for 1.1 percent of the vehicles in the classification count. If single-unit trucks with trailers are also considered (0.2 percent), it is apparent that about 2 percent of the traffic stream can be considered in the low-clearance vehicle category.

TABLE 1 REPRESENTATION OF SELECTED VEHICLES IN TRAFFIC STREAM BASED ON 7-DAY CLASSIFICATION COUNT ON I-79 NEAR PENNSYLVANIA BORDER, MAY 1990

Vehicle Class	Number	Percentage
Passenger Cars	17,411	61.7
Autos with Trailers	47	0.2
Pick-up Trucks and Vans	6,600	23.4
Pick-ups with Trailers	245	0.9
Van Semi-Trailers	1,478	5.2
Twin Trailers	74	0.3
Dry-Bulk Semi-Trailers	139	0.5
Liquid Tankers	110	0.4
Flatbed Semi-Trailers	532	1.9
Single Unit Dump Trucks	87	0.3
SU Trucks with Trailers	63	0.2
Double Drop Low Boys [a]	44	0.2
Double Drop Van Semi-Trailers [a]	25	0.1
Double Drop Livestock Trailers [a]	6	0.0
Auto Transporters [a]	30	0.1
Boat Transporters [a]	5	0.0
Multiple Unit Low Boys [a]	0	0.0
Single Drop Low Boys [a]	85	0.3
Single Drop Van Semi-Trailers [a]	17	0.1

[a] Low-Clearance Trucks

Although the relative percentage of low-clearance vehicles in the traffic stream will vary with the nature of the route (Interstate, arterial, local service, etc.), geographic region of the country, and season of the year, among other things, these results are judged to be important. The data indicate that although low-clearance vehicles perhaps are not a significant proportion of the traffic stream, they do occur with enough frequency to warrant consideration by highway designers and traffic engineers.

The characteristics of these vehicles can be further defined by the wheelbase and ground-clearance data. As noted earlier, the researchers acquired data from 114 vehicles that were measured at a port-of-entry in Oregon. The data demonstrated a great deal of variability; no pattern or relationship between wheelbase and ground clearance could be discerned. To supplement this data base, additional information was collected in West Virginia for 42 vehicles as shown in Table 2. Both data sets were merged and a wheelbase–versus–ground-clearance frequency plot was prepared. This is shown in Figure 4. Again, the lack of any pattern is readily apparent. It is interesting to note that ground clearances as low as 2 in. were identified. Although not included in the data base, it was reported to the researchers that some low-boy trailers operate with as few as 3 in. of ground clearance for a 47-ft wheelbase.

Besides stopping the trucks for the measurements, researchers initiated informal discussions with drivers in an attempt to learn more about the hang-up problem. Results confirmed that the problem is widespread. Virtually every driver either had experienced a hang-up (not necessarily at a grade crossing) or knew a truck driver who had. Several drivers indicated that hang-ups happen quite often. One driver got out of the cab and pointed to a gouge mark on his low-boy trailer. He stated that the damage was the result of striking a railroad track at a steel mill near Chicago. He also said that this damage was minor compared with the 30 ft of track that the rig tore up.

Most of the ground-clearance data presented are for loaded, as opposed to empty, vehicles. It is interesting to note that the loading condition turned out to be not as significant a factor as originally anticipated. A substantial number of rigs, especially automobile transporters, provide the same ground clearance when loaded as they do when empty.

On the basis of the literature review and field data collection, the researchers developed a "first cut" categorization of design low-clearance vehicles to be used in assessing crossing profiles. The general categories are (a) double-drop equipment trailers, (b) automobile transporters, (c) double-drop van trailers, and (d) car- and truck-trailer combinations. Sketches of each type of vehicle are presented in Figure 5. As more field data are collected, design wheelbase and ground-clearance values will be established for each class of design vehicle.

TABLE 2 VEHICLE CLEARANCE AND WHEELBASE DATA OBTAINED IN WEST VIRGINIA SURVEY OF LOW-CLEARANCE VEHICLES (MAY AND JUNE 1990 ON I-79 AND US-48 IN NORTHERN WEST VIRGINIA)

clearance(in)	wheel base (ft)						
9.5	43.8	31.6					
9	35.0	29.7	30.7				
8.5	35.5						
8	40.8	40.0	37.5	32.5	32.4	27.6	
7.25	33.4						
7	38.4	38.0	35.5	34.8	32.7	28.9	26.6
6.75	33.6	28.8					
6	33.5	31.4	31.3	30.5	29.9	29.5	28.2
5.75	26.0						
5.5	35.0	31.8	30.0	28.5			
5	34.6	31.1					
4.75	38.8	35.0					
4.5	32.5	30.6					
4	31.8						
3	36.0						

Implementation of Software

The computer software package developed to simulate the movement of low-clearance vehicles over grade crossings was implemented on a trial basis to check certain potentially hazardous grade crossings for a construction contractor, the Oregon Public Utilities Commission, WVDOH, and a short-line railroad. This section will describe one of these applications to illustrate the capabilities of the software.

A multiple-track mainline crossing in Huntington, West Virginia, came to the attention of the researchers. Mobile homes are often moved across the tracks at this location; a number of hang-ups have been reported. In at least one instance a rig was struck by a train. In an attempt to alleviate the problem, WVDOH engineers had proposed an alternative design. Residential driveways along the road severely limited the changes that could be made to the roadway profile. It was believed that this would be a good place to try to implement the software package.

WVDOH personnel had collected roadway profile data for the approach in question. This information was entered as input data to the model. The program was first run in the manual mode to obtain an animated view of various vehicles negotiating the crossing. As expected, a number of hang-up points were indicated. Figure 6 represents plotter output from a manual-mode run of the model for a truck-trailer combination with a 30-ft wheelbase and a 5-in. ground clearance. The vertical arrows indicate locations where this rig would lodge on the crossing.

To determine which combinations of wheelbase and ground clearance would successfully negotiate the crossing, an automatic-mode run was made. As with the manual mode, these results can be viewed graphically on the screen and via hard-copy output. Figure 7 presents the hard-copy version of the automatic mode output for the crossing in question. Note that the 1s in the table mean the rig will hang up, and the 0s mean that the rig will safely negotiate the crossing.

Program output confirmed that hang-up problems exist for a wide variety of vehicles at this crossing. However, the real value of the package is in allowing the engineer to explore different design alternatives. In this case, the agency solution was to add several inches of asphaltic concrete over about a 70-ft section of the approach. On the basis of the profile data furnished for the improvement, the researchers ran HANGUP again. Figures 8 and 9 present the results. Figure 8 clearly shows that a truck with a wheelbase of 30 ft and 6

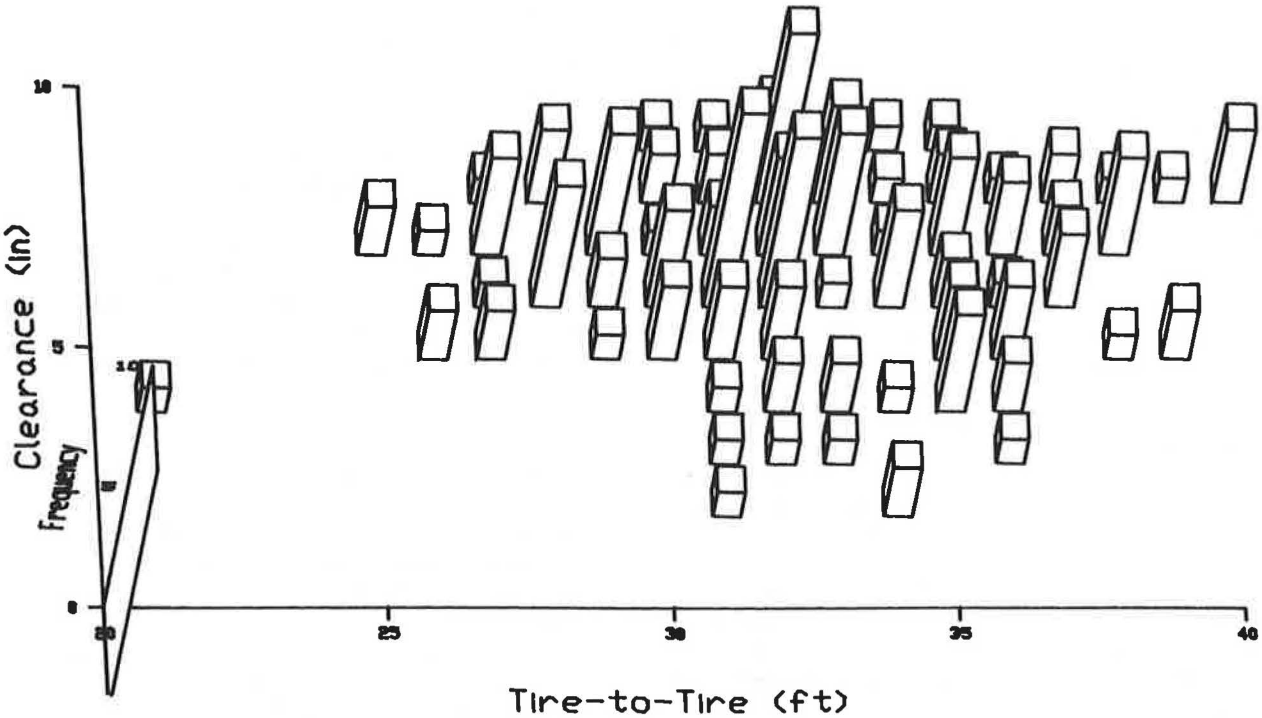


FIGURE 4 Wheelbase–versus–ground-clearance frequency histogram for combined Oregon (114 vehicles) and West Virginia (42 vehicles) data.

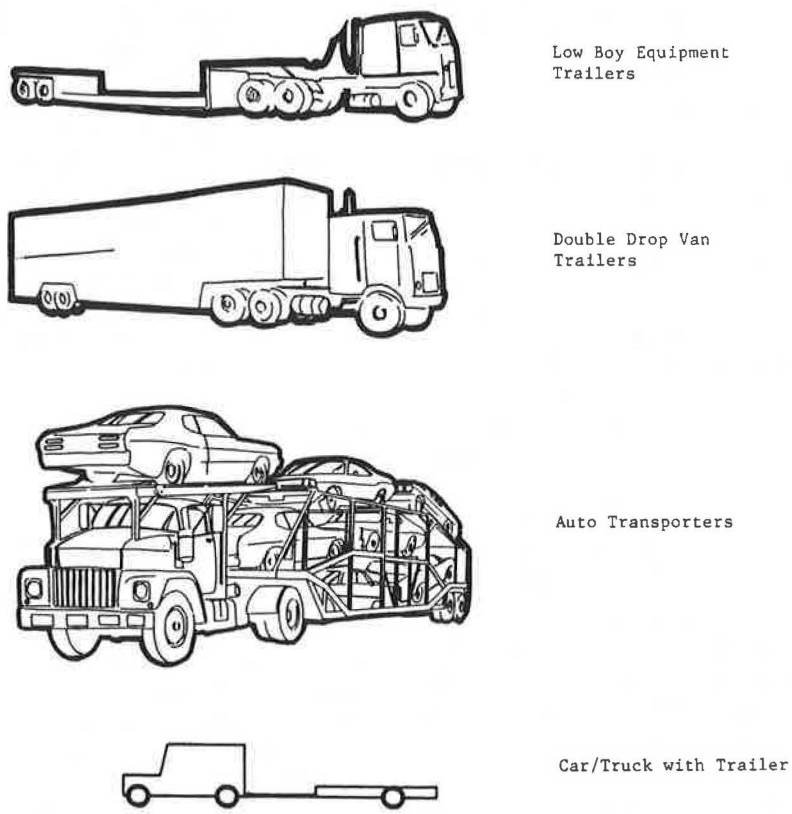


FIGURE 5 Categories of low-clearance vehicles identified in research.

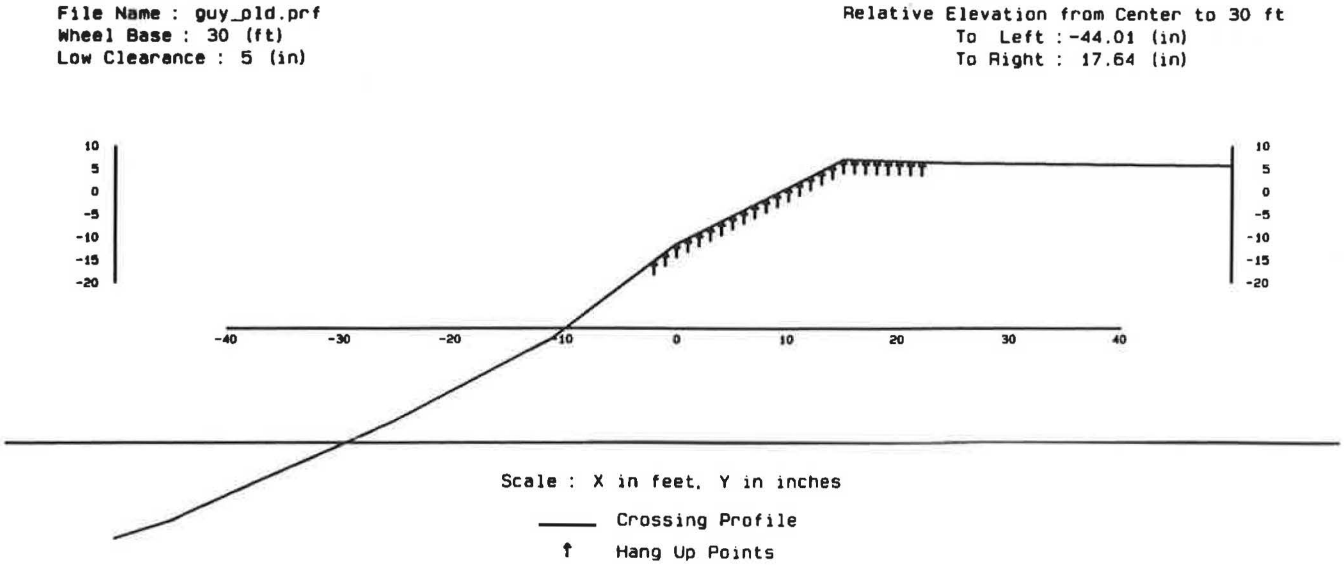


FIGURE 6 Manual-mode output of HANGUP for example crossing in Huntington, W.Va.

H A N G U P

File Name : a:guy_old.prf Date : 07-24-1990

	Ground Clearance (in)									
Wheel Base	1	2	3	4	5	6	7	8	9	10
10 (ft)	1	1	1	0	0	0	0	0	0	0
11 (ft)	1	1	1	0	0	0	0	0	0	0
12 (ft)	1	1	1	0	0	0	0	0	0	0
13 (ft)	1	1	1	1	0	0	0	0	0	0
14 (ft)	1	1	1	1	0	0	0	0	0	0
15 (ft)	1	1	1	1	0	0	0	0	0	0
16 (ft)	1	1	1	1	1	0	0	0	0	0
17 (ft)	1	1	1	1	1	0	0	0	0	0
18 (ft)	1	1	1	1	1	0	0	0	0	0
19 (ft)	1	1	1	1	1	1	0	0	0	0
20 (ft)	1	1	1	1	1	1	0	0	0	0
21 (ft)	1	1	1	1	1	1	0	0	0	0
22 (ft)	1	1	1	1	1	1	1	0	0	0
23 (ft)	1	1	1	1	1	1	1	0	0	0
24 (ft)	1	1	1	1	1	1	1	0	0	0
25 (ft)	1	1	1	1	1	1	1	1	0	0
26 (ft)	1	1	1	1	1	1	1	1	0	0
27 (ft)	1	1	1	1	1	1	1	1	0	0
28 (ft)	1	1	1	1	1	1	1	1	1	0
29 (ft)	1	1	1	1	1	1	1	1	1	0
30 (ft)	1	1	1	1	1	1	1	1	1	1
31 (ft)	1	1	1	1	1	1	1	1	1	1
32 (ft)	1	1	1	1	1	1	1	1	1	1
33 (ft)	1	1	1	1	1	1	1	1	1	1
34 (ft)	1	1	1	1	1	1	1	1	1	1
35 (ft)	1	1	1	1	1	1	1	1	1	1
36 (ft)	1	1	1	1	1	1	1	1	1	1
37 (ft)	1	1	1	1	1	1	1	1	1	1
38 (ft)	1	1	1	1	1	1	1	1	1	1
39 (ft)	1	1	1	1	1	1	1	1	1	1
40 (ft)	1	1	1	1	1	1	1	1	1	1

1 -> Hang up, 0 -> Safe

FIGURE 7 Automatic-mode output of HANGUP for example crossing in Huntington, W.Va.

in. of ground clearance will still experience problems; however, the magnitude of the problem is not as severe as in the existing case. This is shown in Figure 9, which illustrates that more vehicles can safely negotiate the crossing in this situation than in the situation shown in Figure 7. Other alternatives were generated and checked using the model. A profile that used a circular vertical curve was identified that eliminated hang-up problems at this crossing.

HANGUP has also been implemented in several other situations. A grade crossing in an urban area of Morgantown, West Virginia, was analyzed to determine, for a construction contractor, whether a low-boy equipment trailer delivering a crane to a construction site could negotiate the crossing. A rural crossing in Oregon was analyzed to determine whether certain double-drop livestock trailers could use an access road to a ranch. A shortline railroad in Pennsylvania, aware of humpback and potential high-profile crossings, used the results of the package to assess the impact that track raises would have on crossings. Obviously, the package would be helpful in a variety of other circumstances, including identification of crossings at which warning signs would be appropriate. Readers who know of particular problem crossings are invited to send crossing profile and vehicle data to the researchers.

CONCLUSIONS AND RECOMMENDATIONS

Although difficult to quantify, low-clearance vehicle accidents at rail-highway grade crossings have been identified as a problem nationwide by federal, state, and local agencies and railroad and trucking interests. Classification count data collected in West Virginia indicated that low-clearance vehicles make up about 2 percent of the traffic stream. It can be concluded solely on the basis of vehicle population data that the low-clearance vehicle hang-up problem is potentially serious.

Low-clearance vehicles vary greatly in their physical characteristics. Typical wheelbases range from 25 to 40 ft, and ground clearances of 3 in. are not atypical. There is little, if any, standardization; apparently, vehicle manufacturers are guided by the wishes of shippers and vehicle operators. Despite the variability, four general classes of vehicles were identified for the purpose of developing geometric design criteria: (a) low-bed equipment trailers, (b) automobile transporters, (c) double-drop van semitrailers, and (d) car- and truck-trailer

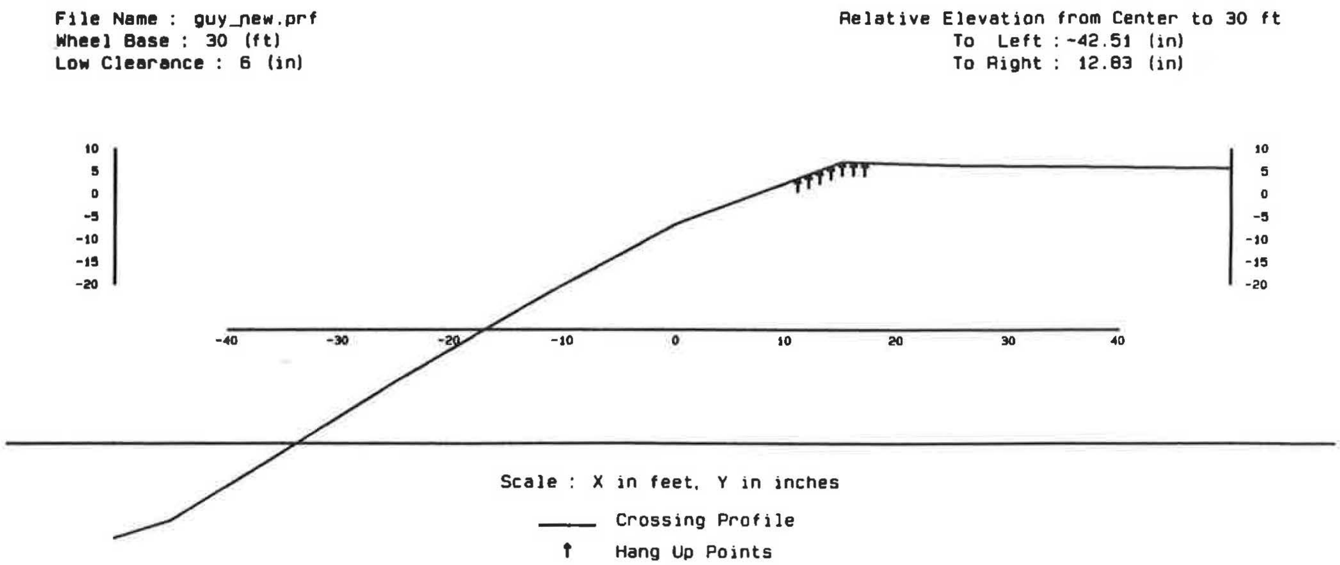


FIGURE 8 Manual-mode output of HANGUP for redesign of example crossing.

H A N G U P

File Name : guy_new.prf Date : 07-24-1990

Ground Clearance (in)

Wheel Base	1	2	3	4	5	6	7	8	9	10
10 (ft)	1	1	0	0	0	0	0	0	0	0
11 (ft)	1	1	0	0	0	0	0	0	0	0
12 (ft)	1	1	0	0	0	0	0	0	0	0
13 (ft)	1	1	1	0	0	0	0	0	0	0
14 (ft)	1	1	1	0	0	0	0	0	0	0
15 (ft)	1	1	1	0	0	0	0	0	0	0
16 (ft)	1	1	1	0	0	0	0	0	0	0
17 (ft)	1	1	1	1	0	0	0	0	0	0
18 (ft)	1	1	1	1	0	0	0	0	0	0
19 (ft)	1	1	1	1	0	0	0	0	0	0
20 (ft)	1	1	1	1	0	0	0	0	0	0
21 (ft)	1	1	1	1	1	0	0	0	0	0
22 (ft)	1	1	1	1	1	0	0	0	0	0
23 (ft)	1	1	1	1	1	0	0	0	0	0
24 (ft)	1	1	1	1	1	0	0	0	0	0
25 (ft)	1	1	1	1	1	1	0	0	0	0
26 (ft)	1	1	1	1	1	1	0	0	0	0
27 (ft)	1	1	1	1	1	1	0	0	0	0
28 (ft)	1	1	1	1	1	1	0	0	0	0
29 (ft)	1	1	1	1	1	1	1	0	0	0
30 (ft)	1	1	1	1	1	1	1	0	0	0
31 (ft)	1	1	1	1	1	1	1	0	0	0
32 (ft)	1	1	1	1	1	1	1	1	0	0
33 (ft)	1	1	1	1	1	1	1	1	0	0
34 (ft)	1	1	1	1	1	1	1	1	0	0
35 (ft)	1	1	1	1	1	1	1	1	1	0
36 (ft)	1	1	1	1	1	1	1	1	1	0
37 (ft)	1	1	1	1	1	1	1	1	1	0
38 (ft)	1	1	1	1	1	1	1	1	1	1
39 (ft)	1	1	1	1	1	1	1	1	1	1
40 (ft)	1	1	1	1	1	1	1	1	1	1

1 -> Hang up, 0 -> Safe

FIGURE 9 Automatic-mode output of HANGUP for redesign of example crossing.

combinations. Additional wheelbase and ground-clearance data must be collected from several geographic regions before specific dimensions can be specified for each of these categories.

Although several agencies have developed geometric design standards for low-clearance vehicles at crossings, they generally are not known to highway designers and therefore are not used. A few jurisdictions have developed warning signs for high-profile grade crossings. However, placement has apparently been based on accident experience or on a seat-of-the-pants approach rather than on a formal analysis of crossing geometry and truck traffic.

The microcomputer package HANGUP was developed to simulate the movement of low-clearance vehicles over grade crossings. The program appears to be useful as an analysis tool for identifying crossings at which hang-ups may be a problem. In this application, it could be used as a tool to identify locations that warrant the low-clearance-crossing warning sign currently under review by the National Committee on Uniform Traffic Control Devices.

In addition, the package can be used as a design tool to assist in the selection of a suitable crossing profile. Although written for application to grade crossings, HANGUP can be used to assess ground-clearance problems at other highway locations, such as driveway entrances and roadway crowns.

Because the research effort is still under way, recommendations can be considered in two categories: those for practitioners and those in which additional research is needed. Recommendations to practitioners will be outlined first.

It is apparent that highway and traffic engineers need to devote more attention to the ground-clearance issue in general and problems at rail-highway grade crossings in particular. Attention in this sense also includes communicating existing design standards to practitioners.

Low-clearance vehicles also must be considered in highway operations. For example, the current permitting process for oversize vehicles considers weight, height, width, and such, but apparently does not consider ground clearance. This oversight should be corrected. The wide variety in dimensions of low-clearance vehicles should be examined with an eye toward possibly establishing reasonable minimum ground-clearance standards for vehicles operating on public highways.

Using HANGUP and low-clearance vehicle data, specific design criteria should be developed for highway vertical alignment at rail-highway grade crossings. Work on this task will begin soon.

Additional work is suggested to adapt HANGUP to the issue of highway ground clearance in general. The package can already be used in this application; however, refinements are needed to facilitate its use in that regard. HANGUP is a two-dimensional model, so it would be desirable to expand the program to include three-dimensional capabilities.

For "designed" roadways, roadway profile input data can be easily obtained from roadway plans. However, many crossings in the United States have no plans or designs. For most of the example applications implemented on HANGUP so far, the researchers used a conventional level and rod to determine centerline elevations approximately every 5 ft along the roadway. Although experienced individuals can do this operation relatively quickly, the procedure does require the use of at least two people (and more on heavily traveled roadways) and thus may not be feasible in many instances. Close-range photogrammetric techniques employing a standard 35-mm camera should be investigated for their applicability to this problem. Such an approach would mean that field work could be done quickly with one individual, a standard 35-mm camera, and simple control equipment such as a length of pipe or a sphere.

ACKNOWLEDGMENTS

Support for the research described in this paper was provided by a grant from the U. S. Department of Transportation's University Transportation Centers Program and by WVDOH. The encouragement and cooperation provided throughout the project by Ray Lewis of WVDOH and Richard Mather of the Public Utility Commission of Oregon are sincerely appreciated. Thanks are also due to the many trucking company representatives, trailer manufacturers, truck drivers, and railroads that provided information.

REFERENCES

1. *Rail-Highway Crossings Study*. FHWA, U. S. Department of Transportation, 1989.
2. *Railroad/Highway Accident Report—Collision of AMTRAK Train No. 88 with Tractor Lowboy Semitrailer Combination Truck, Rowland, NC, August 25, 1983*. Report NTSB/RHR-84/01. NTSB, Washington, D.C., Aug. 1984.
3. *Safety Study—Passenger/Commuter Train and Motor Vehicle Collisions at Grade Crossings (1985)*. Report NTSB/SS-86/04. NTSB, Washington, D.C., Dec. 1986.
4. *Safety Recommendations H-89-6 and H-89-7*. NTSB, Washington, D.C., Feb. 1989.
5. W. A. McConnell. Passenger Car Overhang and Underclearance as Related to Driveway Profile Design, Part I—Vehicle Data. In *Highway Research Board Bulletin 195*, HRB, National Research Council, Washington, D.C., 1958, pp. 14–23.
6. L. A. Bauer. Passenger Car Overhang and Underclearance as Related to Driveway Profile Design, Part II—Street and Highway Design. In *Highway Research Board Bulletin 195*, HRB, National Research Council, Washington, D.C., 1958, pp. 23–29.
7. *Manual for Railway Engineering*. American Railway Engineering Association, Chicago, Ill., 1971.
8. *A Policy on Geometric Design of Highways and Streets*. AASHTO, Washington, D.C., 1984.
9. *Policy for Approach Paving and Widths of Rail-Highway Crossings*. State of Alabama Highway Department, Montgomery, 1984.
10. *Field and Office Manual for Profile Surveys of Highway-Rail At-Grade Crossings on Existing Paved Roadways*. Bureau of Roadway Design, Florida Department of Transportation, Tallahassee, Dec. 1984.
11. *Manual on Uniform Traffic Control Devices for Streets and Highways*. FHWA, U. S. Department of Transportation, 1978.
12. *Report on Level Crossing Protection*. Department of Transport, Her Majesty's Stationery Office, London, England, 1978.

The contents of the paper reflect the views of the authors, who are responsible for the facts and accuracy of the information presented herein.

Publication of this paper sponsored by Committee on Rail-Highway Grade Crossings.

Safety Evaluation of Converting On-Street Parking from Parallel to Angle

TIMOTHY A. MCCOY, PATRICK T. MCCOY, RICHARD J. HADEN, AND VIRENDRA A. SINGH

To increase the supply of parking in the central business district of Lincoln, Nebraska, the city converted parking on several streets from parallel to angle. The conversions were made on streets that had enough room to permit the removal of a traffic lane to provide the additional street width needed for angle parking without increasing traffic congestion. The safety effects of converting on-street parking from parallel to angle were evaluated and the cost-effectiveness of the parking conversion was determined. The conversion resulted in a significant increase in the number of parking-related accidents; the number of parking-related accidents per million vehicle miles also increased significantly. But when the increase in accident exposure caused by the increase in parking activity was considered, there was no significant increase in the parking-related accident rate, nor was there a significant change in the severity of the parking-related accidents. The on-street conversion was cost-effective because the increase in accident costs resulting from the conversion was lower than the cost of providing additional off-street spaces.

To increase the supply of parking in the central business district of Lincoln, Nebraska, the city converted parking on several streets from parallel to angle. The conversions were made on streets that had enough room to permit the removal of a traffic lane to provide the additional street width needed for angle parking. Although the removal of traffic lanes did not increase congestion, the city was concerned about a possible increase in accidents.

The objective of this study was to evaluate the safety effects and cost-effectiveness of converting on-street parking from parallel to angle. The rates and severity of parking-related accidents before and after the conversion were compared to determine the safety effects. In addition, an economic analysis was conducted to determine the cost-effectiveness of the conversion.

PREVIOUS RESEARCH

Several studies (*1*) have compared the accident experience of angle and parallel parking. The studies have reported accident rates for parallel parking to be from 19 to 71 percent lower than those for angle parking. Many of the studies were before-and-after studies involving changes from angle to parallel parking. However, none of them were before-and-after studies of changes from parallel to angle parking, and none of the studies accounted for the change in accident exposure associated with the change in parking configuration. For example, when angle parking is changed to parallel parking, accident exposure is reduced because fewer parking spaces remain after the conversion. Therefore, the reductions in accidents that have been associated with changes from angle to parallel parking may have been caused by changes in accident exposure rather than by differences in the types of parking maneuvers associated with the parking configurations.

Humphreys et al. (*2*) found parking utilization to be a primary factor affecting accident rates. Increased parking utilization resulted in significantly higher accident rates regardless of the type of parking. Type of parking was found to have no effect on accident rates when parking utilization, abutting land use, and street classification were taken into account. Thus, these findings brought into question the conclusion of many studies that parallel parking is safer than angle parking.

PARKING CONVERSION

Since September 1987 27 block faces in downtown Lincoln have been converted from parallel to angle parking. The typical parallel parking space was 22 ft long and 8 ft wide. The minimum width of the adjacent traffic lane was 12 ft. This traffic lane was removed to provide the additional space needed for angle parking.

The typical angle parking space that replaced the parallel parking is 9 ft wide and has a 15-ft stall line extending from the curb. The parking angle is 55 degrees, which is the angle used in Lincoln as a balance between number of spaces and ease of parking. The minimum width of the adjacent traffic lane is 15 ft.

STUDY SITES

Of the 27 block faces, 15 were not included in the study because the parking on them had been converted less than a year ago, and 1 other was excluded because traffic-volume data were not available. Eleven converted block faces were thus evaluated. In addition, to account for any overall change in the parking-related accidents in the central business district during the study period, eight block faces on which the parallel parking had not been converted were used as comparison sites. Thus, 19 block faces were included in the study.

The study and comparison sites are shown in Table 1. All of the sites were on downtown streets that have 25-mph speed

T. A. McCoy and P. T. McCoy, Department of Civil Engineering, University of Nebraska-Lincoln, W348 Nebraska Hall, Lincoln, Neb. 68588. R. J. Haden and V. A. Singh, Lincoln Transportation Department, 233 South 10th Street, Lincoln, Neb. 68508.

TABLE 1 STUDY AND COMPARISON SITES

Block Face	Speed Limit	Street Ways	Street Lanes	Spaces Before	Spaces After	Turn-over	% Use	ADT
Study Sites								
1	25	2	2	1	6	4.21	100	2,300
2	25	2	2	4	10	4.18	92	2,200
3	25	2	2	6	12	3.54	92	3,400
4	25	1	3	4	9	8.05	94	3,100
5	25	1	3	6	14	8.05	94	3,100
6	25	1	3	7	16	7.92	94	2,550
7	25	1	3	5	13	3.24	88	5,330
8	25	1	3	5	11	8.05	94	4,100
9	25	1	3	5	10	8.05	94	4,100
10	25	1	3	9	14	5.96	94	5,730
11	25	2	2	8	16	2.97	85	1,000
Comparison Sites								
1	25	1	4	7	7	8.05	92	12,300
2	25	1	4	1	1	8.05	92	12,300
3	25	1	4	8	8	8.05	94	11,600
4	25	1	4	8	8	8.05	94	11,600
5	25	1	4	5	5	7.92	94	15,200
6	25	1	4	4	4	7.92	94	15,200
7	25	1	4	6	6	7.92	94	13,500
8	25	1	4	10	10	7.92	94	13,500

limits. Four of the eleven study sites were on two-lane, two-way streets; the rest were on three-lane, one-way streets. All the comparison sites were on four-lane, one-way streets.

The study sites had 60 parking spaces before the conversion and 131 spaces after the conversion. The comparison sites had 49 spaces. All the spaces were metered with time limits of 1 or 2 hr. The parking turnover on the study-site block faces ranged from 2.97 to 8.05 parkers per 8-hr parking day between 9:00 a.m. and 5:00 p.m. The turnover on all the comparison-site block faces was about 8 parkers per 8-hr parking day. The average parking utilization during the 8-hr parking day was from 85 to 100 percent on the study-site block faces and from 92 to 94 percent on the comparison-site block faces. The parking turnover and utilization on each block face were obtained from a parking study (*3*) conducted in downtown Lincoln in 1985. These data were believed to apply to the study period because the on-street parking revenues per space on the study- and comparison-site block faces have remained nearly constant since 1985.

The study-site average daily traffic (ADT) counts were much lower than those of the comparison sites. The study-site ADTs ranged from 1,000 to 5,730 vehicles per day (vpd). The comparison-site ADTs ranged from 11,600 to 15,200 vpd. The ADTs were computed from traffic counts provided by the city.

ACCIDENT DATA

The accident report files maintained by the city of Lincoln were reviewed to determine the number of parking-related accidents that occurred on each block face. The parking-related accidents included not only parked-vehicle and parking-maneuver accidents but also any accidents determined to have resulted from the presence of on-street parking. For example, also included were collisions in which one vehicle was trying to avoid a vehicle involved neither in parking nor in the collision. And only those accidents that occurred on weekdays between 9:00 a.m. and 5:00 p.m. were included.

The parking-related accidents before and after the conversion were identified for each study site. Because the parking on all block faces was not converted at the same time, the before and after periods had to be determined individually for each study site. A 3-month adjustment period was used. Therefore, the after period for each study site was the time beginning 3 months after the conversion until the end of 1989. The before period for each study site was an equal number of weekdays before the conversion.

The parking-related accidents for the comparison sites were identified for the years 1986–1989, which covered the duration of longest before and after periods of the study sites. However, the study sites had different before and after periods, ranging from 282 to 518 days. Therefore, the numbers of parking-related accidents on the comparison-site block faces before and after the conversion had to be determined separately for each study site.

The numbers of parking-related accidents for the study sites are shown in Table 2. On most of the study-site block faces, the number of parking-related accidents during the 8-hr parking day increased from none before the conversion to one or two after the conversion. Overall, the number of study-site parking-related accidents increased from two to eleven (450 percent) after the conversion from parallel to angle parking. Meanwhile, the total number of comparison-site parking-related accidents also increased between the before and after periods for each study-site block face. On the average, the

TABLE 2 PARKING-RELATED ACCIDENTS

Study-Site Block Face	Number of Weekdays Before	Number of Weekdays After	Number of Accidents Before	Number of Accidents After	Total Number of Accidents at the Comparison Sites Before	Total Number of Accidents at the Comparison Sites After
1	407	407	0	2	4	8
2	383	383	0	0	4	8
3	374	374	0	1	2	4
4	518	518	0	1	5	9
5	500	500	0	1	5	9
6	518	518	0	1	5	9
7	341	341	0	2	2	5
8	282	282	1	1	5	5
9	337	337	0	0	2	5
10	337	337	1	1	2	5
11	375	375	0	1	3	7
Total	4,372	4,372	2	11	Not Applicable	
Mean	397.5	397.5	.2	1.0	3.5	6.7

total number of comparison-site accidents increased from 3.5 to 6.7 (90 percent). Therefore, the number of study-site accidents expected during the after period if the parking had not been converted would have been two plus 90 percent, or about four parking-related accidents. Thus, the increase in study-site parking-related accidents over the expected increase was from four to eleven (175 percent). According to the Poisson distribution test, all three of these increases are statistically significant at the 5 percent level of significance.

Prevailing traffic conditions and the level of parking activity did not change during the study period. However, mild weather during the winter months of 1987 resulted in an unusually low number of accidents in 1987, which was during the before period of the study sites. The numbers of parking-related accidents during the 8-hr parking day on the comparison-site block faces from 1986 through 1989 are shown in Table 3. There was only one accident in 1987 compared with an average of 3.5 accidents per year. But, as shown in Table 3, this was also the case on 11 other block faces in downtown Lincoln that were not included in the study. The parking on these block faces was angle parking throughout the study period. In 1987 there were only 6 accidents on these block faces compared with an average of 10.5 accidents per year. Thus, the low number of parking-related accidents in 1987 was not limited to the comparison sites but was the general case in downtown Lincoln. The large increase in accidents on the comparison-site block faces was therefore attributed to the mild winter weather in 1987.

ACCIDENT RATES

The accident rate commonly used in comparisons of parallel and angle parking has been expressed in terms of accidents per million vehicle miles (*1*). However, as mentioned previously, this rate does not account for the change in accident exposure caused by the difference in the parking activity between parallel and angle spaces. Parking activity is a function of the turnover and utilization of the parking. In addition to the common accident rate, therefore, an accident rate was computed in terms of accidents per million space-hours per 1,000 parkers per million vehicle miles as follows:

$$\mathrm{AR} = N/(T * P * H) \tag{1}$$

where

AR = accident rate (accidents/10^6 space-hr/10^3 parkers/10^6 vehicle-mi),
N = number of parking-related accidents,
T = vehicle miles of travel (millions),
P = number of parkers (thousands), and
H = number of space-hours of parking used (millions).

TABLE 3 ANNUAL PARKING-RELATED ACCIDENTS

Block Faces	1986	1987	1988	1989
Comparison Sites	5	1	3	5
Others not included in the study	10	6	14	12

The vehicle miles of travel, the number of parkers, and the number of space-hours used in Equation 1 were computed as follows:

$$T = (0.48 * \mathrm{ADT} * D * L)/10^6 \tag{2}$$

$$P = (\mathrm{TO} * S * D)/10^3 \tag{3}$$

$$H = (8 * U * S * D)/10^6 \tag{4}$$

where

ADT = average daily traffic (vpd),
D = number of parking days,
L = length of block face (miles),
TO = turnover rate (parkers per space per parking day),
S = number of parking spaces, and
U = percentage of space-hours used.

In calculating the rates for each study site, the space-hours of parking, number of parkers, and vehicle miles of travel were computed for the 8-hr parking day for the number of weekdays shown in Table 2. The number of spaces, turnover rate, percentage of space-hours used, and ADT shown in Table 1 were used to compute these values for each site. About 48 percent of the ADT at the study sites occurred during the 8-hr parking day.

In addition to the before and after accident rates, adjusted-after accident rates were also calculated for each study site. The adjusted-after accident rates were the after accident rates reduced to account for the increase in parking-related accidents on the comparison-site block faces as follows:

$$R_i' = R_i(B/A) \tag{5}$$

where

R_i' = adjusted-after accident rate for study site i,
R_i = after accident rate for study site i,
B = number of parking-related accidents on the comparison-site block faces during the before period for study site i, and
A = number of parking-related accidents on the comparison-site block faces during the after period for study site i.

The parking-related accident rates computed for each study site are shown in Table 4. Mean parking-related accident rates are also shown. The mean accident rate, based on vehicle miles of travel, increased from 4.6 to 33.6 accidents per million vehicle miles after the conversion. When the increase in the number of parking-related accidents on the comparison sites was considered, the rate still increased to 17.5 accidents per million vehicle miles. But when the effect of the level of parking activity was also considered, the parking-related accident rate increased only from 28.1 to 36.0 accidents per million space-hours per 1,000 parkers per million vehicle miles.

The one-tail paired t-test was used to determine whether the after and adjusted-after mean parking-related accident rates were significantly higher than the before mean parking-related accident rates. The comparison of the before with the after mean parking-related accident rate is shown in Table 5, and the comparison of the before with the adjusted-after mean

TABLE 4 PARKING-RELATED ACCIDENT RATES AT STUDY SITES

Block Face	Accidents/ Million Vehicle Miles			Accidents/ Million Space Hours/ 1,000 Parkers/ Million Vehicle Miles		
	Before	After	Adjusted After	Before	After	Adjusted After
1	0	78.3	39.1	0	389.7	194.9
2	0	0	0	0	0	0
3	0	28.8	14.4	0	54.9	27.5
4	0	22.8	12.7	0	17.3	9.6
5	0	23.6	13.1	0	8.0	4.4
6	0	22.7	15.4	0	6.8	3.8
7	0	40.3	16.1	0	89.9	36.0
8	31.7	31.7	31.7	263.3	54.4	54.4
9	0	0	0	0	0	0
10	19.0	19.0	7.6	46.0	19.0	7.6
11	0	97.7	41.9	0	134.4	56.7
Mean	4.6	33.6	17.5	28.1	70.4	36.0

TABLE 5 COMPARISON OF BEFORE AND AFTER PARKING-RELATED ACCIDENT RATES AT STUDY SITES

Parking-Related Accident Rate	Mean Diff.	Stand. Dev.	Sample Size	t-value	Level of Significance[a]
Accidents/ Million Vehicle Miles	29.0	32.7	11	2.94	< 0.01
Accidents/ Million Space Hours/ 1,000 Parkers/ Million Vehicle Miles	42.3	143.6	11	0.98	> 0.20

[a] One-tail, paired t test.

parking-related accident rate is shown in Table 6. In both cases, the after and the adjusted-after mean rates in terms of accidents per million vehicle miles were significantly higher than the corresponding before mean accident rate at the 5 percent level of significance. But the after and adjusted-after mean accident rates in terms of accidents per million space-hours per 1,000 parkers per million vehicle miles were not significantly higher than the corresponding before mean accident rate at the 5 percent level of significance.

Thus, the conversion from parallel to angle parking resulted in a significantly higher mean parking-related accident rate when the measure of accident exposure was vehicle miles of travel. However, when the measure of accident exposure included the level of parking activity, there was no significant difference in the mean parking-related accident rates.

ACCIDENT SEVERITY

From 1986 through 1989, 13 percent of the parking-related accidents on streets in downtown Lincoln were nonfatal injury accidents, and the rest were property-damage-only accidents. Of the 11 parking-related accidents that occurred at the study sites after the parking had been changed from parallel to angle parking, 2 (18 percent) of them were nonfatal injury accidents, and the other 9 (82 percent) were property-damage-only accidents. There was no statistically significant difference between the percentages of nonfatal injury accidents for the parallel and angle parking according to the binomial distribution test of proportions conducted at the 5 percent level of

TABLE 6 COMPARISON OF BEFORE AND ADJUSTED-AFTER PARKING-RELATED ACCIDENT RATES AT STUDY SITES

Parking-Related Accident Rate	Mean Diff.	Stand. Dev.	Sample Size	t-value	Level of Significance[a]
Accidents/ Million Vehicle Miles	12.9	16.2	11	2.63	0.01
Accidents/ Million Space Hours/ 1,000 Parkers/Million Vehicle Miles	7.9	93.8	11	0.28	>0.20

[a] One-tail, paired t test.

significance. Thus, no significant change in accident severity occurred as a result of the parking conversion.

ECONOMIC ANALYSIS

The conversion of on-street parking from parallel to angle parking provided an average of 6.5 more parking spaces per block face at a nominal cost. The conversion was accompanied by about a 280 percent increase in the number of parking-related accidents per million vehicle miles with no significant increase in accident severity. For the conversion to be cost-effective, the increase in accident costs resulting from it must be less than the cost of providing the additional spaces off-street. The additional accident cost per year is

$$\mathrm{AAC} = (C_a * \mathrm{IAR} * D * P * \mathrm{ADT} * \mathrm{BFL})/(10^8 * S) \quad (6)$$

where

AAC = additional accident cost (dollars per year per space),
C_a = average parking-related accident cost (dollars per accident),
IAR = increase in number of parking-related accidents per million vehicle miles,
D = number of parking days per year,
P = percentage of ADT during parking day,
BFL = average block-face length (miles), and
S = average number of additional spaces per block face.

The accident cost figures currently used by the city of Lincoln (*4*) are $13,600 per fatal or nonfatal injury accident and $1,700 per property-damage accident. Because of the random occurrence of fatal accidents in the city of Lincoln, the same accident cost is used for fatal and nonfatal injury accidents to avoid overemphasizing locations where a fatal accident occurs. The accident cost used for fatal or nonfatal injury accidents is the weighted average cost, which was derived using the National Safety Council costs and the numbers of fatal and nonfatal injury accidents that occurred in Lincoln during 1989. On the basis of these accident costs and the severity of parking-related accidents observed in downtown Lincoln, the average parking-related accident cost is about $3,250.

The adjusted average increase in the number of parking-related accidents per million vehicle miles shown in Table 6 was 12.9, and the average number of additional spaces per block face was 6.5. The number of parking days per year in downtown Lincoln is 307, and 48 percent of the ADT is during the 8-hr parking day. The average block-face length in downtown Lincoln is about 300 ft.

Substituting these values into Equation 6, the annual additional accident cost resulting from the parking conversion in downtown Lincoln is

$$AAC = 0.054 * ADT \quad (7)$$

The present worth of the annual additional accident cost is

$$PAC = 0.054 * ADT * PWF \quad (8)$$

where PAC is the present worth of additional accidents (dollars per space) and PWF is the present worth factor of a uniform series.

The recent cost of developing off-street parking facilities in downtown Lincoln, including land acquisition, demolition, and construction, is about $10,400 per space. Therefore, for the parking conversion to be cost-effective, the present worth of the additional accident costs per space would have to be less than $10,400.

The ADT at which the additional accident cost of the conversion would be equal to the cost of providing the additional spaces off-street can be found by setting Equation 8 equal to the average cost per space of off-street parking and solving for ADT as follows:

$$ADT_{BE} = (18.5 * C_s)/PWF \quad (9)$$

where ADT_{BE} is the break-even ADT (vpd) and C_s is the average cost of off-street parking (dollars per space).

Using a 10 percent interest rate and a 30-year service life, the break-even ADT in downtown Lincoln would be 20,400 vpd. This indicates that the conversion would be cost-effective on streets with ADTs below 20,400 vpd. It would not be cost-effective on higher-volume streets, where lower-cost off-street parking can be provided, or where an increase in traffic congestion may occur. As shown in Table 1, the ADTs were well below 20,400 vpd on the block faces where the conversions were made in downtown Lincoln. Therefore, it is readily apparent that converting on-street parking from parallel to angle in downtown Lincoln was cost-effective.

CONCLUSION

The conversion from parallel to angle parking in downtown Lincoln resulted in a significant increase in the number of parking-related accidents on the converted block faces. In addition, the number of parking-related accidents per million vehicle miles of travel on the block faces also increased significantly. But when the increase in accident exposure due to the increase in the number of spaces was accounted for, there was no significant increase in the parking-related accident rate on the block faces. There was no significant change in the severity of parking-related accidents as a result of the conversion, either.

Although these findings may suggest that there is no difference in the safety effects of parallel and angle parking, this is not so. Instead, the findings indicate that the difference between the safety of the two types of parking is primarily a result of differences in levels of parking activity rather than in the nature of the parking maneuvers associated with them. Therefore, where the supply of parking spaces is sufficient, the conversion of on-street parking from parallel to angle should *not* be considered, because the number of accidents will increase as a result of more parking activity because of more parking spaces. In the same way that no parking is safer than parallel parking, parallel parking is safer than angle parking, because it reduces accident exposure.

The conversion more than doubled the number of parking spaces on the block faces. However, the cost of the additional parking-related accidents caused by the conversion was less than the cost of providing an equivalent number of additional parking spaces off-street in a parking garage or parking lot. Therefore, the conversion was cost-effective. However, it must be remembered that the conversion was made on streets that had enough room to allow the removal of traffic lanes to provide the additional street width needed by the 55-degree angle parking without increasing traffic congestion. Consequently, additional traffic operations and delay costs were not included in the analysis.

The primary conclusion of this study was that converting on-street parking from parallel to angle may be a cost-effective way of increasing the supply of parking in downtown areas, if the streets are wide enough to provide for angle parking without increasing traffic congestion, as was the case in downtown Lincoln. But, even if the streets are wide enough, on-street parking should not be converted from parallel to angle unless the additional parking-related accident cost is compared with the cost of developing alternative off-street parking spaces. In some cases, the development of off-street parking may be so inexpensive that the cost of alternative off-street spaces would be lower than the cost of the additional parking-related accidents that would occur as a result of the conversion. In other cases, the traffic volumes may be so high that the additional parking-related accident cost would exceed the cost of developing alternative off-street parking spaces. In these cases, converting on-street parking would not be cost-effective. The solution of Equation 9 for various off-street parking costs, as illustrated in Figure 1, could provide a guideline for determining the cost-effectiveness of converting on-

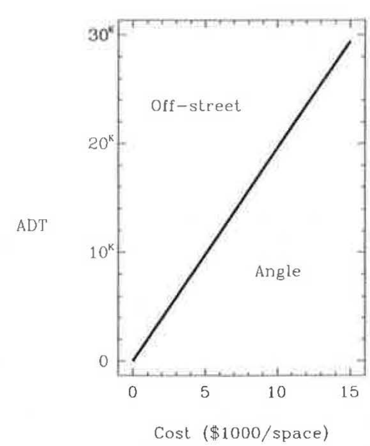


FIGURE 1 Cost-effective guideline.

street parking. Points below the line indicate that the conversion to angle parking would be cost-effective, and points above the line indicate that converting spaces would not be cost-effective when compared with providing the additional spaces off-street.

It should be noted that the findings of this study were based on a limited amount of accident data. Larger sample sizes over a wider range of ADTs are needed to validate the findings. The city of Lincoln has concluded that the conversion of on-street parking from parallel to angle was cost-effective for the ADT levels studied, but it is continuing to study the safety effects of ADTs greater than 6,000 vpd. More research is needed before definitive guidelines for converting on-street parking from parallel to angle can be established. Additional research should be conducted to examine effects of factors such as traffic volumes, speeds, street widths, parking angle, and parking activity on the cost-effectiveness of the conversion.

REFERENCES

1. *Synthesis of Safety Research Related to Traffic Control and Roadway Elements*, Vol. 1. Report FHWA-TS-82-232. FHWA, U. S. Department of Transportation, Dec. 1982.
2. J. B. Humphreys, P. C. Box, T. D. Sullivan, and D. J. Wheeler. *Safety Aspects of Curb Parking.* Report FHWA-RD-79-76. FHWA, U. S. Department of Transportation, Sept. 1978.
3. *Lincoln Parking Study*. Interim Report. Traffic Institute, Northwestern University, Evanston, Ill., May 1985.
4. *1989 Accident Report*. Department of Transportation, Lincoln, Neb., May 1990.

Publication of this paper sponsored by Committee on Parking and Terminals.

Conditional Analysis of Accidents at Four-Approach Traffic Circles

A. AL-BAKRI, MARK R. CONAWAY, AND JAMES STONER

A conditional analysis for relating the number of accidents at four-approach traffic circles to the geometric and flow characteristics of the circles is presented. The conditional analysis takes into account the association among the observations taken at the four approaches within a traffic circle. It also allows for the inherent differences in safety among different circles. The main advantage of the conditional approach is one can use it to estimate the effects of geometric and flow variables without having to specify a distribution to represent the variability between circles. The conditional method is applied to a study of traffic circles in Amman, Jordan.

The relationship between the number of accidents occurring at a traffic circle and the geometric characteristics and traffic flows of the circle is examined in this paper. The data were collected at seven traffic circles in Amman, Jordan. Each had four approaches approximately at right angles, relatively large circular central islands, parallel entries, and yield-sign traffic control. The data were collected by two teams of 30 police officers and 6 graduate students in transportation engineering at the Jordan University of Science and Technology. The teams collected information on the geometric characteristics and traffic flows in each approach in each traffic circle and counts of several types of accidents. The data collection involved designing special forms and training the teams to ensure the accuracy of the data collection. Details of the procedure are given by Al-Bakri (*1*).

The method of analysis is similar to that of Maycock and Hall (*2*) in that it is based on the assumption that the accident counts have Poisson distributions, the mean of the Poisson distribution depending on the geometric and flow variables. The geometric and flow variables to be used in this study, given here, are described in Figure 1. A more comprehensive definition of these variables is contained in the Jordanian Department of Transport's departmental standards.

- CE—Entry-path curvature in meters^{-1} is the shortest straight-ahead vehicle path.
- CA—Approach curvature in meters^{-1} is the reciprocal of the minimum radius of the bend nearest to the traffic circle.
- ICD—Inscribed circle diameter in meters; diameter of the largest circle that can be inscribed in the outline of the traffic circle.
- CID—Central island diameter in meters.
- QC—Average daily traffic volume in vehicles per day circulating the circle.
- QE—Average daily traffic volume in vehicles per day entering the circle.
- E—Entry width in meters, measured at a point in the upstream approach.
- V—Approach half-width in meters.
- θ—Angle in degrees between the approach leg and next approach leg clockwise.
- PED—Pedestrian volume per day.
- Entering Accidents—Collisions between an entering vehicle and a vehicle within the right-of-way.
- Approaching Accidents—Collisions between vehicles on the approach to the circle, such as rear-end impacts and lane-changing accidents.
- Single-Vehicle Accidents—Collisions involving a vehicle and the circle layout, signs, lighting columns, and such.
- Other Accidents—Collisions between circulating vehicles, circulating vehicles and vehicles exiting the circle, and exiting vehicles and entering or exiting vehicles.
- Pedestrian Accidents—Collisions involving pedestrians.

It is also assumed that a random effect is associated with each traffic circle, which affects the variability of the accident counts within the circle. Similar assumptions were made by Maycock and Hall (*2*), who also assumed that the random effects have a gamma distribution. They estimated the effects of the geometric and flow variables and the parameters of the assumed gamma distribution. A conditional likelihood approach to the problem is taken, so the effects of the geometric and flow variables can be estimated without a distribution for the random effects being specified. This is an important distinction, because the effect of misspecifying the distribution for the random effects is not known.

METHODS

In the method used to analyze the data, Y_{ij} denotes the number of accidents in approach j of traffic circle i, and μ_{ij} is the mean of Y_{ij} for $j = 1, \ldots, 4$ and $i = 1, \ldots, 7$. The geometric and flow variables associated with the jth approach in circle i will be denoted by $\underline{x}_{ij}$. A regression approach to the problem would use the model $Y_{ij} = \underline{x}'_{ij}\underline{\beta} + \varepsilon_{ij}$, and the following assumptions: (a) the Y_{ij}s are normally distributed with mean $\mu_{ij} = \underline{x}'_{ij}\underline{\beta}$, and (b) the Y_{ij}s are independent of one another. Because the number of accidents at an approach is a count variable, with possible values 0, 1, 2, 3, . . . , the normal assumption does not seem valid. An alternative analysis, pro-

A. Al-Bakri and J. Stoner, Department of Civil and Environmental Engineering, University of Iowa, Iowa City, Iowa 52242. M. R. Conaway, Department of Statistics and Actuarial Science, University of Iowa, Iowa City, Iowa 52242.

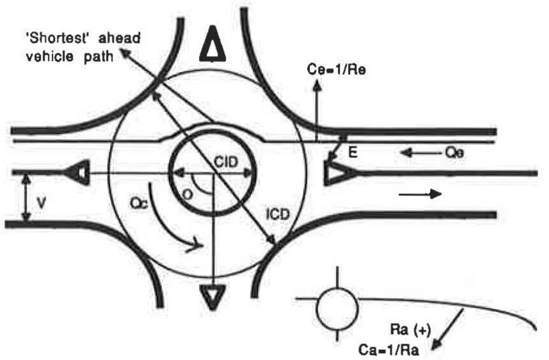


FIGURE 1 Geometric characteristics of traffic circles.

posed by Hauer (*3*), is based on the assumption that the number of accidents at each approach has a Poisson distribution with mean μ_{ij}, with $\ln(\mu_{ij}) = \underline{x}'_{ij}\underline{\beta}$. Hauer fit generalized linear models to estimate $\underline{\beta}$, the effects of the geometric and flow variables. This is based on a more realistic assumption about the distribution of the accident counts than is the regression method, but it ignores the fact that the observations from the same circle are likely to be dependent, even after adjusting for geometric and flow variables. In addition, as noted by Maycock and Hall, one would expect that some traffic circles are inherently more dangerous than others, in the sense that they have more accidents than others with similar geometric and flow characteristics.

Maycock and Hall account for the inherent differences among traffic circles by assuming that the number of accidents has a Poisson distribution with mean q, where q has a gamma distribution with parameters μ and S. The parameter μ depends on the geometric and flow variables through the model $\ln(\mu) = \underline{x}'\underline{\beta}$. Averaging over the assumed gamma distribution, this model states that the number of accidents has a negative binomial distribution with mean μ and variance $\mu(\mu + S)/S$. Maycock and Hall outline a procedure for estimating S then use generalized linear modeling to estimate $\underline{\beta}$.

Several features make the Maycock and Hall method inappropriate for this data. If the differences among traffic circles are assumed to be random, then averaging over the distribution of the random effects induces a dependence among the four observations within a circle. The generalized linear model approach of Maycock and Hall would not take this into account; it would treat all the observations as if they were independent. A second problem that arises in this study, though not in the Maycock and Hall study, is that there are only seven traffic circles to be used in estimating the parameters of the assumed gamma distribution. With such a small sample, it is difficult to check the assumption of a gamma distribution; even if the assumption were correct, it would be difficult to obtain reliable estimates of the parameters of the gamma distribution. In this study, these problems will be solved with a conditional analysis, which will allow inferences to be drawn about the effect of the approach-specific geometric and flow characteristics without requiring the specification of a distribution to represent the inherent differences among traffic circles. This conditional analysis yields valid inferences about the effects of the geometric and flow variables under a wide variety of possible distributions, including common distributions such as the gamma or log-normal distributions. Being able to obtain valid inferences under a variety of distributions may be important, because the effect of misspecifying the distribution—that is, assuming a gamma distribution when the true distribution is not gamma—is not yet known.

The theoretical justification of the conditional approach will be given in the next section. In this section, the basic model underlying the conditional approach will be outlined and the interpretation of the parameters in the model will be discussed. Following the earlier notation, Y_{ij} represents the accident count in approach j of circle i, $j = 1, \ldots, 4$ and $i = 1, \ldots, 7$, and $\underline{x}_{ij}$ represents the geometric and flow variables associated with this approach. To represent the inherent differences among traffic circles, a random quantity (θ_i) is associated with the ith circle. The model is that, given θ_i, the accident counts in the four approaches $(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4})$ are independent Poisson variables, with mean $\theta_i\mu_{ij}$. The parameter μ_{ij} depends on the geometric and flow variables through $\ln(\mu_{ij}) = \underline{x}'_{ij}\underline{\beta}$. With this model it is assumed that the effect of the ith circle is to multiply the mean number of accidents in the four approaches by the same factor (θ_i).

Figure 2 illustrates the model for a traffic circle with the associated effect θ_i. If θ_i were averaged over some assumed distribution, the common factor of θ_i in each of the four approaches would induce an association among the accident counts within the same circle, even after adjustments for the geometric and flow variables.

The multiplicative model yields $E(Y_{ij}) = E[E(Y_{ij}|\theta_i)] = \mu_{ij} E(\theta_i)$, and if an intercept term is included in the model for the $\ln(\mu_{ij})$, then the model can be reparametrized so that $E(\theta_i) = 1$. With this reparametrization, $E(Y_{ij}) = \mu_{ij}$. The unconditional variance of Y_{ij} is given by $\mathrm{Var}(Y_{ij}) = \mu_{ij} + \mu_{ij}^2 \mathrm{Var}(\theta_i)$. The covariance between two accident counts within the same circle is $\mathrm{Cov}(Y_{ij}, Y_{ik}) = \mu_{ij}\, \mu_{ik}\, \mathrm{Var}(\theta_i)$. The multiplicative model is the simplest model that allows for additional variability in the Y_{ij} because of inherent differences in the circle, and it allows observations within the same circle to be associated.

The fundamental idea behind the conditional analysis is to consider the conditional distribution of the four accident counts $(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4})$, given the total number of accidents at that traffic circle $(Y_{i+} = Y_{i1} + Y_{i2} + Y_{i3} + Y_{i4})$. Analyzing Poisson variables by conditioning on their sum is a standard statistical technique used in a number of applications [compare McCullagh and Nelder (*4*)]. One of the attractive features of the

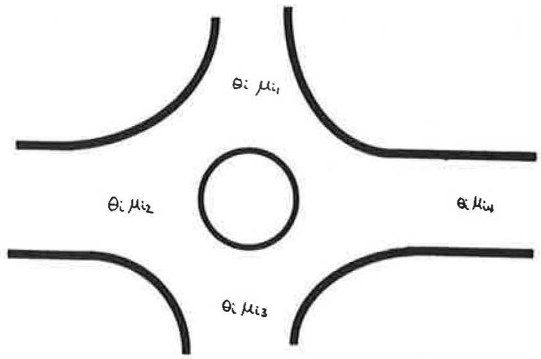


FIGURE 2 Expected number of accidents in each approach.

conditional analysis is that the resulting distribution is multinomial, with parameters y_{i+} and $(p_{i1}, p_{i2}, p_{i3}, p_{i4})$, where

$$p_{ij} = \frac{\exp(\underline{x}'_{ij}\underline{\beta})}{\sum_k \exp(\underline{x}'_{ik}\underline{\beta})} \tag{1}$$

This has the form of a multinomial logit model [Agresti (5)], which facilitates the interpretation of the parameters and allows for the computation of the conditional estimate of $\underline{\beta}$ in a standard statistical package such as GLIM (6). To illustrate the interpretation of the parameters in the conditional analysis, consider Approaches 1 and 2 in a particular traffic circle and suppose that the approaches are identical in all the geometric and flow characteristics except the one measured by the predictor x_{ijm}. In the conditional distribution, $p_{i1}/p_{i2} = \exp[\beta_m(x_{i1m} - x_{i2m})]$, so that the expected number of accidents in Approach 1 would be $\exp[\beta_m(x_{i1m} - x_{i2m})]$ more than expected in Approach 2.

The conditional estimate of $\underline{\beta}$ can be computed in GLIM by specifying that the Y_{ij} have Poisson distributions and using the model $\ln(\mu_{ij}) = \underline{x}'_{ij}\underline{\beta} + \Sigma_i \delta_i I_i$, where I_i is an indicator variable for the ith circle. Note that the model for the mean, μ_{ij}, has the form of a parallel regressions model. This illustrates one of the drawbacks of the conditional analysis. The effect on accidents of changing the geometric and flow variables can be estimated, but the actual value for the mean number of accidents that would occur at a traffic circle with given geometric and flow characteristics cannot be predicted.

DETAILS OF CONDITIONAL ANALYSIS

The theoretical justification for using the conditional estimates will be outlined. As before, let Y_{ij} be the number of accidents in approach j, $j = 1, \ldots, 4$, of circle i, and let Y_{i+} be the total number of accidents at circle i, $i = 1, \ldots, 7$. The random quantity θ_i is associated with the ith circle, and the conditional distributions of the accident counts $(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4})$, given θ_i, are assumed to be independent Poisson random variables with means $\theta_i\mu_{ij}$, $j = 1, \ldots, 4$. In addition, assume that θ_i is sampled from a population with density g.

As in Maycock and Hall (2), the distribution of the accident counts in the ith circle is considered averaging over the distribution of the random effect θ_i. Let $L_i(y_{i1}, y_{i2}, y_{i3}, y_{i4}; \underline{\beta}) = P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, Y_{i3} = y_{i3}, Y_{i4} = y_{i4} \mid \underline{\beta})$ be the likelihood from circle i, averaging over the distribution of the θ_i. The likelihood is

$$\begin{aligned} L_i(\underline{y}_i; \underline{\beta}) &= \int P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, Y_{i3} = y_{i3}, \\ &\quad Y_{i4} = y_{i4} \mid \theta_i) g(\theta_i)\, d\theta_i \end{aligned} \tag{2}$$

$$\begin{aligned} &= \int \sum_{y_{i+}} P(Y_{i1} = y_{i1}, \ldots, \\ &\quad Y_{i4} = y_{i4} \mid Y_{i+}, \theta_i) h(Y_{i+} \mid \theta_i) g(\theta_i)\, d\theta_i \end{aligned} \tag{3}$$

Because Y_{i+} is a sufficient statistic for θ_i, the conditional distribution of $(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4})$ given Y_{i+} is free from θ_i. This conditional distribution is given by

$$\begin{aligned} &P(Y_{i1} = y_{i1}, Y_{i2} = y_{i2}, Y_{i3} = y_{i3}, Y_{i4} = y_{i4} \mid Y_{i+}, \theta_i) \\ &= \prod_j \left[\frac{\exp(\underline{x}'_{ij}\underline{\beta})}{\sum_k \exp(\underline{x}'_{ik}\underline{\beta})}\right]^{y_{ij}} \quad \text{if } \sum_j Y_{ij} = y_{i+} \\ &= 0 \quad \text{if } \sum_j Y_{ij} \neq y_{i+} \end{aligned} \tag{4}$$

Although Equation 3 is written in terms of a sum over the values of Y_{i+}, Equation 4 indicates that only one term in the sum is nonzero. From this expression for the conditional distribution of $(Y_{i1}, Y_{i2}, Y_{i3}, Y_{i4})$ given Y_{i+}, the following expression can be written:

$$L_i(\underline{y}_i; \underline{\beta}) = \prod_j \left[\frac{\exp(\underline{x}'_{ij}\underline{\beta})}{\sum_k \exp(\underline{x}'_{ik}\underline{\beta})}\right]^{y_{ij}} * \int h(Y_{i+} \mid \theta_i) g(\theta_i)\, d\theta_i \tag{5}$$

Because the observations at different traffic circles are assumed to be independent, the likelihood from the sample is the product of the likelihoods from each circle and can be written as

$$\begin{aligned} L &= \prod_i \prod_j \left[\frac{\exp(\underline{x}'_{ij}\underline{\beta})}{\sum_k \exp(\underline{x}'_{ik}\underline{\beta})}\right]^{y_{ij}} \\ &\quad * \prod_i \int h(Y_{i+} \mid \theta_i) g(\theta_i)\, d\theta_i \end{aligned} \tag{6}$$

The distribution of the random effect enters only through the second factor of the likelihood, and an estimate of $\underline{\beta}$ can be computed without a distribution for g being specified, by maximizing the first factor only. The resulting estimates of $\underline{\beta}$ are known as conditional maximum likelihood estimates and possess a number of desirable statistical properties [Andersen (7)]. These properties depend on the size of the Y_{i+}, not on the number of circles in the sample. This makes the conditional approach particularly well suited for these data. There are few traffic circles in the sample, but a fairly large number of accidents were observed at each one.

It should be noted that there is some loss of information about $\underline{\beta}$ in doing the conditional procedure. Because the distribution of Y_{i+} depends on $\underline{\beta}$, ignoring the second factor of the likelihood ignores some of the information about $\underline{\beta}$. Computing the distribution of Y_{i+}, however, reveals that it depends on $\underline{\beta}$ only through the quantity $\mu_{i+} = \Sigma_j \exp(\underline{x}'_{ij}\underline{\beta})$. In trying to estimate "within circle" effects, this should contain little information about the amount of information about $\underline{\beta}$ in the conditional distribution, so that the loss of information from ignoring the second factor should not be large. Precisely how much is lost in the conditioning is difficult to answer, because the amount of information lost depends on the true distribution of the θ_i. A more important feature to note is that, with the conditional procedure, the effects of factors common to all the approaches of a traffic circle cannot be estimated. These disadvantages of the conditional procedure must be balanced against the gains made in estimating the effects of approach-specific factors and in guarding against biases that can result from misspecifying the distribution for the θ_i.

RESULTS OF ANALYSIS

Before the results of the conditional analysis are presented, a brief description of how one can select a model using generalized linear models will be given. A comprehensive treatment of this topic is found in McCullagh and Nelder (*4*). A description of fitting generalized linear models for analyzing traffic data is given in Maycock and Hall (*2*).

Besides providing estimates for the coefficients of the predictors in the model, GLIM also provides a way of checking the fit of the model and of checking whether or not a predictor is a significant addition to a model. These measures are based on a quantity known as the deviance, which, for the Poisson models, equals

$$2 * \sum_{ij} y_{ij} \ln(y_{ij}/\hat{m}_{ij})$$

where $\hat{m}_{ij}$ is an estimate, based on the model being fit, of the expected number of accidents in approach j of circle i. Associated with the deviance is the number of degrees of freedom, which equals the number of observations minus the number of parameters being estimated. A model that describes the data should have a deviance approximately equal to the number of degrees of freedom.

To check if a predictor is a significant addition to a model, compare the deviances from two models: one without the predictor and one after the predictor has been added to the model. For example, suppose a model includes only the logarithm of the entering volume (LNQE) as a predictor. Let d_1 represent the deviance found by fitting this model. To see if another predictor—say, CA—significantly improves the model, obtain the deviance (d_2) that results from using both LNQE and CA as predictors. If the difference ($d_1 - d_2$) is large, compared with a chi-squared distribution with 1 degree of freedom, then it would be concluded that CA is a significant addition to the model. This procedure can be thought of as the generalized linear model version of doing an *F*-test to determine whether a predictor is a significant addition to a regression model.

One of the analyses was performed with total accidents as the dependent variable, where total accidents include entering, approaching, single-vehicle, pedestrian, and other accidents. The most parsimonious model that fit the data included terms for LNQE, CA, and CE. None of the other available predictors significantly improves the model. Deleting the flow or either of the curvature variables results in a model that fits significantly worse than the chosen model. The conditional estimates from the model that includes the predictors CA, CE, and LNQE are given in Table 1. The estimates were computed with and without an outlying observation.

The deviance associated with this model is 26.2 on 18 degrees of freedom ($p = .1$). The fit of this model is adequate, but not particularly good, primarily because of one large residual, corresponding to Arm 3 in the traffic circle R6. This arm has one of the smaller values of flow in the data set, and one would expect fewer accidents here than the 15 that were

TABLE 1 ESTIMATES FOR FINAL MODEL

With case 12			Without case 12		
Estimate	Std. Error	Predictor	Estimate	Std. Error	Predictor
-3.67	1.26	intercept	-5.261	.46	intercept
0.73	0.15	LNQE	.89	.16	LNQE
45.04	17.4	CA	55.86	18.2	CA
159.8	36.9	CE	150.0	37.9	CE

observed. This arm has six single-vehicle accidents, an unusually large number that contributes to the large total number of accidents. No other arm has more than two single-vehicle accidents. Deleting this observation and refitting the model yields a deviance of 17.5 on 17 degrees of freedom.

One way to see how to interpret these estimates is to consider two approaches in the same traffic circle and suppose that Approach 1 has a log of entering flow of $LNQE_1$, an entry curvature of CE_1, and an approach curvature of CA_1. Similarly, Approach 2 has values of $LNQE_2$, CE_2, and CA_2 for these characteristics. Using the estimates from the final model (without Case 12), we have $\hat{p}_1/\hat{p}_2 = \exp[.89 * (LNQE_1 - LNQE_2) + 150.0 * (CE_1 - CE_2) + 55.86 * (CA_1 - CA_2)]$.

To interpret the estimated coefficient (.89) for LNQE, suppose that the two approaches have the same approach curvature and entry curvature, but Approach 1 has an entering flow of α ($\alpha \geq 1$) times the entering flow of the Approach 2. In other words, $LNQE_1 = \ln(\alpha) + LNQE_2$, $CE_1 = CE_2$, and $CA_1 = CA_2$. For these approaches,

$$\hat{p}_1/\hat{p}_2 = \exp[.89 * \ln(\alpha)] = \alpha^{.89} \tag{7}$$

so that the estimate of accidents in Approach 1 would be $\alpha^{.89}$ times the number of accidents in Approach 2. Note that the standard error of the estimate for the coefficient of LNQE is .16, so that the estimate of .89 is not inconsistent with the hypothesis that the parameter equals 1. A value of 1 for the parameter is intuitively appealing, because this means that on average the number of accidents, adjusting for the other factors, changes in direct proportion to the traffic volume.

Similar calculations can be done to interpret the coefficients of the other predictors in the model. CE is measured in units of .001 and has an estimated coefficient of 150.0. This indicates that for two arms with identical flows and approach curvatures, 1.16 times the number of accidents would be expected in an arm with entry curvature CE + .001 as would be expected in an arm with entry curvature CE. CA is also measured in units of .001; its estimated coefficient is 55.86. From this, 1.06 times the number of accidents would be expected in an arm with approach curvature CA + .001 as would be expected in an arm in the same circle with approach curvature CA, if the flow and entry curvature were held constant.

The outlier primarily affects the estimate of the coefficient of LNQE. This might be expected given that Case 12 has a relatively small flow but an unusually large number of accidents. From the estimates that include Case 12, 1.17 times the number of accidents would be expected for an arm with entry curvature CE + .001 as would be expected for an arm with curvature CE. Similarly, 1.05 times the number of accidents would be expected in an arm with approach curvature of CA + .001 as for CA. These estimates are close to those based on deleting Case 12.

These findings are consistent with those of Maycock and Hall (*2*), who also found that CA, CE, and LNQE were important predictors of total accidents.

CONCLUSIONS

In this paper, a conditional analysis was used to relate the number of traffic accidents at traffic circles to the geometric and flow characteristics of the circles. The conditional analysis is based on a Poisson distribution for the accident counts and is particularly suited for these data. It can give valid estimates of the effects without a large number of traffic circles in the sample and without a specified particular distribution for the variability among the circles. This can protect against biases that might result from misspecifying this distribution. All the computations can be done with readily available statistical software, which also provides methods for checking the adequacy of the model and the effect of adding predictors to the model. Applying the method to data collected in Amman, Jordan, it was found that, besides traffic volume, the entry and approach curvatures were important factors in determining the number of accidents.

ACKNOWLEDGMENTS

Part of this work was completed while Mark Conaway was visiting assistant professor at the department of biostatistics at the Harvard School of Public Health and is partially supported by a grant from the National Institutes of Health. Ala Al-Bakri was partially supported by a quarter-time research and teaching assistantship at the University of Iowa and by the Jordanian Royal Endowment for Culture and Education, chaired by Queen Noor Al Hussein, and the Public Security Directorate.

REFERENCES

1. A. Al-Bakri. *Investigation of the Effectiveness of Selected Intersection Transportation System Management Techniques and Traffic Enforcement Levels in a Developing Country: Jordan Case Study.* Ph.D. dissertation. Department of Civil and Environmental Engineering, University of Iowa, Iowa City, 1990.
2. G. Maycock and R. D. Hall. *Accidents at 4-Arm Roundabouts.* Report 1120. U.K. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1984.
3. E. Hauer, J. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988.
4. P. McCullagh and J. A. Nelder. *Generalized Linear Models*, 2nd ed. Chapman and Hall, London, England, 1989.
5. *GLIM 3.77.* Numerical Algorithms Group, Inc., Downers Grove, Ill.
6. A. Agresti. *Categorical Data Analysis.* John Wiley & Sons, New York, N.Y., 1990.
7. E. B. Andersen. Asymptotic Properties of Conditional Maximum Likelihood Estimators. *Journal of the Royal Statistical Society*, Series B, Vol. 32, 1973, pp. 283–301.

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

Estimating Accident Potential of Ontario Road Sections

BHAGWANT N. PERSAUD

The identification of unsafe road locations (blackspots) and the evaluation of treatment effectiveness should be based on the number of accidents expected in the long run (accident potential) rather than on the short-term count. A method for estimating the underlying accident potential of Ontario road sections, using accident and other data, is presented. The method first uses regression models to produce an initial estimate of a section's accident potential on the basis of its traffic and geometric characteristics. This estimate is then refined by being combined with the section's accident count, using an empirical Bayesian procedure. The results indicate that the empirical Bayesian estimates are superior to those based on the accident count or the regression prediction by themselves, particularly for sections that might be of interest in a program to identify and treat unsafe road locations.

The efficient allocation of resources to highway safety programs requires that accident blackspots (unsafe intersections, road sections, etc.) be properly identified and that estimates of the safety benefit of a potential treatment be as sound as possible. Crucial to both aspects is the proper estimation of accident potential of a site being considered for treatment.

For blackspot identification, it has been common to use variants of the "rate-and-number" method (*1*), in which sites registering an unusually high accident potential are selected for detailed examination.

For safety benefits, the safety effect of a treatment is usually estimated from its previous applications by comparing the accident potential of a site before and after it was treated. This estimate of safety effect is then applied to the current estimate of accident potential of a site under consideration for the treatment.

In both cases, recent research (*2,3*) has recognized that accident potential should be the average number of accidents expected in the long run on a site, not the short-term count. For obvious reasons, this underlying long-term accident potential cannot be observed, so it must be estimated.

The primary focus of the research on which this paper is based was to provide estimates of accident potential for Ontario road sections, using data readily available to safety analysts at the Ontario Ministry of Transportation (MTO). The work described is an application of a recently developed empirical Bayesian approach (*3*) that refines prior estimates of accident potential obtained from multivariate regression analysis.

Department of Civil Engineering, Ryerson Polytechnic Institute, 350 Victoria Street, Toronto M5B 2K3 Canada.

DATA

Raw computer data were obtained from two sources at MTO for roads under the ministry's jurisdiction. Inventory information for Linear Highway Referencing System (LHRS) sections and subsections contained information on highway division, category and environment, section length, lane and shoulder widths, roadway and shoulder surface type, number of lanes, speeds and speed limits, and other geometric characteristics.

The other raw data set contained, for each of the years 1983–1986, the total number of accidents and traffic information for LHRS sections and subsections. The traffic information included percentage of commercial vehicles, seasonal and annual average daily traffic (ADT) volumes, and directional split. Details of the accidents were not available, but the absence of more detailed data should not be seen as a shortcoming because the aim of the research was to build functional models that use readily available data.

Models were desired for each of three road clasess. Final data sets were prepared for each road class by combining the raw traffic and inventory data for as many LHRS sections or subsections as possible. Summary information for each final data set is shown in Table 1.

The summary figures in Table 1 show that from 1985 to 1986, Class 1 roads experienced a 20 percent increase in accidents and from 1983 to 1984, a 12 percent increase in traffic. Comparable figures for the other classes show that Class 2 roads had a 7 percent increase in accidents and a 4 percent increase in traffic, while Class 3 roads had a 4 percent increase in accidents and a 3 percent increase in traffic.

METHODOLOGY

The procedure for developing and applying the models for estimating accident potential of road sections closely follows that for estimating accident potential for rail-highway grade crossings (*4*), signalized intersections (*5*), and Ontario drivers (*6*). Further details of the methodology can be obtained from these papers as well as from a recent FHWA publication (*3*). This section summarizes the bare essentials of the methodology.

The procedure provides an estimate of the accident potential of a road section, given (a) its accident history and (b) its traffic, geometric, and other characteristics. The method used to reach this estimate is, in effect, an empirical Bayesian procedure (*7*) that mixes the two sets of information. This procedure is depicted in Figure 1 and summarized in the following as a sequence of two steps.

TABLE 1 SUMMARY INFORMATION FOR FINAL DATA SETS

	CLASS 1 Freeways	CLASS 2 Other Primary	CLASS 3 Secondary Tertiary
Number of sections	404	1680	166
Total length (km)	1583.6	12024.6	2396.5
Average daily traffic (Weighted by length)			
1983	24298	2548	318
1984	25129	2608	326
1985	26219	2610	324
1986	28981	2767	338
Total Accidents			
1983	11612	15853	564
1984	12533	16963	601
1985	14641	18125	662
1986	14420	17033	554

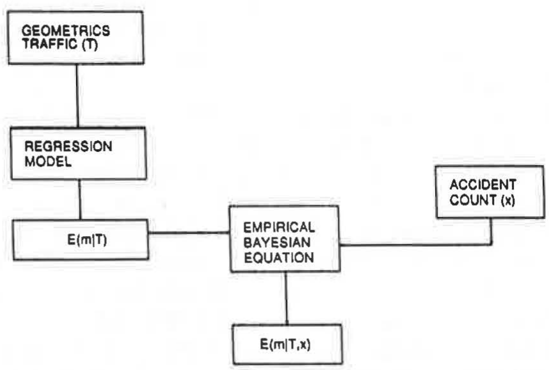


FIGURE 1 Procedure for estimating accident potential.

Regression Models for Initial Prediction of Accident Potential (Step 1)

Generalized linear modeling using the GLIM computer package (*8*) was used to estimate $E(m|T)$, the underlying accident potential of a section for the period 1983–1984, given its characteristics (e.g., traffic volume) for this period. The model form used was

$$E(m|T) = \text{SCL} * a1 * (\text{ADT})^{b1} \tag{1}$$

where

T = set of traffic and geometric characteristics,
SCL = section length (km),
$a1$ and $b1$ = model parameters estimated by GLIM.

The observed accident count on a section in the period 1983–1984 was used as an estimate of the dependent variable $E(m|T)$. GLIM allows the specification of a negative binomial error structure for the dependent variable. This is now known to be more appropriate for accident counts than the traditional normal distribution (*3,5,9*).

The negative binomial error specification follows on assumptions that $(m|T)$ is gamma distributed and that accident occurrence on a section follows the Poisson probability law. Under these conditions (*3,5,9*), the variance of the regression estimates can be estimated from

$$\text{Var}(m|T) = [E(m|T)]^2/k \tag{2}$$

where k is a parameter of the negative binomial distribution.

The procedure for estimating k is iterative. It is first necessary to specify an initial guess of the value of k in order to calibrate the regression model that estimates $E(m|T)$. This guess is then refined by comparing it with the value of k estimated from a maximum likelihood procedure that assumes that each squared residual of the regression model is an estimate of $\text{Var}(m|T)$ and that each count comes from a negative binomial distribution with mean $E(m|T)$ and variance given by Equation 2.

Rearranging the terms of Equation 2 indicates that k can be used as a measure of the variation explained by the regression model, that is, the larger the value of k, the more variation is explained. This is useful because the more traditional R^2 measure is not appropriate when GLIM is used with a negative binomial structure.

Empirical Bayesian Revised Estimates of Accident Potential (Step 2)

In general, two road sections that are similar in all of the independent variables used in the regression model will still be different in true accident potential even though they have the same model predictions according to Equation 1. This is because it is not possible to account in the regression model for all the factors that cause differences in accident potential (e.g., weather).

To account for this shortfall—in effect, to reduce the variation not explained by the regression model—$E(m|T)$ from Equation 1 can be further refined for an individual road section using the accident count (x) on that section to give $E(m|T, x)$, a final revised estimate of accident potential.

It can be shown (*3*) that, under the assumptions stated earlier (i.e., that the variation in $m|T$ can be described by a gamma probability distribution and that accident occurrence on a section follows the Poisson probability law), the revised estimate of accident potential is

$$E(m|T,x) = wE(m|T) + (1 - w)x \quad (3)$$

where

$$w = [1 + E(m|T)/k]^{-1} \quad (4)$$

in which, from Equation 2, $k = E(m|T)^2/\mathrm{Var}(m|T)$.

It can also be shown (*3*) that

$$\mathrm{Var}(m|T,x) = \frac{E(m|T,x)}{1 + [E(m/T)/\mathrm{Var}(m|T)]} \quad (5)$$

The value obtained from Equation 3 is known as an empirical Bayesian estimate. It turns out that the variation in $\{m|T,x\}$ can also be described by a gamma distribution.

Before moving on, it is instructive to explore the meaning of Equation 3. This equation shows that $E(m|T,x)$, the estimated accident potential of a section, is a mixture of what is observed (x) and of $E(m|T)$—what is predicted on the basis of its characteristics (ADT, etc.).

If the prediction has a large variance, that is, there is much unexplained variation, then $\mathrm{Var}(m|T) >> E(m|T)$, w would be small and the accident potential would be close to x. This situation typically arises, according to the hypothesized model, for sections that are long or have a high traffic volume. For such sections, in effect, the accident count could reasonably be used as an estimate of accident potential and the benefit of empirical Bayesian estimation would be marginal.

Conversely, for $\mathrm{Var}(m|T) << E(m|T)$, the accident potential would be close to $E(m|T)$ and a relatively small weight would be given to the observed count (x).

ANALYSIS AND RESULTS

This section is divided into two parts. First, the development of regression models of accident potential for Ontario road sections is described. Next, the empirical Bayesian estimation of accident potential is illustrated.

Regression Models for Initial Estimation of Accident Potential

In this section, the development of the regression models is described. It should be emphasized that these models should not be judged on their ability to explain the causal factors related to accident occurrence. In particular, they should not be used to answer "what if" questions about the impact of altering values of the independent variables.

Equation 1 indicates that the models are of the form

$$E(m|T) = \mathrm{SCL} * a1 * (\mathrm{ADT})^{b1}$$

where $\mathrm{Var}(m|T) = E(m|T)^2/k$. This form is quite common because it ensures that predicted accidents would be zero for an ADT of zero, but it does not, a priori, assume a linear relationship between accidents and traffic volume (*9*).

Whereas GLIM allows the error structure to be specified for $E(m|T)$, it can actually estimate a model for some specified "link" function of $E(m|T)$. In this case GLIM actually estimated models of the log (base e) form

$$\log[E(m|T)] = \log(\mathrm{SCL}) + \log(a1) + b1 * \log(\mathrm{ADT}) \quad (6)$$

In GLIM, the term log(SCL) is specified as an offset that is subtracted from each point estimate of $\log[E(m|T)]$.

The geometric variables selected for the regression models varied with class of road, which should not be surprising because the classes differ significantly in traffic and geometric characteristics. For each class the variable selection was accomplished by using similar statistical procedures to those detailed elsewhere (*8,9*). Models for each road class are described separately later.

Class 1

Freeway sections tend to be all multilane and divided, have the same high geometric standards, and be similar in other such features as speed limit, so it appeared unlikely that a model could incorporate geometric variables and be significantly better than the traffic volume model for this class. Indeed, it turned out that all attempts to add geometric variables—most notably, the number of lanes—were unsuccessful.

The estimates of $a1$ and $b1$ produced by GLIM are given in Table 2 along with other relevant details.

Class 2

The Class 2 road sample was seen as a mixture of urban and rural roads, two-lane and multilane roads, and divided and undivided roadways, giving six possible categories (not eight, because there are no two-lane divided highways). Pertinent details of the road sections in each category are shown in Table 3.

Exploratory analysis revealed that both speed and speed limit varied significantly from category to category but that within each category the variation was quite small. Similarly, for each geometric characteristic that might affect accident occurrence, the variation within individual categories was small. Thus, the use of a road section's category as a variable in an accident prediction model was found to be sufficient to account for variation attributable to geometric and speed factors. The final models for Class 2 roads reflect this reasoning. All attempts at incorporating other variables failed.

The final model form used the full Class 2 data set and allowed the ADT coefficient ($b1$) to vary with the two categorical variables, two-lane/multilane and rural/urban, whereas

TABLE 2 REGRESSION MODEL FOR CLASS 1 ROADS

Model Parameter	Estimated Value
a1 for ADT in thousands	0.6278
b1	1.024

Other statistics: k = 2.95; Observations = 404

TABLE 3 CATEGORIES FOR CLASS 2 ROADS

	RURAL			URBAN*		
	2-lane	Multilane Divided	Multilane Undivided	2-lane	Multilane Divided	Multilane Undivided
Sections	1400	49	90	73	8	60
Length,km	11413.1	188.7	219.9	106.3	5.3	91.3
Average ADT						
1983	2175	8459	11379	4683	16551	12396
1984	2218	8575	11886	4826	18074	13146
1985	2206	8631	12371	4925	19923	13435
1986	2337	9240	13152	5185	20891	14206
Accidents						
1983	12273	570	1429	369	44	1168
1984	13000	724	1631	370	53	1185
1985	13879	740	1660	406	74	1370
1986	12735	653	1767	400	102	1375

* -- The urban category includes all roads with MTO road environment codes 2 and 3 (semi-urban and urban).

the coefficient $a1$ also varied with the categorical variable divided/undivided.

GLIM provided estimates of $\log(a1)$ and $b1$ in Equation 6 for a base category (rural, undivided, two-lane) along with adjustments to be applied to these values for the other five categories. It turns out that the estimated coefficients obtained in this way are the same as those that would have been obtained by separately estimating Level 1 traffic volume models for each category. The estimated values of $a1$ and $b1$ for each category are shown in Table 4.

TABLE 4 REGRESSION MODELS FOR CLASS 2 ROADS

Model Parameter	Estimated value
a1 for ADT in thousands for:	
Rural/undivided/2-lane	1.3392
Rural/undivided/multilane	0.6528
Urban/undivided/2-lane	3.6514
Urban/undivided/multilane	1.4196
Rural/divided/multilane	0.4591
Urban/divided/multilane	0.9984
b1 for:	
Rural/2-lane	0.8310
Rural/multilane	1.3037
Urban/2-lane	0.5588
Urban/multilane	0.8763

Other statistics: k=2.90; Observations=1680

Again, the estimated coefficients are for a 1-km section. Thus, the accident potential of a rural, undivided, multilane road section of length SCL km, for example, could be estimated from

$$E(m|T) = \text{SCL} * 1.3392 * \text{ADT}^{0.8310}$$

Class 3

The modeling exercise for Class 3 roads was similar to that done for Class 2 roads except that now the best explanatory variables turned out to be surface width and surface type. Additional explanatory variables provided little or no improvement.

Exploratory analysis resulted in selection of surface width and surface type as categorical variables, each with two levels, giving four category combinations. Table 5 describes these categories and provides relevant data for each category.

The final model form used the full Class 3 data set and allowed the ADT coefficient ($b1$) and the constant term [$\log(a1)$] to vary with the two categorical variables narrow/wide pavement and low/high class surface. GLIM provided estimates of $\log(a1)$ and $b1$ in Equation 6 for a base category (pavement width $\leq$ 6.1 m, surface type $\leq$ 4) and adjustments to be applied to these values for the other three categories. The

TABLE 5 CATEGORIES FOR CLASS 3 ROADS

	Pavement width ≤ 6.1 m		Pavement width >6.1 m	
	High Class Surface*	Lower Class Surface*	High Class Surface	Lower Class Surface
Sections	11	75	55	25
Length,km	101.2	999.1	771.5	524.7
Average ADT				
1983	625	219	455	246
1984	607	220	477	247
1985	613	220	463	259
1986	637	229	484	271
Accidents				
1983	43	167	257	97
1984	34	186	259	122
1985	41	190	302	129
1986	43	162	248	100

* -- High Class surfaces have MTO surface type codes >4 and consist of concrete, asphalt on concrete pavements and high class bituminous pavements. Lower class surface types have codes ≤ 4.

estimated values of $a1$ and $b1$ for each category are shown in Table 6.

As before, the estimated coefficients are for a 1-km section. Thus, the accident potential of a narrow Class 3 road section with a high class surface and length of SCL km, for example, could be estimated from

$$E(m|T) = \text{SCL} * 0.8988 * \text{ADT}^{0.3884}$$

Illustrating the Method

Consider again a narrow Class 3 road section with a high class surface. Suppose the section is 0.6 km long, has an ADT of 5,900, and recorded six accidents during 1983–1984.

From this example, the relevant regression model gives

$$E(m|T) = 0.6 * 0.8988 * 5.9^{0.3884} = 1.075$$

From Table 6, $k = 2.81$, giving $\text{Var}(m|T) = 1.075^2/2.81 = 0.416$. From Equation 4, $w = 0.72$, which gives, from Equation 3, $E(m|T,x) = 2.46$.

This is the value of accident potential that should be used in the blackspot identification process and in estimating the effect of any treatment that might be applied. The variance of this estimate, $\text{Var}(m|T,x)$, is 0.30 (from Equation 5). Without a model, the best estimate of accident potential is 6, which, it appears, might have been a randomly high accident count.

VALIDATION

It was shown earlier how an initial accident potential estimate from a regression model could be combined with the actual accident count on a road section to yield an empirical Bayesian estimate of accident potential for that section.

The point was made, however, that this estimation process (as opposed to using the accident count as an estimate of accident potential) is not as vital for sections with higher values of $E(m|T)$—the regression model prediction. Thus, it is reasonable that the true test of the method should be in its application to sections with low values of $E(m|T)$, that is, those that are short or have low traffic volume. Because the data obtained were at the level of LHRS sections, many of which tend to be relatively long or have high values of $E(m|T)$, there could be only limited testing of the empirical Bayesian procedure with the current data set.

TABLE 6 REGRESSION MODELS FOR CLASS 3 ROADS

Model Parameter	Estimated value
log(a1) for ADT in thousands for:	
Narrow/low class	0.9961
Narrow/high class	0.8988
Wide/low class	1.2348
Wide/high class	1.3336
b1 for:	
Narrow/low class	0.5844
Narrow/high class	0.3884
Wide/low class	0.7688
Wide/high class	0.6313

Other statistics: k = 2.81; Observations=166

Nevertheless it was possible to demonstrate the usefulness of the models using a sample of Class 2 roads. The sample consists of all 505 rural, two-lane road sections in the data set that are no more than 4 km long. From this sample, the road section having the highest value of accidents per kilometer during 1983–1984 in each of several ADT ranges was identified as a potential blackspot (the width of the range varied to allow for 10 to 15 sections in each range). From the model forms indicated by Equation 1 and Table 2, it can be seen that, because ADT is now the only regression variable in this sample, the same sections would be identified if the selection were properly based on $E(m|T,x)$ per kilometer.

The 39 sections identified are listed in Table 7 along with other items, including the ADT, accidents, regression model predictions, and empirical Bayesian estimates of accident prediction. Period 1 is 1983–1984; Period 2 is 1985–1986.

Several observations follow from the results in Table 7. First, there is a substantial decrease in accident count from Period 1 to Period 2. Assuming that the sections were largely untreated during those 4 years (a reasonable assumption, according to information in the inventory file), the logical conclusion is that this difference occurs because of a random up-fluctuation in the accident count during Period 1. The difference is even more pronounced if the Period 1 count is adjusted to account for the fact that rural Class 2 roads had a 5.3 percent increase in accidents in Period 2.

The second observation is that, because the sections were not selected on the basis of the Period 2 accident count, this count is an unbiased estimate of true accident potential of these sections. On this basis, an unbiased estimate of the total of the true accident potentials of these sections is 503. Thus, using the Period 1 accidents as estimates of accident potential overestimates this value by 34 percent. By contrast, using the regression prediction by itself as an estimate of accident potential underestimates this value (compare 243 and 503).

The third and most important observation is the close correspondence between the total of the empirical Bayesian estimates, $E(m|T,x)$ (487, or the adjusted value of 512), and the total of the estimates of true accident potential (503).

Finally, note that, even though the Period 2 accident count is an unbiased estimate of the true accident potential, the variance of this estimator (equal to the count, assuming a Poisson distribution of the total count on these sections) is substantially higher than that of the empirical Bayesian estimator (shown in the last column of Table 7). This is also the case when the variance of the regression prediction is compared with the variance of the empirical Bayesian estimator.

Though the final observation is of considerable interest, it is immaterial when considering the practical need to use present accident counts to estimate *future* accident potential of road sections. In such an application, the only alternative estimators are the recent accident count, the emperical Bayesian estimator, and perhaps the regression prediction. On the basis of this limited validation exercise, it is apparent that the empirical Bayesian estimator might be preferred.

SUMMARY AND CONCLUSIONS

The research used data available at the level of total accidents on relatively long road sections. Despite this limitation, useful

TABLE 7 ACCIDENT POTENTIAL OF SELECTED RURAL, TWO-LANE CLASS 2 ROADS

	Section km		DATA FOR PERIODS 1, 2: AADT		Accidents					
LHRS#	Start	Length	1	2	1	2	E(m¦T)	Var(m¦T)	E(m¦T,x,n)	Var(m¦T,x,n)
26420	0.0	1.2	195	210	8	5	0.413	0.059	1.359	0.169
43250	0.0	1.3	650	655	12	4	1.217	0.511	4.405	1.302
18170	10.3	1.0	665	655	11	8	0.954	0.314	3.441	0.852
34540	0.7	0.7	1050	1050	10	5	0.976	0.329	3.249	0.818
25080	4.2	1.6	1200	1275	6	6	2.493	2.144	4.114	1.902
10375	0.0	1.3	1375	1475	3	0	2.268	1.774	2.590	1.137
25483	0.0	2.9	1400	1550	8	7	5.137	9.099	6.967	4.453
26810	0.0	2.1	1625	1600	33	24	4.210	6.112	21.258	12.587
21600	1.1	1.1	1675	1750	61	33	2.262	1.764	27.998	12.267
43800	0.0	3.5	1725	2150	15	8	7.374	18.750	12.847	9.221
17400	0.0	1.0	1875	1950	5	4	2.258	1.758	3.458	1.514
11900	0.0	1.6	2000	2000	8	8	3.812	5.010	6.190	3.516
16940	4.4	2.2	2200	2100	11	8	5.673	11.099	9.198	6.087
10360	0.0	1.1	2350	2450	7	4	2.996	3.096	5.031	2.557
11330	0.0	0.6	2500	2550	4	4	1.721	1.021	2.569	0.957
46250	0.0	2.2	2600	2625	16	12	6.518	14.650	13.080	9.053
10350	7.2	1.2	2750	2925	8	6	3.725	4.785	6.129	3.446
32410	5.4	2.0	2800	2925	9	8	6.302	13.695	8.150	5.581
20040	0.0	1.0	2950	3100	10	13	3.291	3.734	6.857	3.645
19410	0.0	0.7	3050	3325	8	9	2.368	1.934	4.900	2.203
14946	0.0	2.1	3225	3775	12	9	7.442	19.096	10.722	7.715
10770	8.4	2.1	3425	3550	12	12	7.823	21.104	10.870	7.931
20715	0.0	2.0	3625	2800	20	12	7.810	21.035	16.699	12.178
14680	0.0	1.2	3700	3725	10	5	4.767	7.835	8.020	4.987
34735	2.7	1.2	3850	6350	10	7	4.927	8.370	8.120	5.111
35240	6.3	1.1	4125	3975	11	9	4.783	7.887	8.653	5.387
23460	0.0	1.4	4225	3575	12	6	6.209	13.295	10.157	6.923
16140	0.0	1.6	4475	5450	1	11	7.444	19.106	10.003	7.198
19435	0.0	2.3	4600	4800	28	21	10.948	41.331	24.429	19.313
23706	4.7	1.9	4900	4550	24	5	9.532	31.328	20.625	15.814
12200	3.9	4.0	5250	5400	37	44	21.251	155.720	35.109	30.893
29500	1.9	0.4	5600	5850	6	0	2.242	1.734	3.881	1.692
28540	0.0	0.6	5900	6050	17	12	3.512	4.254	10.900	5.970
33010	1.8	0.8	6050	6300	8	7	4.782	7.885	6.785	4.224
40330	2.8	1.6	6850	6650	34	27	10.603	38.768	28.975	22.752
35430	0.0	1.5	7525	7150	23	23	10.748	39.834	20.397	16.063
16490	1.9	1.4	8500	8950	26	33	11.100	42.488	22.914	18.167
16100	0.0	2.2	10200	10400	49	44	20.297	142.055	45.412	39.734
10550	0.0	0.8	12450	13650	38	30	8.710	26.162	30.684	23.020
TOTAL		60.5			641	503	230.9	750.9	487.1	338.3
(ADJUSTED)*					(675)		(243)		(513)	(375)

* Adjustments account for a 5.3% increase in accidents from period 1 to 2.

insights were gained into the estimation of accident potential for Ontario road sections. A summary and important conclusions of the research follow:

1. An empirical Bayesian procedure for the estimation of accident potential of road sections was presented. This estimator combines the recent accident count with a regression model prediction of accidents expected on the basis of the section's traffic volume and geometric characteristics.
2. The regression models developed make sense intuitively, but, because they are intended for use in initial accident prediction from variables whose values are readily available, they should not be judged on their ability to explain the causal factors related to accident occurrence. In particular, they should not be used to answer "what if" questions about the impact of altering values of the independent variables.
3. When the regression predictions were combined with the accident count information for a set of potential blackspot sections, using the resulting empirical Bayesian estimates was observed to be superior to using the accident count or the regression prediction alone. The results confirm that the accident count of a road section is, by itself, not a good estimator of accident potential, particularly for road sections that might be of interest when considering and applying treatment in a blackspot identification and treatment process. The accident count is a reasonable estimator of accident potential if sections are identified at random, that is, without regard to the accident count itself, and then only, according to the hypothesized models, for sections with that have relatively high traffic volume, section length, or accident frequency.
4. The procedures presented can be used for estimating accident potential for a number of applications, including (a) estimating the safety effect of improvements, (b) identifying potential blackspots, and (c) estimating the safety benefits (or disbenefits) of potential treatments at blackspots or other sites being considered in a resource-allocation procedure.
5. The procedures presented are currently useful for Ontario road sections, but it is still desirable to test their applicability outside Ontario and to seek to build improved regression models of accident potential. For example, separate models could be constructed for single-vehicle accidents and collisions, and for intersection and nonintersection accidents. Sim-

ilar models should be constructed for roads and intersections under the jurisdiction of municipalities.

ACKNOWLEDGMENTS

This paper resulted from research performed under a contract with the research and development branch of MTO. This support, along with the guidance of Alex Ugge, is gratefully acknowledged.

REFERENCES

1. J. I. Taylor and H. T. Thompson. Identification of Hazardous Locations. Report FHWA-RD-77-83. FHWA, U.S. Department of Transportation, 1977.
2. J. L. Higle and J. M. Witkowski. Bayesian Identification of Hazardous Locations. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 24–36.
3. E. Hauer. Empirical Bayesian Approach to the Estimation of Unsafety—The Multivariate Approach. Report FHWA-RD-90-006. FHWA, U.S. Department of Transportation, 1990.
4. E. Hauer and B. N. Persaud. How To Estimate the Safety of Rail-Highway Grade Crossings and the Safety Effects of Warning Devices. In *Transportation Research Record 1114*, TRB, National Research Council, Washington, D.C., 1987, pp. 131–140.
5. E. Hauer, J. C. N. Ng, and J. Lovell. Estimation of Safety at Signalized Intersections. In *Transportation Research Record 1185*, TRB, National Research Council, Washington, D.C., 1988, pp. 48–61.
6. E. Hauer, B. N. Persaud, A. Smiley, and D. Duncan. Estimating the Accident Potential of Ontario Drivers. *Accident Analysis and Prevention* (in press).
7. C. N. Morris. Parametric Empirical Bayesian Inference: Theory and Applications. *Journal of the American Statistical Association*, Vol. 78, No. 3, 1983, pp. 47–65.
8. R. J. Baker and J. A. Nelder. *Generalized Linear Interactive Modelling. The GLIM System—Release 3*. Rothamsted Experimental Station, Harpenden, Hertfordshire, England, 1978.
9. G. Maycock and R. D. Hall. *Accidents at 4-Arm Roundabouts*. Report 1120. Transport and Road Research Laboratory, Crowthorne, Berkshire, England, 1984.

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

Conflicts at Traffic Circles in New Jersey

AHMAD SADEGH, JAYESH MEHTA, AND MARK SMITH

Traffic circles create irregular traffic patterns characterized by continuous vehicle weaving and lane changing. Because of this, circles are a potential source of confusion to drivers. Therefore, a study was conducted to determine whether improvement of guidance signing used at traffic circles would lessen confusion and increase safety. Five circles were selected as test sites, and diagrammatic guidance signs were installed at the approaches under study. The five circles were individually videotaped for 2 days before and 2 days after installation of the signs. The results of conflict analysis at all five circles indicated that the diagrammatic signs effectively reduced the number of confusion-oriented conflicts.

In 1925 New Jersey became the first state to develop and construct traffic circles. Various other states and cities soon followed its lead. Today, some 70 of these are still in operation in New Jersey.

Traffic circles, by nature, create irregular traffic patterns, characterized by continuous vehicle weaving and lane changing and attended by a large variance in vehicle speeds. Because of this, circles are a potential source of confusion to drivers.

Traffic circles as originally designed worked well at low volume and low speed flow, but the growth of traffic over the years has reduced their effectiveness. Improving circle effectiveness has often required major changes, such as changing a regular circle to a cut-through circle by continuing a major road entering the circle through the central island and producing two at-grade signalized intersections at the points where this road crosses the original circle. However, in some instances, traffic conditions at circles may be improved by minor, less expensive changes, such as improving the motorist information system.

STUDY OBJECTIVES

In 1985 the New Jersey Department of Transportation (NJDOT) initiated an in-house study in which diagrammatic guidance signs similar to the ones used at traffic circles in the United Kingdom (with minor modifications for adopting U.S. standards) were placed at five circles in New Jersey (see Figure 1). The purpose of this study by NJDOT was to determine the effectiveness of diagrammatic signs on the basis of percentage of vehicles making preferred maneuvers. The result of this study concluded that diagrammatic guidance signing was more effective than conventional signing in reducing driver confusion at circles (*1*).

A. Sadegh, Department of Civil and Environmental Engineering, New Jersey Institute of Technology, Newark, N.J. 07012. J. Mehta, Port Authority of New York and New Jersey, One World Trade Center, 72N1, New York, N.Y. 10048. M. Smith, New Jersey Department of Transportation, 1600 North Olden Avenue, Trenton, N.J. 08625.

Before-and-after studies were conducted, consisting of data collected with existing conventional signing (before) and data collected after diagrammatic guidance signs were installed (after).

The preferred maneuvers were determined on the basis of safety and effectiveness. Percentages of vehicles making the preferred maneuvers were measured for both before and after the sign changes and then compared to determine the effectiveness of the new signing.

This paper is a supporting study on the study by NJDOT. The main objective of this research was to determine the effect of improved guidance signing at traffic circles on conflicts at these circles. Conflicts can be used as additional measures of effectiveness for determining the effects of the new signing on safety and effectiveness of traffic circles.

LITERATURE SEARCH

A Highway Research Information Service search by NJDOT turned up 112 articles related to traffic circles. However, most of the articles concerned either the design of traffic circles or the calculation of capacity for traffic circles, neither of which was pertinent to this study. Only six articles addressed the topic of either guidance signing at circles or measures of effectiveness useful for analyzing traffic flow at circles.

Two of the three articles about guidance signing described a study that compared two improved methods of signing—diagrammatic and modified stack (*2,3*). However, the improved signing was not compared with conventional signing. The third article was an analysis of the effect of diagrammatic signing on traffic at one circle in Washington, D.C. (*4*). It used a driver survey rather than measures of effectiveness to determine sign effectiveness.

In regard to measures of effectiveness, two of the articles were concerned mainly with vehicle paths through circles (*5,6*). The other studied the use of traffic conflict techniques to assess the safety of road design elements, which eventually became the major measure of effectiveness used in this study. The article contained some information about the use of this technique at a miniroundabout (small circle). Although several articles addressed some pertinent aspects, the fact remains that very little work has been done concerning guidance signing at traffic circles.

STUDY PROCEDURES

Traffic Conflicts

The term "conflict" in traffic engineering was introduced by Perkins and Harris in 1967. They defined traffic conflict as

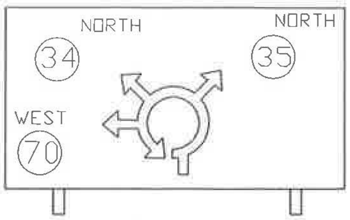


FIGURE 1 Diagrammatic guidance sign.

any potential accident situation that leads to evasive actions such as braking or swerving (7). In studies the criterion for evasive actions is simply to determine brakelight indications or lane changes under various sets of circumstances.

A traffic conflict is an observable situation in which two or more road users approach each other in space and time to such an extent that there is a risk of collision if their movements remain unchanged. In other words, a conflict may not lead directly to a collision, but it is an event parallel with a collision.

For conflict studies to serve as an analytical tool, they should be made before and several months after the device is implemented. If outside influences are held relatively constant, the effects of the device change can then be estimated by the observed changes in conflict risks. Many problems related to analysis of accidents can be solved by estimating risks by a conflict technique. The traffic conflicts technique is a device for indirectly measuring safety. It requires a count of conflicts occurring, which gives the basis on which the conflict rate is estimated. The conflict method is especially suitable when the effects of safety devices and measures are to be investigated.

The results of the literature search and traffic observation indicated that measures of effectiveness based on various types of conflicts would be useful for determining changes in the level of driver confusion at a circle.

In this study, brakelight indication was used as one of the measures of traffic conflicts at circles before and after installation of new guidance signs. Brake applications can be identified and counted easily. Because brakes are applied in almost all categories of conflicts, subjectivity in data collection can be avoided. Using brake application as a measure has the following disadvantages:

1. Braking habits vary from driver to driver. Some drivers are very cautious and may apply brakes on entering a circle, regardless of hazard present; others may not brake even when a hazardous situation is present.
2. Braking information does not give us information about the severity of a conflict situation.
3. Common procedure of observing brake application by only one of the vehicles involved in a conflict situation would not consider the information describing the actions of the other vehicles.
4. Brakelights may not be visible because of mechanical failures.

Characteristics of Conflicts

For the purpose of this study, all types of conflicts observed at regular or cut-through circles were divided into the following two categories:

Confusion-Oriented Conflicts

This type of conflict can be characterized as one resulting from a driver's difficulty in making the right decision in time. The following types of conflicts are categorized as confusion-oriented conflicts:

Left-turn conflict A left-turn conflict is a situation in which a left-turning vehicle crosses directly in front of an opposing through vehicle. The criterion of the conflict is the evasive action, braking, or lane changing of the through vehicle.

Lane-change conflict Lane-change conflict is defined as a situation in which a vehicle changes lanes into the path of another vehicle. The offended vehicle must brake to avoid a collision.

Cross-traffic-from-left conflict It is defined as a situation in which a vehicle crosses or turns into the path of a through vehicle, causing the through vehicle to brake to avoid a collision.

Erratic maneuvers A more severe form of confusion conflict is an erratic maneuver, which is any sudden, unexpected vehicle movement that could cause an accident. An erratic maneuver usually involves only one vehicle making an unsafe move independent of other vehicles. Such a maneuver may 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.

An erratic maneuver can also be defined as any movement that involves a sudden disruption in the continuity of direction or speed of a vehicle or a deviation from the traveled path intended by design and traffic engineers responsible for geometric configuration and marking in the area of interest.

Following are different erratic maneuvers found at circles:

Use of painted gore area Examples include

- *Cross painted gore*: Vehicle traverses the gore pavement marking while either exiting or continuing through.
- *Stop in painted gore*: Vehicle comes to a complete stop in any part of gore before exiting or continuing through.
- *Back up*: Vehicle passes the gore area, stops, and then backs up to change direction.

Lane change Vehicle traverses one or more full lanes within the deceleration lane area in order to exit.

Traffic-Oriented Conflicts

Traffic-oriented conflicts can be characterized as conflicts that are caused by existing roadway geometry and traffic conditions at a particular moment. These conflicts are primarily due to heavy traffic rather than confusion. The following types of conflicts are categorized as traffic-oriented conflicts:

Cross-traffic-from-right conflict This type of conflict occurs when vehicles entering the circle obstruct the path of circulating traffic and vehicles exiting.

Red-light–violation conflict It is a situation in which a vehicle enters the intersection on a red signal. Vehicles that have entered the intersection legally and complete their movements after signal changes are not considered violators.

Vehicle-passing-on-amber conflict When a vehicle enters the circle intersection after the traffic signal has changed from green to amber, it might confuse the driver of the following vehicle in deciding to cross or to stop at intersection.

Rear-end conflict When a vehicle stops unexpectedly and causes a following vehicle to take evasive action to avoid a rear-end collision, it is defined as a rear-end conflict. Such conflicts are primarily due to heavy traffic rather than confusion. Rear-end conflicts are further divided into the following subcategories:

Stop-on-amber–rear-end conflict This occurs when a vehicle stops unexpectedly because of an amber traffic signal, causing the following vehicles to apply brakes.

Slow-vehicles–rear-end conflict A slow-moving vehicle causes the following vehicle, which is moving at regular speed, to apply brakes to avoid collision.

Slow-for-traffic conflict It happens when a vehicle slows or stops because of a traffic conflict and causes a following vehicle to take evasive action to avoid a rear-end collision.

SITE SELECTION

A guideline consisting of four criteria was set up for the purpose of determining test sites for guidance signing. The four criteria were

1. The traffic circle should be expected to have a significant number of unfamiliar drivers;
2. The approaches to the circle should be state highways, if possible;
3. The circle should have high weaving volumes; and
4. There should be some evidence of driver confusion that is susceptible to correction by improved signing.

Using these guidelines, five circles were selected as test sites, three of which are regular circles and two, cut-through circles. The three regular circles are Freehold (US-9 and NJ-33), Lakehurst (NJ-70 and NJ-37), and Brielle (NJ-34, NJ-35, and NJ-70). The two cut-through circles are Marlton (NJ-70 and NJ-73) and Livingston (NJ-10 and Eisenhower Parkway).

Freehold Circle has Freehold Raceway located just off the circle on one of its legs, and four out of its five approaches are state highways. This should account for a fair amount of drivers, many of whom are expected to be unfamiliar with this circle, traveling through the circle.

The next two circles, Brielle and Lakehurst, are on major routes leading to New Jersey–shore resort areas. All the approaches to these two circles are state highways. These circles, particularly during the summer months, experience heavy recreational traffic going to and from the New Jersey shore. It was assumed that a good percentage of this traffic is composed of drivers who are unfamiliar with these circles.

The Marlton Circle is a good representative of the typical cut-through design currently in use. The Livingston Circle, although exhibiting a typical cut-through design, is a five-legged circle. However, because anticipated cut-throughs include several circles that have more than four legs, it was decided to include Livingston Circle as a test site.

After completing before-installation traffic studies, the diagrammatic guidance signs were installed at the approaches under study for all traffic circles.

The new signs were placed on the approach at least 250 ft from the circle, along with circle warning signs and standard road junction signs.

DATA COLLECTION

Traffic studies were used to document traffic conditions preceding and following sign installation. Data were collected for 2 days at each circle during before and after periods.

The time selected for videotaping by NJDOT was based on its pilot study. The main aim was to obtain a high number of unfamiliar drivers. Therefore, the 12:00–1:00 p.m. period was excluded from the study because of the high incidence of lunch trips associated with this hour. So that the variation in traffic patterns could be studied, data were collected for selected weaving areas from videotapes recorded for morning and afternoon periods. The morning period ran from 10:30 a.m. to 12:00 p.m., and the afternoon period from 1:00 to 2:30 p.m. The after condition studies were conducted on the same day of the week and for the same day of the year as the before condition studies.

A minimum of 1 month was allowed to elapse between installation of the signs and the after condition studies to allow drivers to familiarize themselves with the new signing.

During the 3-hr observation periods for two different days of week, data were recorded for traffic flow in 15-min intervals. Fifteen-min periods appear to be quite representative for collecting data from videotapes.

After preliminary viewing of the tapes, different types of conflicts found at circles were defined. It was not possible to measure the distances between conflicting vehicles or the relative speeds of conflicting vehicles, so it was decided to observe vehicles with brakelight indication criteria. Because data were collected for only two legs out of four or five at the circles, all conclusions were made based on data collected for those two legs.

Conflicts were classified on basis of traffic flow. Two types of traffic flows were found at circles: entering traffic and circulating traffic. Entering traffic flow consisted of vehicles entering the circle from an approach, and circulating traffic flow consisted of vehicles already on the circle.

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, traffic volumes, and different driving attitudes of the observers.

During data collection, initially all the videotapes were observed to classify different conflicts as they occurred and to help to ensure consistency during final observation of conflict counts.

CONFLICT ANALYSIS

After studying traffic patterns, traffic conflicts were classified in various categories. At circles, conflicts due to stopping or slowing of vehicles, cross-traffic movement, sudden lane change, and use of painted gore area for right turns or lane changing were most frequent.

Because of differences between traffic operations at cut-through and regular circles, different types of conflicts had to be collected at each circle. During the initial period of observation, it was necessary to classify conflicts between those based on confusion and those that were based on traffic.

These conflicts would be useful in showing any changes in the level of driver's confusion at a circle.

All five circles were individually studied for 2 days before and 2 days after diagrammatic signs were installed. The following section describes each site and lists different types of conflicts observed in each site.

Freehold Circle

Freehold Circle is a typical regular circle with five approaches. It is in the town of Freehold, New Jersey (Figure 2). With four state highway approaches and Freehold Raceway just off on one of its legs, it receives a fair volume of unfamiliar drivers. At Freehold, two routes—US-9 and NJ-33—were studied. There was concentration on two approaches, US-9 northbound (NB) and NJ-33 eastbound (EB). After reviewing the before-and-after studies, the following types of conflicts were found:

1. *Confusion-oriented conflicts*: (a) cross traffic from left and (b) lane changes;
2. *Traffic-oriented conflicts*: (a) cross traffic from right, (b) rear-end, and (c) sudden slowing of circulating traffic.

Lakehurst Circle

Lakehurst Circle is a three-legged circle. All three approaches are state highways (Figure 3). The diagrammatic sign was installed on NJ-70 EB approach, and NJ-70 and NJ-37 were studied. The concentration was on two approaches: NJ-70 EB and NJ-37 westbound (WB). The following types of conflicts were found at this circle:

1. *Confusion-oriented conflicts*: (a) cross traffic from left and (b) erratic maneuvers (lane changes);
2. *Traffic-oriented conflicts*: (a) cross traffic from right and (b) rear-end.

Brielle Circle

Brielle Circle is a four-legged regular circle connecting NJ-34, and NJ-35, and NJ-70 (Figure 4). Diagrammatic signs were installed on the NJ-70 EB and NJ-35 NB approaches, and the study concentrated on NJ-35 NB and NJ-35 southbound (SB) approaches. The following types of conflicts were found at this circle:

1. *Confusion-oriented conflicts*: (a) cross traffic from left, (b) erratic maneuvers (including sudden change of lane due to confusion and entering or leaving circle from a wrong lane);
2. *Traffic-oriented conflicts*: (a) cross traffic from right and (b) rear-end.

Marlton Circle

Marlton Circle is a typical cut-through circle with four approaches; it connects NJ-70 to NJ-73 (Figure 5).

At Marlton Circle, diagrammatic signs were located on NJ-70 EB and NJ-70 WB approaches. The study concentrated on the NJ-70 EB and NJ-73 approaches. Erratic maneuvers

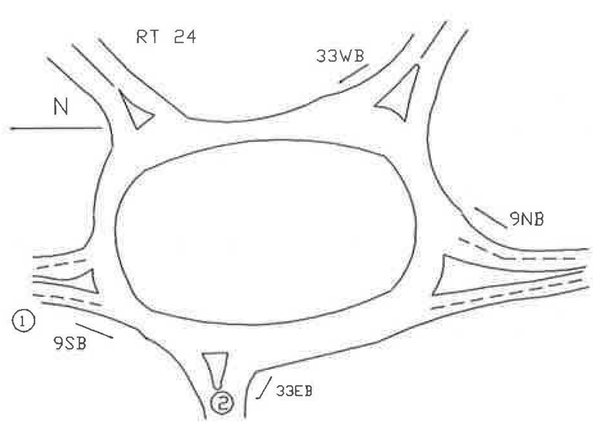


FIGURE 2 Freehold Circle.

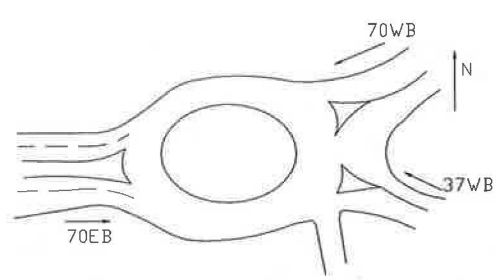


FIGURE 3 Lakehurst Circle.

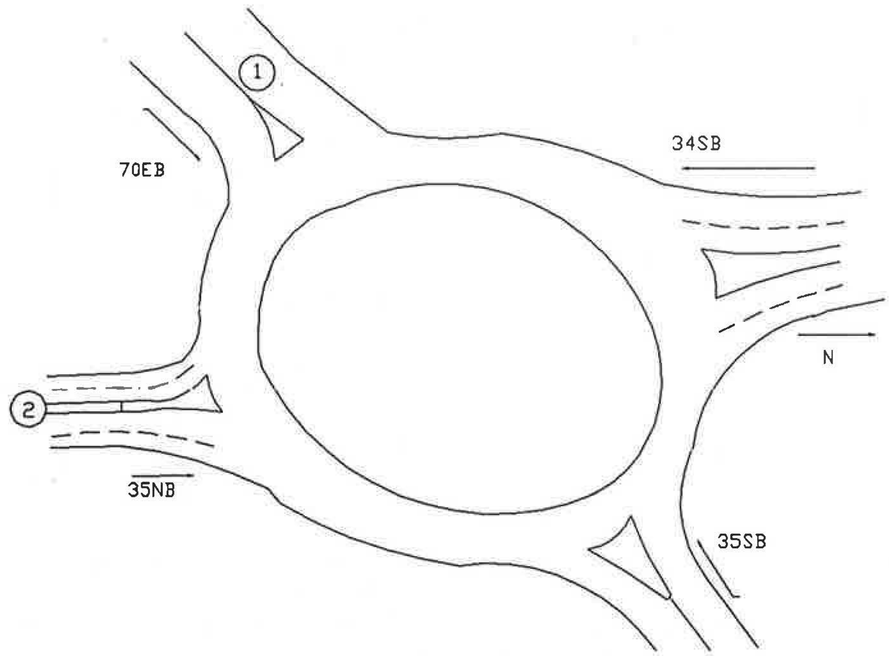


FIGURE 4 Brielle Circle.

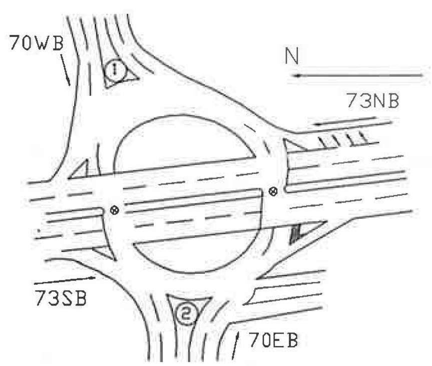


FIGURE 5 Marlton Circle (cut-through).

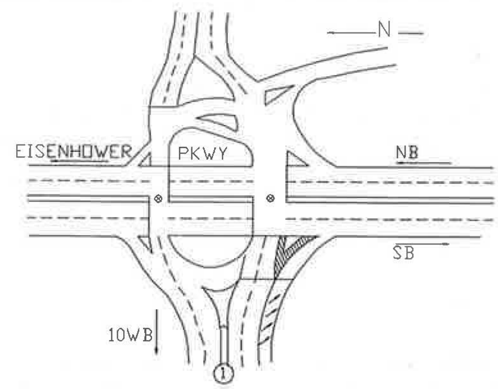


FIGURE 6 Livingston Circle (cut-through).

and conflict occurrence by vehicles were obtained through videotape. The following types of conflicts were compared:

1. *Confusion-oriented conflicts*: (a) cross traffic conflicts, (b) vehicles stopped for right turn, and (c) lane change;
2. *Traffic-oriented conflicts*: (a) vehicle passing on amber and (b) use of painted gore area.

Livingston Circle

Livingstone Circle is a five-legged cut-through circle connecting NJ-10 to the Eisenhower Parkway (Figure 6). The diagrammatic sign was placed on the NJ-10 EB approach, and the study concentrated at the NJ-10 EB and Eisenhower Parkway approaches. The following types of conflicts were observed at Livingston Circle:

1. *Confusion-oriented conflicts*: (a) use of wrong lane for left turn and (b) erratic maneuvers (lane change);
2. *Traffic-oriented conflicts*: (a) rear-end and (b) use of painted gore area.

DISCUSSION OF RESULTS

Freehold Circle

The results of conflict analysis at Freehold Circle before and after diagrammatic sign installation are shown in Table 1. Two cases were compared in which each case included 1 day before and 1 day after installation of the signs. Volumes shown for this and all other circles were taken in the circle weaving areas where the conflict data was collected.

The results indicated that the number of confusion-oriented conflicts was reduced in both cases after installation of guidance signs: a 40 percent reduction in Case 1 and a 30 percent reduction in Case 2 were realized after sign installation.

During the two cases of before-and-after studies, there were reductions of 70 and 85 percent in lane-change conflicts after installation of the signs.

The reduction in the number of conflicts caused by sudden slowing of circulating traffic was countered by the increase in

TABLE 1 RESULTS OF CONFLICT STUDIES AT FREEHOLD CIRCLE

	CROSS TRAFFIC FROM LEFT	CROSS TRAFFIC FROM RIGHT	REAR END	LANE CHANGE	SUDDEN SLOWING OF TRAFFIC	TOTAL VOLUME	CONFUSION ORIENTED	TRAFFIC ORIENTED
BEFORE I (10/01/85)								
10:30 - 12:00	37	21	15	23	26	2465	60	62
13:00 - 14:30	41	28	30	40	63	3206	81	121
TOTAL	78	49	45	63	89	5671	141	183
AFTER I (09/30/86)								
10:30 - 12:00	32	38	17	10	17	2618	42	72
13:00 - 14:30	36	49	36	10	19	3301	46	104
TOTAL	68	87	53	20	36	5919	88	176
BEFORE II (10/11/85)								
10:30 - 12:00	46	25	15	30	30	2750	76	70
13:00 - 14:30	51	35	38	55	76	3278	106	149
TOTAL	97	60	53	85	106	6028	176	219
AFTER II (10/10/86)								
10:30 - 12:00	47	33	29	5	24	3033	52	86
13:00 - 14:30	73	47	48	9	35	3519	82	130
TOTAL	120	80	77	14	59	6552	134	216

the number of cross-traffic-from-right and rear-end conflicts that resulted in reductions of only 7 and 9 percent in the number of traffic-oriented conflicts after installation of the guidance signs.

Lakehurst Circle

The results of conflict analysis at Lakehurst Circle and after diagrammatic sign installation are shown in Table 2.

The results indicate that the number of confusion-oriented conflicts was reduced in each case after installation of the signs—by 24 percent in Case 1 and 26.4 percent in Case 2. These reductions were based mainly on 53 percent reduction in Case 1 and 52.3 percent reduction in Case 2 in the number of lane-change conflicts after installation of the signs, which indicated that drivers were more aware of their routes from the information obtained from the guidance signs.

Traffic-oriented conflicts, which were composed of cross-traffic-from-right and rear-end conflicts, showed 28 and 19 percent increases in the numbers of conflicts after installation of the signs. These increases were caused mainly by the growth of 22 percent and 26 percent in the number of rear-end conflicts, which was expected because of the increase in the volume of traffic.

Brielle Circle

Table 3 shows the results of conflict analysis at Brielle Circle. The results show reductions of 7 percent (Case 1) and 13.3 percent (Case 2) in the numbers of confusion-oriented conflicts.

The numbers of lane-change conflicts were reduced by 36 percent and 35 percent in both cases after installation of the signs, which supports the previous findings that drivers were

TABLE 2 RESULTS OF CONFLICT STUDIES AT LAKEHURST CIRCLE

	CROSS TRAFFIC FROM LEFT	CROSS TRAFFIC FROM RIGHT	REAR END	LANE CHANGE	TOTAL VOLUME	CONFUSION ORIENTED	TRAFFIC ORIENTED
BEFORE I (07/03/85)							
10:30 - 12:00	29	2	105	30	1070	59	107
13:00 - 14:30	13	1	125	18	496	31	126
TOTAL	42	3	230	48	1566	90	233
AFTER I (07/03/86)							
10:30 - 12:00	29	1	165	10	1108	39	166
13:00 - 14:30	20	1	150	14	558	34	151
TOTAL	49	2	315	24	1666	73	317
BEFORE II (07/25/85)							
10:30 - 12:00	20	1	121	23	1046	43	122
13:00 - 14:30	42	1	116	17	1024	59	133
TOTAL	62	2	237	40	2070	102	255
AFTER II (07/24/86)							
10:30 - 12:00	29	0	147	13	1064	42	147
13:00 - 14:30	29	0	162	7	1088	36	169
TOTAL	58	0	309	20	2152	78	316

TABLE 3 RESULTS OF CONFLICT STUDIES AT BRIELLE CIRCLE

	CROSS TRAFFIC FROM LEFT	CROSS TRAFFIC FROM RIGHT	REAR END	LANE CHANGE	TOTAL VOLUME	CONFUSION ORIENTED	TRAFFIC ORIENTED
BEFORE I (07/12/85)							
10:30 - 12:00	33	2	118	37	1886	70	120
13:00 - 14:30	30	2	141	65	2473	95	143
TOTAL	63	4	259	102	4359	165	263
AFTER I (07/11/86)							
10:30 - 12:00	40	0	138	35	2139	75	138
13:00 - 14:30	52	2	193	33	2405	85	195
TOTAL	92	2	331	68	4544	160	333
BEFORE II (07/01/85)							
10:30 - 12:00	27	7	71	34	2050	61	78
13:00 - 14:30	37	1	125	62	2228	99	126
TOTAL	64	8	196	96	4278	160	204
AFTER II (07/21/86)							
10:30 - 12:00	31	1	102	33	2309	64	103
13:00 - 14:30	44	4	151	29	1909	73	155
TOTAL	75	5	253	62	4218	137	258

more aware of their routes from information obtained from the guidance signs.

A 23 percent and 30 percent increase in the number of rear-end conflicts after installation of the signs largely contributed to the increases of 22 percent and 28 percent in the number of traffic-oriented conflicts.

Marlton Circle

The results of conflict analysis at Marlton Circle before and after diagrammatic sign installation are shown in Table 4.

The results indicate that the numbers of confusion-oriented conflicts were reduced by 10 percent (Case 1) and 17 percent (Case 2). These reductions were more obvious in the number of lane-change conflicts (reduction of 29 percent for Case 1 and 67 percent for Case 2).

The results also indicate that the numbers of traffic-oriented conflicts were increased by 69 percent (Case 1) and 37 percent (Case 2) after installation of the signs. The increase in the number of traffic conflicts was mainly caused by the sharp increases of 98 percent (Case 1) and 142 percent (Case 2) in the number of conflicts caused by the drivers using the painted gore area. These increases are due to the increase in traffic. Because of extensive queueing in the two lanes at the intersection, right-turning vehicles were forced to use the painted gore area to reach the right-turn slot.

The numbers of rear-end conflicts were reduced by 16 percent (Case 1) and 27 percent (Case 2), and the numbers of cross-traffic-from-right conflicts were increased in both cases.

TABLE 4 RESULTS OF CONFLICT STUDIES AT MARLTON CIRCLE

	CROSS TRAFFIC LEFT	CROSS TRAFFIC RIGHT	REAR END	LANE CHANGE	USE OF GORE AREA	VEHICLE PASSING ON AMBER	TOTAL VOLUME	CONFUSION ORIENTED	TRAFFIC ORIENTED
BEFORE I (07/25/85)									
10:30 - 12:00	13	17	17	9	24	12	1655	22	70
13:00 - 14:30	14	15	27	16	36	20	1661	30	98
TOTAL	27	32	44	25	60	32	3316	52	168
AFTER I (07/24/86)									
10:30 - 12:00	13	30	13	7	65	15	1843	20	123
13:00 - 14:30	20	64	29	13	70	36	1902	33	199
TOTAL	33	94	42	20	135	51	3745	53	322
BEFORE II (10/18/85)									
10:30 - 12:00	18	17	29	10	26	24	1889	28	96
13:00 - 14:30	17	34	34	23	38	34	1831	40	140
TOTAL	35	51	63	33	64	58	3720	68	236
AFTER II (10/16/86)									
10:30 - 12:00	25	36	29	6	70	24	2057	31	159
13:00 - 14:30	25	43	22	6	102	27	2053	31	194
TOTAL	50	79	51	12	172	51	4110	62	353

TABLE 5 RESULTS OF CONFLICT STUDIES AT LIVINGSTON CIRCLE

	REAR END	USE OF GORE AREA	USE OF WRONG LANE	LANE CHANGE	TOTAL VOLUME	CONFUSION ORIENTED	TRAFFIC ORIENTED
BEFORE I (01/08/86)							
10:30 - 12:00	199	20	6	29	1229	35	219
13:00 - 14:30	225	19	14	25	1829	39	244
TOTAL	424	39	20	54	3058	74	463
AFTER I (12/16/86)							
10:30 - 12:00	165	15	12	26	1336	38	180
13:00 - 14:30	227	13	6	24	2366	30	240
TOTAL	392	28	18	50	3702	68	420
BEFORE II (01/09/86)							
11:00 - 12:00	125	5	6	25	1415	31	130
13:00 - 14:30	231	16	9	22	2056	31	247
TOTAL	356	21	15	47	3471	62	377
AFTER II (01/14/87)							
11:00 - 12:00	74	6	2	11	1018	13	80
13:00 - 14:30	228	22	10	18	1917	28	250
TOTAL	302	28	12	29	2935	41	330

Livingston Circle

Table 5 shows the results of the conflict analysis at Livingston Circle. The results indicate 24.4 percent and 23 percent reductions for both cases, which were caused mainly by the reduction of 23 percent (Case 1) and 28 percent (Case 2) in the numbers of lane-changes conflicts (erratic maneuvers).

The results of the conflict analysis for the first case indicate a reduction of 23.5 percent in the number of rear-end conflicts and a reduction of 25 percent in the number of traffic-oriented conflicts. The results of Case 2 indicate no changes.

CONCLUSIONS AND RECOMMENDATIONS

The conclusions drawn from the results of the study of conflict analysis conducted before and after placing the diagrammatic signs are as follows:

- The numbers of confusion-oriented conflicts at regular and cut-through circles were reduced after installation of diagrammatic signs.
- The reduction in the numbers of confusion-oriented conflicts indicates that drivers were much more aware of the required route because of information obtained from the guidance signs. These signs, thus, helped drivers to make the right decisions in time.
- The numbers of lane-change conflicts at regular and cut-through circles were significantly reduced after installation of the signs.
- The numbers of traffic-oriented conflicts for all circles, except for Livingston Circle, were increased after installation of the signs. This result was expected because of the increase in traffic volume after installation of the signs, which caused more friction in the traffic flow.
- After installation of the signs, the numbers of rear-end conflicts increased for all the regular circles and reduced for the two cut-through circles.
- Traffic conflicts (brakelight indication, etc.) are not as useful measures of effectiveness at circles as confusion conflicts (maneuvers). This is because under heavy traffic conditions these types of conflicts are more the rule than the exception and appear not to be related to actual driver confusion.

The uniform results obtained at regular and cut-through circles indicated that the diagrammatic signs are effective in reducing confusion-oriented conflicts.

REFERENCES

1. M. Smith. *Improved Signing for Traffic Circles*. Draft Report. New Jersey Department of Transportation, Trenton, Dec. 1989.
2. A. Meerman, E. R. Hoffman, W. A. MacDonald. A Comparison of Advance Direction Signs for Use at Irregular Intersections in Terms of Drivers' Route Negotiation Behavior. *Proc., 11th Australian Road Research Board Conference*, 1982.
3. A. Meerman, E. R. Hoffman, W. A. MacDonald. Rapid Detection of Destination Direction on Stack and Diagrammatic Advance Direction Signs. *Proc., 11th Australian Road Research Board Conference*, 1982.
4. M. A. Kraft. *Diagrammatic Signs at Traffic Circles*. Research Report 33. Department of Highways and Traffic, Bureau of Traffic Engineering and Operations, Washington, D. C., 1973.
5. R. J. Salter and J. A. Alalawi. Vehicle Movements and Interactions at Roundabout Intersections. *Proc., Institution of Civil Engineers*, Vol. 73, No. PT2, Bradford University, West Yorkshire, England, Dec. 1982.
6. L. W. Ackroyd and A. J. Madden. Vehicle Speeds and Paths at Rural Diamond Roundabout Motoring Interchange. *Traffic Engineering and Control*, Vol. 16, No. 5, May 1975.
7. J. Harris and S. Perkins. *Traffic Conflict Characteristics: Accident Potential at Intersections*. New Jersey Department of Transportation, Trenton, 1967.

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

Innovative Evaluations of Traffic System Management Measures for Postearthquake Projects in Oakland, California

KEITH G. MEYER AND BALA M. RAJAPPAN

On October 17, 1989, the upper deck of the I-880 elevated freeway structure between 18th and 34th streets in Oakland, California, collapsed because of the Loma Prieta earthquake. The collapse of this Cypress Viaduct severed the major artery, disrupting local and regional transportation. An interim replacement for the Cypress Viaduct could not be put in place because of community opposition, so the California Department of Transportation (Caltrans) had to develop interim traffic system management (TSM) measures to relieve severely congested freeways. The innovative network-based methodology used to analyze the interim TSM alternatives proposed by Caltrans is described, and the alternatives' projected effectiveness in relieving congestion on freeways and arterials in Oakland is examined. The analysis used performance-based ranking to select the alternatives that would qualify for FHWA emergency relief funding. The examination of results from various combinations of alternatives is included, and the phenomenon called synergistic effect is introduced as a measure of effectiveness to rank the combinations. The need for simultaneous improvement of arterials, ramp metering, and freeways to achieve optimum delay reduction and freeway congestion relief is also discussed. The synergistic effects of the combined alternatives are compared, and the projects approved by FHWA are listed.

On October 17, 1989, the upper deck of the I-880 elevated freeway structure between 18th and 34th streets in Oakland, California, collapsed because of the Loma Prieta earthquake. I-880 is the principal north-south freeway along the east shore of the San Francisco Bay, connecting I-80 in Oakland to I-280 in San Jose. The collapse of the freeway structure, commonly called the Cypress Viaduct, severed this major artery, disrupting local and regional transportation. In addition, a section of the San Francisco–Oakland Bay Bridge (I-80) upper-level deck also failed, causing the temporary closure of the Bay Bridge. The California Department of Transportation (Caltrans) quickly restored the Bay Bridge, but Caltrans has yet to install an interim replacement of the collapsed Cypress Viaduct. A permanent replacement is under development but will not be finished for 4 or 5 years. An at-grade expressway was proposed in this corridor to alleviate congestion until a permanent replacement facility could be built. This proposal met strong local public opposition. The public's concern was that the expressway would become permanent, disrupting the community and interfering with local traffic circulation. This created a need to quickly develop interim traffic system management (TSM) measures to relieve severely congested freeways until the permanent replacement freeway is completed. Figure 1 shows the location of the subject freeways.

DeLeuw, Cather & Company, 100 Park Center Plaza, Suite 450, San Jose, Calif. 95113.

A network-based methology is described that was used to analyze interim TSM alternatives proposed by Caltrans and the projected effectiveness of the alternatives in relieving congestion on freeways and arterials in Oakland. The analysis ranked alternatives on the basis of their performance to select those that would qualify for FHWA emergency relief funds.

This paper also includes the examination of results from various combinations of alternatives and introduces the phenomenon called synergistic effect as a measure of effectiveness for ranking the combinations. The need for simultaneous improvement of arterials, ramp metering, and freeways to reduce delay and relieve freeway congestion is also discussed. The paper ends with the comparison of synergistic effects due to various combinations of arterial and freeway improvement projects, and it lists the projects that FHWA approved.

PRE-EARTHQUAKE CONDITIONS

Before the Loma Prieta earthquake, traffic flows were very congested on all routes leading to and from the Bay Bridge toll plaza in the peak periods. I-80 and I-580 regularly operated at capacity leading into the I-580 distribution structure and into the toll plaza. Eastbound I-80 traffic was heavily congested because of weaving constraints in the I-80/I-580 distribution structure. Delay during peak periods ranged from 15 min to 1 hr; it was highly variable. Figure 2 shows the estimated pre-earthquake average daily traffic (ADT) for 1988. The Cypress Viaduct was estimated to carry up to 165,000 vehicles, including 25,000 trucks, per day. An estimated 40 percent of northbound I-880 traffic traveled to and 20 percent of the southbound traffic traveled from the Bay Bridge during the morning peak hour.

Bay Area Rapid Transit (BART) transbay service parallels I-880 and the Bay Bridge. Before the earthquake, about 105,000 passengers used BART's transbay service every day. Arterial streets near the Cypress Viaduct, including West Grand Avenue and 14th Street leading into the industrial sections west of I-880, were used lightly. I-980, constructed in the early 1980s, was not heavily used as a peak-period bypass, although it provided convenient access to and from Route 24 and was the major route into and out of downtown Oakland. Truck volumes on I-880 were very high on the Cypress Viaduct, reaching 15 percent of daily traffic and 12 percent of peak-

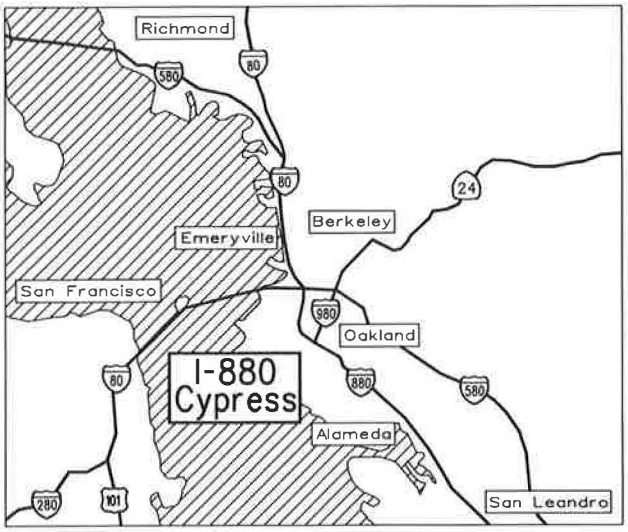


FIGURE 1 Project location.

hour traffic. This was because of the freeway's regional connections and the high concentration of major trucking centers along the I-880 corridor within Alameda County. In addition, the Cypress Viaduct was adjacent to the Port of Oakland, Army Supply Depot, U. S. Post Office Distribution Center, Naval Supply Center, and major distribution warehouses. Furthermore, the city of Oakland prohibited truck traffic on I-580 east of Grand Avenue, which forced more truck traffic onto I-880. The collapse of the Cypress Viaduct resulted in a severe deficiency in corridor capacity. As indicated in Figure 3, some of the vehicular demand was shifted to other modes and routes, but an estimated 90,000 vehicle equivalents per day was not accounted for in the transportation system.

POSTEARTHQUAKE CONDITIONS

Immediately after the earthquake, most streets and highways within the I-880/Cypress area were closed because of unstable structures, demolitions, or reconstruction activities. When the Bay Bridge reopened, only the I-580 and I-80 approaches to the bridge could carry its 240,000 daily vehicles. Because of the loss of the I-880 Cypress Viaduct capacity, Caltrans has pursued up to 15 alternative configurations for the interim replacement of the Cypress Viaduct. None of these was acceptable to the city of Oakland or the West Oakland neighborhood. Cypress Street is now only partially open to traffic, and its ramp connections to freeways are incomplete.

IMPACTS ON THE NETWORK

Corridor Travel

Vehicle queues now regularly cause much congestion within the I-580/I-980/Route 24 distribution structure, thus affecting all freeway routes. Vehicle queues also back up on northbound I-880 to the south for traffic using I-980 and on southbound Route 24 for traffic using I-580. Traffic congestion along I-80 east of the I-580 interchange has significantly worsened because of limited ramp capacities. Weekend travel along this route is equally congested.

Trucking Impacts

Many of the 25,000 trucks per day (15 percent of ADT) that had used the Cypress Viaduct now use I-980 and I-580, resulting in up to 12 percent truck traffic mix on these freeways during peak hours. Local morning and afternoon deliveries have been delayed during peak periods, requiring that they be rescheduled to different time periods. Off-peak truck travel now significantly affects the level of service of freeway connectors on I-880, I-980, I-580, and I-80.

Impacts On Oakland

The absence of an interim Cypress replacement has reduced accessibility and mobility within Oakland and throughout the region. Traffic has sought alternative routes. Because substantial replacement highway capacity has not been provided in the Cypress Street corridor, traffic has attempted to filter into the local street system, especially into residential neighborhoods. Figure 4 shows recent changes in traffic volume on local streets. Poor signal progression and low operating speeds have limited the amount of traffic diversion to neighborhood streets. Caltrans thus proposed interim TSM projects to provide additional arterial corridor capacity until a permanent replacement can be built.

INTRODUCTION TO METHODOLOGY

The objectives of the proposed interim TSM alternatives were specifically to reduce freeway congestion caused by the loss of I-880 on the corridors adjacent to the Bay Bridge and to improve mobility within the city of Oakland. These alternatives were classified as interim measures because they were designed to relieve congestion temporarily until the replacement facility can be completed. The permanent facility is planned to be open to traffic within 5 years, and it will eventually divert excess traffic off the freeways that are now congested.

Because the prime objective of the interim TSM measures was to reduce congestion on the freeways, the principal measure of effectiveness used to screen alternatives was reduction in vehicle hours traveled (VHT). Another measure used effectively was the cost-effectiveness of the alternatives, represented as the VHT reduced per $1,000 of implementation cost. As the analysis proceeded, a measure called synergistic effect was also introduced to compare various combinations of TSM alternatives. The measures of effectiveness for each alternative were estimated on the basis of output by a computer travel-forecasting model. The model produced raw output such as vehicle miles traveled (VMT) and VHT by facility types for the area under study. This output was then refined to produce effectiveness ratings.

The evaluation of TSM measures was unique for several reasons. First, the main goal was to reduce congestion spe-

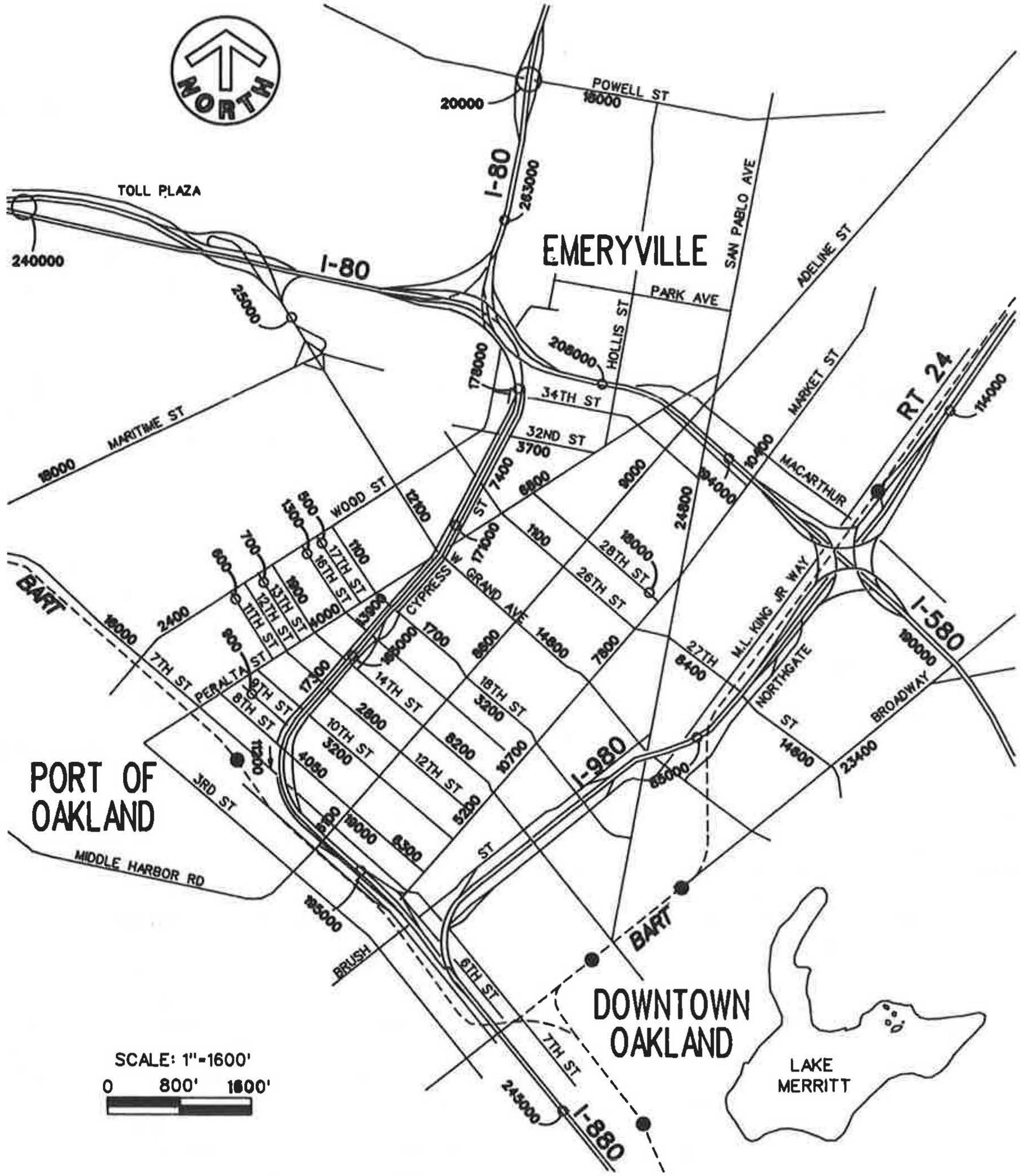


FIGURE 2 Estimate of pre-earthquake ADT.

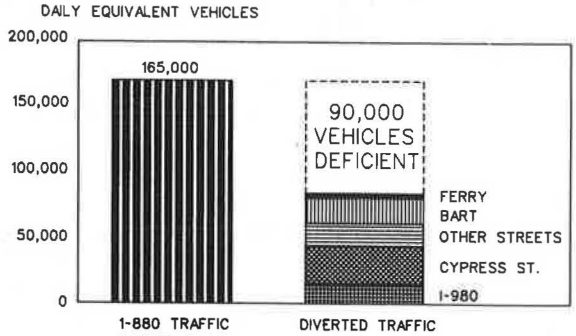


FIGURE 3 Deficient corridor capacity: where has I-880 traffic gone?

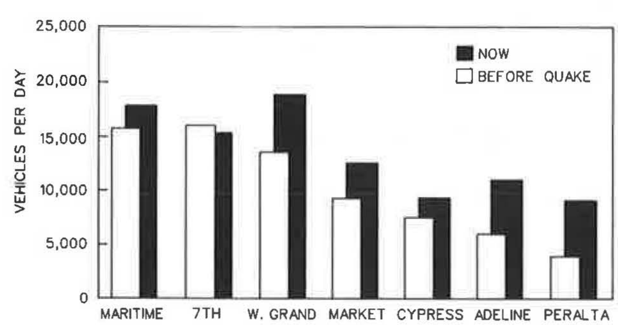


FIGURE 4 Postearthquake traffic volume changes on selected Oakland streets.

cifically on freeways rather than on arterials. Second, project alternatives needed to be operational by April 1991 in order to meet the early implementation requirements of FHWA. Third, low-cost and high-cost alternatives were considered together, low-cost alternatives being arterial improvements and high-cost alternatives being improvements to freeway connectors and mainlines. Fourth, measures of effectiveness included only VHT reduction, cost-effectiveness, and synergistic effects. Other measures commonly used in low-cost TSM studies, such as user costs, energy use, and parking, were not considered because they were not related to critical project issues, and a full-scale benefit-cost analysis was not prepared. Finally, time constraints required shortcut methods to evaluate projects quickly.

Transit measures were evaluated by other agencies. Many transit measures were implemented as emergency measures just after the earthquake to cope with the loss of the Cypress facility and the collapse of part of the Bay Bridge. Transit measures included introducing ferry service between the East Bay and San Francisco, increasing capacity on BART and Alameda–Contra Costa Transit, adding park-and-ride lots, and using carpool services. Some of the measures initially implemented have caused some permanent diversion from automobile trips to transit trips.

Caltrans proposed other options, such as traffic surveillance and motorist information systems, changeable message signs at weigh stations, and additional park-and-ride facilities. But because these options were not part of the freeway or arterial construction projects, they were excluded from this analysis. Caltrans, however, assessed these options independently and forwarded their recommendations to FHWA.

Travel Forecasting Methodology

At the time of the earthquake, a travel demand forecast model was being developed for Alameda County by Barton-Aschman Associates, Inc., of San Jose, California. Its principal purpose was to evaluate future transportation and land use alternatives as part of a countywide transportation plan. Fortunately, the highway network and corresponding traffic zone system were already developed at a level of detail sufficient for analyzing link-level traffic volumes on the freeways, expressways, major arterials, and selected minor arterials in the Oakland area. In order to simulate current travel patterns for the analysis of alternative interim TSM measures, the 1987 Metropolitan Transportation Commission (MTC) Bay Area person-trip tables (home-based work, home-based shop/other, home-based social/recreational, and nonhome based) were converted and applied to the Alameda County zone structure. County-to-county specific adjustment factors were applied to each of the person-trip tables to account for intercounty transit use and county-to-county specific vehicle occupancy rates. MTC's regional morning peak-hour factors for each trip purpose were applied and adjusted to create a morning peak-hour vehicle trip table.

At the time this study was conducted, the development of the Alameda County modeling procedure was not at a stage that allowed carpooling to be addressed. Alternatively, MTC's high-occupancy vehicle (HOV) travel demand estimates were used to account for carpool traffic. Existing HOV travel is reported to make up approximately 12 percent of the morning peak-hour vehicle trips on the Bay Bridge and much less elsewhere.

After the morning peak-hour trip table was assigned to the highway network, a validation procedure was used to measure the accuracy of the assignment results against observed travel patterns and volumes. Pre-earthquake ground-count data on freeway sections (Bay Bridge, I-80 at University, I-880 at 66th, I-980 at 18th, I-580 at 35th and Kellar and at the Caldecott Tunnel) and several arterial streets (Cypress Street, Peralta Avenue, and Grand Avenue) were compared with the model results. Minor modifications were made to the network and to transbay volumes so that the model represented vehicle trips on the network. The adjustments ultimately allowed the model to estimate pre-earthquake link volumes to within approximately 10 percent of the observed counts. An evening trip table was quickly developed by inverting the morning trip table and adjusting external stations, such as the Bay Bridge and the Caldecott Tunnel.

The postearthquake roadway conditions ("no-build" alternative) defined and coded in the network included some freeway improvements implemented by Caltrans soon after the earthquake. The existing ground-level Cypress Street between 8th and 34th streets (formerly under the Cypress Viaduct) that was partially reopened was also included in the network.

Model Representation of Interim TSM Measures

The TSM alternatives proposed for the interim included improving signal coordination on major arterials parallel to congested freeways, adding arterial lanes by removing onstreet parking and restriping, and improving some intersections on selected arterials. Alternatives on freeways included upgrading connectors at major interchanges and adding lanes in congested sections. Connector upgrades at some locations involved physically widening structures and at other locations, just restriping and reducing shoulder widths. Freeway upgrades included restriping and widening roadways.

To estimate the VHT reduction for each TSM measure, the network description of each alternative was coded into the travel forecasting model as closely as possible. The model then provided regional travel assignments on all roadway facilities represented in the network, from minor arterials to freeways. Some innovative techniques were developed to represent improved arterial signal progression and freeway on-ramp metering. Improvements to signal progression on arterials were represented in the model as increased arterial operating speed. Input operating speeds on all segments of arterial streets with improved signal progression were increased by 10 mph, the maximum improvement practicably attainable. This increase shortened running time in the model, which equated to the improved progression in the field. Widened, added, and restriped lanes were represented as increased capacity for given segments.

To test systemwide ramp metering, metered freeway ramps were coded with reduced capacity. A single-lane ramp with a typical capacity of 1,500 vehicles per hour was reduced 40 percent to 900 vehicles per hour (*1*). This resulted in significantly fewer vehicles' entering the freeway system and diverted traffic from the on-ramps to the arterial system. It was understood that optimum metering rates could not be developed by this method, but the order-of-magnitude effectiveness rating was reasonable.

Another useful output from the travel forecast model was the volume-difference plots of the network. Volume plots were prepared to show the differences in peak-hour volumes for each link of the network between a specific TSM alternative and the no-build scenario. The plots were prepared in color: green showed decreases in volumes and red showed increases. Each plot provided a networkwide comparison of individual TSM measures with the no-build scenario, so the effects of the proposed TSM measures could be identified by just looking at the plot. The new traffic circulation patterns caused by the improvements were also observed and analyzed from the plots. Analyses of traffic patterns revealed that almost every TSM alternative considerably affected the freeways.

Other output from the travel forecast model included VHT differences between alternatives. It appeared that the VHT values reasonably reflected the TSM measures and were sensitive to minor modifications. Furthermore, the model provided VHT by facility type and level of service. This allowed the examination of VHT reductions on freeways and major and minor arterials by level of service. All of these measures were obtained from the model for morning and evening peak hours.

INTERIM TSM ALTERNATIVES CONSIDERED

This section describes the alternatives analyzed in this project, which are also summarized in Table 1. Several of these were dropped from further consideration because of opposition from public groups.

1. *Improvements along Adeline, Market, West Grand, Castro, and Brush streets.* This project would replace signal controllers along West Grand Avenue and Adeline, Market, Brush, and Castro streets. It would allow time-based coordination of signals along routes that could serve as alternatives to congested freeways. Some striping and parking controls during

TABLE 1 SUMMARY DESCRIPTION OF INTERIM TSM MEASURES

ALT.	JURISDICTION	PROJECT DESCRIPTION	COST ESTIMATE	TIME TO IMPLEMENT
1	Caltrans/Cities of Oakland and Emeryville	Installation and coordination of signal controllers on local city streets (W. Grand, Adeline, Market, Brush, Castro	$170,000	4-6 months
2	Caltrans and City of Oakland	Replacement of exsisting signal controllers along W. Grand Ave. with new TCT type controllers	$40,000	2-4 months
3	Caltrans/Cities of Oakland,Emeryville and Berkeley	Upgrade San Pablo Ave with new signals, new controllers, left turn lanes and signals	$2.1 million	3-12 months
4	Caltrans and City of Emeryville	Upgrade existing signal equipment and install new signal on Hollis St.	$370,000	9 months
5	Caltran and City of Oakland	Restripe W Grand Ave. to provide a bus and vanpool lane only	$80,000	3-4 months
6	Caltrans/Cities of Emerville and Oakland	Extend West MacArthur Blvd., Shellmound and Ettie St. ramps	5 million	24-36 months
7	Caltrans	Widen eastbound mainline I-580 and modify I-580/980 and I-580/80 connectors	$10 million	24 months
8	Caltrans	Widen west bound I-980 between I-580 and I-880 by striping and restriping of I-980 to I-880 connectors	$300,000	16 to 18 months
9	Caltrans and City of Berkeley	Improvements to westbound I-80 Powell St. ramps plus auxiliary lane	$9.8 million	24 months
10	Caltrans	Widen EB on-ramp of Powell St. Interchange	$150,000	6 months
11	Caltrans	Ramp Metering along Ala-24/80/880/980	$5.25 million	9-18 months

peak periods were proposed for Market Street. The estimated costs for improvements along each route would be as follows: Adeline, $35,000; Market, $60,000; and Castro and Brush, $75,000. Implementation was estimated to take between 4 and 6 months.

2. *Signal improvements to West Grand and Northgate avenues.* This project would replace signal controllers along West Grand and Northgate avenues. It would allow time-based coordination of signals along routes that could serve as alternatives to congested freeways. This project was estimated to cost $40,000 and take 2 to 4 months for implementation.

3. *Signal Improvements along San Pablo Avenue.* This project would upgrade signal equipment, install left-turn phasing, and add left-turn bays on San Pablo Avenue between Ashby Street and the I-580 and I-980 freeways. The cost was estimated to be $2.1 million, and the time to implement was estimated to be 3 months to 1 year. Proposed improvements to San Pablo Avenue from Ashby Street to the I-580 and I-980 freeways were originally developed as part of the transportation management plans for reconstructing I-80 east of I-580.

4. *Signal improvements on Hollis Street.* This project would upgrade signal equipment and install new signals on Hollis Street between Ashby Avenue and Yerba Buena Street. The cost was estimated to be $370,000. Time to implement was estimated to be 9 months.

5. *HOV lanes on westbound West Grand Avenue.* This project would provide an HOV lane on West Grand Avenue from Campbell Street to the San Francisco–Oakland Bay Bridge toll plaza. Project cost was estimated to be $80,000. Implementation time was estimated to be 3 to 4 months. This HOV lane would replace one of the current mixed-flow lanes during peak periods. The right lane on the West Grand Avenue viaduct, beginning at Campbell Street, would be restricted to buses and vanpools.

6. *MacArthur Boulevard extension and Ettie Street ramps.* This project would extend Ettie Street underneath I-580, extend MacArthur Boulevard to join Ettie Street, and construct on- and off-ramps to provide local traffic access to and from the Bay Bridge. The cost was estimated to be $6 million, and the project could be implemented in 2 to 3 years. Currently, the MacArthur Boulevard on-ramp to westbound I-580 is closed because of operational problems on I-580. The MacArthur Boulevard off-ramp from eastbound I-580 was a left-hand off-ramp, so it was proposed to close this ramp as well as the 32nd Street on- and off-ramps to the Cypress corridor.

7. *Widen eastbound mainline I-580 and modify I-580/I-980 and I-580/I-80 connectors.* This project would widen eastbound mainline I-580 between the distribution structure and Hollis Street, widen the I-580 eastbound branch connector to westbound I-980 from two lanes to three by restriping and additional paving, and restripe the connector from westbound I-580 to westbound I-980 from two lanes to one. It would involve major construction work and would cost approximately $10 million. It was estimated that these projects could be implemented in 2 years. When implemented, they would provide increased capacity within the existing I-580 and I-980 corridor.

8. *Widen westbound I-980 between I-580 and I-80.* This project would widen the mainline westbound I-980 and the direct connector from westbound I-980 to southbound I-880 to provide three 12-ft lanes. The estimated cost would be $300,000 and the estimated implementation time, 16 to 18 months.

9. *Improvements to westbound I-80/Powell Street ramps.* Currently, a project to reconstruct the Powell Street interchange on I-80 is in the design phase. It involves the construction of two hook ramps and a new lane from the on-ramp to the distribution structure. The cost estimate was $9.8 million and the time to fully implement this project was estimated at 2 years. This project would provide operational improvements at the Powell Street interchange on the west side of the freeway only.

10. *Widen eastbound I-80/Powell Street on-ramp.* This project would widen the eastbound on-ramp to I-80 from Powell Street and provide a free right turn from westbound Powell Street onto the ramp. The cost was estimated to be $150,000 and the implementation time, 6 months.

11. *Ramp metering along Route 24 and Interstates 80, 580, 880, and 980.* Ramp metering was proposed as part of an overall traffic operations system aimed at improving the operation of the area freeway system. For this project, it was proposed that 35 on-ramps be metered on I-880 south of I-980/I-880, on I-80 east of the distribution structure, on Route 24 east of the Route 24/I-580 interchange, I-980, and I-580 in Oakland. About 35 of the metered ramps would be located within the core of the study area, and the rest distributed within a few miles on I-880, I-80, Route 24, and I-580. Ramp metering would require geometric modification at many ramps to provide storage for ramp queues, additional lanes to accommodate high volumes, HOV bypass lanes, and enforcement areas. Caltrans estimated the cost of metering each ramp at $150,000, including geometric modifications.

EVALUATION OF INDIVIDUAL ALTERNATIVES

The results of the alternatives analysis are summarized in Tables 2–4. Table 2 shows the VHT differences of TSM alternatives from no-build by facility type. Alternatives 1–11 are the individual TSM measures as proposed by Caltrans. Various combinations of individual measures are referred to as Combinations 1–7.

1. *Improvements along Adeline, Market, West Grand, Castro, and Brush streets.* Increased speed and capacity on Brush and Castro streets would provide easier access to West Grand Avenue for vehicles on I-980 and I-880 during peak-period congestion. Table 2 indicates that this alternative would reduce 865 VHT per day on freeways. VHT would increase on the arterials, however, because of increased congestion. This delay due to increased congestion on arterials would override the benefits gained on freeways and thus increase VHT by 1,260 on the network. This alternative would achieve the objective of reducing congestion on freeways, but it would have negative impacts on the arterials. Table 4 shows that this alternative would perform fairly well, in terms of cost-effectiveness, in comparison with others.

2. *Signal improvements to West Grand and Northgate avenues.* Merely enhancing signalization on West Grand Avenue would increase westbound traffic flow anywhere from 80 to 1,100 vehicles per hour during both peak periods. The largest increase, 1,100 vehicles per hour, would occur between I-980

TABLE 2 1990 VHT DIFFERENCES OF TSM ALTERNATIVES FROM NO-BUILD BY FACILITY TYPE

ALTERNATIVE DESCRIPTION	ALTERNATIVE	DAILY VHT DIFFERENCES		
		FREEWAYS	MAJOR	MINOR
W. Grand, Market, Adeline, Brush, Castro	1	- 865	1590	535
W. Grand & Northgate	2	- 1240	135	- 50
San Pablo Avenue	3	- 520	195	- 1560
Hollis Street	4	- 420	- 395	485
HOV Lane on WB Grand	5	- 445	- 300	- 25
MacArthur, Ettie Street Extension	6	- 435	- 75	- 5
Widen EB I-580	7	305	250	830
Widen WB I-980	8	95	- 295	50
Improvments to WB I-80 Powell	9	255	1110	980
Widen EB I-80 Powell	10	0	0	0
Ramp Metering	11	- 2260	3970	2700

NOTE: EB = eastbound, WB = westbound.

TABLE 3 SYNERGISTIC EFFECT OF COMBINATION ALTERNATIVES FOR INTERIM CYPRESS REPLACEMENT

COMBINATIONS	ALTERNATIVE	MAJOR ARTERIALS			MINOR ARTERIALS			FREEWAYS			ALL FACILITIES		
		VHT			VHT			VHT			VHT		
		COMB	SUM OF INDIV ALTS	SYN	COMB	SUM OF INDIV ALTS	SYN	COMB	SUM OF INDIV ALTS	SYN	COMB	SUM OF INDIV ALTS	SYN
Alts. 1, 2, 4, & 5	Comb 1	- 710	1030	- 1740	- 730	945	- 1675	- 1335	- 2970	1635	- 2775	- 995	- 1780
Alts. 1, 2, 3, 4, & 5	Comb 2	1365	1225	140	- 2780	- 615	- 2165	- 815	- 3490	2675	- 2230	- 2880	650
Alts. 1 & 2	Comb 3	705	1725	- 1020	310	485	- 175	- 1255	- 2105	850	- 240	105	- 345
Alts. 7, 8, & 9	Comb 4	- 1935	1065	- 3000	- 1095	1860	- 2955	1445	655	790	- 1585	3580	- 5165
Alts. 1, 2, 3, 4, 5, 6, & 11	Comb 5	2845	5120	- 2275	- 1625	2080	- 3705	- 4880	- 6185	1305	- 3660	1015	- 4675
Alts. 7, 8, 9, & 11	Comb 6	3640	5035	- 1395	425	4560	- 4135	- 1825	- 1605	- 220	2240	7990	- 5750
Alts. 1 to 11	Comb 7	2005	6185	- 4180	- 6460	3940	- 10400	- 3335	- 5530	2195	- 7790	4595	- 12385

SYN = Synergy Effect

TABLE 4 SUMMARY OF ALTERNATIVE RANKINGS AND FHWA APPROVAL

Alternative	Daily Systemwide Delay Reduction (VHT/Day)	Daily Freeway Delay Reduction (VHT/Day)	Daily Systemwide Synergy Effect (VHT/Day)	Cost Per Yearly VHT Reduced ($/VHT per Yr)	FHWA Approval
1	0	865	n/a	0.75	Yes
2	1155	1240	n/a	0.12	Yes
3	1885	520	n/a	4.29	Yes
4	330	420	n/a	3.39	Yes
5	770	445	n/a	0.39	Yes
6	515	435	n/a	45.45	No
7	0	0	n/a	n/a	No
8	150	0	n/a	7.69	Yes
9	0	0	n/a	n/a	No
10	0	0	n/a	n/a	Yes*
11	0	2260	n/a	8.93	Yes
Comb 1	2775	1335	1780	0.88	Yes
Comb 2	2230	815	0	1.74	Yes
Comb 3	240	1255	345	0.58	Yes
Comb 4	1585	0	5165	47.62	No
Comb 5	3660	4880	4675	9.71	Yes w/o Alt 6
Comb 6	0	1825	5750	52.63	No**
Comb 7	7790	3335	12385	59.17	No**

* This alternative was approved to alleviate local ramp congestion

** Some alternatives in these combinations were approved, but they were low cost compared to the high cost ones which were not approved and thus the high "cost per yearly VHT reduced."

and Cypress in both directions along West Grand Avenue. The West Grand Avenue segment between Cypress Street and I-80 would experience an increase of about 580 vehicles per hour in both directions. Even larger increases would occur during the evening peak period. The increase in traffic flow on West Grand Avenue and Northgate would be primarily due to diversion of traffic from I-580, I-980, and other adjacent minor arterials.

Overall, this alternative would reduce 1,155 VHT daily, of which 1,240 VHT would be from freeways. On the basis of this measure, this alternative ranked second among all TSM alternatives. However, it would rank first in cost-effectiveness, costing 12 cents per VHT reduced, as shown in Table 4. Although this would be the most cost-effective individual TSM alternative of those considered, it would cause a minor increase in congestion on major arterials.

3. *Signal improvements along San Pablo Avenue.* San Pablo Avenue signal improvements would cause a systemwide reduction of 1,885 daily VHT. This made it the best alternative in terms of systemwide VHT reduction, but it ranked fourth for freeway improvement. Increases of traffic on San Pablo, West Grand, Powell, Brush, and Castro would apparently be due to diversion of traffic from adjacent major and minor arterials and from freeways such as I-980, I-580, and I-80.

Unfortunately, when cost was taken into account, the cost per yearly VHT reduced would be $4.29, among the lowest as shown in Table 4. However, improvements on San Pablo Avenue would contribute to significant reductions in freeway traffic volumes across several freeways. As shown later, this alternative would greatly help to reduce congestion if it were implemented in combination with other TSM measures.

4. *Signal improvements on Hollis Street.* Improvements along Hollis Street would have some effect on arterial traffic paralleling I-880 without major changes in freeway ramp movements on I-580, I-880, and I-980. Hollis Street could be a truck route for port-related traffic. Any reduction of truck traffic on Interstates 880, 580, and 980 would increase operating capacity for passenger cars on these freeways.

This alternative would rank fourth in terms of systemwide cost-effectiveness, costing $3.39 per yearly VHT reduced. Improvements on Hollis Street thus would provide easier, less-congested access to the north of the distribution structure for trucks and other vehicles from Peralta Street instead of via Cypress Street and I-80 eastbound, which were already congested during morning and evening peak periods. The effectiveness of this alternative would be limited to areas adjacent to Peralta Street, Hollis Street, and some of San Pablo Avenue. Some reduction in traffic on freeways would be noted on I-980, Route 24, I-880, and I-580. This alternative would rank similarly to improvements to San Pablo Avenue in terms of systemwide cost-effectiveness but better in terms of freeway cost-effectiveness. Furthermore, it could be used effectively in combination with other TSM measures.

5. *HOV lanes on westbound West Grand Avenue.* This proposal to provide an HOV lane on westbound Grand Avenue from Campbell Street to the San Francisco–Oakland Bay Bridge toll plaza would reduce daily systemwide VHT by 770. Table 2 indicates that the majority of reduction would be on freeways. This systemwide and freeway cost-effectiveness of this alternative given in Table 4 would put it second behind improvements to West Grand and Northgate. The introduction of an HOV lane on West Grand Avenue would be most successful in combination with other TSM measures.

6. *MacArthur Boulevard extension and Ettie Street ramps.* Although this alternative would involve extensive construction, the VHT reduced systemwide would not be as substantial as Alternatives 2, 3, and 5. However, travel on freeways would be reduced. This alternative would also cause minor VHT reductions on major and minor arterials. Systemwide and freeway cost-effectiveness would be low because of the high cost of implementation. This alternative would rank sixth in cost-effectiveness and in VHT reduction on freeways.

7. *Widen eastbound mainline I-580 and modify I-580/I-980 and I-580/I-80 connectors.* This alternative would cause similar changes in travel patterns for morning and evening peak periods. This was also one of the alternatives that would not cause any reduction in VHT on freeways or on minor and major arterials. Furthermore, it would divert traffic from West Grand Avenue to I-580 and I-980 and thus contradict the objective of the TSM measures. This was not desirable, although the alternative could be used in combination with other TSM measures to produce more desirable results.

8. *Widen westbound I-980 between I-580 and I-80.* This alternative would have little effect by itself. The introduction of capacity to the connectors and to I-980 would allow more traffic into the freeway system, increasing congestion on bottleneck sections of I-980. Because of the low systemwide VHT reduction, the cost-effectiveness would not be significant. It would cost only $300,000 to implement this alternative, so it still may be combined with other TSM measures to produce better results.

9. *Improvements to westbound I-80/Powell Street ramps.* This proposal would construct two buttonhook ramps and a new lane from the Powell Street on-ramp to the distribution structure in the westbound direction of I-80. It would cause a significant increase in daily VHT, 10 percent of which would be due to an increase on freeways. Because the I-80 freeway segment between Powell Street and the distribution structure were already severely congested, the introduction of an extra lane and the reconstruction of hook ramps to provide longer weaving distance would only allow more traffic to get into the freeway system and be delayed at bottleneck points. Thus, the cost-effectiveness would be zero because of the high cost of implementation and no VHT reduction systemwide. This alternative would not meet the TSM objectives.

10. *Widen eastbound I-80/Powell Street on-ramp.* This proposal would widen the eastbound on-ramp to I-80 from Powell Street and provide a free right turn from westbound Powell Street onto the ramp. No changes in freeway volumes or daily VHT were predicted. It can be concluded that this improvement would be insignificant in affecting the network.

11. *Ramp metering along Route 24 and Interstates 80, 580, 880, and 980.* This alternative would achieve desirable results in terms of reducing freeway traffic volumes on I-980, I-580, and I-80. It would reduce daily VHT on freeways by almost 2,260, but it would increase VHT on surface streets by an equivalent amount. This was the alternative that would offer the highest reduction in VHT on freeways.

It appears that this alternative would also reduce traffic congestion on the distribution structure, especially during the evening peak and some during the morning peak, and thus reduce VHT on freeways and for the whole system. During the evening peak period, this alternative would have a more widespread effect on arterials such as San Pablo, Seventh, MacArthur, Northgate, and Broadway. All would experience an increase in traffic volumes that were diverted from freeways by ramp metering.

With the cost of each ramp meter estimated at $150,000, the total cost for this alternative was estimated to be $5.25 million. This measure would rank among the lowest in cost-effectiveness, but it would achieve the overall objective of reducing traffic on the freeway system and reducing VHT.

ANALYSIS OF COMBINATION ALTERNATIVES

In reviewing the volume-difference plots for each alternative, it was theorized that certain characteristics of each alternative could be combined to achieve better results. The interim TSM alternatives analyzed were thus categorized into three main sets. The first set included Alternatives 1–6, which primarily involved improvements to arterials and signal progression. The second set included Alternatives 7–10, which were mainly capacity increases to selected freeway segments and ramp connectors. Alternative 11, the freeway ramp metering alternative, was the only additional alternative in the third set.

Developing combined TSM alternatives involved consideration of freeway congestion relief in selected directions. The results of combining freeway measures with arterial improvements produced remarkable results. These combinations displayed a synergistic effect, in which the coordinated benefits from a combination of alternatives were greater than the sum of benefits from the individual alternatives—the whole was greater than the sum of its parts.

Synergistic Effect of Combination Alternatives

Combinations 1–7 were various combinations of Alternatives 1–11. Various combinations of street improvements were incorporated in Combinations 1–3. Combination 4 was a combination of freeway improvements. Ramp metering was combined with arterial projects in Combination 5 and with freeway projects in Combination 6. Combination 7 was the "chef's special": it included all street and freeway improvements, including ramp metering.

Table 3 compares the benefits of various combinations in terms of VHT reduction by facility type and includes benefits of each combination and the sum of benefits from individual measures in that combination. The increase in combined benefits compared with the sum of benefits gained by individual alternatives is the synergistic effect.

Synergistic effects reducing up to 12,835 VHT were reached in this evaluation, and in some cases no effect was shown on certain facility types (which indicates incompatibility). Overall, all the combinations would have synergistic effects except Combination 2. Combining street improvements and ramp metering would have systemwide synergy for reducing 4,675 VHT, four to five times more combined VHT reduced than the sum of the VHT reduced by individual alternatives.

Synergistic Effect of Arterial Combinations

All the TSM measures that included improvements to arterials would relieve congestion on freeways. Most of them would also reduce delay on minor arterials and throughout the system. Because of their low costs of implementation, these alternatives would perform better than freeway improvements in terms of cost-effectiveness. The performances of these individual alternatives would be, however, restricted by low-level improvements on selected arterials. These selected improvements would create some bottlenecks and increased congestion where projects ended. Thus, these alternatives would not produce optimum networkwide benefits, as had been anticipated. Under such circumstances, a need was identified to combine some alternatives to prevent bottlenecks and harness the favorable synergistic effects.

Combination 1, which included Alternatives 1, 2, 4, and 5, is a good example of the synergistic effect. Alternatives 1, 2, and 4 individually would cause increased delay on major and minor arterials combined, minimizing systemwide benefits. Alternative 5, however, would cause significant delay reduction on all facility types, so the net benefit produced by the combination of these alternatives would be a significant systemwide synergistic effect, as shown in Table 4.

Combination 2 includes Alternatives 1–5. This combination would significantly reduce congestion on minor arterials but the increased congestion on San Pablo Avenue would be more significant than the gains made by combining the alternatives. This is the only combination among the seven considered that would not produce a systemwide synergistic effect. One solution would be to include some geometric improvements at intersections in addition to signal coordination and to extend the limits of improvements at both ends of the project.

Analysis of the results from Combination 4 and a comparison with the results from Alternative 11 produced an important conclusion about ramp metering. Alternative 11 indicated that ramp metering would reduce congestion on freeways but cause significant diversion to major and minor arterials, resulting in no systemwide benefit. Because no improvements would be made to the arterial system in conjunction with the ramp metering in this combination, the level of service of the arterials would decrease because of increases in traffic.

A solution proposed to reduce arterial congestion included improvement of local arterials along with ramp metering. Combination 4 would accomplish this with considerable benefits to freeways, minor arterials, and all facilities. Major arterials, however, would require further improvements, such as extending the project limits and geometric modifications at intersections. The systemwide benefit of 4,675 VHT in delay reduction in Combination 5 would be the highest among the combination alternatives. This indicates that ramp metering should be accompanied by simultaneous improvements to the local arterial system to relieve systemwide congestion.

Synergistic Effect of Freeway Combinations

The second set of TSM alternatives analyzed were the freeway improvement alternatives, including Alternatives 7–10. Surprisingly, these improvements would increase congestion on all types of facilities in the network. The main reason for the poor showings is apparently that the increase in capacity on the freeways would cause some diversion of traffic from the arterial system to the freeway system. The improvements would also create capacity for through traffic that previously used alternative freeway routes. The freeway improvements would not relieve congestion on the arterial system.

The synergistic effect of Combination 4 (Alternatives 7–9), however, is encouraging. Combination 4 shows a synergistic effect of 5,165 VHT reduced on all facilities combined. The freeways do not show any synergistic effect, though, which means that the combination would not help to relieve congestion on freeways. It can be concluded at this stage that individual freeway improvements would help the arterial system but hurt the freeways. However, arterial benefits could be realized only by combining the various freeway improvements, not individually. It is then concluded that the arterial improvements would be more effective than freeway alternatives in relieving freeway congestion. As for a synergistic effect, a combination of freeway and arterial improvements would be worth exploring.

The combinations of freeway improvements with ramp metering is shown on Combination 5. This combination displays similar characteristics to Alternative 11, which is freeway ramp metering without arterial improvements. This combination would also produce results that are directly opposite of those that would be produced by Combination 3 (freeway improvements combination without ramp metering). In Combination 3, congestion would be relieved on the freeways but worsened on the arterial system. Synergistic effect on this combination would be negligible.

Synergistic Effect of Freeway and Arterial Combinations

Maximum systemwide benefit in congestion relief would be obtained in Combination 7, which includes all arterial and freeway improvements and ramp metering. This combination would have an additional systemwide synergistic effect of more than 12,300 VHT reduced.

FHWA REVIEW OF ALTERNATIVES

Caltrans submitted an interim traffic relief plan to FHWA to qualify for emergency relief funding. The plan included the alternatives in this paper and others proposed by transit agencies. The freeway and arterial alternatives in the package included information presented in this paper. FHWA responded to the package within weeks, and its responses to freeway and arterial improvements are summarized in this section.

FHWA reviewed each freeway and arterial improvement project for cost-effectiveness, as well as for the probability that each project would be fully operational by the target date. FHWA approved local city street improvements such as West Grand, Adeline, Market, Brush, Castro, Northgate, San Pablo, and Hollis. These streets were approved as Alternatives 1–4.

Alternative 5 was first conditionally approved but later dropped because of feasibility reasons. Alternative 6 was not approved as proposed because the implementation time would be too long and the project would conflict with using the 32nd Street on-ramp as an HOV entrance. However, a lesser project of extending MacArthur to Hollis was to be considered if it could be implemented within the established time limit.

Alternative 7 was partially approved. The ramp connector modifications were approved as an excellent example of a low-cost operational improvement. However, the mainline I-580 widening was conditionally approved: it must be split-funded using system and emergency relief funds. The use of emergency relief funds is considered necessary to expedite the design and construction of the project. The ramp connector from westbound I-580 to eastbound I-80 was not approved at this time because of the implementation time and staging requirements for building the distribution structure.

Alternative 8 was approved and considered another good example of a low-cost operational and capacity improvement. Alternative 9 was not approved because the interchange would not serve the I-880/Cypress traffic and the project would take too long. Alternative 10 was approved as long as signing is provided at Hollis Street and San Pablo Avenue to advise eastbound I-80 travelers of its availability.

FHWA supported Alternative 11. The ramp meters at critical ramps that can be implemented by the set deadline were approved. Because ramp metering is a new concept in this area, Caltrans must take appropriate marketing measures as part of the implementation process.

CONCLUSIONS

Several conclusions can be based on the results of this analysis. First, low-cost improvements on arterials paralleling the congested Oakland freeways would be more effective in relieving freeway congestion than selected low-cost improvements to the freeways themselves. This is because minor freeway improvements attract traffic to the freeways whereas arterial improvements divert traffic from the freeways. Second, metering freeway on-ramps without improving adjacent arterials would worsen congestion on the arterial system much more than it would lessen congestion on freeways. Third, the most effective TSM alternative to relieve congestion on freeways would include minor freeway improvements with arterial signal coordination, along with freeway ramp metering at bottleneck locations. The final recommendation is that the synergistic effect of project combinations be included as one of the measures of effectiveness to rank low-cost TSM alternatives. Table 4 summarizes the benefits, cost-effectiveness, and FHWA response for all the alternatives studied. FHWA clearly approved alternatives with low implementation costs and significant contribution reducing freeway delay. Synergistic effect was also high for combination of approved alternatives.

Although the methodology described in this paper was applied to projects resulting from the Loma Prieta earthquake in order to qualify for the FHWA emergency relief funds, the computer travel forecast model and measures of effectiveness employed are typical analysis tools used for most urban transportation projects. The process used in the methodology was independent of the cause that initiated the work. Furthermore, every effort was made for the methodology and measures of effectiveness to adhere to FHWA and Caltrans guidelines on project evaluation, and the results were accepted by both agencies. The results presented in this paper have tremendous applications for TSM studies nationwide that evaluate arterial improvements, freeway improvements, and ramp metering.

REFERENCE

1. *California Department of Transportation Highway Design Manual of Instructions*, 4th ed. Central Publication Distribution Unit, California Department of Transportation, Sacramento, Jan. 1987, pp. 100-1–100-2.

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

Fuel Savings Through Traffic Signal Hardware Improvements

SOO BEOM LEE AND ROBERT L. SMITH, JR.

The Wisconsin Fuel Efficient Transportation (FET) program was funded with $1.5 million in "oil overcharge" funds to reduce fuel consumption by implementing computer-optimized traffic signal timing plans. The FET program provided funds to the 24 participating communities to be used only for traffic signal hardware improvements. Data from 28 signal networks were used to develop regression models of the fuel savings generated by the hardware improvements. The results of various network-level benefit-cost measures are also presented. Finally, the potential for adoption of the FET program by other states is explored. Under the FET program, optimal signal timing plans were developed using the TRANSYT-7F microcomputer program. TRANSYT-7F provided estimates of fuel and travel time savings and stop reduction. Most of the communities required significant hardware improvements to achieve full interconnection with three-dial capability. On average the program resulted in fuel savings of 4,350 gal/year per intersection and total annual savings of $28,450/intersection. Considering only the hardware improvement costs, the overall benefit-cost ratio for a 10-year project life and a 10 percent discount rate was 44.0. The most significant independent variables for the regression models of fuel savings per signal were cost per signal, average volume, population, percentage difference in interconnection, and percentage actuated signals. The best multivariate model included cost per signal and population.

Microcomputer programs for optimizing traffic signal timing are now readily available in user-friendly formats. The latest generation of microprocessors enables the timing of large signal networks—of even 50 more signals—to be optimized in minutes rather than hours. Studies of signal timing projects in a dozen states have demonstrated the cost-effectiveness of signal timing improvements. Most of these studies, however, provide little more than overall estimates of effectiveness. More information is needed on the factors that are important in determining the benefits of signal timing improvements for individual networks.

In this study, data from 28 signal networks were used to develop regression models of the factors that best explain fuel savings from traffic signal timing optimization. The results of various benefit-cost measures are also presented. In addition, the potential for adoption by other states of the signal timing improvement program used in Wisconsin is explored.

PREVIOUS STUDIES

From 1973 to 1981 U. S. oil consumers were systematically overcharged by domestic oil companies in violation of the Emergency Petroleum Allocation Act of 1973. Funds from suit settlement against the oil companies have been placed in the U. S. Department of Energy's Petroleum Violation Escrow Account ("oil overcharge funds"). Because an estimated 60 percent of the oil overcharge was for automotive fuel, a logical means of providing restitution would be to fund signal timing programs that reduce automotive fuel use. At least 12 states have initiated traffic signal timing programs in recent years, funded in most cases with oil overcharge money. Arnold (*1*) has categorized 10 of these programs into four groups: (a) training and technical assistance by lead agency with local responsibility for signal timing, (b) grant program for local governments, (c) contracts with consultants, and (d) state transportation agency responsibility. Information on the programs for all 12 states is summarized in Table 1.

Program-level estimates of benefits are available for five states. Fuel savings per intersection ranged from 930 gal/year in Iowa to 12,400 gal/year in North Carolina. Annual benefit-cost ratios ranged from 7 to 143 for the same two states. Estimates of benefits and benefit-cost ratios vary widely—in part because of different assumptions about the value of travel time savings, but more importantly because of the types of expenditures allowed and the different mixes of project types and city sizes. Only signal retiming, not hardware improvements, was included in the North Carolina program. In contrast, hardware improvements were incorporated in all of the signal retiming projects in Iowa. The lower overall initial benefits from the Iowa program hardware improvements should be partly offset by the much longer duration of the benefits compared with only signal timing improvements.

The potential for wide variations in benefit-cost ratios is illustrated by the Iowa program. One of the arterial system retimings in Des Moines produced a benefit-cost ratio of 300 whereas the second arterial system had negligible benefits (*2*). Overall, for the 15 projects in which significant benefits were found, benefit-cost ratios ranged from 1.75 to 55.6.

Most of the 19 signal timing and hardware upgrade projects in the Iowa program can be classified into three categories of hardware improvements: (a) upgrade pretimed controls at isolated intersections to fully actuated, (b) interconnect arterial controllers with time-based coordinators (TBCs), and (c) upgrade arterial controllers with full closed-loop interconnected controllers. Benefit-cost ratios for the Iowa projects were tabulated by city population category (very small, small, medium, and large) for each type of hardware improvement. The fully actuated and TBC projects were concentrated in the smaller categories, so no estimates of the effects of city size on benefit-cost ratios could be made. For the closed-loop projects, however, the benefit-cost ratios tended to increase with city size.

Department of Civil and Environmental Engineering, University of Wisconsin, 2204 Engineering Building, Madison, Wis. 53706.

TABLE 1 STATE TRAFFIC SIGNAL TIMING PROGRAMS

State	Training & Technical Assistance Only	Grant Program for Local Governments	Contracts with Consultants	State Transportation Agency
California-FETSIM		X		
Florida-GASCAP -STSRP	X		 X	
Illinois-SCAT			X	
Iowa		X	X	
Maryland-STSSP				X
Michigan-TSOP -TSMP			X X	
Missouri-TRANSYT-7F	X			
New York-STOP		X		
North Carolina-TSMP				X
Pennsylvania		X		
Virginia			X	
Wisconsin-FET		X		

FET PROGRAM GOALS

This research is based on data from the Wisconsin Fuel Efficient Transportation (FET) program. The FET program was initiated by the Wisconsin legislature in April 1987 with a $1.5 million grant of oil overcharge funds. The overall goal of the program was to reduce fuel consumption by implementing computer-optimized signal timing plans. The program was modeled after the highly successful FETSIM program in California (*3*). The California program provided training and technical assistance to local community staff in the use of the signal timing optimization program, TRANSYT-7F. Participating communities were reimbursed, at a rate of $1,000/intersection, for the staff time required to collect field data, apply TRANSYT-7F, and implement new timing plans. A fully interconnected traffic signal network was required for participation in the FETSIM program. No funds were available for any hardware improvements.

In contrast with the FETSIM program, the Wisconsin FET program focused on funding hardware improvements. Communities participating in the FET program received a grant of $1,000/intersection that could be used only for traffic signal equipment and installation expenses. A microcomputer to run the TRANSYT-7F program could also be purchased under the program. Additional funds for more extensive hardware improvements were to be allocated on the basis of fuel savings-to-hardware cost ratios.

Three secondary goals of the FET program were (a) to train local staff and consultants to use TRANSYT-7F, (b) to provide a wide distribution of the available hardware funds among the participating communities, and (c) to maximize the fuel saving effectiveness of the hardware improvements. The only constraint placed on participation in the program was that the traffic signals be reasonably interconnected. Communities were also encouraged to keep the data collection work load within the capability of their own staffs unless local funds were used to hire a consultant.

FET PROGRAM IMPLEMENTATION

Invitations to all Wisconsin communities with traffic signals to participate in the FET program resulted in contracts with 24 communities covering 518 intersections. The participants included all but two of the largest communities in the state: one of the two had recently upgraded and retimed its signal system; the other declined to participate because of staff limitations.

All of the communities were required to use TRANSYT-7F to develop optimal signal timing plans and to evaluate alternative hardware improvements for three time periods: morning peak, midday, and evening peak. Before-and-after travel time field studies were also required. TRANSYT-7F was used because of the need to model signal networks as well as single arterial systems. The computer program PASSER could be used to determine optimum phasing on individual arterial systems, but for consistency in developing performance measures, TRANSYT-7F analysis was also required.

The initial grants of $1,000/intersection left about $725,000 in unallocated hardware funds. These funds were set aside in a separate supplemental hardware grant program based on a comparison of hardware cost and the effectiveness of the hardware improvements in saving fuel. To estimate the hardware-generated fuel savings, three TRANSYT-7F runs were required for each time period (see Figure 1). The simulation run was used to calibrate the model so that it accurately reproduced existing conditions. Next, the calibrated model was optimized given the limitations of the existing hardware. Finally, the signal network was optimized on the basis of the capabilities of the new hardware alternatives.

If a FET community's existing signal system was fully interconnected and capable of three-dial, multiphase operation, little or no additional fuel savings would result from new hardware. The only additional hardware required might be the signal heads and rewiring needed for additional turn phases. Most of the FET communities, however, did not have fully

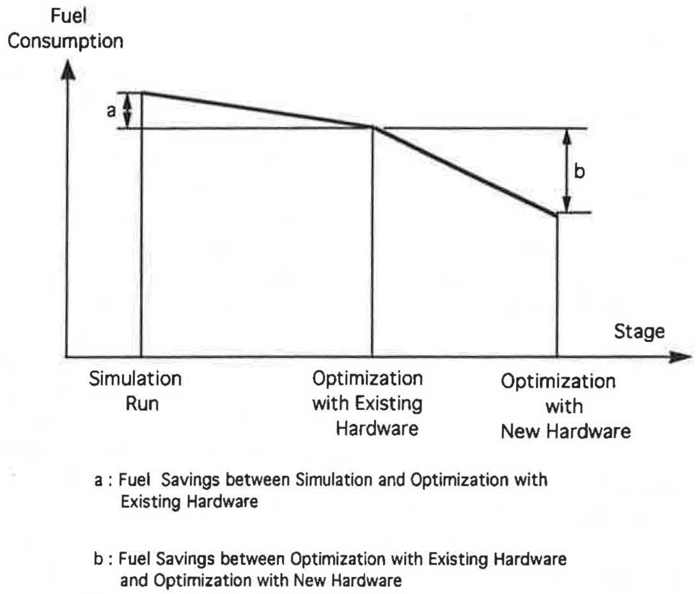


FIGURE 1 Fuel savings from optimizing traffic signal timing.

interconnected signal systems, and many did not have full three-dial, multiphase capability. Thus, the communities proposed a wide range of hardware improvements. Some communities were satisfied with adding TBC rather than more costly hardware to achieve interconnection. A few communities wanted the flexibility and other operating benefits of the even more expensive closed-loop systems. Nearly all communities needing to upgrade to three-dial capability chose to replace electromechanical controllers with solid-state controllers. The communities were not required to select the hardware alternative that would provide the highest fuel saving–to–hardware cost ratio, but they did run the risk of not being competitive with other communities in the allocation of the supplemental hardware grant funds if their fuel-effectiveness ratio was too low.

Application of TRANSYF-7F to simulate and optimize existing conditions was complicated by the lack of full interconnection for most of the networks. Noninterconnected signal systems can be modeled with TRANSYT-7F by "delinking" the signals; that is, the input flows to the intersections are assumed to be uniform—to arrive randomly rather than in platoons. This should be a reasonable assumption unless offsets among the signals are maintained continuously by signal technicians. Traffic-actuated signals can also be modeled by delinking and using average phase lengths.

Preliminary results for the TRANSYT-7F estimated benefits of the hardware improvements are available for 404 of the 518 signals. The average benefits per intersection include (a) fuel savings of 4,350 gal/year, (b) savings from fewer stops of $2,800/year, and (c) travel time savings of $21,300/year (based on value of travel time of $6.00/hr). Using fuel costs of $1.00/gal, the total annual benefits equal $28,450/intersection compared with hardware improvement costs of about $4,000/intersection. Thus, the benefit-cost ratio considering only 1 year's benefits is 7.1. For a 10-year project life and a 10 percent discount rate, the total benefit-cost ratio is 44.0.

The annual fuel savings generated by the FET program intersections are near the middle of the 930- to 12,400-gal range cited earlier (Iowa and North Carolina programs, respectively). The benefit-cost ratios are highly dependent on the costs required to generate the benefits. The North Carolina benefit-cost ratio is very high in part because no hardware improvements were made and relatively small amounts of staff time were used.

Overall, the FET program succeeded in meeting its overall goal of reducing fuel consumption, and the program effectiveness measures compare favorably with measures for similar programs in other states. The secondary goals of the program were also met, although much higher levels of fuel savings per dollar of hardware expenditures could have been achieved if only TBC and not more costly hardwire and closed-loop systems had been funded.

ESTIMATION OF FUEL SAVINGS

Fuel savings effectiveness ratios (fuel saved per dollar of program expenditure) can be increased by limiting expenditures, but for hardware-based programs, selection of the least-cost hardware, such as TBC, may not be possible because local communities have other objectives to consider. For example, TBC will not minimize staff costs for system operation and maintenance. At the state program level, selection of the most fuel effective projects might better focus on the characteristics of traffic signal networks that are likely to affect fuel savings. Data from 28 networks in the FET program are available for developing models to estimate fuel savings from unconstrained hardware improvements.

Fuel savings are the difference between the fuel consumption for the TRANSYT-7F optimized timing plan with existing hardware and the TRANSYT-7F optimized timing plan with the hardware improvement. TRANSYT-7F estimates of fuel consumption are used in both cases. Possible dependent variables for fuel savings models follow:

- FUEL/SIG—Fuel savings per signal (gallons per day),
- FUEL/TVM—Fuel savings per thousand vehicle miles (gallons per 1,000 vehicle miles), and
- FUEL-DAY—Fuel savings per day (gal).

Of the three possible dependent variables, fuel savings per signal was selected because it is not biased by network size and is generally more highly correlated with the available independent variables (Figure 2).

Possible independent variables are

- GRID-ART—Binary code for grid (=1) versus arterial (=2) network;
- NET-SIGN—Number of traffic signals in the network;
- POPULATN—Population of the community (thousands);
- SPACING—Average distance between intersections (ft);
- %-ONEWAY—Percentage of the link miles that are one-way;
- B/C-RATI—Fuel savings-to-hardware cost ratio;
- AVVOL-TO—Average approach volume summed over three peak hours (morning, midday, and evening);
- PER-DIF—Percentage difference in the number of interconnected signals (IS) after versus before the hardware improvement, $[(IS_a - IS_b)/IS_a] \times 100$ percent;
- PER-ACT—Percentage actuated signals before the hardware improvement; and
- COST/SIG—Average cost per signal for the hardware improvement.

On the basis of the correlations shown in Figure 2, four independent variables should initially be considered for explaining fuel savings per signal: (a) cost per signal (COST/SIG), (b) average volume (AVVOL-TO), (c) percentage actuated (PER-ACT), and (d) percentage difference in interconnection (PER-DIF).

All four possible independent variables have a logical relationship with fuel savings. Higher levels of expenditure per signal should be associated with larger urban areas, which tend to have higher traffic volumes and more complex phasing arrangements. These higher demands on the signal system should provide the potential for greater fuel savings. As shown in Figure 3, fuel savings per signal does tend to increase with cost per signal. Although there is substantial variance in the relationship, the variance is relatively constant. And four outliers exist that have much higher fuel savings than indicated by the general relationship.

One of the outliers shown in Figure 3 is a miscode. For Neenah, the fuel savings should be zero because the basic network was fully interconnected. No additional hardware was required to implement the optimal timing plan that generates fuel savings of 30 gal/day per signal compared with the initial timing plan. Thus, the incremental fuel savings that are the result of new hardware are zero. In contrast, the La Crosse–Western and New Berlin networks both required relatively low cost hardware improvements on moderate-volume arterials but generated large fuel savings. The La Crosse–Western network was nearly fully interconnected, but it required upgrading from two to three dials. The New Berlin network initially had only traffic-actuated signals. Interconnection was achieved in the hardware upgrade with low-cost TBC. The Madison–East Washington network had the second highest traffic volumes of all the networks. The results show that even very high cost, closed-loop systems can be extremely cost-effective.

Average traffic volume has the next highest correlation with fuel savings per intersection. As shown in Figure 4, there is

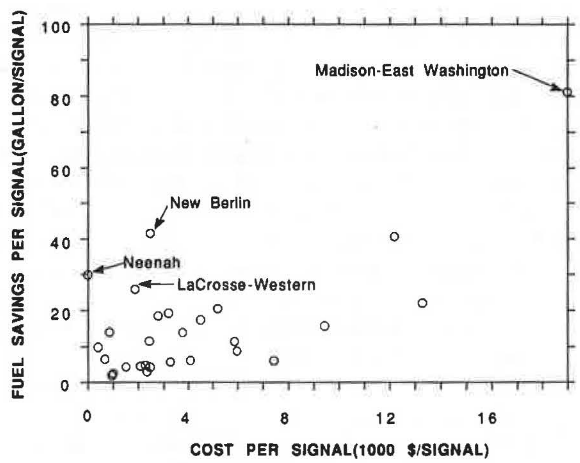


FIGURE 3 Fuel savings per signalized intersection versus hardware cost per signalized intersection.

	FUEL/STG	FUEL/TVM	FUEL-DAY	GRID-ART	NET-SIGN	POPULATN	SPACING
FUEL/TVM	0.710						
FUEL-DAY	0.710	0.728					
GRID-ART	0.263	0.086	0.029				
NET-SIGN	-0.135	-0.023	0.405	-0.505			
POPULATN	0.026	-0.174	0.035	0.272	-0.099		
SPACING	-0.074	-0.384	-0.280	0.224	-0.254	0.382	
%-ONEWAY	0.123	0.384	0.448	0.082	0.344	-0.118	-0.346
B/C-RATI	0.204	0.220	0.075	0.036	0.027	-0.188	-0.022
AVVOL-TO	0.656	0.468	0.761	0.399	0.197	0.198	-0.180
PER-DIF	0.437	0.250	0.266	0.359	-0.237	0.049	0.248
PER-ACT	0.590	0.423	0.375	0.476	-0.282	-0.082	-0.046
COST/SIG	0.671	0.440	0.550	0.398	-0.118	0.290	-0.007

	%-ONEWAY	B/C-RATI	AVVOL-TO	PER-DIF	PER-ACT
B/C-RATI	0.159				
AVVOL-TO	0.388	0.234			
PER-DIF	0.011	-0.211	0.349		
PER-ACT	0.088	-0.066	0.467	0.769	
COST/SIG	0.025	-0.326	0.594	0.517	0.590

FIGURE 2 Correlation matrix for dependent and independent variables.

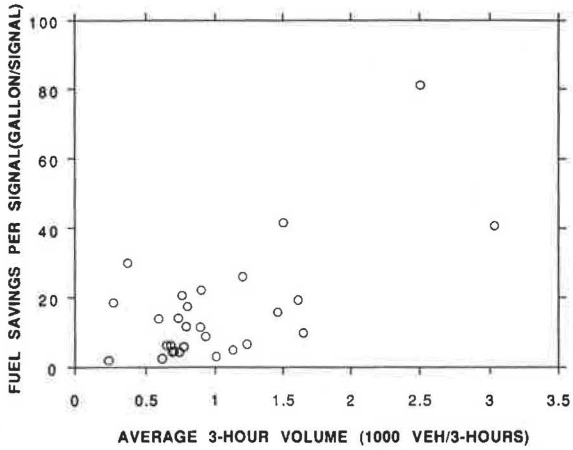


FIGURE 4 Fuel savings per signalized intersection versus average traffic volume per three peak hours.

a generally linear relationship between the two variables, but the variance increases with increasing volume. Thus, a regression model for the relationship will tend to overestimate the goodness of fit. Also, there are few data points for networks with high traffic volumes.

The last two possible independent variables, percentage actuated and percentage difference in interconnection, are highly correlated (r = .769). Consequently, only percentage difference will be examined in detail. After the hardware improvements were implemented all but two of the networks were fully interconnected. Each of those two networks included one isolated intersection that functioned better as a separate node. An increase in the percentage difference indicates a greater change in the extent of interconnection as a result of the hardware improvement. In general, a greater degree of interconnection change should result in a greater potential for fuel savings. As shown in Figure 5, there is a general but small trend toward higher fuel savings as the percentage difference in interconnection increases. The trend is more evident if the three outliers are removed. Three of

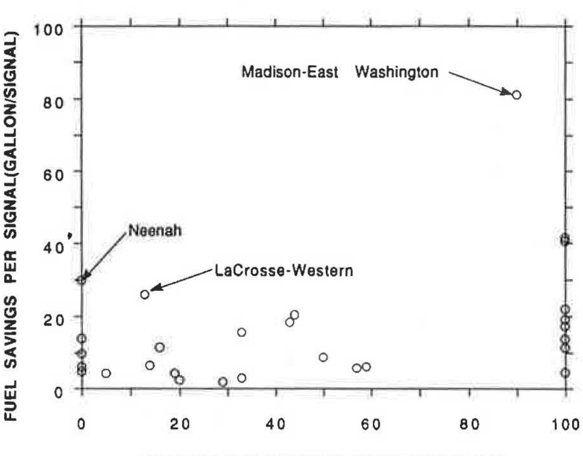


FIGURE 5 Fuel savings per signalized intersection versus percentage difference in interconnection.

the four networks found to be outliers for the cost per signal relationship are also outliers here (Neenah, La Crosse–Western, and Madison–East Washington). Only the Neenah outlier was deleted in subsequent analysis. Without the data points for Neenah and Milwaukee County (see later discussion), the correlation between fuel savings and percentage difference increases modestly, from .437 to .506. As with the average traffic volume relationship, the variance in fuel savings increases moderately with increasing percentage difference.

Population should have a reasonably high correlation with fuel savings per signal, but the correlation shown in Figure 2 is only .026. Population is also essentially uncorrelated with average traffic volume (r = .198). The reason for the lack of correlation is shown clearly in Figure 6. The three Milwaukee County networks (population 608,000) all have traffic volumes generally found in cities with populations of less than 100,000. In fact, the Milwaukee County networks are located in suburban communities with populations less than 50,000. Clearly, county-level population cannot be compared with city-level population. Consequently, the Milwaukee County networks are deleted from subsequent analysis when population is included as an independent variable.

REGRESSION MODELS OF FUEL SAVINGS

Single-variable regression models for fuel savings per intersection are presented as follows; the Neenah network is deleted from each regression model.

$$\text{FUEL/SIG} = \underset{(t\,=\,0.97)}{3.17} + \underset{(t\,=\,5.27)}{2.76}\ (\text{COST/SIG})$$
$$R^2_{\text{adj}} = 50.7\%,\ n = 27$$

$$\text{FUEL/SIG} = \underset{(t\,=\,-1.48)}{-6.90} + \underset{(t\,=\,5.50)}{0.0209}\ (\text{AVVOL/TO})$$
$$R^2_{\text{adj}} = 53.9\%,\ n = 26$$

$$\text{FUEL/SIG} = \underset{(t\,=\,-1.16)}{5.34} + \underset{(t\,=\,2.84)}{0.207}\ (\text{PER/DIF})$$
$$R^2_{\text{adj}} = 21.4\%,\ n = 26$$

The relatively low explanatory power of the models is consistent with the large variation in the basic data. The variability can be explained in part by the wide range of initial traffic signal hardware capabilities ranging from one-dial to multidial and two-phase to multiphase and the varying degrees of interconnection. Clearly, fuel savings generated by such a wide range of improvements are likely to be highly variable. Nevertheless, reasonably consistent relationships are found between fuel savings and the independent variables cost per signal, average traffic volume, and percentage difference in interconnection.

When the data for the three Milwaukee County networks are deleted, the correlations of the key independent variables with fuel savings per signal increase somewhat, as shown in Figure 7. And, as expected, population is now highly correlated with fuel savings per signal. Population now provides a

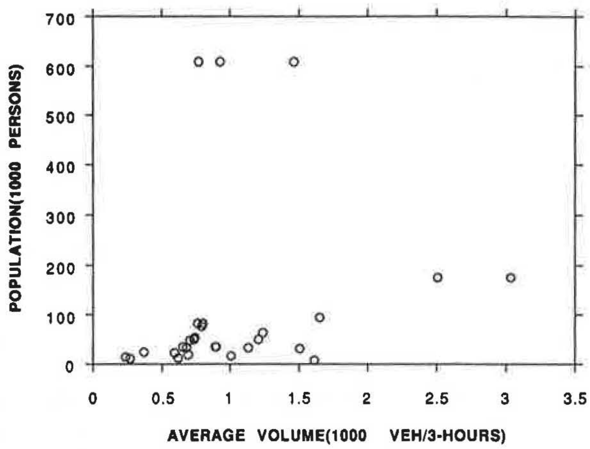


FIGURE 6 Population versus total traffic volume per three peak hours.

	FUEL/SIG	POPULATN	AVVOL-TO	PER DIF
POPULATN	0.694			
AVVOL-TO	0.718	0.771		
PER-DIF	0.506	0.229	0.336	
COST/SIG	0.756	0.627	0.569	0.533

FIGURE 7 Correlation matrix of variables for multiple-variable regression models.

reasonable single-variable model and the best multivariate model includes population and cost per signal, as shown in the following:

$$\text{FUEL/SIG} = \underset{(t\,=\,0.49)}{2.03} + \underset{(t\,=\,4.52)}{0.272}\ (\text{POPULATN})$$

$$R^2_{\text{adj}} = 45.8\%,\ n = 24$$

$$\text{FUEL/SIG} = \underset{(t\,=\,0.02)}{0.05} + \underset{(t\,=\,2.19)}{0.141}\ (\text{POPULATN}) + \underset{(t\,=\,3.20)}{2.05}\ (\text{COST/SIG})$$

$$R^2_{\text{adj}} = 61.8\%,\ n = 24$$

Because population and cost per signal are highly correlated ($r = .627$), the values of the regression coefficients are highly interdependent. The relative importance of population could have changed easily as the result of minor changes in population or cost per signal values.

The multivariate model provides a reasonable basis for estimating fuel savings per signal that can be generated by traffic signal hardware improvements. Higher population values reflect more traffic and congestion, which lead directly to greater fuel savings when signal timing is optimized. The positive coefficient for cost per signal reflects the need for more complex and sophisticated signal hardware in larger communities. The equation models the choices made by traffic engineers when reducing fuel consumption was not necessarily the most important objective. In many cases, similar levels of fuel savings could have been achieved with less costly hardware improvements, such as TBC.

An attempt was also made to identify one or more stratifying variables that would provide additional explanation for the variation in fuel savings per signal. Average intersection spacing appeared to provide the best stratification with a breakpoint of 1,800 ft. Only the constant term was statistically significant for the long average spacing regression model. For the short average spacing model, only one independent variable, cost per signal, was significant with an adjusted R^2 of 74 percent.

BENEFIT-COST RESULTS

The results of the benefit-cost analysis for the 28 Wisconsin networks are presented in Table 2. The total benefits include savings in travel time ($6.00), reduced stops ($0.01/stop), and fuel ($1.00/gal). The benefit-cost ratio considering only fuel savings for 1 year ranges from 0.25 to 5.0. Thus, even the network with the lowest level of fuel savings has a payback period of only 4 years. When the value of all savings is included, the benefit-cost ratios increase substantially. All of the network improvements are now cost-effective with a 1-year benefit-cost range of 1.78 to 40.0.

The hardware improvements will have a useful life of at least 10 years. The long-term benefits of the fuel savings alone during that time using a 10 percent discount rate are substantial: all the network hardware improvements can be justified on fuel savings alone. The lowest network benefit-cost ratio is 1.51. For the overall program, the long-term benefit-cost ratio for fuel savings alone is 6.73; considering all long-term benefits, the benefit-cost ratio is 44.0.

SUMMARY AND CONCLUSIONS

The $1.5 million Wisconsin FET program clearly was able to meet its primary goal of reducing fuel consumption through implementation of computer-optimized signal timing plans. The annual fuel savings of 4,350 gal/year per intersection compare favorably with the results of signal timing programs in other states (savings from 930 to 12,400 gal/year for Iowa and North Carolina, respectively). The FET program was also highly cost-effective: the overall benefit-cost ratio was 1.1 considering only fuel savings for the first year. When all first-year savings are included (adding travel time and stop reduction benefits), the benefit-cost ratio jumps to 7.2. Over a 10-year period with a 10 percent discount rate, the total benefit-cost ratio is 44.0.

The FET program focused on improving traffic signal timing through traffic signal hardware improvements. Traffic signal timing plans were optimized and the additional fuel savings attributable to the hardware improvements were modeling by using the TRANSYT-7F microcomputer program. Data for 27 traffic signal networks were used to develop regression models of the fuel savings. Fuel savings per signal in the network was selected as the independent variable because it is independent of network size. Initially, cost per signal and average volume were found to produce the best single-variable models. When illogical population data for Milwaukee County were deleted, population also produced a good single-variable model and the best multivariate model together with cost per signal.

TABLE 2 NETWORK-LEVEL BENEFITS AND COSTS

NETWORK NAME	TRAVEL TIME SAVINGS PER YEAR ($/YEAR)	STOP SAVINGS PER YEAR ($/YEAR)	YEAR-FUEL SAVINGS (GAL/ YEAR)	TOTAL COST ($)	ONE YEAR B/C[a] RATIO (FUEL)	ONE YEAR B/C[b] RATIO (ALL)	TEN YEAR B/C[c] RATIO (FUEL)
APPLETON	574,200	50,049	96,900	33,970	2.85	21.23	17.53
BELOIT- HENRY	64,836	9,108	13,647	23,500	0.58	3.73	3.57
BELOIT-PRAIRIE	351,882	32,330	39,618	79,750	0.50	5.31	3.05
CUDAHY	86,400	53,682	38,700	19,500	1.98	9.17	12.19
DePEER	63,000	22,350	7,950	21,145	0.38	4.41	2.31
KENOSHA	288,000	-13,530	110,400	78,000	1.42	4.93	8.70
MADISON-EW	1,396,764	82,612	243,357	190,000	1.28	9.07	7.87
MADISON-JSN	1,460,466	371,550	390,093	390,000	1.00	5.70	6.15
MIL CTY-GH	230,400	-948	28,200	56,600	0.50	4.55	3.06
MIL CTY-P WASH	50,400	13,728	12,000	22,940	0.52	3.32	3.21
MIL CTY-76TH ST	111,600	27,558	36,600	83,540	0.44	2.10	2.69
RACINE-DURAND	135,000	10,794	31,200	27,000	1.16	6.56	7.10
RACINE-16 NODES	462,600	62,694	98,100	83,000	1.18	7.51	7.26
SHEBOYGAN	10,800	60,399	35,100	54,500	0.64	1.95	3.96
WAUKESHA	270,000	-8,208	20,400	24,400	0.84	11.57	5.14
WAUSAU	181,134	21,039	34,353	55,000	0.62	4.30	3.84
WEST BEND	330,300	19,347	45,600	41,466	1.10	9.53	6.76
WISCONSIN RAPID	82,080	21,465	25,377	49,400	0.51	2.61	3.16
BEAVER DAM	34,200	-804	3,900	6,726	0.58	5.55	3.56
BELOIT-CBD	214,470	36,036	40,485	90,000	0.45	3.23	2.76
GREEN BAY	153,000	32,127	43,800	6,200	7.06	36.92	43.41
MANITOWOC	135,900	24,015	25,500	104,000	0.25	1.78	1.51
MARINETTE	10,800	4,014	3,600	5,000	0.72	3.68	4.42
NEENAH	428,400	75,684	107,400	0	****	****	****
NEW BERLIN	469,800	49,203	74,700	14,860	5.03	39.95	30.89
SHAWANO	172,800	28,146	28,800	15,990	1.80	14.37	11.07
LA CROSSE-CBD	396,846	29,870	58,581	12,000	4.88	40.44	30.00
LA CROSSE-WES	421,704	18,598	62,133	15,000	4.14	33.50	25.45
TOTAL	8,587,782	1,132,908	1,756,494	1,603,487	1.10	7.16	6.73

[a]Benefit cost ratio considering only one year fuel savings
[b]Benefit cost ratio considering all one year savings
[c]Benefit cost ratio considering only ten year fuel savings with 10% discount rate

The regression models for fuel savings per signal must be interpreted in view of the way in which the hardware improvement decisions were made. Initial hardware grants of $1,000/intersection were not constrained by considerations of fuel savings effectiveness. Subsequent supplemental hardware grants, however, were allocated on a competitive basis using fuel savings–to–hardware cost-effectiveness ratios. Some communities chose low-cost hardware in order to maximize their opportunity for receiving additional hardware funds. Other communities, particularly larger communities that expected high fuel savings, chose more costly hardwire interconnect and closed-loop hardware. Thus, the cost of the hardware improvements reflects the multiple objectives of communities for traffic signal hardware improvements rather than simple cost minimization.

Despite the wide latitude given to the FET communities in making their hardware improvements, all the network improvements were cost-effective using fuel savings alone over a 10-year period. When travel time and stop reduction savings are included, all the network improvements were cost-effective during the first year: the lowest first-year benefit-cost ratio was 1.78. Thus, at least for communities that need substantial hardware improvements—typically, an upgrade from partial or no interconnection to full interconnection—hardware improvements combined with signal timing optimization using TRANSYT-7F are highly effective in saving fuel and generating other benefits to motorists.

The FET program methodology should be highly effective when applied to similar communities in other states. Whereas signal timing improvements alone have been shown in other states to be effective in saving fuel, much greater fuel savings can be realized by funding the hardware improvements needed to achieve full interconnection with three-dial operation. For many smaller communities with simple linear networks, PASSER should be used for signal timing optimization instead of TRANSYT-7F. In either case substantial staff time is required to learn how to use the computer programs, to collect the required traffic count and other field data, and to apply the computer models. Providing training and technical assistance to the participating communities is essential. Direct funding for local staff time and funding for consultant support may be as important as funding hardware improvements in many communities.

ACKNOWLEDGMENTS

This research was funded in part by a grant from the Wisconsin Department of Administration, Bureau of Energy and Intergovernmental Relations. The authors are grateful for the continuing support and encouragement provided by James O'Neal. The opinions presented in this paper do not necessarily represent the views of the funding agency.

REFERENCES

1. E. D. Arnold, Jr. State Signal Timing Optimization Programs. *ITE Journal*, Feb. 1989, pp. 33–35.
2. T. H. Maze, N. Hawkins, and M. Elahi. *Iowa Motor Vehicle Fuel Reduction Program: Final Report*, Iowa State University, Ames, 1989.
3. *Fuel Efficient Traffic Signal Management Program*. Institute of Transportation Studies, University of California, Berkeley, 1984.

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

Analysis of Left-Turn-Lane Warrants at Unsignalized T-Intersections on Two-Lane Roadways

SHINYA KIKUCHI AND PARTHA CHAKROBORTY

At an unsignalized T-intersection, where a major two-lane roadway intersects a minor roadway, criteria that justify a left-turn lane on the major roadway are analyzed. Three criteria are considered: (a) probability that one or more waiting through vehicles are present on the approach; (b) delay (average delay to the "caught" through vehicles, average delay to all through vehicles, and delay savings due to the left-turn lane); and (c) degradation of the level of service. The volume combinations (through, left-turn, and opposing flow) that would justify a left-turn lane under each of the criteria are presented. The current AASHTO guidelines are based on the probability that one or more through vehicles are in the queue behind a waiting left-turn vehicle. The original mathematical formulation of the AASHTO guidelines is examined and corrected, and a new set of volume warrants is developed. A simulation model of the movements of the vehicles on the approach is simulated, and delays to through vehicles with and without a left-turn lane for different traffic volumes are computed. Finally, a set of traffic volumes at which the level of service of the approach changes from A to B is developed. The warrant volumes based on the three criteria are different. Delay and the level of service are more easily understandable measures of traffic performance than probability, so the volume combinations based on these two criteria should also be considered. The result provides a range of volume combinations within which an engineering judgment should be made. Discussions of other considerations for justification of a left-turn lane are also provided.

A rapid increase in the number of residential developments, shopping centers, and professional centers in the suburbs has added many unsignalized T-intersections on two-lane roadways. The left-turn movements made from a major roadway to a minor roadway create various negative effects on the flow of the through movements on the two-lane roadways. They include delay, reduction of capacity, accident potential, increased fuel consumption due to deceleration and acceleration, and the general annoyance associated with the possibility of delay. Many states require that developers prepare traffic impact reports that evaluate the effects of the left-turn movements on the existing through traffic and the need for a left-turn lane. This paper examines the left-turn-lane warrants practiced in different states and develops and compares different warrant criteria for installing a left-turn lane on the major roadway approach at an unsignalized T-intersection as shown in Figures 1 and 2.

The 1984 and 1990 AASHTO Green Books (*1,2*) provide a set of traffic volumes to be used as a guide when installing a left-turn lane at an unsignalized intersection. The guide is based on the probability that one or more through vehicles are present in the queues created by vehicles waiting to turn left. The guide provides the combinations of traffic volumes consisting of advancing, opposing, and left-turn percentages for the given probability.

However, it does not give a clue about the corresponding delay and delay savings with the lane, nor does it provide the level of service on the approach at the traffic volume. If delay is known, the warrant would be more meaningful to the public as well as to engineers and planners, because delay is an easily understood measure of inconvenience. Furthermore, as a comprehensive measure of the efficiency of the approach, the reduction of the level of service can be used as a criterion for justifying a left-turn lane.

Other criteria, such as hazard and energy consumption, must be taken into account in justifying a left-turn lane at an unsignalized T-intersection. Some studies, such as one by Failmezger (*3*), attempted to use empirical equations to quantify hazards caused by left-turning vehicles. Although these aspects are important, site-specific elements—such as geometric characteristics of the intersection—affect the relative importance of these factors. A more general discussion of design considerations at an unsignalized intersection is found in a study by Kimber (*4*).

PURPOSE

This study focuses on criteria that are quantitative and basic to all intersections. Its purpose is to evaluate different warrant criteria for justifying a left-turn lane, those that are currently used, and those that can be considered. First, the existing criteria used in different states are examined. Second, three different criteria are examined:

1. Probability that a queue containing one or more through vehicles is present on the approach lane,
2. Average delay experienced by all through vehicles; average delay experienced by through vehicles caught by the queue; potential delay savings with the left-turn lane, and
3. Level of service on the approach lane.

Based on a threshold value given to each of these criteria, the traffic volumes that warrant the left-turn lane are calculated and compared. Possible problems with applying any of

Department of Civil Engineering, University of Delaware, 342 Du Pont Hall, Newark, Del. 19716.

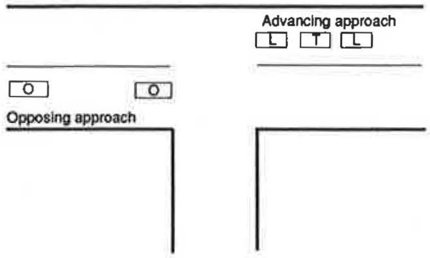


FIGURE 1 Schematic diagram of T-intersection with no left- turn lane.

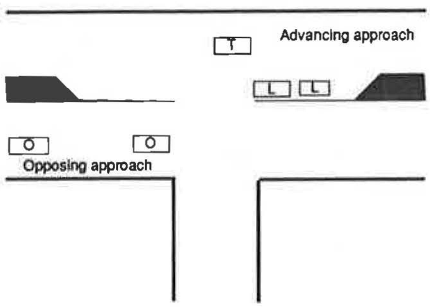


FIGURE 2 Schematic diagram of T-intersection with left-turn lane.

these criteria are discussed, and the general ranges of the traffic volume combinations for which a left-turn lane should be considered are presented.

To prepare for the analysis of the three criteria, this study (a) reviews the criteria practiced in different states, (b) reviews the current AASHTO warrants, (c) develops a simulation model that estimates delays to through vehicles, and (d) reviews a procedure for calculating the shared lane capacity on an approach to an unsignalized T-intersection (*5*).

CRITERIA FOR JUSTIFYING A LEFT-TURN LANE

The current criteria for justifying a left-turn lane are presented. They are the AASHTO guidelines and the warrants practiced by transportation departments in the United States and Canada.

AASHTO Green Book Guidelines

The 1984 and 1990 AASHTO Green Books (*1,2*) provide combinations of three traffic volumes (through, left-turn, and opposing) as a guide for installing a left-turn lane at an unsignalized T-intersection (Table 1). Sets of volume combinations for approach speeds of 40, 50, and 60 mph are given. According to Table 1 if, for example, the opposing volume is 400 vehicles per hour (vph) and the percentage of left-turn vehicles in the advancing flow is 10 percent, a left-turn lane is justified when the total advancing volume exceeds 380 vph for 40-mph approach speed. The source of the AASHTO guide is a study published by Harmelink (*6*) in 1967. The values are also adopted in an NCHRP report (*7*, p. 51). Detailed discussions of Harmelink's work are presented in the next section.

TABLE 1 AASHTO'S GUIDE FOR LEFT-TURN LANES ON TWO-LANE HIGHWAYS (*2*)

	Advancing Volume			
Opposing Volume	5% Left Turns	10% Left Turns	20% Left Turns	30% Left Turns
40-mph Operating Speed				
800	330	240	180	160
600	410	305	225	200
400	510	380	275	245
200	640	470	350	305
100	720	575	390	340
50-mph Operating Speed				
800	280	210	165	135
600	350	260	195	170
400	430	320	240	210
200	550	400	300	270
100	615	445	335	295
60-mph Operating Speed				
800	230	170	125	115
600	290	210	160	140
400	365	270	200	175
200	450	330	250	215
100	505	370	275	240

Warrants Used by Different Departments of Transportation

A survey was conducted to examine the types of criteria different states use to justify a left-turn lane. Inquiries were sent in 1989 to the transportation departments of all states in the United States and provinces in Canada. A total of 25 responses were obtained. Sixteen departments responded that they did not have a specific volume warrant for installing a left-turn lane. Accident experiences, public complaints, and engineer judgments were cited as the bases for these decisions. Among the states that use volume warrants, most cited the AASHTO criteria (Table 1). Others cited volume criteria different from AASHTO's; these include daily volume and one of the three volumes only. None of the states reported that delay, delay savings to the through vehicles, or the reduction of the level of service were used to justify a left-turn lane.

PROBABILITY-BASED MODELS

This section examines and conducts a critical review of the criterion based on the probability that through vehicles are delayed.

Discussion of AASHTO Guidelines (Harmelink's Model)

The AASHTO warrants (those proposed by Harmelink) are based on the probability that one or more through vehicles are present in queues formed by left-turning vehicles waiting for gaps in the opposing flow. The values of the maximum allowable probabilities were determined on the basis of the judgment of a panel of traffic engineers. The values of the probability are different depending on the approach speeds; they are as follows:

Approach Speed (mph)		
Design	*Operating*	*Probability*
50	40	.02
60	50	.015
70	60	.01

For each value of the probability, the combination of three volumes (opposing, left-turn, and through) that result in the value is computed assuming a queueing system.

The original queueing model is based on the following parameters:

- Advancing volume (V_A),
- Percentage of left-turn volume in the advancing volume (L),
- Opposing volume (V_O),
- Critical gap (G_c),
- Time required for a left-turning vehicle to clear itself from the advancing stream (t_e), and
- Time taken to complete a left-turn maneuver (t_1).

The queueing system as defined by Harmelink assumes that the arriving units are the through vehicles behind the vehicles waiting to turn left and that the service is the departure of the left-turning vehicles. More specifically, the arrival and service rates are defined as

$$\lambda = L \cdot (1 - L) \cdot V_A \cdot \frac{t_w + t_e}{(2/3)t_A} \tag{1}$$

and

$$\mu = \frac{\text{unblocked time/hr}}{t_1} \tag{2}$$

where λ is the arrival rate and μ is the service rate.

The equation for the arrival rate (λ) is derived based on the following:

- Each left-turning car blocks the intersection for $t_w + t_e$ sec, where t_w is the average time a left-turning vehicle waits to find a suitable gap in the opposing flow. It is given by

$$t_w = \frac{3{,}600}{V_O} \cdot \left(e^{\frac{V_O}{3{,}600} \cdot G_c} - \frac{V_O}{3{,}600} \cdot G_c - 1 \right) \tag{3}$$

- The total time the advancing approach is blocked by left-turning vehicles is

$$T_B = (L * V_A)(t_w + t_e) \tag{4}$$

- The number of advancing cars that arrive during the time period T_B is

$$C_A = \frac{(L \cdot V_A)(t_w + t_e)}{(2/3)t_A} \tag{5}$$

where $(2/3)t_A$ is the median headway of the advancing stream.

- Out of these C_A advancing cars, the number of through vehicles is

$$(1 - L) \cdot C_A = (1 - L) \cdot \frac{L \cdot V_A(t_w + t_e)}{(2/3)t_A} \tag{6}$$

The expression of the service rate (μ) is derived on the basis of the following:

- The unblocked time in Equation 2 is the total amount of time during which left turns can be made. This is equivalent to the sum of headways greater than G_c in the opposing flow minus an adjustment factor.
- Therefore, the number of left turns that can be made per hour is derived by dividing the unblocked time by t_1 as seen in Equation 2.

Given λ and μ in the queueing system, the probability of k units in the system is derived by

$$P(k) = \left(\frac{\lambda}{\mu}\right)^k \cdot \left(1 - \frac{\lambda}{\mu}\right) \tag{7}$$

From Equation 7, $1 - P(0)$ represents the probability that one or more units are in the system. The criterion for installing a left-turn lane is based on the probability that one or more units in the system will be less than a given value α. Therefore,

$$1 - P(0) = \frac{\lambda}{\mu} \leq \alpha \tag{8}$$

where the value of α is the preset probability defined in Table 2. The probability can be restated as the proportion of the time during which through vehicles are present in the queueing system or the probability that a through vehicle is delayed due to the left-turn vehicles.

Critical Evaluation and Limitation of Harmelink's Model

There are two problems in Harmelink's formulation. They are (a) inconsistent definitions of λ and μ, and (b) incorrect representation of the total number of possibilities of making a left turn in μ.

Problem 1

In Harmelink's model, λ denotes the arrival rate of through vehicles while one or more left-turning vehicles are waiting and μ denotes the rate at which vehicles can make left turns per unit of time. In queueing theory, the arrival rate and the service rate must refer to the same units in the system. In this case, λ refers to the through vehicles, but μ does not refer to the discharge rate of the through vehicles. This apparent inconsistency can be explained with the help of an example. Suppose that in 1 hr there are 10 possibilities of making a left turn, and assume that there are 10 left-turning vehicles. Assume also that every time a left-turning vehicle is waiting, three through vehicles arrive. Then λ, which counts each of the through vehicles as separate units, would take on the value

TABLE 2 VOLUME COMBINATIONS JUSTIFYING A LEFT-TURN LANE ON BASIS OF MODIFIED HARMELINK'S MODEL

	Advancing Volume			
Opposing Volume	5% Left Turns	10% Left Turns	20% Left Turns	30% Left Turns
40-mph Operating Speed				
800	434	300	219	189
600	542	375	272	234
400	682	472	343	293
200	863	600	435	375
100	946	679	493	424
50-mph Operating Speed				
800	366	257	185	162
600	460	320	234	202
400	577	403	294	255
200	735	513	373	324
100	830	576	424	365
60-mph Operating Speed				
800	294	207	154	146
600	365	259	187	165
400	461	324	238	206
200	586	414	303	263
100	663	468	344	297

of 30, whereas μ would represent the discharge rate of the leading left-turning vehicle and be equal to 10. This means that the system would never reach a steady state, because λ is greater than μ. But this is not correct, because every time a left-turning vehicle is discharged, more than one through vehicle can be discharged.

The inconsistency in the definitions of λ and μ in Equations 1 and 2 ceases to be critical only when there is no more than one through vehicle waiting behind a left-turning vehicle. Under this condition, the number of opportunities for turning left equals the discharge rate of through vehicles, and thus the units represented in λ and μ become consistent. Alternatively, the definitions are consistent only when the probability that two or more through vehicles waiting behind a left-turning vehicle is very small. This condition could occur when the proportion of through vehicles in the advancing stream is very small. Under such conditions, however, the question of installing a left-turn lane is not relevant, because the approach is essentially used as the left-turn lane.

Problem 2

The value of μ represents the total number of possibilities (per hour) of making left turns based on the available gaps in the opposing flow. It is an aggregated value in Harmelink's model; in other words, μ is derived by dividing the sum of gaps that are greater than the critical gap by the time required to make a left turn (t_1). A problem in this derivation is that the residual gaps (the remainder of individual gap size divided by t_1) are added and the sum is also considered to be part of the time available for making left turns. This would make μ represent more left-turn opportunities than are actually available. For example, if there were four consecutive 6-sec gaps, the value of μ would be 6: $4 \times 6 \div t_1$, where t_1 equals 4 sec. In reality, however, there are only four left-turning possibilities because each 6-sec gap can accommodate one left turn, assuming t_1 equals 4.

Thus, even when λ and μ are consistent, as pointed out, the value of λ, derived from Equation 8, is overestimated because of this definition of μ.

Modified Formulation of Harmelink's Model

In this section Harmelink's model is modified so that the definitions of λ and μ are consistent and μ represents real-world conditions more closely. One left-turning vehicle followed by one or more through vehicles is considered as an arriving unit. The modified arrival rate of units (λ^*) is

$$\lambda^* = L \cdot V_A \left[1 - e^{\frac{-(1-L)V_A}{3{,}600}(t_w + t_e)} \right] \tag{9}$$

where the term in brackets represents the probability of one or more through vehicles' arriving behind a waiting left-turning vehicle.

The corresponding service rate (μ^*) should then be the total number of left-turning possibilities. It is assumed that in a headway between $[G_c + (\eta - 1) \cdot G_s]$ and $(G_c + \eta \cdot G_s)$, η left turns are possible, where G_s is the follow-up gap size, assumed to be 3 sec. This assumption is based on a suggestion made by Baass in his 1987 paper (*8*). Therefore, μ^* can be expressed as

$$\mu^* = \left[1 - e^{-3\left(\frac{V_O}{3{,}600}\right)} \right] \cdot V_O \cdot \sum_{\eta=1}^{N} \eta \left\{ e^{-\frac{V_O}{3{,}600}[G_c + 3(\eta - 1)]} \right\} \tag{10}$$

where the value of N is the maximum number of left-turning opportunities per single headway. It is approximated by solving the following for N:

$$\text{Probability } \{\text{headway} \geq G_c + N \cdot G_s\} \approx 0 \tag{11}$$

Based on the modified arrival rate (λ^*) and service rate (μ^*) and the threshold probability shown earlier, the volume combinations that warrant a left-turn lane are computed using Equation 8. The results are presented in Table 2, in the same format as the current AASHTO guide (Table 1).

Tables 1 and 2 provide warrant volumes based purely on probability and do not provide a reference to the delays experienced by through vehicles.

DELAY-BASED MODELS

In this section, expressions are derived that compute delays to the through vehicles under different volume combinations. Savings in time accrued by providing a separate left-turn lane are also computed. To compute the delays, a simulation model is developed. The approach is to build the simulation model and, from many runs of the model, develop a set of regression equations that expresses delay as a function of the volume combination. The simulation model, its validation, and the values of delays are presented in the following.

Simulation Model and its Validation

Before the development of the simulation model, TRAF-NETSIM was tested to determine whether it could be used to derive delay for this problem. However, the TRAF-NETSIM model did not provide a reasonable and consistent set of delay values. It is believed that TRAF-NETSIM may not be suited for computing delay to the vehicles in the advancing stream of an isolated, unsignalized T-intersection on a two-lane roadway (as shown in Figure 1). Because the problem with TRAF-NETSIM could not be resolved, a separate simulation model was developed. The assumptions of the model are as follows:

1. The arrivals of all three types of vehicles follow the Poisson distribution.
2. The basic time unit for simulation is 1 sec. For each time unit the arrival of vehicles is checked according to the Bernoulli experiment.
3. The acceptable gap in the opposing flow for making a left turn is 6 sec.
4. A waiting left-turning vehicle initiates the turning movement only when the first acceptable gap becomes available. The time at which the vehicle initiates the movement is referred to as the "departure time."
5. If a gap is long, more than one left-turning vehicle can use the same gap. In this case, the next left-turning vehicle in the queue makes the turn 1 sec after the preceding left-turning vehicle departs.
6. The difference in departure time between a left-turning vehicle and the succeeding through vehicle, which is in the queue, is 1 sec.
7. The difference between the departure times of consecutive through vehicles in the queue is 1 sec.
8. The delay to a left-turn or through vehicle is measured by the time difference between the time of arrival at the intersection and the time of departure from the intersection.
9. The time loss to the through vehicles is computed based on a linear acceleration and deceleration pattern (all vehicles are assumed to travel at a designated speed before and after the delay). The rates of acceleration and deceleration used are 5.4 ft/sec^2 and 8 ft/sec^2, respectively.

The output of the simulation model includes the following:

- Total hourly delay (TD),
- Average delay to left-turning vehicles (ALTD),
- Average delay to through vehicles caught in the queue (ACTHD),
- Average delay to all through vehicles, caught and not caught (ATD),
- Total delay savings per hour as a result of the left-turn lane (DS), and
- Distributions of queue lengths, times of queue dissipation, and frequency of queue formation.

The performance of the model was verified by checking the values of selected parameters. For those parameters, the values computed from previously developed equations are compared with the values obtained from the simulation model. The selected parameters are ALTD and NTVC, which is the number of through vehicles caught behind waiting left-turn vehicles per hour.

For NTVC, the data from the model were compared with the value of λ in Equation 1, and the comparison is shown in Figure 3. In this figure, NTVC and λ are calculated for different volume combinations and compared. If the values of NTVC and λ were an ideal match, the plot would be a 45-degree line through the origin.

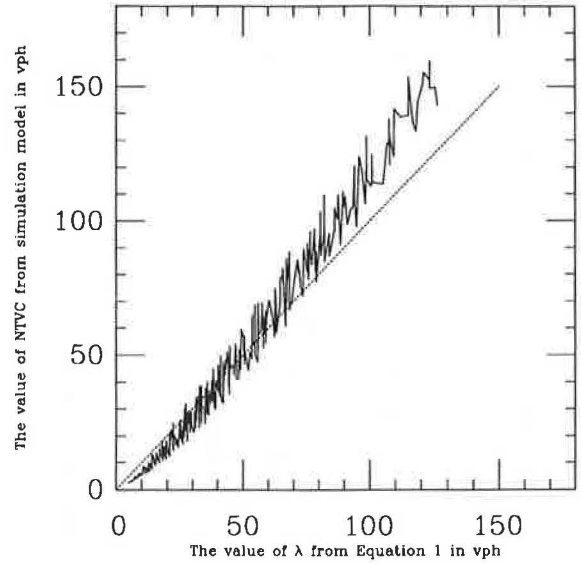


FIGURE 3 Comparison of λ (Equation 1) and NTVC (simulated result).

For ALTD, the data from the model were compared with the results from the expression of vehicle waiting time at the merge point of two traffic streams, Equation 3. This equation has been presented by many studies, among them Drew (*9*) and Tanner (*10*). The comparison is shown in Figure 4. In this figure, t_w from Equation 3 and ALTD are plotted against the opposing volume (V_O). The value of ALTD presented in this figure corresponds to the situation in which the proportion of through vehicles in the advancing approach is zero in the simulation model.

Development of Equations on Delay and Delay Savings

For the same set of traffic volumes, the simulation model was run many times to attain the average value of delay. A regression analysis was conducted to develop the general relationships between the volume combinations and TD, ACTHD, ATD, and DS.

The regression equations express these delays in terms of the three input volumes (opposing, left-turn, and through). Total delay (TD) refers to the sum of all delays faced by vehicles in the advancing stream, and it is expressed in seconds per hour. The average delay to the through vehicles caught in the queue (ACTHD) refers to the average time the through vehicles must wait in the queue; it is expressed in seconds per vehicle. The average delay to the left-turn vehicles (ALTD) is the expected delay to any left-turning vehicles, including those that do not have to wait for a gap in the opposing stream and hence suffer no delay. It too is expressed in seconds per vehicle. The average delay to the through vehicles (ATD) is the average time the through vehicles spend in the system, irrespective of their being caught in the queue. It is measured in seconds per vehicle. Delay savings (DS) refers to the delay that would be eliminated by providing a left-turn lane; it is

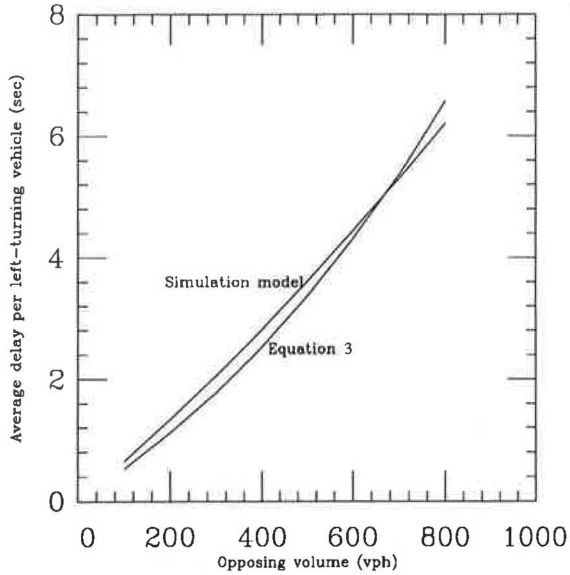


FIGURE 4 Comparison of t_w (Equation 3) and ALTD (simulated result).

expressed in seconds per hour. All of these delays account for the time loss due to deceleration and acceleration.

In formulating each equation, the influencing factors for the delay are identified and arranged in a polynomial form. Once the basic forms of the equation are determined, regression analyses are conducted to determine the coefficients of the equations. The regression equations and the R^2 values obtained are

$$TD = 0.087 \cdot (V_O/100)^2 \cdot (V_l/10)^2 \cdot (V_t/100) + 3.147 \cdot (V_O/100)^2 \cdot (V_l/10) + T_l \cdot \text{PTHC}\, V_t \qquad R^2 = .88 \tag{12}$$

$$ACTHD = 0.016 \cdot (V_O/100)^2 \cdot (V_l/10) + 1.39 \cdot (V_O/100) + T_l \qquad R^2 = .9 \tag{13}$$

$$ALTD = 22.86 \cdot [(V_O/3{,}600) + (V_O/3{,}600)^2] + 222.65 \cdot PT \cdot (V_O/3{,}600)^2 \qquad R^2 = .83 \tag{14}$$

where

V_l = left-turn volume,
V_t = through movement volume, and
T_l = time loss due to deceleration and acceleration.

Once these equations are developed, ATD and DS are developed as follows:

$$ATD = ACTHD \cdot PTHC \tag{15}$$

where PTHC is the proportion of the through vehicles caught in the queue, and its value is again computed from a regression equation of the form

$$PTHC = 14.19 \cdot 10^{-10} V_O \cdot V_l \cdot V_t + 46.511 \cdot V_O^2 \cdot V_l \tag{16}$$

DS consists of two elements:

1. The total delay to the through vehicles. This is given as ACTHD · NTVC, where NTVC = PTHC · V_t, and
2. The delay experienced by the left-turning vehicles because of missed gaps. This is due to the presence of through vehicles in the queue. ALTD, given by Equation 14, consists of two terms. The first term represents the delay due to waiting for acceptable gaps in the opposing stream. For this reason it is dependent only on the opposing volume. The second term represents the delay caused by the presence of through vehicles in the queue. Hence, this term depends on the proportion of through vehicles in the advancing stream. This latter delay will be saved when a separate left-turn lane is provided.

Thus, the delay savings is expressed as

$$DS = ACTHD \cdot NTVC + 222.365 \cdot (1 - L) \cdot (V_O/3{,}600)^2 \cdot V_l \tag{17}$$

Delays and Delay Savings at Warrant Volumes Based on Probability Models

By using the regression equations derived in the previous subsection, delays and delay savings for volume combinations of the AASHTO guidelines and of the modified Harmelink's model are now computed. For each volume combination ACTHD, ATD, and DS are computed and are shown in Tables 3 and 4.

Delays at AASHTO Guidelines

Table 3 shows the ACTH, ATD, and DS for each volume combination shown in Table 1. ACTHD ranges from 10 to 28 sec, ATD from less than 0.1 to 4 sec, and DS from 22 sec/hr to nearly 500 sec/hr. The existence of these variations indicates that the installation of a left-turn lane based on the AASHTO warrant volumes would not result in the consistent reduction of delay. It is particularly interesting to see that the total delay savings vary more than 20 times for the same threshold probability (α).

Delays at Warrant Volume Combinations Based on Modified Harmelink's Model

Table 4 shows ACTHD, ATD, and DS for the volume combinations shown in Table 2, which is derived after modifying the original formulation of Harmelink's model. When comparing Tables 3 and 4, the values of delay are higher in Table 4 than in Table 3. This reflects the fact that the warrant volume conditions based on the modified formulation are more re-

TABLE 3 DELAYS AT VOLUME COMBINATIONS OF AASHTO'S GUIDE FOR LEFT-TURN LANES

Opposing Volume	Advancing Volume: 5% Left Turns	10% Left Turns	20% Left Turns	30% Left Turns
40-mph Operating Speed				
800	330 (22/1.2/550)	240 (22.6/1.8/615)	180 (23.8/2.7/704)	160 (25/3.7/787)
600	410 (18.6/0.8/418)	305 (19/1.1/475)	225 (20/1.6/518)	200 (21/2.2/573)
400	510 (15.3/0.4/259)	380 (15.6/0.6/284)	275 (16/0.8/289)	245 (16.5/1.0/316)
200	640 (12/0.1/105)	470 (12.1/0.2/103)	350 (12.3/0.2/102)	305 (12.4/0.3/104)
100	720 (10.5/0/43)	575 (10.5/0.1/46)	390 (10.5/0.1/34)	340 (10.6/0.1/33)
50-mph Operating Speed				
800	280 (24/1.1/439)	210 (25/1.7/519)	165 (26/2.7/642)	135 (27/3.3/626)
600	350 (21/0.7/339)	260 (21.2/1.0/386)	195 (22/1.6/434)	170 (22.6/2.0/464)
400	430 (17.5/0.4/206)	320 (17.7/0.5/227)	240 (18.1/0.7/248)	210 (18.5/1.0/263)
200	550 (14.3/0.1/87)	400 (14.4/0.2/84.4)	300 (14.5/0.2/86)	270 (14.6/0.3/93)
100	615 (12.8/0/35)	445 (12.8/0/30)	335 (12.8/0.1/29)	295 (13/0.1/28)
60-mph Operating Speed				
800	230 (26/1.0/331)	170 (26.5/1.4/385.5)	125 (27.4/2.1/431)	115 (28.3/3.0/506)
600	290 (23/0.6/260)	210 (23.2/0.9/286)	160 (24/1.4/332.5)	140 (24.4/1.8/358)
400	365 (19.7/0.3/165)	270 (20/0.5/182)	200 (20.3/0.7/196)	175 (20.6/0.9/209)
200	450 (16.6/0.1/64)	330 (16.7/0.1/65)	250 (16.8/0.2/68.2)	215 (16.9/0.2/68.5)
100	505 (15.1/0/25)	370 (15.1/0/23.2)	275 (15.2/0.1/22)	240 (15.2/0.1/22)

Notes: Delays (a/b/c)
a: Average delay per through vehicle caught in the queue (in sec/veh)
b: Average delay per through vehicle (in sec/veh)
c: Total delay savings with a left-turn lane (in sec/hour)

TABLE 4 DELAYS AT VOLUME COMBINATIONS OF MODIFIED HARMELINK'S MODEL

Opposing Volume	Advancing Volume: 5% Left Turns	10% Left Turns	20% Left Turns	30% Left Turns
40-mph Operating Speed				
800	434 (22.5/1.7/922)	300 (23.4/2.3/922)	219 (24.8/3.4/988)	189 (26.1/4.6/1047)
600	542 (19.1/1.1/724)	375 (19.7/1.5/697)	272 (20.7/2.1/725)	234 (21.6/2.7/892)
400	682 (15.6/0.6/472)	472 (15.9/0.7/432)	343 (16.5/1.0/430)	293 (17.0/1.3/434)
200	863 (12.2/0.2/209)	600 (12.3/0.3/173)	435 (12.5/0.3/155)	375 (12.6/0.4/152)
100	946 (10.6/0.1/86.1)	679 (10.6/0.1/69)	493 (10.7/0.1/56)	424 (10.7/0.1/51)
50-mph Operating Speed				
800	366 (24.4/1.5/716)	257 (25.2/2.1/740)	185 (26.3/3.1/779)	162 (27.5/4.2/847)
600	460 (21.1/1.0/568)	320 (21.6/1.3/561)	234 (22.4/1.9/592)	202 (23.2/2.5/618)
400	577 (17.7/0.5/371)	403 (18/0.7/350)	294 (18.5/0.9/355)	255 (19/1.2/365)
200	735 (14.4/0.2/166)	513 (14.5/0.2/141)	373 (14.7/0.3/129)	324 (14.8/0.4/128)
100	830 (12.9/0.1/73)	576 (12.9/0.1/55.1)	424 (12.9/0.1/47)	365 (13.0/0.1/43)
60-mph Operating Speed				
800	294 (26.3/1.3/508)	207 (26.9/1.8/537)	154 (28/2.7/602)	146 (29/3.7/677)
600	365 (23.1/0.8/395)	259 (23.5/1.1/409)	187 (24.2/1.6/428)	165 (24.9/2.2/467)
400	461 (19.8/0.5/259)	324 (20.1/0.6/253)	238 (20.5/0.8/263)	206 (20.8/1.1/272)
200	586 (16.7/0.2/113)	414 (16.7/0.2/101)	303 (16.9/0.3/96)	263 (17/0.3/97)
100	663 (15.1/0.1/48.3)	468 (15.1/0.1/39)	344 (15.2/0.1/34)	297 (15.2/0.1/32)

Notes: Delays (a/b/c)
a: Average delay per through vehicle caught in the queue (in sec/veh)
b: Average delay per through vehicle (in sec/veh)
c: Total delay savings with a left-turn lane (in sec/hour)

laxed. Although delays are higher in Table 4, the ranges of individual delays and delay savings are similar between the two cases (i.e., AASHTO and modified Harmelink). Discussions on the variation in the values of delays are presented in more detail later.

Delay as Warrant Criterion

It is now attempted to establish a set of volumes that can be considered as warrants on the basis of a given value of delay. A given value of ACTHD, ATD, or DS can be selected, and the regression equations 12, 14, or 16, respectively, can be used to compute the volume combinations for the selected value of the parameters. Shown in Figure 5, as an example, are the volume combinations that would result in 14, 19, 24, and 29 sec of ACTHD. These delay values include the acceleration and deceleration time loss, which for a 40-mph approach speed is 9 sec. If for a given value of ACTH the volume combination points to the upper right of the line in Figure 5, a left-turn lane is justified; if it points to the lower left of the line, the lane is not justified.

LEVEL OF SERVICE AS WARRANT CRITERION

In this section, the reduction of level of service (LOS) from A to B on the advancing approach is considered as the criterion for justifying a left-turn lane. On the basis of the transition from LOS A to LOS B, a set of volumes (V_O, V_A, L) are computed and presented as warrant conditions.

The procedure to determine the level of service of an approach on the major road where through and left-turn movements share the same lane is not clearly explained in the *Highway Capacity Manual* (5). The shared lane capacity of an approach lane at an unsignalized intersection (C_{SH}) is defined in the manual as

$$C_{SH} = \frac{V_l + V_t + V_r}{\frac{V_l}{c_{ml}} + \frac{V_t}{c_{mt}} + \frac{V_r}{c_{mr}}} \tag{18}$$

where V_l, V_t, and V_r are left-turn, through, and right-turn flow rates, and c_{ml}, c_{mt}, and c_{mr} are the movement capacities for left-turn, through, and right-turn flows. (in this case, $V_r = 0$).

It is not clear if c_{mt} can represent the capacity of the advancing through movements that have no conflicting flow but are affected by the presence of the left-turning movement. In this analysis, it is assumed that c_{mt} represents the capacity of the through movement. Assuming V_r equals 0, Equation 18 can be written as

$$C_{SH} = \frac{1}{\frac{L}{c_{ml}} + \frac{(1 - L)}{c_{mt}}} \tag{19}$$

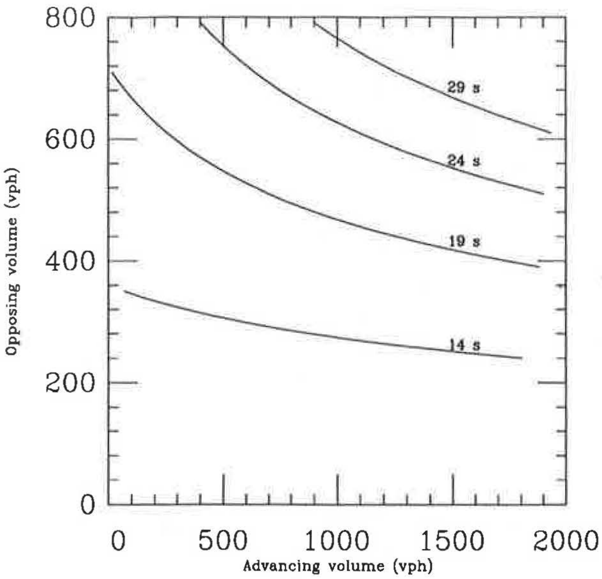


FIGURE 5 Volume combinations (V_O and V_A) at different values of ACTHD for L = 10 percent.

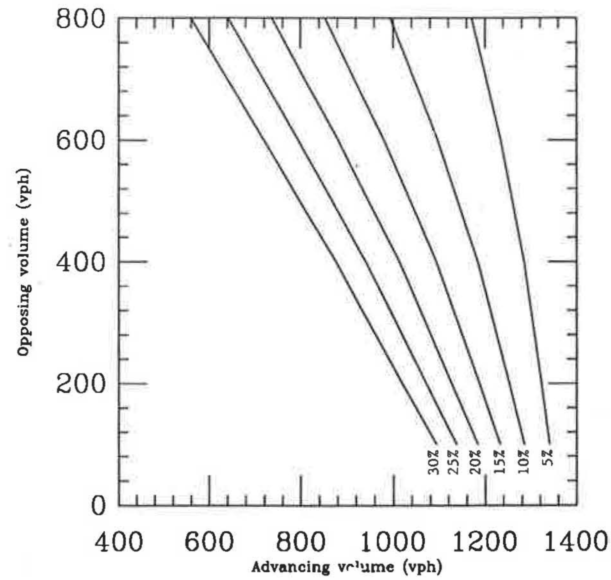


FIGURE 6 Volume combinations (V_O and V_A) at boundaries of LOS A and LOS B for different values of L.

In this equation, C_{SH} is computed as the weighted average of the minimum headways of the through and left-turn movements. It is, however, very difficult to decide what value of c_{ml} should be used. It would depend on many factors and, as such, can take on a wide range of values. This study presents the warrant conditions for c_{ml} = 1,800 vph. This value is chosen because it is the maximum capacity that can be attained on such a roadway.

For LOS A, the minimum reserve capacity is 400 vph. Thus, the combination of three volumes that results in a reserve capacity 400 vph can be computed from

$$C_{SH} - V_A = 400 \tag{20}$$

The volume combinations that satisfy Equation 20 are plotted in Figure 6. As the percentage of left-turn volume increases, the effect of opposing volume on the level of service becomes more pronounced. This is not surprising: with the increase of L, the effect of c_{ml} on C_{SH} becomes greater, so, as mentioned earlier, C_{SH} becomes more strongly dependent on the opposing volume. In the figure, if the volume combination (V_O, V_A, L) points to the upper right of the line (corresponding to L), then a left-turn lane should be provided; if the combination points to the lower left of the line, the left-turn lane is not required.

It should be noted that peak hour factor (PHF) can be included in this analysis by dividing the hourly volumes by the PHF and then using these values as V_O and V_A.

DISCUSSION OF CRITERIA FOR LEFT-TURN-LANE WARRANTS

The characteristics and problems are discussed of using each of the three criteria for justifying a left-turn lane based on (a) a given probability that waiting through vehicles are present on the approach, (b) a given value of delay, and (c) the reduction of the level of service from A to B.

Comparison of Probability- and Delay-Based Criteria

The probability-based criterion does not take into account how long individual vehicles must wait. For the same value of probability depending on the combination of V_O, V_A, and L, the delay to the through vehicles can vary significantly. This can be seen from the delay values (ACTHD, ATD, and DS) calculated at the warrant volume based on the probability and presented in Tables 3 and 4. As seen in the tables, the values of ACTHD, ATD, and DS have large variations for the same probability of .02 (for the 40-mph approach speed). For the modified Harmelink's model, for example, at an approach speed of 40 mph, ACTHD varies for 10 to 25 sec, ATD from 0.1 to 5 sec, and DS from 51 to 1,050 sec/hr. The wide variation in the DS, in particular, suggests that if the probability-based warrant were applied, the economic justification for installing a left-turn lane would not be consistent for different volume combinations.

Comparison of Probability-Based Criterion and Level of Service–Based Criterion

When the reduction of the level of service from A to B is used as a criterion, the values of volume combination at which a left-turn lane is justified are much greater than the ones for the probability-based criterion. For example, as seen in Table 2, at V_O equals 600 vph and L equals 10 percent, V_A is 375 under the probability criterion of .02; under the level of service-based criterion, V_A is 1,100, as seen in Figure 6. A possible explanation for this discrepancy is that the level of service is a macroscopic analysis, considering the average condition during 1 hr, whereas the probability-based criterion is a more microscopic analysis of flow characteristics. The volume combination that corresponds to the level of service criterion should be considered as the minimum limit.

Comparison of Three Criteria

To compare the volume combinations for the three criteria, Figure 7 is provided. It shows the volume combinations when the probability is .02; ACTHD is 19 sec and level of service changes from A to B at c_{mt} = 1,800 vph; in all cases the percentage of left turns in the advancing flow (L) is 10 percent.

The volume combinations developed on the basis of these three criteria provide a range in which a left-turn lane can be considered under the threshold values stated above. The volume combination based on the level of service is perhaps the minimum acceptable criterion, and the volume combination based on the probability (as seen in AASHTO or the modified Harmelink's model) is the most luxurious criterion; in other words, the latter is an ideal criterion. The volume combinations based on the delay criterion fall between the volume combinations for the other two criteria. This is applicable when the advancing volume is between 500 and 1,250 vph.

Justifying a left-turn lane on the basis of a given probability is difficult to comprehend. Justification based on delay is easier to understand; however, depending on which criterion is used (the average delay or the total delay savings), the volume combination that justifies the left-turn lane will be significantly different. Justification based on the degradation of the level of service can also be a reasonable concept that the public understands.

The volume combinations based on these three criteria should provide the general volume range for which the left-turn lane should be considered. The precise limits should vary based on the standards of the community and other factors, such as the accident experience and the number of buses included in the through vehicles. The delay experienced by the persons rather than the vehicles involved should also be an important consideration. Thus, if the percentage of transit vehicles is large, the more stringent considerations should be used.

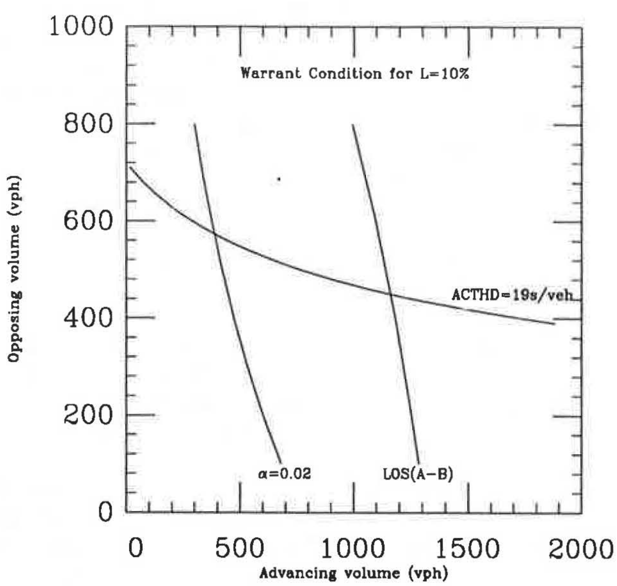


FIGURE 7 Volume combinations for justification of left-turn lane for the three criteria.

CONCLUSIONS

This paper has examined the criteria that should be considered when justifying a left-turn lane on the major approach of an unsignalized T-intersection. They are (a) probability that one or more waiting through vehicles exist on the approach, (b) delay to the through vehicles and delay savings as a result of the left-turn lane, and (c) degradation of the level of service. For each case, combinations of three volumes (opposing, left-turn, and through movements) that result in a given condition are computed and presented. During the process of developing the volume combination, the mathematical model on which the existing AASHTO guidelines are based is reviewed, and modifications to the model are made. Furthermore, a set of regression equations is developed that represents delay to the through vehicles and delay savings. A computation procedure for the level of service on a shared lane approach to an unsignalized T-intersection is examined.

The problem of the left-turn-lane justification will continue to be a matter of engineering judgment; however, this study should help the decision-making process. In addition to the volume warrant, particular attention should be paid to (a) an appropriate value of the threshold values for probability and delay, (b) delay based on the number of passengers in vehicles, in the case of large percentages of transit vehicles among the through vehicles, (c) the length of time for which the warrant conditions exist, and (d) environmental and energy issues.

ACKOWLEDGMENTS

The authors are grateful to the Delaware Department of Transportation for its assistance in collecting survey responses. The authors are also thankful to Mohammad Dzulkifli for his help with the development of the simulation model.

REFERENCES

1. *A Policy on Geometric Design of Highways and Streets*. AASHTO, Washington, D.C., 1984.
2. *A Policy on Geometric Design of Highways and Streets*. AASHTO, Washington, D.C., 1990.
3. R. W. Failmezger. Relative Warrant for Left-Turn Refuge Construction. *Traffic Engineering*, April 1963, pp. 18–20.
4. R. M. Kimber. The Design of Unsignalized Intersections in the U.K. *Proc., International Workshop on Intersections Without Traffic Signals*, Bochum, Germany, 1988, pp. 111–153.
5. *Special Report 209: Highway Capacity Manual*. TRB, National Research Council, Washington, D.C., 1985.
6. M. D. Harmelink. Volume Warrants for Left-Turn Storage Lanes at Unsignalized Grade Intersections. In *Highway Research Record 211*, HRB, National Research Council, Washington, D.C., 1967, pp. 1–18.
7. T. R. Neuman. *NCHRP Report 279: Intersection Channelization Design Guide*. TRB, National Research Council, Washington, D.C., 1985.
8. K. G. Baass. The Potential Capacity for Unsignalized Intersections. *ITE Journal*, Oct. 1987, pp. 43–46.
9. D. R. Drew. *Traffic Flow Theory and Control*. McGraw-Hill, New York, N.Y., 1968.
10. J. C. Tanner. The Delay to Pedestrians Crossing a Road. *Biometrika*, Vol. 38, 1953, pp. 383–392.

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